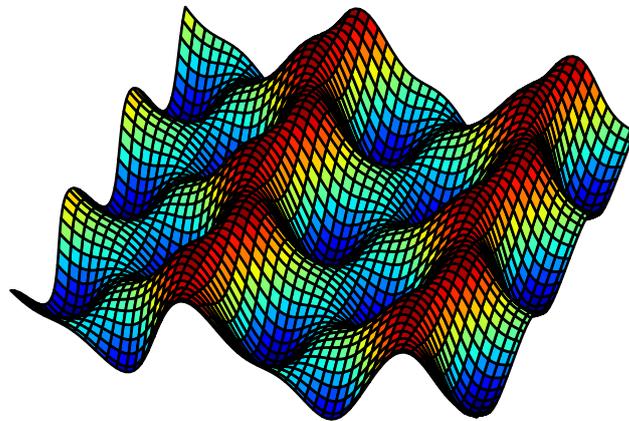


Cold p -band atoms in the zig-zag optical lattice:
implementing a quantum simulator of next-nearest
neighbor spin chains



Thesis submitted for the degree of Master of Science

Author: Axel Gagge

Supervisor: Jonas Larson



Stockholm University

March 1, 2016

Abstrakt

Optiska gitter är en typ av kvantmekaniskt laboratoriesystem som under de senaste tjugo åren spelat en central roll inom kvantoptik och den kondenserade materiens fysik. Atomer som kyls till temperaturer nära den absoluta nollpunkten kan hållas fångna med elektriska och magnetiska fält. Med laserstrålar är det sedan möjligt att placera atomerna i en periodisk elektromagnetisk potential. Sådana kvantmekaniska mångpartikelsystem delar många egenskaper med fasta tillstånd, exempelvis metaller. Flexibiliteten hos optiska gitter gör att de är lämpliga som *kvantsimulatorer*: modellsystem som kan efterlikna egenskaperna hos andra kvantmekaniska system men är lättare att mäta och manipulera.

Denna avhandling berör en speciell typ av gitter: sick-sack-potentialen, vilken kan sägas ha en serie minima i ett diagonalt mönster och tillåter atomer att tunnla såväl diagonalt (mellan närmsta minima) som horisontellt (mellan näst närmsta minima). Det fullständiga problemet formuleras i termer av en andrakvantiserad Hamiltonian av Hubbardtyp, där kopplingskonstanterna ges som överlappsintegraler av Wannierfunktioner i gittret. En exakt lösning av problemet presenteras för det fall att interaktionen kan försummas. Genom en kvantmekanisk medelfältteori undersöks ordningen i den superflytande fasen, som visar sig ha en längre period än det underliggande gittret. Periodiciteten är möjlig att variera genom att manipulera graden av horisontell tunnling. I Mott-fasen används störningsräkning för att ta fram en effektiv Hamiltonian och genom en spin-boson-mappning omvandlas problemet till en ekvivalent beskrivning i termer av spinn-1/2 partiklar. Medelfältteori med spin-koherenta tillstånd visar att denna spinn-modell innehåller en polariserad fas, en fas som är anti-ferromagnetisk i y -komponenten av ordningsparametern samt en som uppvisar en fördubblad periodicitet. Avhandlingen ger också den nödvändiga bakgrunden i kvantfältteori, det fasta tillståndets fysik, fasövergångar och optiska gitter.

Aknowledgements

First, I want to thank my supervisor Jonas for the trust, support, inspiring discussions and proof-reading which made this thesis possible.

I would like to thank Pil for a crucial proof-reading and discussion of my thesis, and the whole KomKo group at Stockholm University, for the warm welcome and the stimulating discussions. I want to thank Fernanda for figure 7.5.

I want to mention my fellow students and friends Douglas, Mihae and Johan. Our discussions were crucial for me in learning to become a somewhat more systematic person.

I want to thank my friends and loved ones - involved in far more important things than physics - for constantly shaking me out of any pattern. Without you, I would have been a lesser person and a lesser student. I especially want to thank C, J, K, my mother, sister and father for their constant love and support.

Contents

1	Introduction	6
2	Quantum field theory	9
2.1	Motivation for quantum field theory	9
2.2	Second quantization	10
2.3	Operators in second quantization	12
2.4	Normal ordering	12
2.5	Mean-field approximations	13
3	Many-body physics	17
3.1	The solid state	17
3.1.1	Models of the solid state	17
3.1.2	Crystal lattices	19
3.1.3	Bloch functions	20
3.1.4	Wannier functions	21
3.1.5	Hamiltonians in the Wannier basis	24
3.2	Classical and quantum phase transitions	25
3.3	The Bose-Einstein Condensate	28
3.4	The Bose-Hubbard Model	29
4	Optical lattices	33
4.1	Trapping	33
4.2	Cooling	34
4.3	Relevant scales	35
4.4	Measurement and manipulation	36
5	From potentials to overlap integrals	40
5.1	The square lattice	40
5.1.1	Wannier functions in the square lattice	40
5.1.2	The harmonic oscillator approximation	41
5.1.3	Analytic overlap integral	44
5.1.4	Overlaps in the quadratic lattice	47
5.2	The zig-zag lattice	48
5.2.1	Constructing the zig-zag lattice	48
5.2.2	Wannier states in the zig-zag lattice	50
5.2.3	Overlap integrals	51
5.2.4	The second-quantized Hamiltonian	54
6	Exact diagonalization	56

7	Mean-field approximations	61
7.1	Toy model of next-nearest neighbor tunneling	62
7.2	Mean-field treatment of the quantum Ising model	66
7.3	Mean-field derivation of vortex-antivortex structure	66
7.4	Mean-field approximation of the zig-zag chain	68
8	The Mott insulator phase	71
8.1	Perturbation theory	71
8.2	The effective Hamiltonian	73
8.3	Schwinger spin-boson mapping	75
8.4	The spin Hamiltonian	77
8.5	Mean-field approximation of the spin model	80
9	Conclusions	85

Chapter 1

Introduction

As long as we stick to things and words we can believe that we are speaking of what we see, that we see what we are speaking of, and that the two are linked.

Gilles Deleuze

In most physical problems, and even in our daily lives, we are confronted with complicated things which behave in simple ways. Gases, networks or solids are complicated in detail, but often follow simple laws in some or most aspects. It is remarkable in itself that it is often possible to make such simplified models without knowing the microscopic details of why they work. The reason is that most correlations between constituents “average out” at a large scale, so that only some microscopic effects really matter. As an example: ordinary objects around us contain more than 10^{23} particles, but most of the measurable properties at our scale in length, time and energy can be computed from a few numbers such as temperature, heat capacity and electrical conductivity. This averaging behavior and the macroscopic laws it leads to is the reason why thermodynamics is so powerful [J.11, p. 113, p. 116].

To really understand *how* the macroscopic average laws arise, however, is the real challenge. This is the subject of statistical mechanics, which relates microscopic events to macroscopic. The reality is quantum mechanical, but the quantum mechanical nature of the parts does not always influence the state of the whole system. In such cases, it may be possible to describe it by a semi-classical model. But in some cases - when interactions are strong and thermal fluctuations small - the quantum mechanics matter for global properties. This is the case for most solids and some exotic liquids and gases. Global properties such as conductivity are accessible in solids, but it is complicated to measure the state of single atoms due to the small lattice spacing and strong interactions.

Such truly quantum mechanical systems may be impossible to even simulate with a computer. The reason is that quantum mechanics requires us to account for any possible state of the system. The number of possible states increases exponentially as we add particles to the system. For even a moderate number of particles, the computation time could surpass what can be computed during the age of the Universe! To simulate such a quantum mechanical system, it is necessary to use a different

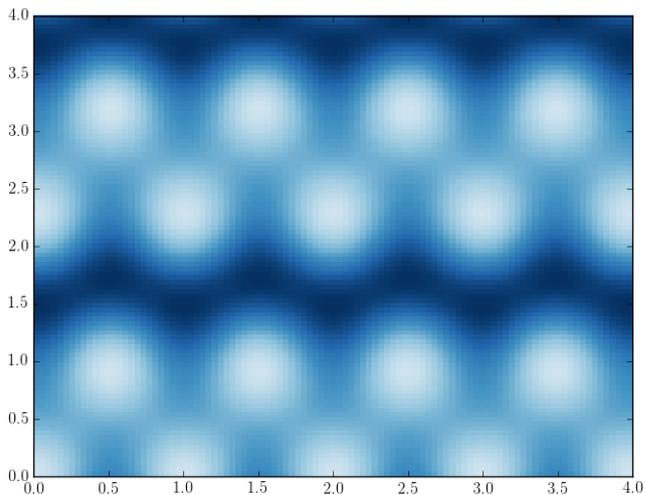


Figure 1.1: The zig-zag lattice. Minima of the potential are shown in white and maxima in dark blue. All lengths are rescaled by the lattice constant ℓ .

kind of hardware. Two quantum mechanical systems may possess similar physics but different characteristic length scales, lifetimes et cetera. Therefore, it is in principle possible to “simulate” quantum systems with *other* quantum systems which are simpler to manipulate and measure, an idea first proposed by Feynman [Fey81, MAV12, p. 66]. Such a model system is called a *quantum simulator*.

During the later half of the 20th century, a number of key technologies were developed which lead to an increasing level of control over quantum systems in the lab. The ion trap was one of the techniques used for single-particle quantum systems. Such technology enabled the Bell experiment, atomic clocks and the development of quantum information theory. Methods of cooling, trapping and manipulation of atoms were developed which led to better control of many-body systems [MAV12, p. 3]. In 1995, a group of experimental physicists successfully prepared a cool dilute gas of atoms in the same quantum state, creating a *Bose-Einstein condensate* (BEC) [PS08, p. 1]. For this they were awarded the Nobel prize in 1997. The creation of *optical lattices* made it possible to trap dilute alkali gases in periodic potentials: such potentials resemble the effective potentials for valence electrons in solids, but possesses an adjustable lattice spacing about three orders of magnitude wider [PS08, p. 409]. The rate of tunneling between sites can be controlled by adjusting the amplitude of the laser, and even the strength of atom interactions can be controlled through a technique called Feshbach resonances [PS08, p. 143, p. 148]. This flexibility means that optical lattices are useful for quantum simulation of many different problems in quantum optics, condensed matter physics and even fundamental particle physics. Some of the most important experimental results were the observation of a superfluid-to-Mott quantum phase transition by [MOT⁺02] and the observation of topological order in condensates [RPC⁺99]. One dramatic example of interest to both particle physicists and the popular science press: in some aspects, atoms in optical lattices near certain phase transitions behave like black holes [JMB10, GC14]!

In isolated atoms, the bound electrons occupy discrete energy levels of increasing angular momentum. These levels can be labelled by the *azimuthal* quantum num-

bers $0, 1, 2, 3, 4, 5, \dots$, which are also called s, p, d, f, g, h, \dots for historical reasons. The energy levels correspond to the different *orbitals* of the atom. In solids, the interactions between atoms lead to dense bands of discrete energy levels which are often referred to as *energy bands*. While atoms in the s band have been studied extensively, for example in [W.03] or [AS05], the physics of p -band atoms is less explored: such systems are accessible and good candidates for simulating quantum magnetism [TSAI07]. In [PML12] and [PBML13], the physics of the p band was explored in a quadratic, two-dimensional lattice. Such lattices are *separable*, which leads to some limitations of how atoms can tunnel. They can be used to simulate for example the *Heisenberg* model, but *non-separable* lattices open up the possibility to study a broader range of spin models. Non-separable lattices may also have *topological* properties [MAV12, p. 429].

For this thesis, I have studied the physics of *bosonic* atoms in the p band, loaded in a *zig-zag* lattice, shown in figure 1.1. The minima (white in the figure) form zig-zag chains separated by strong maxima (dark blue) which effectively prevents atoms from tunneling between different chains. The system can therefore be viewed as quasi-one dimensional. Quantum mechanical many-body systems require new theoretical tools for their solution: in chapter 2, I will therefore give an introduction to non-relativistic *quantum field theory* and some approximation methods related to it. With these tools, many-body systems are then investigated in chapter 3 where *Bloch* and *Wannier* states are introduced - the wavelike and localized quantum mechanical states of a periodic potential. In this chapter, the concept of *phase transitions* is also introduced and exemplified by the classical and quantum *Ising* phase transitions, *Bose-Einstein condensation* and the *superfluid-Mott phase transition*. Chapter 4 gives a brief survey of how optical lattices are created and how measurement and manipulation is performed.

In chapter 5, the many-body Hamiltonian is rewritten in terms of quantum field operators by computing the *overlaps* of Wannier wavefunctions. These overlaps describe the energy associated with atoms *tunneling* between minima of the potential. In chapter 6, the Hamiltonian is diagonalized exactly in the absence of interactions. In chapter 7, the ground state of the system is found by a variational calculation called *mean-field theory*. For strong interactions, the atoms are expected to localize with constant density on each site. I derive an effective Hamiltonian in chapter 8, describing the relevant dynamics in this phase. Furthermore, via a *Schwinger spin-boson* mapping, the dynamics in the Mott phase is shown to be equivalent to a spin model with properties beyond the common *Heisenberg* models. Particularly, the model shows next-nearest neighbor couplings and a *Dzyaloshinskii-Moriya* or anti-symmetric exchange.

The text is meant to be accessible to students of physics at an advanced undergraduate or graduate level. I have tried to avoid technical jargon when possible and italicized new terms to highlight them. The quantum field theory is perhaps the most inaccessible part of the text. Since a thesis for a master's degree is somewhat less formal, I have chosen to address the reader with the pronoun "I".

Chapter 2

Quantum field theory

An inexperienced reader of the most interesting physical papers is often left in a vacuum about the precise meaning of the most common terms. . . . What is a current algebra, a supersymmetry transformation, a topological field theory, a path integral, finally? They are very open concepts, and it is precisely their openness that makes them so interesting.

Yuri Manin

2.1 Motivation for quantum field theory

The research areas of particle physics and condensed matter physics both rely strongly on *quantum field theory*. This theory can be seen as the type of quantum mechanics which is suitable and natural to many-body physics. Since this thesis only concerns Bose particles, I will not discuss the theory for fermions, which differs in some important aspects. We can in principle construct multi-particle states by forming products of single-particle states by hand. But to write down the combined state of just two bosons in states $|\psi_1\rangle$, $|\psi_2\rangle$ with no special quantum numbers, we have to be careful. By the laws of quantum mechanics, the total wavefunction¹ of two identical Bose particles must be *symmetric* under the exchange of the particle coordinates. It must therefore be of the form

$$\psi_{12}(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}} (\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2) + \psi_1(\mathbf{x}_2)\psi_2(\mathbf{x}_1)). \quad (2.1)$$

For a N -particle state, this sum would include $N!$ terms, one for every permutation of the N coordinates [AS10, p. 42].

¹The wave function can be seen as the function defined by the inner product $\psi(\mathbf{x}) = \langle \mathbf{x} | \psi \rangle$, where $|\mathbf{x}\rangle$ is an eigenstate of the translation operator, with the position eigenvalue \mathbf{x} . In this thesis, I will assume the reader to be familiar with the concepts of states, wavefunctions, Hilbert spaces and the bra-ket formalism for standard quantum mechanics.

However, there are two problems with this description: first, it is unfeasible to even write down the state for $\mathcal{O}(10^{23})$ electrons - even more so for the interactions. Second, the theory fails to describe situations when particles are created or annihilated, for example the annihilation of an electron-hole pair. Every theory is also a statement about the “realness” and relevance of that particular description: standard quantum mechanics have problems with the dynamics of many particles because it works with “states” which describe the state of the whole system. Only in special cases, such as for product states, is it easy to discern single particles. Since many-body quantum mechanics starts by treating the *particles* as relevant, then tries to understand their dynamics, it needs another flavor of quantum mechanics.

So far, I have described a many-particle wave function in terms of the unique state of every single particle, but this is a highly redundant description if many particles occupy the same states, as would be probable for bosons. Another redundancy is that general states are not assumed to be symmetric and instead symmetrized “by hand”.

2.2 Second quantization

It is better to work with a different picture, the *number representation*, also called *Fock* representation, invented by Anish Bose in 1924 [PB11, p. XXV] to explain black-body radiation. The number representation does not ask for the state of each particle, but for the number of particles in each single-particle state².

$$|s_1, s_2, \dots, s_N\rangle = |n_1, n_2, \dots, n_M\rangle, \quad (2.2)$$

where N is the number of electrons and M is the number of states. States such as (2.2) are called number or Fock states. We say that many-body quantum mechanics takes place in (the symmetrized subspace of) *Fock space*:

$$\mathcal{F} = \bigotimes_{i=0}^{\infty} \mathcal{H}_i, \quad (2.3)$$

which is the tensor product of a (possibly infinite) number of copies of single-boson Hilbert spaces, one for each particle. The Fock states form an orthonormal basis in Fock space. The number representation is more “natural” because the indistinguishability of quantum states is built into it, if certain *commutation rules* are obeyed. Define the *boson creation operator* as an operator which adds a particle to the state:

$$\hat{a}_i^\dagger |n_1, n_2, \dots, n_i, \dots, n_M\rangle = \sqrt{n_i + 1} |n_1, n_2, \dots, n_i + 1, \dots, n_M\rangle. \quad (2.4)$$

The index i here stands for a set of quantum numbers in some single-particle basis of choice. Depending on basis the creation operators could be labeled by momentum, position, site, orbital or spin. The *boson annihilation operator* is similarly defined by:

$$\hat{a}_i |n_1, n_2, \dots, n_i, \dots, n_M\rangle = \sqrt{n_i} |n_1, n_2, \dots, n_i - 1, \dots, n_M\rangle. \quad (2.5)$$

²Meaning that particles in the same single-particle state share position, momentum and all quantum numbers.

Note that the creation operator is the hermitian conjugate of the annihilation operator. These operators additionally satisfy the important *commutation relation*³:

$$[\hat{a}_i^\dagger, \hat{a}_j] = \hat{a}_i^\dagger \hat{a}_j - \hat{a}_j \hat{a}_i^\dagger = \delta_{ij}. \quad (2.6)$$

The *vacuum state* is defined as the state for which

$$\forall i : \hat{a}_i |0\rangle = 0. \quad (2.7)$$

Defined like this, the annihilation and creation operators construct symmetrized states:

$$\hat{a}_i^\dagger \hat{a}_j^\dagger |0\rangle = \hat{a}_j^\dagger \hat{a}_i^\dagger |0\rangle \quad (2.8)$$

Any Fock state can be constructed by repeated application of creation operators on the vacuum state as

$$|n_1, n_2, \dots, n_M\rangle \propto (a_1^\dagger)^{n_1} \dots (a_M^\dagger)^{n_M} |0\rangle. \quad (2.9)$$

Furthermore, any operator may be described by combining creation and annihilation operators, as should be obvious since they can be used to construct any basis state.

The field of all these quantized degrees of freedom for a kind of particle, labeled by position, momentum, spin or other quantum numbers, is called a *quantum field* [FG10, p. 7]. Single particles and, for example, the process of electron-hole-annihilation, may be described as processes involving low-energy excitations of this quantum field⁴. This formalism for writing a quantum mechanical system in terms of field operators, is a procedure known as *second quantization*. “First quantization” refers to the fact that simultaneous measurement of some variables, like position and momentum, is impossible in principle [Mah90, p. 1]. “Second quantization” refers to the fact that fields are created and destroyed in *quantas*, called particles.

It is common to work with an orthonormal functional basis which is able to describe any state, and expand the Fock states in this basis. One common such functional basis is the position basis⁵, consisting of Dirac delta distributions [Aue94, p. 4] and labeled by the position quantum number. Let $\hat{\psi}_s^\dagger(\mathbf{x})$ be the creation operator of a state at position \mathbf{x} with some set of other quantum numbers s :

$$\langle \mathbf{x}' s' | \hat{\psi}_s^\dagger(\mathbf{x}) | 0 \rangle = \delta_{s's} \delta(\mathbf{x} - \mathbf{x}'). \quad (2.10)$$

Suitable bases in condensed matter applications are the Bloch or Wannier states, which will be described closer in section 3.1.5.

³For fermions, the commutation relation has to be replaced by an *anti-commutation* relation $\{\hat{a}^\dagger, \hat{a}\} = \hat{a}^\dagger \hat{a} + \hat{a} \hat{a}^\dagger = 1$. This relation guarantees that fermion creation operators create anti-symmetric states (states which changes sign when permuting the coordinates and quantum numbers of any particle).

⁴Quantum fields are very abstract. They do not necessarily have to describe elementary particles such as electrons. In solids, the *relevant excitations* (the best description of what kind of particles are interacting) are not the “bare” electrons found in vacuum, since all electrons in a solid interact in such a way that their effective mass and charge is altered. This is sometimes referred to as “dressed” electrons or “quasi-particles” forming a *Fermi liquid*. Quantum fields can also describe *collective excitations*, such as sound vibration (phonons) or collective electronic dynamics (plasmons) [AS10, p. 210] [Mah90, p. 893]

⁵The Dirac delta functions $\delta(x)$ (position basis) and plane waves e^{ikx} (momentum basis) are common in quantum mechanics. This is confusing to new students of quantum mechanics, since these functions are not square-integrable ($\int_{-\infty}^{\infty} e^{-ikx} e^{ikx} dx = \infty$) and hence do not fit into Hilbert space. The solution to this mathematically is to either extend Hilbert space to a larger space (*rigged* Hilbert space) where these distributions are included, or to formulate quantum mechanics with projection operators [Bal10, p. 20, p. 28]

2.3 Operators in second quantization

Since it is possible to construct any state in Fock space with the creation and annihilation operators, it is also possible to write any operator on Fock states in terms of them. Since quantum field theory is concerned with the relation between the full quantum system and the individual “particles” of which it consists, it is natural to classify Fock space operators after the number of individual particles involved. *One-body operators* are operators only acting on one particle at a time. In the basis where the operator is diagonal [AS10, p. 47]:

$$\hat{\mathcal{O}}_1 = \sum_n o_n \hat{a}_n^\dagger \hat{a}_n. \quad (2.11)$$

Now do a transformation back to a general Fock space basis and use the spectral resolution of the first-quantized one-body operator:

$$\begin{aligned} \hat{a}_n &= \sum_\mu \langle n | \mu \rangle \hat{a}_\mu \Rightarrow \\ \hat{\mathcal{O}}_1 &= \sum_{\mu\nu} \sum_n o_n \langle \mu | n \rangle \langle n | \nu \rangle \hat{a}_\mu^\dagger \hat{a}_\nu = \\ &= \sum_{\mu\nu} \langle \mu | \sum_n (|n\rangle o_n \langle n|) | \nu \rangle \hat{a}_\mu^\dagger \hat{a}_\nu = \sum_{\mu\nu} \langle \mu | \hat{o} | \nu \rangle \hat{a}_\mu^\dagger \hat{a}_\nu, \end{aligned} \quad (2.12)$$

where \hat{o} is the local operator acting only on particle n , which follows from the fact that the single-particle basis is orthogonal. The set of numbers $\{\langle \mu | \hat{o} | \nu \rangle\}$ can be viewed as a “matrix” representation of the operator in Fock space.

Two-body operators are similarly described by

$$\mathcal{O}_2 = \sum_{\mu\mu'\nu\nu'} \langle \mu\mu' | \hat{\mathcal{O}}_2 | \nu\nu' \rangle \hat{a}_\mu^\dagger \hat{a}_{\mu'}^\dagger \hat{a}_\nu \hat{a}_{\nu'}. \quad (2.13)$$

This expression is not fundamentally different from the one-body operator. It only involves the sum over four sets of quantum numbers instead of two.

2.4 Normal ordering

Quantum field theory is plagued by mathematical subtleties, which often appear in the form of infinities⁶. One prominent example is related to expectation values, which are in turn related to observation. The expectation value of the Hamiltonian operator is what is commonly referred to as the *energy* of a state. For the vacuum state $|0\rangle$, one would *expect* $\langle 0 | \hat{H} | 0 \rangle$ to be zero (the vacuum is empty!). This is often referred to as the *vacuum expectation value* (VEV) of an operator. If the Hamiltonian has a non-zero VEV, the vacuum contains an unphysical energy. This vacuum energy has no physical meaning since only energy differences matter physically by the equation $\mathbf{F} = -\nabla V$ which holds for any conservative force in classical as well as quantum mechanics. There is a canonical procedure to make sure that all VEV:s are zero, known as *normal ordering* [FG10, p. 42]. Normal ordering means “manually”

⁶Many of these problems are related to the formulation in operators on infinite-dimensional spaces, and do not appear in the “equivalent” *path integral formalism* [MD95, p. 275,p. 306] in which one can derive expectation values and correlation functions without having to deal with Hilbert spaces.

moving all creation operators to the left of all annihilation operators, and is often denoted by $::$ like

$$:\hat{a}^\dagger\hat{a}: = \hat{a}^\dagger\hat{a}, \quad (2.14)$$

$$:\hat{a}\hat{a}^\dagger: = \hat{a}^\dagger\hat{a}. \quad (2.15)$$

A normal-ordered VEV of any operator is always zero:

$$\langle 0 | : \hat{O} : | 0 \rangle = 0. \quad (2.16)$$

2.5 Mean-field approximations

Quantum fluctuations become less important for a system above a certain temperature, time scale, length scale *et cetera*. In this limit, classical statistical mechanics is a sufficient description. For bosons, a high occupation number in each Fock state leads to weaker relative fluctuations. The collective degree of freedom of the bosons may then be approximated as classical. But even in systems where quantum effects are important, classical limits may be useful as a first estimate, to gain intuition about the full solution.

The *variational theorem* of quantum mechanics [Sak94, p. 332]

$$\bar{E} = \langle \psi | : \hat{H} : | \psi \rangle \geq E_0, \quad (2.17)$$

where E_0 is the true ground state energy, $: \hat{H} :$ is the normal ordered Hamiltonian and $|\psi\rangle$ is *any* normalized state. By choosing a *trial* wavefunction ψ_λ dependent on a set of parameters $\lambda = \lambda_1, \lambda_2, \dots$ and finding the minimum of \bar{E} by varying λ , an upper bound for E_0 can be found. This bound can be further improved by minimizing with a larger class of functions. An educated guess of trial wavefunction can also provide a way of approximating the true ground state, although \bar{E} is mostly sensitive to short-range correlations. Some guessing is involved in finding the long-range order by a variational method [Aue94, p. 39].

One very simple variational ansatz⁷ in many-body quantum mechanics is to assume nearly all particles to be in the ground state and use a trial wavefunction which is a *product state*⁸ of identical single-particle wavefunctions:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \prod_{i=0}^N \phi(\mathbf{r}_i), \quad (2.18)$$

where N is the total number of bosons. This is sometimes called a *Hartree* approximation [PS08, p. 160] and is an important model of the Bose-Einstein condensate which will be described in section 3.3. In a manner similar to the approximations

⁷“Ansatz” is a complex german word roughly translatable to “starting point”, “beginnings” or “attempt”. Physicists and mathematicians use it as a name for any attempt to rewrite or guess the structure of the solution to a problem, especially a differential equation.

⁸I will often write product states as $|\Psi\rangle = \prod_j |\psi\rangle_j$ for example. It would be more mathematically correct to write $|\Psi\rangle = \otimes_j |\psi\rangle_j$ since what is really meant is a *tensor product* of state vectors in different subspaces. As it is common to write product states as $|xy\rangle = |x\rangle|y\rangle$, I will use the product symbol in this text.

leading to Sommerfeld theory, first-order effects of the interactions can be included⁹ in an effective contact interaction $U(\mathbf{r}, \mathbf{r}') = U_0 \delta^{(3)}(\mathbf{r} - \mathbf{r}')$. The upper bound on the ground state energy is then [PS08, p. 159]

$$\begin{aligned} \bar{E} &= \langle \Psi | : \hat{H} : | \Psi \rangle = \\ &= N \int d\mathbf{r} \left[\frac{\hbar^2}{2m} |\nabla \phi(\mathbf{r})|^2 + V(\mathbf{r}) |\phi(\mathbf{r})|^2 + \frac{(N-1)}{2} U_0 |\phi(\mathbf{r})|^4 \right], \end{aligned} \quad (2.19)$$

which has to be minimized functionally [PS08, p. 162] to find the optimal function $\phi(\mathbf{x})$. Note that this has to be done under any constraints on the system, such as a fixed temperature or particle number. Particle number can be held fixed with the help of a Lagrange multiplier μ , commonly referred to as the *chemical potential*. Functional minimization with respect to variation in $\psi^*(\mathbf{r})$ (treated as a field independent of its complex conjugate) yields

$$\begin{aligned} \frac{\delta}{\delta \psi^*(\mathbf{r})} (E - \mu N) = 0 \Rightarrow \\ \frac{\delta}{\delta \psi^*(\mathbf{r})} \int d\mathbf{r} \frac{\hbar^2}{2m} |\nabla \psi(\mathbf{r})|^2 + \frac{\delta}{\delta \psi^*(\mathbf{r})} \int d\mathbf{r} \left[(V(\mathbf{r}) - \mu) |\psi(\mathbf{r})|^2 + \frac{1}{2} U_0 |\psi(\mathbf{r})|^4 \right] = \\ \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) + U_0 |\psi(\mathbf{r})|^2 \right] \psi(\mathbf{r}) = \mu \psi(\mathbf{r}), \end{aligned} \quad (2.20)$$

where order of functional differentiation and integration was interchanged, which is legitimate because of absolute convergence of the integral. Furthermore, I used the chain rule for functional differentiation

$$\frac{\delta}{\delta \psi^*(\mathbf{r})} |\psi(\mathbf{r})|^\alpha = \alpha |\psi(\mathbf{r})|^{\alpha-1} \psi(\mathbf{r}), \quad (2.21)$$

as well as partial integration of the kinetic term

$$\begin{aligned} \int d\mathbf{r} |\nabla \psi(\mathbf{r})|^2 &= \int d\mathbf{r} (\nabla \psi(\mathbf{r})) (\nabla \psi^*(\mathbf{r})) = \\ (\nabla \psi(\mathbf{r})) (\nabla \psi^*(\mathbf{r}))|_{\partial V} - \int d\mathbf{r} (\nabla^2 \psi(\mathbf{r})) \psi^*(\mathbf{r}) &= - \int d\mathbf{r} (\nabla^2 \psi(\mathbf{r})) \psi^*(\mathbf{r}). \end{aligned} \quad (2.22)$$

The boundary term vanishes because the wavefunction and its gradient goes to zero at infinity. $\psi(\mathbf{r}) = \sqrt{N} \phi(\mathbf{r})$. Note that I have also assumed N to be large enough that $N-1 \approx N$. Equation (2.20) is called the *Gross-Pitaevskii* equation and has the form of a Schrödinger equation with a non-linear density term. In this sense, the Hartree approximation is a mean-field theory, where the field $\psi(\mathbf{r})$ is the mean field which has to be solved for self-consistently¹⁰. It can be viewed as a kind of average, macroscopic wavefunction.

⁹This is the Born approximation of scattering theory [PS08, p. 119], which estimates the scattering length as $a = (m_r/2\pi\hbar^2) \int d\mathbf{r} U(\mathbf{r}) \Rightarrow \int d\mathbf{r} U_{\text{eff}}(\mathbf{r}) = \frac{2\pi\hbar^2 a}{m_r} = U_0$, where m_r is the effective mass. In coordinate space this means that $U_{\text{eff}}(\mathbf{r}) = U_0 \delta^{(3)}(\mathbf{r})$.

¹⁰The concept of a mean-field solution is borrowed from statistical mechanics [MB06, p. 64]. The interaction of one boson with wavefunction $\psi(\mathbf{r})$ with every other boson is accounted for by the term proportional to $|\psi(\mathbf{r})|^2$. This is of course a simplification, since boson interactions are not instantaneous. In the method of mean-field theory, the field is then determined from some

Another choice of trial states is the *coherent state* ansatz, which assumes a product state in the form

$$|\Psi\rangle(\{\alpha_j\}) = \prod_i^N |\alpha_i\rangle \Rightarrow \langle\Psi|\hat{a}_i|\Psi\rangle = \langle\alpha_i|\hat{a}_i|\alpha_i\rangle = \alpha_i \quad (2.23)$$

$$|\alpha_j^x, \alpha_j^y\rangle_j = \exp\left(-\frac{|\alpha_j^x|^2 + |\alpha_j^y|^2}{2}\right) \sum_{n_x, n_y} \frac{(\alpha_j^x)^{n_x} (\alpha_j^y)^{n_y}}{\sqrt{n_x! n_y!}} |n_x n_y\rangle_j \quad (2.24)$$

where the full set of complex eigenvalues α are the parameters to be varied. The set $\{\alpha_i\}$ is commonly referred to as the *order parameters*¹¹ of the system. The states $|\alpha_i\rangle$ are *coherent states*, eigenstates of the annihilation operator \hat{a}_i , with a complex eigenvalue α_i :

$$\hat{a}_i |\alpha_i\rangle = \alpha_i \hat{a}_i |\alpha_i\rangle \Rightarrow \langle\alpha_i|\hat{a}_i|\alpha_i\rangle = \alpha_i. \quad (2.25)$$

The set of coherent eigenstates of an annihilation operator \hat{a}_i are in correspondence with the set of complex numbers, since the above state can be formed for any complex number α_i . For a state to be the eigenstate of an annihilation operator, it cannot simultaneously be an eigenstate of the number operator¹². Such a state can be written as [Bal10, p. 544]

$$|\alpha_i\rangle = e^{-\frac{1}{2}|\alpha_i|^2} \sum_{n_i=0}^{\infty} \frac{\alpha_i^{n_i}}{\sqrt{n_i!}} (\hat{a}_i^\dagger)^{n_i} |0\rangle = e^{-\frac{1}{2}|\alpha_i|^2} \sum_{n_i=0}^{\infty} \frac{\alpha_i^{n_i}}{\sqrt{n_i!}} |n_i\rangle \quad (2.26)$$

The set of states $\{|\alpha_i\rangle\}$ (where i here label all the possible Fock states, for example site number and orbital quantum number), forms an *overcomplete basis* in the complex plane, since

$$\int |\alpha\rangle \langle\alpha| d\alpha^2 = \pi \hat{I} \neq \hat{I}, \quad (2.27)$$

and $\pi > 1$. This means that arbitrary states can be specified in a coherent state basis, but generally not in one unique way. By a similar calculation as above, the coherent state ansatz also leads to a Gross-Pitaevskii equation, which is why it is referred to as a mean-field theory. In section 7.3, I will use the coherent state ansatz to find a mean-field solution for the zig-zag lattice.

One interesting question is if mean-field theories are useful for understanding systems close to a quantum phase transition. While a “classical” phase transition

kind of self-consistency condition: *assuming* the mean-field solution to be correct, all expectation values and correlators must be consistent with the assumptions. In the mean-field solution of the classical Ising model $H = -J \sum_{\langle ij \rangle} s_i s_j$ where $s_i = \pm 1$ and the sum is over nearest neighbors. The statistical average $m = \langle s_j \rangle$ is assumed to hold for all sites j , which may be inserted as a mean field so that $H \approx -J \sum_i q m s_i$ (q is the number of nearest neighbors for each site) if the fluctuation around the mean value is small. The self-consistent condition (reminiscent of the Gross-Pitaevskii equation) is then simply $m = \langle s_j \rangle$ for all sites j

¹¹The name refers to the order parameters in the Landau-Ginzburg theory of phase transitions and fluctuations. The numbers $\{\alpha_i\}$ characterizes the mean-field order of the quantum system in the same way, although we have to keep in mind that this is only a semi-classical approximation

¹²This is just an awkward way of noting that a coherent state cannot be a state of fixed particle number. If it were, annihilation would be able to change the number and thus change the state. Such a state would not be an eigenstate.

is driven by thermal fluctuations, a quantum phase transition is driven by quantum fluctuations and strictly only defined at zero temperature [S.11]. In both cases, the correlation length diverges as the system approaches the critical point, indicating that the system becomes equally sensitive to *all possible length scales*.

Chapter 3

Many-body physics

If you don't know ladders, don't
play go

Go Proverb

This chapter will give a brief survey of topics which are important to understand many-body quantum mechanical problems. Many-body Hamiltonians and important mathematical aspects of periodic potentials will be discussed in section 3.1, in terms of fermionic systems such as valence electrons in solids. Section 3.2 reviews the physics of *phase transitions*, which arise in the thermodynamic limit $N \rightarrow \infty$. Sections 3.3 and 3.4 finally give two examples of bosonic many-body systems where phase transitions occur.

3.1 The solid state

3.1.1 Models of the solid state

Historically, the first working microscopic model of a solid was the *Drude theory of metals*. It treats the solid as a collection of *cores* and valence electrons. The cores can be viewed as consisting of the nucleus and the tightly bound core electrons [AM76, p. 3]. The forces binding the core electrons to the nucleus and the nuclear forces are so strong in comparison to the interaction with the conduction electrons, that the core may be viewed as a rigid unit. For crystalline solids, the cores can be assumed to form a periodic lattice. Drude theory assumes the cores to be inert and neglects interaction between the conduction electrons (*independent electron approximation*) and interaction between conduction electrons and cores (*free electron approximation*) except for a type of hard scattering of the conduction electrons off the cores. The resulting model is similar to a free kinetic gas, but collisions between the gas particles is here replaced by the collisions between electrons and cores. The frequency of scatterings is assumed to happen with a mean probability density $1/\tau$ and assumed to scatter the electrons in a random direction and with a new speed related to the temperature. While Drude theory gives successful predictions for heat and electron conductivity, it fails to describe the heat capacity of common metals, among other things. It turns out that the main reason for this failure is the assumption of a classical, random interaction between electrons and cores.

The most relevant improvement of Drude theory is to replace the random electron-core scattering by a quantum mechanical system of independent electrons moving in a background potential representing the cores. The Schrödinger equation for time-independent potentials

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \hat{\mathcal{H}} |\psi\rangle = \left(-\frac{\hbar^2}{2m} \hat{\nabla}^2 + V(\mathbf{x}) \right) |\psi\rangle, \quad (3.1)$$

is commonly solved by first finding the eigenstates of the time-independent Schrödinger equation

$$\hat{\mathcal{H}} |\psi\rangle = E |\psi\rangle. \quad (3.2)$$

For a many-body problem the Hamiltonian $\hat{\mathcal{H}}$ can always be split up into one part which is a sum of single-body potentials and a part which captures interaction between particles:

$$\mathcal{H} = \mathcal{H}_0 + \frac{1}{2} \sum_{i \neq j} U(\mathbf{x}_i, \mathbf{x}_j), \quad (3.3)$$

$$\mathcal{H}_0 = \sum_i \left(-\frac{\hbar^2}{2m} \nabla_i^2 + V(\mathbf{x}_i) \right) = \sum_i H_0[\nabla_i, \mathbf{x}_i]. \quad (3.4)$$

One particularly simple approximation is to ignore all interactions between electrons, and model the solid as unbound conduction electrons moving independently in a periodic “background” potential. This is possible because Coulomb interaction between negative and positive charges, for example, can be expected to “screen” particles far away from each other [Aue94, p. 4] so that each particle feels an average “background field” plus a weaker, truly inter-particle interaction U^1 . If the true interaction is weak enough, it may be a good approximation to ignore it and just assume independent electrons.

The Schrödinger equation for \mathcal{H}_0 separates into a sum of independent single-electron Schrödinger equations:

$$\left(-\nabla^2 + V(\mathbf{x}) \right) \phi(\mathbf{x}) = E \phi(\mathbf{x}), \quad (3.5)$$

$$V(\mathbf{x} - \mathbf{X}) = V(\mathbf{x}), \quad (3.6)$$

where \mathbf{X} denotes the position of any core. Here the equation have been written in a rescaled form by making a variable change to some rescaled coordinates ($\mathbf{x} \rightarrow \ell \mathbf{x}$ for the simplest case of scaling all coordinates by some factor), dividing by $\hbar^2/2m$ and redefining the potential as $V \rightarrow 2m\ell V/\hbar^2$ and the energy as $E \rightarrow 2m\ell E/\hbar^2$. The rescaling makes the Schrödinger equation dimensionless in length, time, momentum and energy. In this thesis, I will often rescale the Hamiltonian to bring concepts to front and keep equations from getting cluttered. (3.6) simply states that the potential of the periodic array of cores is also periodic. Surprisingly, it turns out that this approximation, called *Sommerfeld* theory or the *independent electron approximation*, is enough to explain many properties solids. That said, Sommerfeld theory

¹Screening is the reason why thermodynamics works and the Earth doesn’t explode: if there were no screening, the forces between charges would have too long range and overpower the weak gravitational force holding the planet together.

fails for example in explaining conductance/resistance of many common elements and the heat capacity of e.g. iron [AM76, p. 59]. In this thesis, the independent electron approximation will often be taken as a starting point for a better estimate taking the interaction potential in (3.3) into account.

3.1.2 Crystal lattices

To understand crystalline solids, it is important to understand the mathematical properties of the idealization of an infinite, periodic lattice. While a true solid has a finite extent and irregularities, an ideal infinite lattice forms a starting point for understanding bulk effects of crystalline solids. In d dimensions, consider an infinite set of vectors with d elements², $\{\mathbf{X}\}$, where \mathbf{X} can always be written as [AM76, p. 65]

$$\mathbf{X} = n_1 \mathbf{a}_1 + \cdots + n_d \mathbf{a}_d, n_i \in \mathbf{Z}. \quad (3.7)$$

Thus any point in the lattice can be identified by d integers n_i . The vectors \mathbf{a}_i are called the *primitive vectors*, and are not unique in the sense that any linear combination with integer coefficients would be another, valid set of primitive vectors. The set $\{\mathbf{X}\}$ is called a *Bravais lattice* and the individual vectors \mathbf{X} are the lattice points. The set of points which are closer to a specific lattice point \mathbf{X} than to any other, forms a volume which is called the *primitive cell* or *Wigner-Seitz cell*. The primitive cell describes a *tessellation*, or tiling, of space, since the primitive cells partition the space. Remarkably, any crystal structure can be written as a Bravais lattice, which describes the tessellation of the volume, if the positions of the individual atoms within each tile are also specified. Figure 3.1 shows some Bravais lattices in two dimension with one atom per cell.

The fact that some lattice geometries lead to Hamiltonians which are *separable* is of importance: separable Hamiltonians have the property that they are reducible to a set of one-dimensional Schrödinger equations by an ansatz $\psi(\mathbf{x}) = \prod_{j=1}^d \psi_j(\tilde{x}_j)$, where $\tilde{\mathbf{x}}_j$ are some set of possibly transformed coordinates. Such an ansatz is successful only if the potential is of the form

$$\begin{aligned} V(\mathbf{x}) = \sum_{j=1}^d V_j(\tilde{x}_j) &\Rightarrow H = \sum_{j=1}^d \left(- \left| \frac{\partial \tilde{\mathbf{x}}}{\partial \mathbf{x}} \right|^2 \tilde{\nabla}^2 + V_j(\tilde{x}_j) \right) = \sum_{j=1}^d \tilde{H}_j \\ \Rightarrow 0 = \left(\sum_{j=1}^d \tilde{H}_j - E \right) \psi(\mathbf{x}) &= \sum_{j=1}^d \left(\prod_{i=1, i \neq j}^d \psi_i(\tilde{x}_i) \right) (\tilde{H}_j - E_j) \psi(\tilde{x}_j), \end{aligned} \quad (3.8)$$

which has to vanish term by term.

Most of the properties we are interested in will also be periodic in the lattice. This means that their Fourier expansions will only contain modes related to the lattice constant. More precisely, I can define the *reciprocal lattice* as the set of wave vectors which yield plane waves with the periodicity of the Bravais lattice. The condition can be written as [AM76, p. 86]

$$e^{i\mathbf{K} \cdot (\mathbf{x} + \mathbf{X})} = e^{i\mathbf{K} \cdot \mathbf{x}} \Leftrightarrow \mathbf{K} \cdot \mathbf{X} = 2\pi n, n \in \mathbf{Z}. \quad (3.9)$$

²I have mentioned a dimension d of the crystal: in actual crystals $d = 1, 2$ or 3 . $d < 2$ can be realized if the system is limited to movement in one direction by a harmonic potential in one or two directions, or possibly if it is translation invariant perpendicular to the lattice.

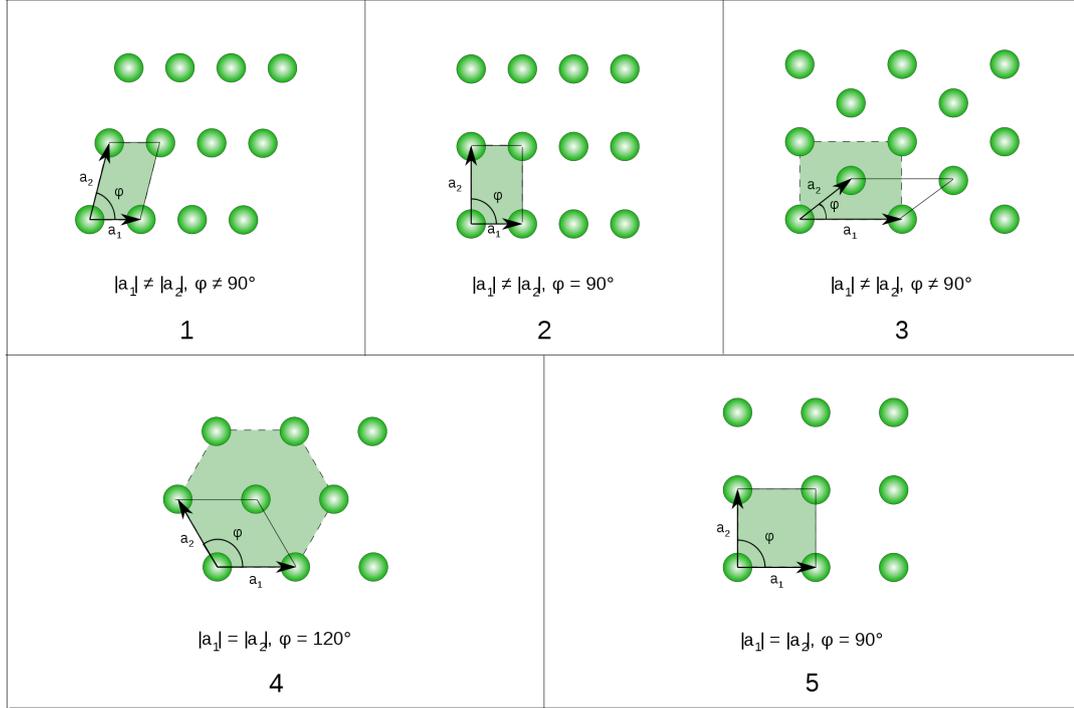


Figure 3.1: Different types of Bravais lattices in two dimensions. The three in the first row, and the quadratic lattice on the second row, have the property that the 2D Schrödinger equation is separable in a potential with these structure. The triangular lattice in the bottom left is not separable. Source: "2d-bravais" by Prolineserver - Own work. Licensed under CC BY-SA 3.0 via Commons - (link).

The lattice defined by (3.9) is also in fact a Bravais lattice with d primitive vectors \mathbf{b}_i . The primitive cell around $\mathbf{k} = 0$ of a reciprocal lattice is known as the *first Brillouin zone* [AM76, p. 89]. It is surrounded by the second, third Brillouin zone *et cetera*. I finally note that (3.9) implies the relation between vectors in space and reciprocal space:

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}. \quad (3.10)$$

Note that this should not be interpreted as an actual scalar product, since the vectors involved live in different spaces. In two dimensions, the primitive reciprocal vectors are easily found:

$$\mathbf{a}_i \cdot \mathbf{b}_j = a_i b_j \cos \theta = 2\pi\delta_{ij} \Rightarrow \mathbf{b}_j = \frac{2\pi}{a_i^2} R(\pm\pi/2)\mathbf{a}_i, \text{ for } i \neq j, \quad (3.11)$$

so that the direction of the primitive vectors of the reciprocal lattice are just given by the $\pi/2$ rotated real space primitive vectors.

3.1.3 Bloch functions

The *Bloch theorem* states that Schrödinger equations with periodical potentials have solutions of the form [AM76, p. 133]

$$\phi_{n\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}} u_{n\mathbf{k}}(\mathbf{x}), \quad (3.12)$$

where u has the same periodic structure as the lattice: $u_{n\mathbf{k}}(\mathbf{x} - \mathbf{X}) = u_{n\mathbf{k}}(\mathbf{x})$.

Such a solution is called a *Bloch function*, *Bloch wavefunction* or *Bloch wave*. The vector \mathbf{k} is called *quasimomentum* or *crystal momentum*. It is related to, but not identical with, the true momentum of the electron. The quasimomentum is a quantum number, a conserved quantity, of the Bloch function, just like its energy. The proper momentum is not conserved for the electron as it moves in a varying potential (i.e. since it is not invariant under arbitrary translations). The label n denotes the n separate energy levels $E_n(\mathbf{k})$ of the solution, but note that the energies can be said to form *bands* since they are functions of \mathbf{k} . The reason is that the Bloch waves are extended and are therefore in a sense intermediate between the bound (which have separate energy levels) and free (continuous energy spectrum).

The conserved crystal momentum is a consequence of the Hamiltonian having *discrete* translation symmetries. This manifests itself in that a set of *translation operators* $\{\hat{T}_d\}$ all commute with the Hamiltonian ($[\hat{T}_d, \hat{H}] = 0$). Assuming for the moment periodic boundary conditions (I will let $N \rightarrow \infty$ after the calculation) $\phi(x + Nd) = \hat{T}_{Nd}\phi(x) = \phi(x) \Rightarrow \hat{T}_{Nd} = 1$. The translation operators form a representation of the *additive group* (modulo N) because $\hat{T}_a\hat{T}_b = \hat{T}_{a+b}$. This suggests writing $\hat{T}_d = e^{i\hat{k}d}$ and since the Hamiltonian and the translation operator are simultaneously diagonalizable:

$$\phi_q(x + d) = e^{i\hat{k}d}\psi_q(x) = \langle x | e^{i\hat{k}d} | q \rangle = e^{iqd} \langle x | q \rangle = e^{iqd}\phi_q(x), \quad (3.13)$$

where $|q\rangle$ is an eigenstate of \hat{k} with eigenvalue q . Hence the solutions consist of a part which is periodic in the lattice, multiplied by a phase proportional to q . The result holds in the limit where $N \rightarrow \infty$ since the argument above holds for any N .

The condition for a wavefunction to be a Bloch function may be rewritten as

$$\psi(\mathbf{x} + \mathbf{X}) = e^{i\mathbf{k}\cdot\mathbf{X}}\psi(\mathbf{x}), \quad (3.14)$$

but for a quasimomentum which is translated by a reciprocal lattice vector $\mathbf{k}' = \mathbf{k} + \mathbf{K}$, equation (3.9) yields

$$e^{i\mathbf{k}'\cdot\mathbf{X}}\psi(\mathbf{x}) = e^{i(\mathbf{k}+\mathbf{K})\cdot\mathbf{X}}\psi(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{X}}\psi(\mathbf{x}). \quad (3.15)$$

Hence

$$\phi_{n(\mathbf{k}+\mathbf{K})}(\mathbf{x}) = \phi_{n\mathbf{k}}(\mathbf{x}). \quad (3.16)$$

Because of this “tiled” property of Bloch functions in the quasimomentum, it is possible to restrict \mathbf{k} to the first Brillouin zone. Figure 3.2 shows a plot of the energy levels, or the spectrum, as a function of the quasimomentum. These energy levels are sometimes referred to as *energy bands* [AM76, p. 141], although the actual energy levels of the whole solid might differ due to interaction between conduction electrons.

3.1.4 Wannier functions

Just like for any quantum mechanical problem, the shape and strength of the lattice potential $V(\mathbf{x})$ will determine the kind of solution. For a weak potential, the

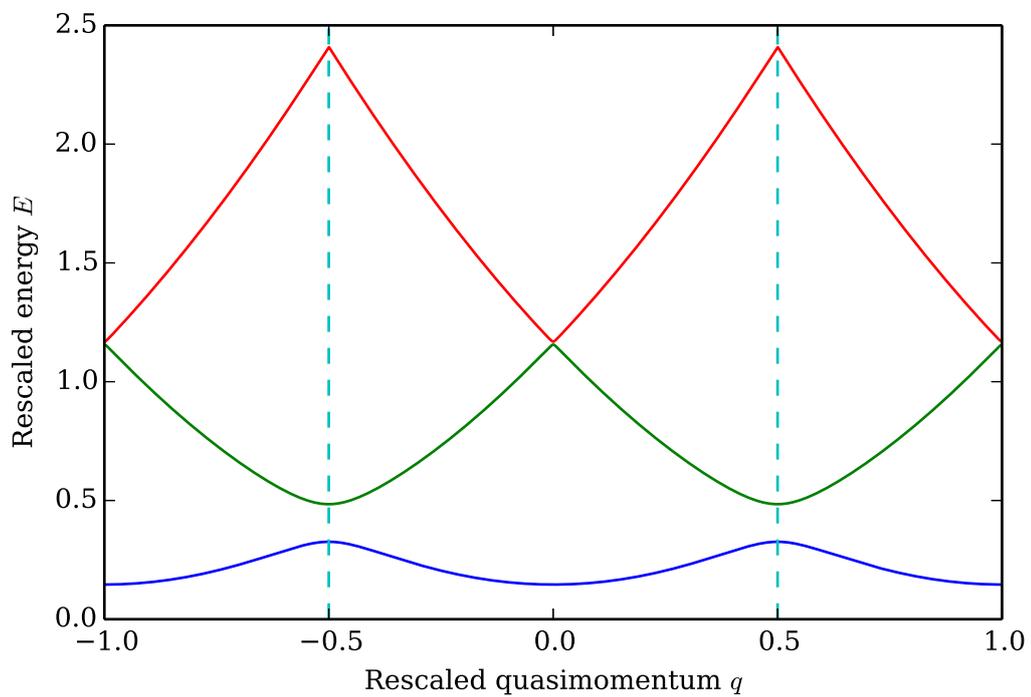


Figure 3.2: Energy-quasimomentum plot for a separable Bloch wave, for the s band (blue), p band (green) and d band (red). Here the spectrum has been extended to ± 1 to show the periodicity in quasimomentum. The extent of the first Brillouin zone is marked by the dotted lines.

conduction electrons are essentially free particles. The Bloch waves are then a suitable description of the electron. But if the “conduction” (non-core) electrons are bound by a strong potential well, the solutions should resemble bound states such as harmonic oscillator or atomic states which are localized [AM76, p. 176]³. In fact, irregardless of the strength of the potential, it is always mathematically justified [AM76, p. 187] (and sometimes useful) to work with a localized single-particle basis of *Wannier functions*:

$$\psi_{n\mathbf{x}}(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbf{BZ}} e^{-i\mathbf{k} \cdot \mathbf{x}} \phi_{m\mathbf{k}}(\mathbf{x}), \quad (3.17)$$

where the sum is over all energy bands and all quasimomenta in the first Brillouin zone. The sum could in principle be extended to all \mathbf{k} , but this is not necessary because of the repeating property of the Bloch function in reciprocal space. Intuitively then, the Bloch functions are localized in momentum space while the Wannier functions are localized in position space. Figure 5.2 shows a Wannier function in the first excited band, in two dimensions for a sinusoid, quadratic lattice. Note that it resembles the p orbital state of an atomic wavefunction.

The Wannier functions are orthogonal on the lattice and in band index, so that

$$\int d\mathbf{x} \psi_{n\mathbf{x}}^*(\mathbf{x}) \psi_{n'\mathbf{x}'}(\mathbf{x}) = \delta_{\mathbf{x}\mathbf{x}'} \delta_{nn'}, \quad (3.18)$$

Notice that the definition is similar to a Fourier transform of the Bloch functions, but the sum is over the allowed quasimomenta in the first Brillouin zone. The asymptotic behaviour of the Wannier states is $|\psi_{n\mathbf{x}}(\mathbf{x})| \sim |\mathbf{x}|^{-3/4} e^{-h_n|\mathbf{x}|}$, where h_n is a decay constant which decreases with increasing band index [BDZ08, p. 14].

In those situations relevant to this thesis, the energy gap to the first excited band can be assumed to be large. This can be used to simplify the Hamiltonian of a system by ignoring interactions between energy bands, among other things [AS10, p. 70]. This is sometimes called a *single-band approximation*.

In the definition of the Wannier function, it is possible to insert a complex phase in each term of the sum. Therefore, there is a local $U(1)$ gauge symmetry of the Bloch functions [NAJ⁺12, p. 4]. However, the transformation affects the shape of the Wannier functions which are thus not defined unambiguously without gauge fixing. In this thesis, I will choose to work in a gauge where the s -band Wannier functions are purely real and the p -band Wannier functions are purely imaginary.

Another interesting idea is to choose a gauge such that the Wannier functions are “maximally localized” in some sense. Define the *localization functional* as

$$\Omega = \sum_n \left[\langle \mathbf{0}n | r^2 | \mathbf{0}n \rangle - \langle \mathbf{0}n | \vec{r} | \mathbf{0}n \rangle^2 \right], \quad (3.19)$$

where $|\mathbf{0}n\rangle$ is the Bloch state with zero quasimomentum, it is possible to work out the conditions to minimize this functional and in this sense minimize the spread. The localization functional is chosen to minimize the sum of the squares of the zero-momentum states in all bands, which is equivalent to the *Foster-Boys criterion* in quantum chemistry [NAJ⁺12, p. 7].

³This is the case if the lattice spacing is larger than the *scattering length*, see section 4.3.

3.1.5 Hamiltonians in the Wannier basis

In the following chapters, I will often work in the Wannier basis and it is therefore appropriate to mention and explain the form of many-body Hamiltonians in this basis. The calculations follow equations (2.12) and (2.13) applied to the many-body Hamiltonian introduced in equations (3.3) and (3.4). In the Wannier basis, where the basis states are described by lattice position and orbital degrees of freedom, I find

$$\begin{aligned} -\langle \mathbf{X}_i\alpha | \hat{\mathcal{H}}_0 | \mathbf{X}_j\beta \rangle &= -\sum_{k\gamma} \langle \mathbf{X}_i\alpha | \mathbf{X}_k\gamma \rangle H_0[\nabla_j, \mathbf{x}] \langle \mathbf{X}_k\gamma | \mathbf{X}_j\beta \rangle = \\ &= -\int d\mathbf{x} \psi_{\mathbf{x}_i\alpha}^*(\mathbf{x}) \left[-\nabla^2 + V(\mathbf{x}) \right] \psi_{\mathbf{x}_j\beta}(\mathbf{x}) = t_{ij}^{\alpha\beta}, \end{aligned} \quad (3.20)$$

where $*$ denotes complex conjugation. $t_{ij}^{\alpha\beta}$ is often called the *tunneling* or *hopping matrix*, since it determines the rate of tunneling from one site and state to another. It is sometimes also called an *overlap integral* since the integrand can be interpreted as the overlap between the two functions $\psi_{\mathbf{x}_i\alpha}^\dagger(\mathbf{x})$ and $\hat{H}_0\psi_{\mathbf{x}_j\beta}(\mathbf{x})$. The second-quantized representation of \mathcal{H}_0 , often called the “kinetic” part, can now be written as

$$\hat{\mathcal{T}} = \sum_{ij} \sum_{\alpha\beta} t_{ij}^{\alpha\beta} \hat{a}_{i\alpha}^\dagger \hat{a}_{j\beta}, \quad (3.21)$$

$t_{ii}^{\alpha\alpha}$ is sometimes called the *onsite energy*. The corresponding string of operators $\hat{a}_{i\alpha}^\dagger \hat{a}_{i\alpha} = \hat{n}_{i\alpha}$ is called the *number operator* or *density operator* and has the action of counting the number of particles in state i, α : $\hat{n}_{i\alpha} |n_{i\alpha}\rangle = n_{i\alpha} |n_{i\alpha}\rangle$. If the onsite energy is the same for all orbital states α and the particle number is conserved, the onsite energy is just a constant contribution to the energy, and the states are called “degenerate”. If it is not, the difference in onsite energy acts like a chemical potential, favouring density in one state over the other. Note that, for a periodic potential, there will be a set of tunnelings which repeat in some period over site number i related to the number of cells in the primitive lattice. This is because of the way that the tunnelings functionally depend on the potential.

The interaction term is

$$\sum_{\alpha\beta\gamma\deltaijkl} U_{ijkl}^{\alpha\beta\gamma\delta} \hat{a}_{i\alpha}^\dagger \hat{a}_{j\beta}^\dagger \hat{a}_{k\gamma} \hat{a}_{l\delta}, \quad (3.22)$$

where

$$U_{ijkl}^{\alpha\beta\gamma\delta} = \langle \mathbf{X}_i\alpha \mathbf{X}_j\beta | \hat{u} | \mathbf{X}_k\gamma \mathbf{X}_l\delta \rangle = \frac{1}{2} \int d\mathbf{x} d\mathbf{x}' U(\mathbf{x}, \mathbf{x}') \phi_{\alpha i}^*(\mathbf{x}) \phi_{\beta j}^*(\mathbf{x}') \phi_{\gamma k}(\mathbf{x}') \phi_{\delta l}(\mathbf{x}). \quad (3.23)$$

As mentioned earlier, I will apply the single-band approximation and assume that different energy bands do not couple in the interaction terms (hence the interaction between p_x, p_y states and for example s states will be assumed to be negligible). I will also neglect everything but *onsite interaction*. In the process, the interaction between different sites is ignored, which is appropriate if the system is far from a density instability and has a lattice spacing larger than the scattering length, which is referred to as the *atomic limit* [Aue94, p. 23]. Then the interaction is dominated by interactions on-site involving the density:

$$\sum_{i\alpha\beta\gamma\delta} U_{iiii}^{\alpha\beta\gamma\delta} \hat{a}_{i\alpha}^\dagger \hat{a}_{i\beta}^\dagger \hat{a}_{i\gamma} \hat{a}_{i\delta}, \quad (3.24)$$

3.2 Classical and quantum phase transitions

Phase transitions are properties of many-body systems which are defined in the *thermodynamic limit* $N \rightarrow \infty$: taking the limit of infinitely many particles while keeping density and other intensive properties in fixed proportion. Phase transitions are related to breaking of symmetries in a very general sense, and physical systems can be divided in abstract *universality classes* from the nature of their phase transitions.

Closed systems can be described by Hamiltonians, a mathematical structure shared by classical and quantum mechanics. A general Hamiltonian can be written in the form

$$\hat{H}(\{K\})/k_B T = - \sum_n K_n \hat{\Theta}_n, \quad (3.25)$$

where $\hat{\Theta}_n$ represent all possible combinations of the degrees of freedom, which may be all possible products of an infinite number of classical/quantum fields, or perhaps just the square of momenta of the particles in a non-interacting gas. K_n are called the *coupling constants* and represent such properties as temperature, strength of magnetic field or chemical potential. Note that the Hamiltonian can be viewed as a function of the coupling constants. The theory of phase transitions leads us to consider the space of coupling constants. Since this thesis is concerned with systems where the degrees of freedom reside on a lattice, let me restrict the discussion to such systems and let N denote the number of sites in the system. In classical mechanics [N.92, p. 23], the *partition function* $Z(\{K\}) = \text{Tr} e^{-H/k_B T}$ and the *free energy per site* is defined as

$$f_b(\{K\}) = \lim_{N \rightarrow \infty} F(\{K\})/N = \lim_{N \rightarrow \infty} (-k_B T \ln Z(\{K\}))/N, \quad (3.26)$$

where certain properties such as density are kept in a fixed proportion in the limiting process. This limit, called the *thermodynamic limit*, is not always well defined, but will be defined “almost everywhere”. A *phase* is a region in the space of coupling constants $\{K\}$, where the free energy per site is an analytic function. A phase may end in a point or set of points where the free energy becomes non-analytic. If this set of points forms a boundary between two regions of analyticity, it is called a phase transition. There are two mutually exclusive possibilities: either some first derivative $\partial f_b / \partial K_i$ is discontinuous across the boundary, or some higher derivative is⁴. The first case is referred to as a *discontinuous* or *first order* phase transition and the latter a *continuous* phase transition.

As an illustration of a classical phase transition, consider the Hamiltonian for the Ising model of a magnetic solid, can be written as

$$H = -J \sum_{\langle ij \rangle} s_i s_j + h \sum_i s_i, \quad (3.27)$$

where the sum is over all pairs of *nearest neighbor* sites in the lattice, which can in principle have any geometry or number of dimensions. I will assume it to be hypercubic lattice of dimension d . $s_i = \pm 1$ and the coupling constants are the *spin-spin interaction* J , the *external field* h and the temperature T . At $T = 0$,

⁴ f_b cannot have a discontinuity across the boundary because of the definition of the free energy. Later, I will instead work with other energies, which may be discontinuous at a phase transition.

the free energy per site f_b is just the internal energy per site by the thermodynamic relation [J.11, p. 108] $F = E - TS$. The solution will try to *minimize* the total energy (3.27). Assuming $J > 0$, the energy is minimized by a *polarized* or *ferromagnetic* configuration where

$$s_i = \begin{cases} +1 & (J > 0, h < 0) \\ -1 & (J > 0, h > 0), \end{cases} \quad (3.28)$$

for all sites i . In this thesis, the polarized/ferromagnetic and *anti-ferromagnetic* configuration $s_i = (-1)^i$ will often be encountered. The name *polarized* is often used when the configuration results from a uniform magnetic field like the second term in the Ising Hamiltonian (3.27). The name *ferromagnetic* should be properly used when referring to a state ordered by spin-spin interactions. The free energy per site is just

$$f_b = H/N = \begin{cases} (-J + h) & (J > 0, h < 0) \\ (-J - h) & (J > 0, h > 0), \end{cases} \quad (3.29)$$

which is continuous but has a *cuspl* (a sharp angle) at $h = 0$, indicating a discontinuous (first order) derivative and thus a discontinuous phase transition. The concept of an *order parameter* comes from Landau and Landau-Ginzburg theory [MB06, p. 83] and is a parameter that contains information about the nature of the phase and which changes continuously or discontinuously across a phase transition. A suitable order parameter for this phase transition is precisely the first derivative of the energy or the *magnetization* per site $M = \sum_i s_i/N$. At the transition, the order parameter *jumps* and the phase changes abruptly, which is typical for a first-order phase transition

$$M = \frac{\partial f_b}{\partial h} = \begin{cases} +1 & (J > 0, h < 0) \\ -1 & (J > 0, h > 0). \end{cases} \quad (3.30)$$

How should the spins align for $h = 0$? Both configurations have the same energy, so the system has to “choose” one of them - which is often referred to as a *spontaneous symmetry breaking* [J.11, p.193]. The Hamiltonian for $h = 0$ is completely symmetric with respect to the global transformation $s_j \rightarrow -s_j$, but the solution has a lower symmetry. In practice the “choice” depends on small perturbations and the “history” of the system.

At $T > 0$, the free energy also has contributions from disorder. It can be shown that creation of *domains*, regions of flipped spins, leads to a lower free energy if $d = 1$. Then the system will break up into a *paramagnetic* state of random domains for any $T \neq 0$. For $d \geq 2$ however, the energy cost of creating domains is positive for a certain $T < T_c$. Therefore, for a square or cubic lattice the ferromagnetic phase is stable for some nonzero temperatures. Close to T_c , the *critical temperature*, domain walls of flipped spins starts to form and eventually disorder the system completely. In this phase transition, the order parameter M changes continuously which is illustrated in figure 3.3. This is the sign of a continuous phase transition.

A useful concept for phase transitions is the *correlation length* ξ , defined as the characteristic rate of decay of the two-point correlation function [MB06, p. 82]

$$\langle s_i s_j \rangle = e^{-|i-j|/\xi}. \quad (3.31)$$

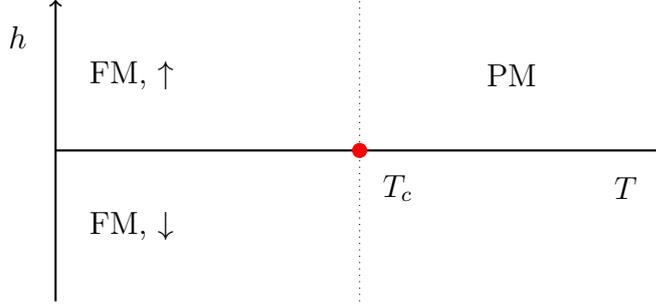


Figure 3.3: Phase diagram in the (h, T) plane for the classical Ising model in one dimension and $J > 0$. For $T < T_c$, the phase is ferromagnetic (FM), all spins pointing up or all down. There is a discontinuous phase transition and a spontaneous symmetry breaking on the line $h = 0$ and $T < T_c$. For $T > T_c$ the spins break up into domains until the phase is disordered (paramagnetic), which is a continuous phase transition.

As the Ising system approaches the critical coupling $T = T_c$, the correlation length can be shown to *diverge* and the two-point correlator no longer falls off exponentially with distance. The physical implications of this is profound: close to the phase transition, every site “communicates” with every other! This behaviour holds for any phase transition [N.92, p. 31]. The divergence of correlation length is, to leading order, asymptotically, $\xi \propto |T - T_c|^{-\nu}$, where ν is a *critical exponent* [N.92, p. 131]. The critical exponent is not an integer or even a rational number, and systems which on the surface appear entirely unrelated show the exact same critical exponents. The Ising model, for example contains both the Ising model and models with liquid-gas phase transitions. This phenomenon is referred to as *universality*. Universality is explained by the theory of the *renormalization group* (RG) [N.92, p. 236] [MB06, p.237], which considers the process of reducing the degrees of freedoms of a system (a procedure applied nearly everywhere in this thesis) in the abstract, as a semi-group transformation. Repeated application of such transformations leads to flows in the *space of all possible Hamiltonians* which partition them into universality classes.

The actual occurrence of a classical phase transition is possible because of *thermal fluctuations*. Above zero temperature there are always microscopic movements which acts as “seeds” from which new phases with lower energy can grow, if the system has passed a phase boundary. A *quantum phase transition* can be defined as a phase transition driven by *quantum fluctuations* instead. Since thermal fluctuations will always dominate for non-zero temperature, quantum phase transitions are never observed directly. However, the existence of a quantum phase transition at $T = 0$ is important for properties at $T \neq 0$.

As an example of a quantum phase transition, I choose the *quantum Ising model* in one dimension [S.11, p. 8], which has the Hamiltonian

$$\hat{H} = -J \sum_j \hat{S}_j^z \hat{S}_{j+1}^z + h \sum_j \hat{S}_j^x, \quad (3.32)$$

where $\hat{S}_j^\alpha = (\hbar/2) \hat{\sigma}_{\alpha i}$ are the Pauli matrices. In the basis where \hat{S}_j^z is diagonal they can be written as

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (3.33)$$

Assume $J > 0$ and $h \ll J$. In this case the ground state should be close to the product states $\prod_j |\uparrow\rangle_j$ or $\prod_j |\downarrow\rangle_j$, much like in the classical case. It is important to remember that this is a quantum mechanical problem and the full ground state has to be obtained through diagonalization of the Hamiltonian. The product states are exact ground states for the first term in the Hamiltonian since

$$-J\hat{\sigma}_{zj}\hat{\sigma}_{z(j+1)}|\uparrow\rangle_j|\uparrow\rangle_{j+1} = -J(\pm 1)^2|\uparrow\rangle_j|\uparrow\rangle_{j+1}. \quad (3.34)$$

Because the up and down states have the same energy, there will again be a spontaneous symmetry breaking and the solution will have lower symmetry than the Hamiltonian.

Assume instead $J > 0, h \gg J$. In the $\hat{\sigma}_x$ basis, depending on the sign of h , the exact solution of the Hamiltonian is the *Néel states* $\prod_j |\leftarrow\rangle$ or $\prod_j |\rightarrow\rangle$ where $|\leftrightarrow\rangle = 2^{-1/2}(|\uparrow\rangle \pm |\downarrow\rangle)$. In between, there has to be some kind of phase transition. In section 7.2 I will use a mean-field approximation to show that this is a second-order quantum phase transition⁵. The quantum Ising model also has a (classical) phase transition to a disordered state. By *Mermin and Wagner's theorem*[Aue94, p. 62], this transition only occurs at a finite temperature for $d \geq 3$.

3.3 The Bose-Einstein Condensate

The Hamiltonian (3.4) is useful to describe any non-interacting many-body system. Since bosonic atoms can occupy the same single-atom state, the ground state at zero temperature of a many-body system of non-interacting bosons can be assumed to be a product state of identical single-atom ground states. This state is called a *Bose-Einstein condensate* (BEC) and is an important bosonic many-body system. It is described here to illustrate many of the properties which will be investigated in this thesis. The density distribution $n(\mathbf{r}) = N |\phi(\mathbf{r})|^2$, where N is the number of bosons, of the condensate reflects the single-particle wavefunction squared.

For finite temperature, thermal fluctuations will excite atoms from the ground state. The *critical temperature* is defined as the highest temperature for which a large fraction of the atoms occupy the ground state. It can be determined by considering the the number of particles in excited states, which is given by [PS08, p. 22]

$$N_{ex} = \int_0^\infty d\epsilon g(\epsilon) f^0(\epsilon), \quad (3.35)$$

where $g(\epsilon)$ is the density of states and $f^0(\epsilon)$ the mean occupation number given by the Bose distribution as

$$f^0(\epsilon_\nu) = \frac{1}{e^{(\epsilon_\nu - \mu)/k_B T} - 1}. \quad (3.36)$$

Equating the number of excited particles with the total number of particles $N = N_{ex}(T_c)$ at $\mu = 0$ gives the condition for the critical temperature. For a Bose gas trapped in a parabolic potential $V(x, y, z) = \frac{1}{2}(\omega_x x^2 + \omega_y y^2 + \omega_z z^2)$, the density of states is found to be proportional to the area of the Fermi surface [PS08, p. 21]

$$g(\epsilon) = \frac{\epsilon^2}{2\hbar^3 \bar{\omega}^3}, \quad (3.37)$$

⁵In fact, the one dimensional quantum Ising model can be solved exactly for any value of g by a *Jordan-Wigner transformation*. [S.11, p. 46].

where $\bar{\omega} = (\omega_x \omega_y \omega_z)^{1/3}$. Equation (3.35) can be rewritten by a variable change $x = \epsilon/kT_c$ as

$$N = \frac{(kT_c)^3}{2\hbar^3 \bar{\omega}^3} \int_0^\infty dx \frac{x^2}{e^x - 1}. \quad (3.38)$$

By expanding the denominator in powers of x , the integral can be computed as

$$\begin{aligned} \int_0^\infty dx \frac{x^2}{e^x - 1} &= \int_0^\infty dx e^{-x} x^2 \sum_{n=0}^\infty e^{-xn} = \sum_{n=0}^\infty \frac{1}{n^2} \int_0^\infty dx (nx)^2 e^{-(n+1)x} = \\ &= \Gamma(3) \sum_{n=1}^\infty \frac{1}{n^2} = \Gamma(3) \frac{\pi^2}{6}, \end{aligned} \quad (3.39)$$

where the known series for $\pi^2/6$ and an integral representation of the Gamma function were used. For (3.35) I obtain

$$N = \frac{(kT_c)^3}{2\hbar\bar{\omega}^3} \Gamma(3) \frac{\pi^2}{6} \Leftrightarrow kT_c = \left(\frac{12N}{\pi^2} \right)^{1/3} \hbar\bar{\omega} \approx 0.94 \hbar\bar{\omega} N^{1/3}. \quad (3.40)$$

This expression shows that large number of particles and a tight confining potential lead to higher transition temperatures. Figure 3.4 shows the density distribution of a dilute gas for various temperatures, measured by time-of-flight techniques as discussed in section 4.4.

3.4 The Bose-Hubbard Model

A model which has been seminal for developing intuition about many-body systems is the *Hubbard model*, which is best viewed as an approximation of an interacting many-body system, and the first model described in this thesis which does not neglect the particle-particle interaction. For Bose particles which are chargeless, the interaction term in the Hamiltonian simplifies to a *contact interaction* [PS08, p. 119] $U(\mathbf{x}_i, \mathbf{x}_j) = U_0 \delta^{(3)}(\mathbf{x}_i - \mathbf{x}_j)$. U_0 was already defined in section 2.5. While the Hubbard model in itself starts from a second-quantized Hamiltonian, which can result from many physical processes, this thesis is concerned with realizations in optical lattices. For cold atoms in the s band, when a single-band approximation and a tight-binding approximation (defined below) are applicable, we obtain the Bose-Hubbard model. As described in sections 3.1.4 and 3.1.5, for such systems a Wannier basis is suitable. In the Wannier basis, the interaction is *on site*, because of the tight-binding approximation. The interaction strength is

$$\begin{aligned} U_{iiii} &= \frac{1}{2} \int d\mathbf{x} d\mathbf{x}' U_0 \delta^{(3)}(\mathbf{x} - \mathbf{x}') |\psi_{\mathbf{x}_i}(\mathbf{x})|^2 |\psi_{\mathbf{x}_i}(\mathbf{x}')|^2 = \\ &= \frac{U_0}{2} \int d\mathbf{x} |\psi_{\mathbf{x}_i}(\mathbf{x})|^4 = U > 0, \end{aligned} \quad (3.41)$$

which should be interpreted as an effective interaction. Note that $\psi_{\mathbf{x}_j}(\mathbf{x})$ here are the separable Wannier functions in the s band. The Bose-Hubbard Hamiltonian can be written as

$$\hat{H}_{\text{Hubbard}} = -t \sum_{\langle ij \rangle} (\hat{a}_i^\dagger \hat{a}_j + \hat{a}_j^\dagger \hat{a}_i) + U \sum_{i=0}^{N-1} \hat{n}_i (\hat{n}_i - 1) - \mu \sum_{i=0}^{N-1} \hat{n}_i. \quad (3.42)$$

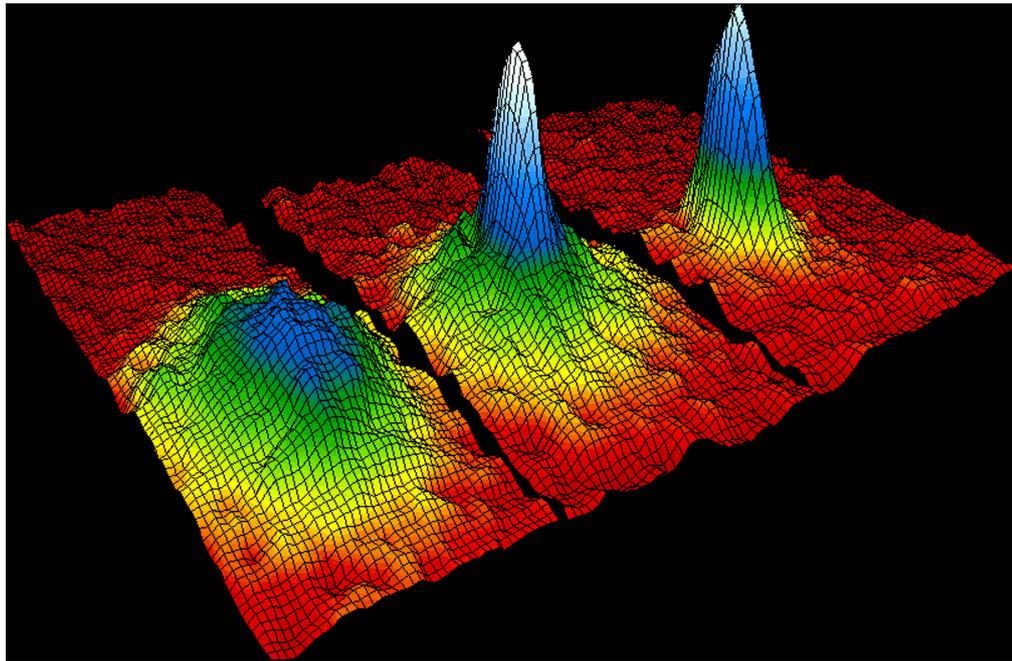


Figure 3.4: Bose-Einstein condensation manifest in the velocity density distribution $n(\mathbf{k})$ of a dilute Bose gas. The leftmost picture is taken above the critical condensation temperature and shows only a thermal density distribution. The middle and right pictures are for a condensed state: there, a BEC state is visible in middle picture as a sharp peak surrounded by thermal fluctuations. In the rightmost picture, temperature is low enough that only the BEC remains. Source: “Bose Einstein condensate” by NIST/JILA/CU-Boulder - NIST Image. Licensed under Public Domain via Commons ([link.](#))

The notation $\langle ij \rangle$ stands for a sum over the nearest neighbors only, and all other tunneling coefficients are neglected. Such a truncation is often called a *tight-binding approximation* and will be used frequently in this thesis. The tunneling t is the same between all sites in this model and the lattice is taken to be linear or quadratic. I have also introduced a chemical potential μ as a way of studying the effect of particle number. This corresponds to a Legendre transform of the Hamiltonian, analogous to the transformation to grand free energy in statistical mechanics.

The Bose-Hubbard model has a quantum phase transition occurring at critical values of t/U and μ/U [BDZ08, p. 23][DCI⁺98]. The hopping term, proportional to t , favors *delocalization* or a Bloch-wave-like state. Because of this delocalization, the system has a long correlation length. This can be seen by noting that for $U = 0$, the Hamiltonian is periodic in the lattice. It is therefore natural to look for solutions with the same periodicity, like Bloch waves. By a Fourier transformation, it is easy to find this solution. The relevant excitations in this limit are therefore delocalized, and in particular the spectrum is gapless which is a characteristic for a symmetry broken phase. This limit, where $t/U \gg 1$ is called the *superfluid* (SF) phase and is really a Bose-Einstein condensed phase since all bosons will tend to occupy the same wavelike state.

In the opposite limit of $t/U \ll 1$, the interaction term dominates and suppresses tunneling between sites. Note that the interaction will be minimized term by term if the occupation number at each site is zero or one. The interaction generally penalizes high occupation numbers and spatially varying densities, since the individual terms grow as $\sim n^2$ with occupation number and not with n . For $t/U \ll 1$ then, the bosons are localized and confined to each site with uniform density. Correlations are short-ranged because of the localization. This phase is called the *Mott insulator* phase, since the flow of bosons is halted⁶. Because the individual bosons are localized at each site, the Mott phase can only have *integer occupation* $n = 1, 2, 3, \dots$. There are therefore a series of Mott phases with different number of bosons per site, denoted as $\text{Mott}_1, \text{Mott}_2, \text{Mott}_3 \dots$. Adding a single particle costs a finite energy, which may be minimized by delocalizing it, effectively creating a superfluid state “on top of” the Mott phase. The important point is that this energy cost remains finite in the thermodynamic limit, $N \rightarrow \infty$, keeping the density n/N constant. In condensed matter lingo: the phase is *gapped*.

The interaction part of the energy difference per site is roughly $\Delta U \approx U((n+1)^2 - n^2)/N \sim Un/N$ which is constant. An increasing chemical potential will therefore not increase particle number inside the Mott phase, $\frac{\partial}{\partial \mu} \langle \hat{n}_j \rangle = 0$ and the Mott phase is said to be *incompressible* [MAV12, p. 50]. At the phase boundary to the superfluid phase, the energy difference of adding a particle vanishes in the thermodynamic limit (the energy gap closes), and there is a continuous quantum phase transition to the superfluid phase where $\frac{\partial}{\partial \mu} \langle \hat{n}_j \rangle \neq 0$. Hence the phase is gapless and the Hamiltonian is invariant under a global U(1) transformation: the corresponding *Goldstone* modes are long-wavelength phase fluctuations. Goldstone modes are the low-energy excitations of a continuous symmetry. A suitable set of order parameters for the phase transition are the expectation number $\langle n \rangle_j$: these also change continuously in the transition.

Figure 3.5 shows a sketch of the phase diagram [BDZ08, p. 23]. Since occupation is regulated by the chemical potential, the Mott phases form “lobes” in

⁶“Normal” insulation of electronic currents is related to a different mechanism.

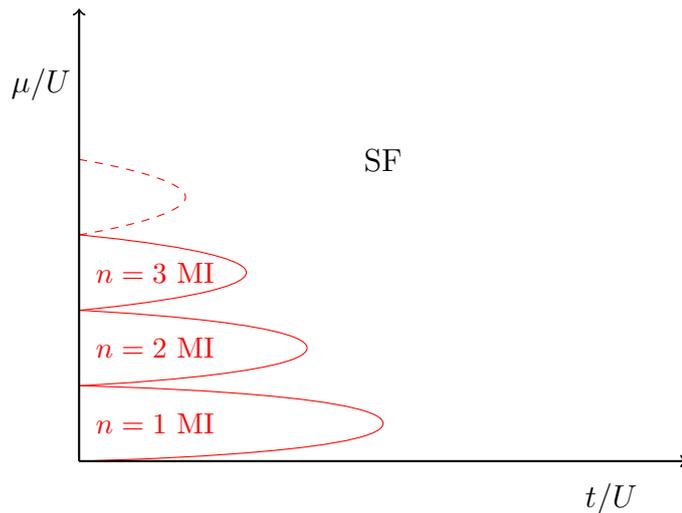


Figure 3.5: Sketch of a phase diagram for the Hubbard model. The red “lobes” are regions in a Mott phase with integer fillings. The superfluid (SF) or BEC phase in between and for higher values of t/U has a continuously varying occupation number.

the $(t/U, \mu/U)$ -plane. In one dimension, the transition is of *Berezinskii-Kosterlitz-Thouless* type, in higher dimensions, continuous.

In cold atom realizations of the Bose-Hubbard model, the atoms are always confined by a potential, which is wide enough to vary slowly over the scale of the optical lattice spacing. In the *local density approximation* [PSanB⁺13], the potential is viewed as a local shift in the chemical potential $V(\mathbf{x}) \approx -\sum_j \mu(j) \hat{n}_j$, which is permissible since the spatial variation between sites is negligible. At the center of the trap, the chemical potential is large, favoring a large occupation number, while $\mu(j)$ is small at the edges of the condensate. Radially, the system then describes vertical line through the phase diagram 3.5 with concentric regions of Mott phases with integer density, with a compressible superfluid phase in between. The occupation number as a function of μ forms integer plateaus in the Mott phases, with the SF phase interpolating in between. An image of this “wedding cake” structure is shown in figure 4.2.

Chapter 4

Optical lattices

When it comes to atoms, language
can be used only as in poetry.

Niels Bohr

In this chapter, I describe the experimental techniques involved in optical lattices. The description is useful both for understanding optical lattices and for understanding the constraints that the theoretical physicist has to work with.

4.1 Trapping

Atoms are electrically neutral but can interact with electric and magnetic fields through the distribution of its electrons. The simplest example of this is the electrical dipole moment \mathbf{d} . The magnitude of this vector describes the degree to which the atom has a positive and negative pole and the direction describes the orientation of the poles. In an external electric field, the resulting *dipole interaction* [PS08, p. 51] can be written as

$$H_d = -\mathbf{d} \cdot \vec{\mathcal{E}}, \quad (4.1)$$

where $\vec{\mathcal{E}}$ is the external electric field. A spatial variation of the electrical field shifts the ground-state energy of the atoms differently at different positions, leading to a lower ground state energy at certain positions. This effectively acts as a potential, pushing the atom states into minima or maxima of the field, depending on the sign of the dipole vector. By perturbation theory, the shift in the ground state energy E_g for a weak oscillating field $\vec{\mathcal{E}}(\mathbf{x}, t) = \vec{\mathcal{E}}_\omega e^{-i\omega t} + \vec{\mathcal{E}}_{-\omega} e^{i\omega t}$ is [Sak94, p. 285]

$$\begin{aligned} \Delta E_g &\approx \sum_e \langle g | \mathbf{d} \cdot \vec{\mathcal{E}}_\omega | e \rangle \frac{1}{E_g - E_e + \hbar\omega} \langle e | \mathbf{d} \cdot \vec{\mathcal{E}}_{-\omega} | g \rangle + \\ &\quad + \sum_e \langle g | \mathbf{d} \cdot \vec{\mathcal{E}}_{-\omega} | e \rangle \frac{1}{E_g - E_e - \hbar\omega} \langle e | \mathbf{d} \cdot \vec{\mathcal{E}}_\omega | g \rangle = \\ &= \sum_e |\langle e | \mathbf{d} \cdot \vec{\mathcal{E}} | g \rangle|^2 \left(\frac{1}{E_g - E_e - \hbar\omega} + \frac{1}{E_g - E_e + \hbar\omega} \right) |\mathcal{E}_\omega|^2 = \\ &= -\frac{1}{2} \alpha(\omega) \langle \mathcal{E}(\mathbf{x}, t)^2 \rangle_t, \end{aligned} \quad (4.2)$$

where the angle brackets denote a time average of the electrical field, ω is the frequency of oscillation, $|e\rangle, |g\rangle$ denotes the first excited and ground states and E_e, E_g the energies of these states. ΔE_g is sometimes called the *alternating-current Stark shift*. Note that ΔE_g has a spatial variation. The atom, “dressed” in its perturbations by the oscillating field, experiences a force towards the regions where ΔE_g is minimal. To create a standing-wave potential, it suffices to superimpose two lasers in opposite directions, for example by reflecting on a mirror:

$$\mathcal{E} = \mathcal{E}_0 \hat{\mathbf{z}} [\cos(qx - \omega t) + \cos(-qx - \omega t)], \quad (4.3)$$

$$\langle E^2 \rangle_t = \int_0^{2\pi/\omega} dt \mathcal{E}(x, t)^2 = 2\mathcal{E}_0^2 \cos^2(qx), \quad (4.4)$$

$$\Rightarrow V(\mathbf{x}) = V_0 \cos^2(\pi x/d), \quad (4.5)$$

for $V_0 = -\alpha(\omega)\mathcal{E}_0^2$, $d = \pi/q = \lambda/2$. This potential varies spatially in the x -direction. By superposition of such standing waves, it is possible to create many kinds of periodic potentials in one, two or three dimensions.

The *Zeeman effect* is superficially similar to the Stark effect, but due to a *magnetic* field, which shifts the energy levels of the atom [PS08, p. 45]. A spatial variation of the energy levels acts as an effective potential for the atoms. By creating a region with lower strength in the magnetic field it is possible to trap atoms magnetically¹. Magnetic trapping potentials are often symmetric and weak enough that they can be described as roughly harmonic $V(\mathbf{x}) \approx V_0(1 + \omega_x x^2 + \omega_y y^2 + \omega_z z^2)$. It is important to note that even if the optical lattice is perfectly regular, the magnetic trap distorts it somewhat. Unless tunneling is completely suppressed, this leads to a spatial structure known as the “wedding cake” of concentric layers of isolating and superfluid phases [PS08, p. 439] [BDZ08, p. 24]. An image of such a structure in an optical lattice is shown in figure 4.2, see section 4.4 for details of how the image was created. The effect has already been discussed in section 3.4.

4.2 Cooling

The lattice potential will only be relevant if the thermal excitation energy of the atoms is less than the potential barrier height V_0 . Historically, optical lattice systems could not be realized until efficient cooling techniques were developed. The cooling is done in steps, with more and more refined techniques being used in later stages. First, a beam of atoms is emitted from a hot oven and slowed down from 700 K to 1 K by applying a laser beam in the opposite direction [PS08, p. 78] [Mey01, p. 27]. Atoms spontaneously absorb and then re-emit photons in random directions. This irreversible process may be seen as an entropy pump which transfer heat from the atoms by the randomly re-emitted radiation².

Then, a magneto-optical trap (MOT) traps the atoms through an EM field created from magnetic fields and lasers. Further laser cooling slows down the atoms to

¹According to the mean value theorem, it is not possible to create a local extremum of a static magnetic field \mathbf{B} . It is however possible to create a local extremum of the absolute field strength $|\mathbf{B}|$, and such a magnetic potential can trap atoms with a negative magnetic moment [PS08, p. 62]

²The lowered temperature of course also leads to lower average speeds for the atoms. Such a deceleration leads to a Doppler effect in the beam, which means that transition frequencies are different along the beam as the atoms are successively cooled. This is remedied by, e.g., use of the Zeeman effect which shifts the transition frequencies so as to counteract the Doppler shift.

100 μK . At this stage they can be confined by a pure magnetic trap. The atoms are then cooled by further techniques (such as *evaporative cooling* or *Sisyphos cooling*, [Mey01, p. 30]) until $k_B T \sim E_r$. Lasers and mirrors can then be used to create the optical lattice, as described in the last section.

4.3 Relevant scales

The optical lattice system has some characteristic energy scales for trapping and interactions [PS08, p. 407]. The *recoil energy* is defined as $E_r = 2\hbar^2\pi^2/m\ell^2 = (\hbar q)^2/2m$, which may be interpreted in two separate ways: either as the energy of a particle trapped in a box potential of length ℓ (the lattice spacing), or as the energy which is transferred when a light quantum is absorbed by an atom. The second interpretation comes from considering the “naked” atom interacting with the photon field, while the first comes from considering the “dressed” atom in an effective potential from the Stark shift of the optical lattice. Often the lattice potential has an amplitude $V_0 \sim 1 - 30E_r$. V_0/E_r small leads to a weak effective potential, with a higher rate of tunneling between potential wells. The tunneling can be expected to become “weak” if the width of the near-isolated wavefunctions at each site is smaller than the lattice spacing. In this situation the overlap of wavefunctions on neighboring sites is small [PS08, p. 408].

To first order, a well-isolated wavefunction can be considered Gaussian, since $V(x) = V_0 \cos(2\pi x/\ell) \approx (V_0/2)(1 - (2\pi x/\ell)^2)$, where ℓ is the lattice spacing. This potential is parabolic and has solutions in the form of a Gaussian multiplied by a Hermite polynomial. Since $H_0 = 1$, to a first approximation the ground state wavefunction is Gaussian with a characteristic width³

$$a_{\text{osc}} = \left(\frac{\hbar}{m\omega_{\text{osc}}} \right)^{1/2}, \quad (4.6)$$

where the oscillations in reciprocal space have a frequency

$$\omega_{\text{osc}} = \left(\frac{1}{m} \frac{d^2 V}{dx^2} \right)^{1/2} \approx \left(\frac{2V_0}{m} \right)^{1/2} \frac{\pi}{\ell}, \quad (4.7)$$

Combining the above two equations, I obtain

$$\frac{a_{\text{osc}}^2}{\ell^2} = \frac{1}{\pi^2} \left(\frac{V_0}{E_r} \right)^{-1/2}. \quad (4.8)$$

For ^{87}Rb and a 852 nm laser, the condition that the left hand is small is given by $V_0 > 10E_r$.

The effective two-body interaction energy [PS08, p. 408] determines the types of processes which will be dominant in the lattice. In terms of the recoil energy it is

$$E_{\text{int}}/E_r = \frac{8}{\pi} n a \ell^2, \quad (4.9)$$

where a is the *scattering length* and n is the atom density. The scattering length gives the intercept of an asymptotic approximation of the wavefunction of a scattered particle at large separations and low scattering energies. The wavefunction

³This *harmonic oscillator* approximation of the wavefunction will be described in section 5.1.2.

of two particles one scattering length apart, has a wavelength comparable to that separation. At smaller distances, the wavefunction oscillates more rapidly, while it becomes that of two separate atoms for large separations [PS08, p. 111, p. 115]. a is of order 10^{-6} cm while $\ell \sim 10^{-4}$ cm and the atom density $n = 10^{13} - 10^{15}$ cm $^{-3}$, so that $E_{\text{int}} = 10^{-3} - 10^{-5} E_r$. Since three-body interactions are characterized by energies with leading terms which are higher order in a , it is usually safe to ignore anything beyond two-body interactions in optical lattices [PS08, p. 109].

4.4 Measurement and manipulation

An important measurement technique in optical lattices is *time-of-flight measurements* [BDZ08, p. 14]. If the trapping potential and optical lattice are suddenly turned off, the atoms will not have time to change adiabatically. Because it is no longer confined, the cloud will experience free fall and expand from its trapped state. Neglecting any interaction during the fall, the expansion will be *ballistic*, so that $\hbar\mathbf{k}t = M\mathbf{x}$ in the reference frame of the cloud. Because of this relation, the density distribution of the expanding cloud will be proportional to the momentum distribution in the lattice. A Bloch state with crystal momentum q can be described under expansion by a Fourier expansion of plane waves with momenta $p_n = \hbar q + 2n\hbar k$, where n is an integer, q is the crystal momentum of the lattice and k is a primitive vector of the reciprocal lattice. This is due to the fact that Bloch states consist of Fourier components periodic in the lattice multiplied by plane waves with the crystal momenta. After letting the cloud expand for a certain time (the “time of flight”), the density distribution is measured using absorption imaging (measuring the absorption of a laser through the cloud by a *CCD chip*). The resulting data is a density profile $n(\mathbf{x})$ which (under the current approximations) is given by

$$n(\mathbf{x}) = \langle \hat{n}(\mathbf{x}) \rangle_{\text{tof}} = \langle \hat{a}_{\text{tof}}^\dagger(\mathbf{x}) \hat{a}_{\text{tof}}(\mathbf{x}) \rangle \approx \langle \hat{a}^\dagger(\mathbf{k}) \hat{a}(\mathbf{k}) \rangle_{\text{trap}} = \langle \hat{n}(\mathbf{k}) \rangle_{\text{trap}}, \quad (4.10)$$

where $\hat{a}_{\text{tof}}(\mathbf{k})$ annihilates a state in mode \mathbf{k} of the cloud at the time of measurement, and $\hat{a}_{\text{trap}}(\mathbf{x})$ annihilates a state at position \mathbf{x} before expansion. (4.10) shows that the expectation value of the number operator in the trap (before expansion) can be approximated by the expectation value of the number operator after the expansion, which again is due to the ballistic expansion of the cloud. Momentum-momentum correlations also manifest themselves in noise correlations of the cloud:

$$\begin{aligned} \langle \hat{n}_\rho(\mathbf{x}) \hat{n}_\rho(\mathbf{x}') \rangle_{\text{tof}} &\approx \langle \hat{a}^\dagger(\mathbf{k}) \hat{a}(\mathbf{k}) \hat{a}^\dagger(\mathbf{k}') \hat{a}(\mathbf{k}') \rangle_{\text{trap}} = \\ &= \langle \hat{a}^\dagger(\mathbf{k}) \hat{a}^\dagger(\mathbf{k}') \hat{a}(\mathbf{k}') \hat{a}(\mathbf{k}) \rangle_{\text{trap}} + \delta_{\mathbf{k}\mathbf{k}'} \langle \hat{n}(\mathbf{k}) \rangle_{\text{trap}}. \end{aligned} \quad (4.11)$$

Figure 4.1 shows absorption images of a series of time-of-flight measurements for increasing lattice depth which is a striking example of the usefulness of this measurement technique [BDZ08, p. 24]. With the potential off, all atoms lie close to the zero momentum state (at the origin in the figure). As the potential is gradually increased, interference peaks become visible around $\mathbf{k} = 0$. The peaks are related to the long-range order in the SF phase. The momentum density distribution (related to the density distribution in figure 4.1 through $\hbar\mathbf{k}t = M\mathbf{x}$) is given by

$$n(\mathbf{k}) \propto |\tilde{\psi}(\mathbf{k})|^2 \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} G^{(1)}(\mathbf{R}) \quad (4.12)$$

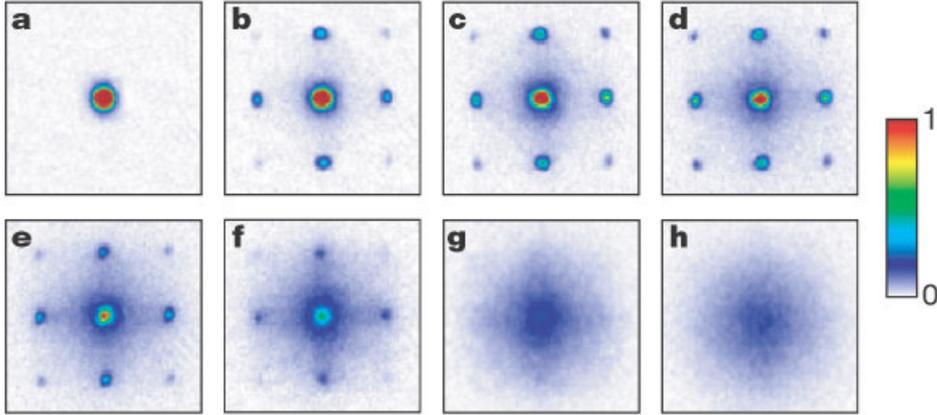


Figure 4.1: Absorption imaging of a cubic lattice with amplitude $V_0 = 0, 3E_r, 7E_r, 10E_r, 13E_r, 14E_r, 14E_r, 16E_r$ and $20E_r$ in order from left to right, top to bottom. Source: [BDZ08].

where $\tilde{\psi}(\mathbf{k})$ is the Fourier transform of the Wannier function of the atoms, $\{\mathbf{R}\}$ are the full set of lattice vectors and $G^{(1)}(\mathbf{R})$ is the *one-particle density matrix* which is related to the correlation function at radius \mathbf{R} . As mentioned in section 3.4, the superfluid phase is characterized by long-range correlations so that $\lim_{|\mathbf{R}| \rightarrow \infty} G^{(1)}(\mathbf{R}) \neq 0$. For $V_0 = 0$, the Fourier transformation of the Wannier function should be almost like a Dirac delta function, since only the long-wavelength Bloch states are occupied. For $V_0 \neq 0$ but less than some critical value, the Fourier transform is non-zero in a larger region, around $\mathbf{k} = 0$ since the particles have a somewhat higher density at lattice minima but are still not fully localized. $G^{(1)}(\mathbf{R})$ also has a peak close to the origin reflecting short-range correlations, but for \mathbf{R} large enough it will approach a constant. The sum in (4.12) will then look much like a delta distribution if the one-particle density matrix approaches its constant limit fast enough. Qualitatively then, one can understand that the momentum distribution will have peaks at $\mathbf{k} \cdot \mathbf{R} = 2\pi n$ for integers n , which leads to the observed interference peaks. For large enough lattice depth, the system leaves the BEC phase and undergoes a phase transition to the Mott phase where atoms are isolated in their sites⁴. Since there is no long-range correlation in the Mott phase, $G^{(1)}(\mathbf{R})$ goes to zero exponentially in \mathbf{R} , leaving only a peaked Gaussian around $\mathbf{k} = 0$. The Gaussian distribution in the last sub-figure is the Fourier transform of the Gaussian shape of the atom wavefunctions at each site.

A useful tool for manipulation is *stimulated Raman transitions*. With the help of two counter-propagating lasers, it is possible to couple the populations of two atomic levels such as the s and p bands. Each laser will drive a cyclic *Rabi process*, consisting of absorption and re-emission of photons, tuned in such a way as to couple the s and p bands with a third state, far detuned from the rest of the system. The combined process leads to a Raman coupling between s and p states which constitutes an inelastic scattering process (since the atoms absorb and re-emit photons of different frequency). Stimulated Raman transitions was used in [TSAI07] to load atoms in the p -band of a one-dimensional optical lattice. By tuning the two laser beams with

⁴The Mott phase can thus be reached in two ways: by increasing the onsite interaction strength and by increasing the lattice depth.

wavevectors $\mathbf{k}_1, \mathbf{k}_2$, they obtained a spatially dependent effective Rabi frequency

$$\Omega_{sp, \text{eff}} = \frac{\Omega_1 \Omega_2^*}{2\Delta} \left| \langle 2 | e^{i(\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{x}} | 1 \rangle \right|^2, \quad (4.13)$$

where Ω_1, Ω_2 denotes the Rabi frequencies of the two lasers and Δ the detuning of the lasers to the atomic transitions which they drive. $|1\rangle, |2\rangle$ denote two Raman levels of a ^{87}Rb atom, which is held isolated in a single site during loading by a strong potential $V_0 \sim 40E_r$, which is lowered after the loading process. The authors found that, for lattice depths of $V_0 \leq 25E_r$, such p -band lattices had long lifetimes, on the order of hundreds of characteristic tunneling times. The system then has time to “delocalize” and reach a stable state before decaying. If not, there would be no p -band orbital physics to talk of, at least not any equilibrium. Raman processes may also be used to tune the value of coupling constants via *laser-assisted tunneling*. Since the tunnelings are functionally dependent on the potential and the Wannier functions, it is also possible to change the tunnelings by directly modifying the potential. Consider a part of the expression for the tunneling in the separable square lattice, in the harmonic oscillator approximation

$$V_0 \int d\mathbf{x} (x - X_i)^\alpha (x - X_j)^\beta \sin^2(\pi x) e^{-\frac{1}{2}\omega(\mathbf{x} - \mathbf{X}_i)^2 - \frac{1}{2}\omega(\mathbf{x} - \mathbf{X}_j)^2}. \quad (4.14)$$

By the coordinate transformation $\mathbf{x} \rightarrow \mathbf{y} = \mathbf{x} + (\mathbf{X}_i + \mathbf{X}_j)/2 = \mathbf{x} + \tilde{\mathbf{X}}$, the integral can be rewritten as

$$\begin{aligned} & \frac{V_0}{2} e^{-\omega\Delta^2/4} \int d\mathbf{y} (y^\alpha y^\beta - \Delta^\alpha \Delta^\beta / 4) (1 - \cos(2\pi(y - \tilde{X}))) e^{-\omega\mathbf{y}^2} = \\ & = \frac{V_0}{2} e^{-\omega\Delta^2/4} \int d\mathbf{y} (y^\alpha y^\beta - \Delta^\alpha \Delta^\beta / 4) (1 - \cos(2\pi\tilde{X}) \cos(2\pi y)) e^{-\omega\mathbf{y}^2}. \end{aligned} \quad (4.15)$$

Since [II15] $\int_0^\infty x^n e^{-\omega x^2} dx \sim \omega^{-(n+1)/2}$ and ω is large (it has the dimension of length $^{-2}$), the largest term in this integral will be proportional to $\Delta^\alpha \Delta^\beta$ and some power of V_0 . By squeezing, the distance between sites $\Delta = \mathbf{X}_i - \mathbf{X}_j$ may be changed. The shape of the potential further affects the value of the tunneling. While this is only an approximation and a particular kind of lattice, the argument holds in general. In section 5.2.3, this is shown by computing the tunnelings for zig-zag lattices of varying shapes.

Feshbach resonances [PS08, p. 143, p. 148] may be used to tune the effective interactions between atoms. It is even possible to make atom-atom interactions attractive. This has been tried in an experiment and led to the collapse of the condensate [PS08, p. 151]! In chapter 8, it will be shown that physics in the Mott phase can be described by an effective Hamiltonian, the coupling constants of which depend on the tunnelings and interaction terms, the strengths of which may be varied through Feshbach resonance.

One of the most important features of the optical lattice technology is the ability to control individual atoms in the lattice. Such *single-site addressing* involves using lenses to focus laser beams onto a specific lattice site [TSAI07]. Single-site imaging involves stimulating fluorescent emissions from one site at a time, which are then captured by a *CCD chip* (i.e. a camera). A direct image of an optical lattice can then be composed from the individual pieces for each site. An interesting limitation in single-site imaging is that the laser pulse ejects atoms in pairs from each site, so

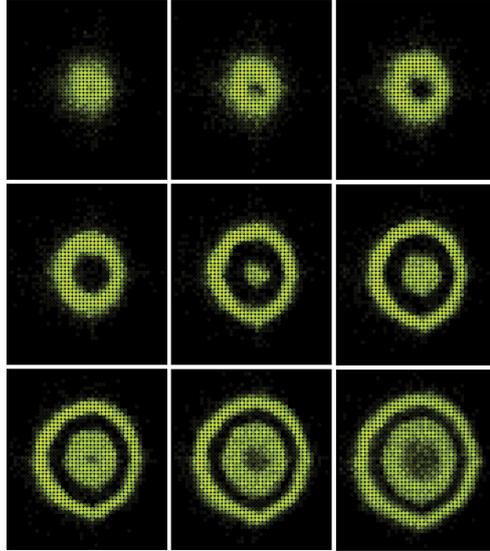


Figure 4.2: Single-site images of a wedding-cake structure as described in section 3.4. The pixels show occupation number modulo 2 of each lattice site in a quadratic lattice. Source: the Greiner group website ([link](#))

that only information about the parity (even or odd) can be gathered. The technique may be seen as a projection of the many-particle wavefunction onto number states of a site [WAM⁺10]. Figure 4.2 has a number of pictures compiled with single-site imaging, showing the *wedding cake structure* mentioned in section 3.4. Regions with exact integral occupation number appear in black and green concentric circles. In principle there should be a superfluid phase “interpolating” between the Mott levels, but here the number of atoms is too small for the effect to be seen. If the number of atoms were large enough, regions with some odd and some even occupation would appear between the Mott phases (remember that only eigenstates of the number operator can be measured!).

Chapter 5

From potentials to overlap integrals

The only simplicity to be trusted is the simplicity to be found on the far side of complexity.

Alfred North Whitehead

A quantum mechanical many-body problem can be said to be specified by the potential function, the particle masses and the interaction potential between the particles. There may also be some internal states of the particles. In the process of second quantization, this information is encoded in the *coupling constants* of the problems - the coefficients which sets the strength of each term in the second-quantized Hamiltonian. In this chapter, I will describe and perform the procedure of how to derive them for two types of optical lattice potentials: the square lattice and the zig-zag lattice. The resulting many-body Hamiltonian for the second potential is the subject of the rest of this thesis.

5.1 The square lattice

5.1.1 Wannier functions in the square lattice

The square lattice potential is, in rescaled variables, defined as

$$V(x, y) = V_0 \left(\sin^2(\pi x) + \sin^2(\pi y) \right). \quad (5.1)$$

This potential is *separable*, since it is a sum of functions of each coordinate. This holds in any number of dimensions for a hypercubic lattice. For a separable potential, the single-particle Schrödinger equation is reduced to a one-dimensional problem by the ansatz¹ $\psi(\mathbf{x}) = \psi_x(x)\psi_y(y)$. The Schrödinger equation, in its rescaled form, is

$$\left(-\partial_x^2 - \partial_y^2 + V_0 \sin^2(\pi x) + V_0 \sin^2(\pi y) \right) \psi(\mathbf{x}) = E\psi(\mathbf{x}), \quad (5.2)$$

¹This also implies that the wavefunction is *constant* in z , and all dependence on this coordinate disappears from the Schrödinger equation, which can therefore be treated as two-dimensional.

which can be separated into two *Mathieu equations* as

$$\begin{aligned} (-\partial_x^2 - \partial_y^2 + V_0 \sin^2(\pi x) + V_0 \sin^2(\pi y)) \psi_x(x) \psi_y(y) &= (E_x + E_y) \psi_x(x) \psi_y(y) \\ \Rightarrow \begin{cases} (-\partial_x^2 + V_0 \sin^2(\pi x)) \psi_x(x) = E_x \psi_x(x) \\ (-\partial_y^2 + V_0 \sin^2(\pi y)) \psi_y(y) = E_y \psi_y(y) \end{cases} &, \quad (5.3) \end{aligned}$$

which can be converted into the standard form by the substitution $\pi x = z$:

$$\psi'' + [a - 2q \cos(2z)] \psi = 0, \quad (5.4)$$

where

$$a = \left(\frac{1}{\pi}\right)^2 \left(E - \frac{V_0}{2}\right) \quad (5.5)$$

$$q = -\left(\frac{1}{\pi}\right)^2 \frac{V_0}{4}. \quad (5.6)$$

The solutions can then be found in terms of plane waves by employing a Fourier transformation. This yields a recursion relation for the modes:

$$\psi = e^{irz} \frac{1}{N} \sum_k c_k e^{ikz} \quad (5.7)$$

$$\Rightarrow \sum_k \left\{ -(r+k)^2 + a - q(e^{2iz} + e^{-2iz}) \right\} c_k e^{ikz} = 0 \quad (5.8)$$

$$\Rightarrow [a - (r+k)^2] c_k - q(c_{k+2} + c_{k-2}) = 0 \quad (5.9)$$

$$\Rightarrow \begin{pmatrix} \dots & \dots & -q & a-(r+1)^2 & -q & \dots & \dots & \dots \\ \dots & 0 & -q & a-r^2 & -q & 0 & \dots & \dots \\ \dots & 0 & 0 & -q & a-(r-1)^2 & -q & \dots & \dots \\ \dots & \dots \end{pmatrix} \begin{pmatrix} \dots \\ c_2 \\ c_0 \\ c_{-2} \\ \dots \end{pmatrix} = 0. \quad (5.10)$$

Only the even modes have been represented in the matrix, since only these are coupled. Truncating the Fourier expansion yields a finite band-diagonal matrix which can be diagonalized numerically for each crystal momentum r . The coefficients c_n of the resulting eigenvectors give the Bloch eigenstates of the problem via equation (5.7). Furthermore, the one-dimensional Wannier functions can then be found by summing Bloch functions from the first Brillouin zone, according to the definition (3.17). For a system with very many lattice sites, the number of allowed quasimomenta, by the definition (3.11), is very large. It is then permissible to approximate the sum with an integral. This integral may in turn be approximated numerically by sampling. The ground state is non-degenerate and can be written as $\psi_s(\mathbf{x}) = \psi_{xs}(x) \psi_{ys}(y)$, a product of one-dimensional s orbitals. In this case of an isotropic cubic lattice, the first excited band is doubly degenerate (the degeneracy is d for a d -dimensional hypercubic lattice), and the states denoted as p_x , p_y in analogue with the atomic orbital states. Such wavefunctions are of the form $\psi_{p_x}(\mathbf{x}) = \psi_{xp}(x) \psi_{ys}(y)$, i.e. the product of one p - and one s -orbital wavefunction. Figure 5.1 shows the one-dimensional Wannier functions in the p band and figure 5.2 the separable two-dimensional Wannier functions p_x, p_y for $V_0 = 25E_r$.

5.1.2 The harmonic oscillator approximation

Since $\sin^2(\pi x) = (\pi x)^2 + \mathcal{O}(x^4)$, the potential is approximately parabolic around each site. It is therefore reasonable to expect that the solutions should resemble the

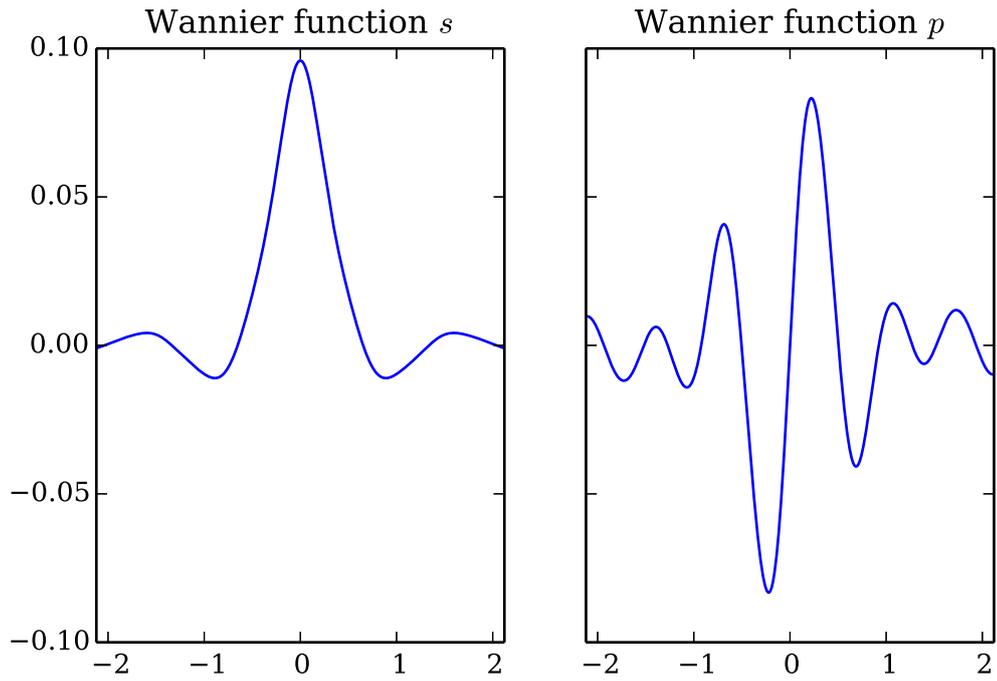


Figure 5.1: One-dimensional Wannier functions in one dimension, calculated for $V_0 = 25E_r$. Note that they are not positive-definite.

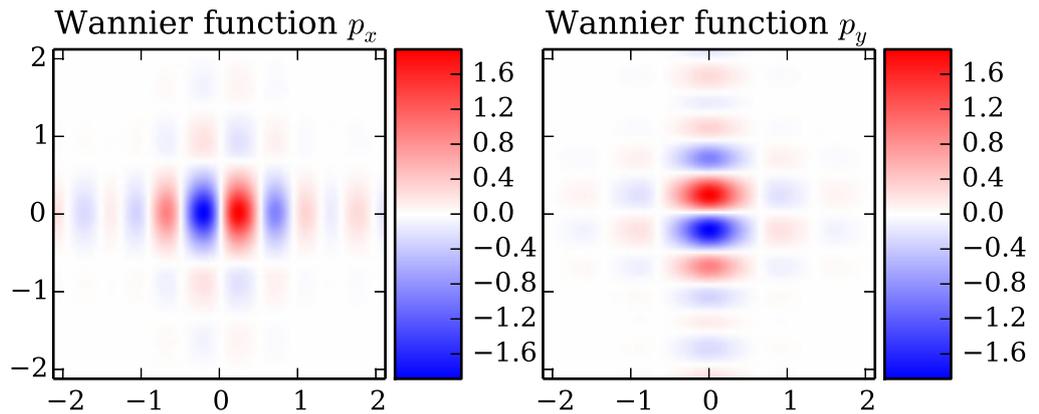


Figure 5.2: Separable Wannier functions in two dimensions, calculated for $V_0 = 25E_r$.

solutions for a parabolic potential in the vicinity of the sites [PS08, p. 411]. The Schrödinger equation for a wavefunction in a parabolic potential is the *quantum harmonic oscillator* (QHO) and have solutions in terms of the Hermite polynomials $H_n(x)$:

$$\left(-\frac{\partial^2}{\partial x^2} + V_0(\pi x)^2\right)\psi_x(x) = E_x\psi_x(x) \Rightarrow \psi_{nx}(x) = NH_n(x)e^{-\omega_x x^2/2}, \quad (5.11)$$

where $\omega_x = \sqrt{V_0}\pi$, $H_0 = 1$, $H_1 = 2x$ and N is a normalizing factor. Higher Hermite polynomials are only relevant for higher bands. One common measure of how well the harmonic approximation works is the *fidelity*, defined as

$$F[\phi, \psi] = \sqrt{\left|\int d\mathbf{x}\phi^*(\mathbf{x})\psi(\mathbf{x})\right|}, \quad (5.12)$$

where ϕ, ψ are assumed to be normalized functions and the integration is over the whole domain of ϕ and ψ . Note that both the normalization and the fidelity involve *overlap integrals*, integrals which are inner products in function space:

$$O(\phi, \psi) = \int d\mathbf{x}\phi^*(\mathbf{x})\psi(\mathbf{x}) = \langle\phi, \psi\rangle. \quad (5.13)$$

The fidelity can be seen as a measure of how big the projection of the function ϕ on function ψ is, but this has a somewhat more complicated meaning in infinite-dimensional spaces [Bal10, p. 26]. Figure 5.3 shows the fidelity of the harmonic oscillator approximation with the true Wannier function, calculated for a range of potential depths. Note that for $V_0 = 25E_r$, the fidelity is already > 0.99 for the s -band, while the p -band Wannier function requires a potential depth of almost $V_0 = 100E_r$ for the harmonic approximation to have such a fidelity. This lower fidelity for the higher orbital derives from the fact that it is much less ‘bound’ to a lattice site, i.e. its width is larger. It can also be understood from the fact that for a higher energy, the state ‘feels’ the anharmonicity of the sinusoid potential more than for the ground state.

Figure 5.4 shows that, even for $V_0 = 100E_r$, the oscillatory character of the “tail” of the Wannier function is significant. The behavior is not captured by the harmonic approximation and the fidelity apparently provides no information about this error. This is an indication of that the true Wannier functions *do not* converge uniformly to the harmonic oscillator approximation even as $V_0/E_r \rightarrow \infty$. Instead, some oscillatory character always remains [BDZ08, p. 14] (remember that the asymptotic behaviour is exponential and not gaussian). This means that tunnelings in particular cannot be reliably calculated with the harmonic approximation even in the limit of very deep lattices.

One striking example of how the harmonic oscillator approximation can fail is in

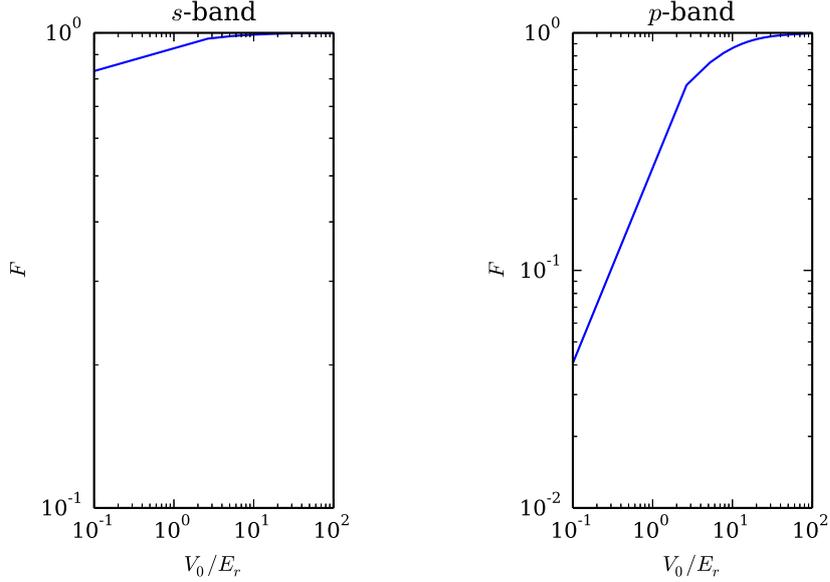


Figure 5.3: Logarithmic plot of the fidelity for a range of V_0/E_r for the s and p orbital states. Note that the fidelity is almost 1 for $V_0 = 25E_r$ while the p band requires much deeper potential wells.

the approximation of the onsite interactions, which are given by

$$\begin{aligned}
 U_{\alpha\beta} &= U_0 \int d\mathbf{x} |\psi_{p_\alpha}(\mathbf{x})|^2 |\psi_{p_\beta}(\mathbf{x})|^2 \\
 \Rightarrow U_{xx} = U_{yy} &= U_0 N_0^4 N_1^4 \int dx x^4 e^{-2\omega x^2} \int dy e^{-2\omega y^2} = \\
 &= U_0 \frac{4\omega^4}{\pi^2} \left(\frac{3}{16\omega^2} \sqrt{\frac{\pi}{2\omega}} \right) \left(\sqrt{\frac{\pi}{2\omega}} \right) = \frac{3\omega}{8\pi} \quad (5.14)
 \end{aligned}$$

$$\begin{aligned}
 U_{xy} &= U_0 N_0^4 N_1^4 \left(\int dx x^2 e^{-2\omega x^2} \right)^2 = \\
 &= U_0 \frac{4\omega^4}{\pi^2} \left(\frac{1}{4\omega} \sqrt{\frac{\pi}{2\omega}} \right)^2 = U_0 \frac{\omega}{8\pi} = U_{xx}/3, \quad (5.15)
 \end{aligned}$$

where $\omega = \sqrt{V_0/2\pi}$. For $V_0 = 25E_r$, the harmonic approximation yields $U_{xx}/U_0 \approx 1.30$, $U_{xy} \approx 0.43$. A computation with the full Wannier functions instead gives $U_{xx}/U_0 \approx 1.44$, $U_{xy}/U_0 \approx 0.60$, which has a ratio of $U_{xx}/U_{xy} \approx 2.4$ and not 3. The ratio 3 turns out to be close to a critical value, leading to a qualitatively different prediction [PSanB⁺13, CLM10].

5.1.3 Analytic overlap integral

As an illustration of why the harmonic approximation gives the wrong answers when computing overlap integrals, let me compute

$$O = \int_{-\infty}^{\infty} dx f_k(x+d) \partial_x^2 f_k(x) \quad (5.16)$$

$$f_k(x) = \mathcal{N} \cos(kx) e^{-ax^2/2}, \quad (5.17)$$

Wannier function and harmonic approximation for $V_0 = 100E_r$, $F=0.99$

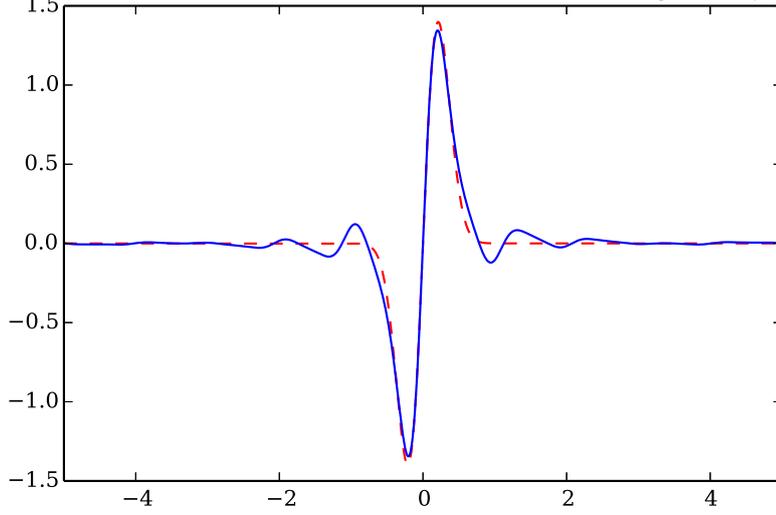


Figure 5.4: The Wannier function (solid blue line) and its harmonic approximation (dashed red line), for the p orbital, at $V_0 = 100E_r$. Note the periodic “wiggles” in the tail, which is the feature which makes the Wannier functions orthogonal between different lattice sites.

which here serves as a model of an oscillatory integrand. While this model decays in a gaussian fashion for large x , this should not affect the discussion about non-uniform convergence here. Proceeding, we have

$$\begin{aligned} \partial_x^2 f_k(x) &= \left\{ (a^2 x^2 - a - k^2) \cos(kx) + 2akx \sin(kx) \right\} e^{-ax^2/2} \\ \Rightarrow O &= \mathcal{N}^2 \int_{-\infty}^{\infty} dx \cos(k(x+d)) \left\{ (a^2 x^2 - a - k^2) \cos(kx) + 2akx \sin(kx) \right\} e^{-a((x+d)^2+x^2)/2}. \end{aligned} \quad (5.18)$$

A change of variables is appropriate: $x + d/2 = x + \lambda = y$. I made use of the trigonometric identities

$$\cos(a+b) \cos(a-b) = \frac{1}{2} (\cos(2a) + \cos(2b)), \quad (5.19)$$

$$\cos(a+b) \sin(a-b) = \frac{1}{2} (\cos(2a) - \cos(2b)), \quad (5.20)$$

so that

$$\begin{aligned} O &= \mathcal{N}^2 e^{-a\lambda^2} \int_{-\infty}^{\infty} dy \cos(k(y+\lambda)) \left\{ (a^2(y-\lambda)^2 - a - k^2) \cos(k(y-\lambda)) + \right. \\ &\quad \left. + 2ak(y-\lambda) \sin(k(y-\lambda)) \right\} e^{-ay^2} = \\ &= \mathcal{N}^2 \frac{1}{2} e^{-a\lambda^2} \int_{-\infty}^{\infty} dy \left\{ (a^2 y^2 - 2a^2 \lambda y + a^2 \lambda^2 - a - k^2) (\cos(2ky) + \cos(2k\lambda)) + \right. \\ &\quad \left. + 2ak(y-\lambda) (\sin(2ky) + \sin(2k\lambda)) \right\} e^{-ay^2}. \end{aligned} \quad (5.21)$$

Every part of this integrand is either even or odd. The odd parts vanish when integrated over. The remaining four integrals are computed using integral tables

from [III15]. Consider

$$\int_{-\infty}^{\infty} dy e^{-ay^2} = \sqrt{\frac{\pi}{a}} \quad (5.22)$$

$$\int_{-\infty}^{\infty} dy y^2 e^{-ay^2} = \sqrt{\frac{\pi}{a}} \frac{1}{2a} \quad (5.23)$$

$$\int_{-\infty}^{\infty} dy \cos(2ky) e^{-ay^2} = \frac{1}{2} \int_{-\infty}^{\infty} dy \left[e^{-ay^2+2iky} + e^{-ay^2-2iky} \right] = \sqrt{\frac{\pi}{a}} e^{-k^2/a} \quad (5.24)$$

$$\int_{-\infty}^{\infty} dy \cos(2ky) y^2 e^{-ay^2} = \frac{1}{2} \int_{-\infty}^{\infty} dy \left[e^{-ay^2+2iky} + e^{-ay^2-2iky} \right] y^2 = \sqrt{\frac{\pi}{a}} \frac{(a-2k^2)}{2a^2} e^{-k^2/a} \quad (5.25)$$

$$\int_{-\infty}^{\infty} dy \sin(2ky) y e^{-ay^2} = \frac{1}{2i} \int_{-\infty}^{\infty} dy \left[e^{-ay^2+2iky} - e^{-ay^2-2iky} \right] = \sqrt{\frac{\pi}{a}} \frac{k}{a} e^{-k^2/a}. \quad (5.26)$$

Using these identities, the overlap integral can be written as

$$O(\lambda) = \frac{\mathcal{N}^2}{2} \sqrt{\frac{\pi}{a}} e^{-a\lambda^2} \left\{ (a^2\lambda^2 - a/2) e^{-k^2/a} + (a^2\lambda^2 - a/2 - k^2) \cos(2k\lambda) + 2ak\lambda \sin(2k\lambda) \right\}. \quad (5.27)$$

The oscillations can be assumed to be periodic in the lattice spacing, so that $2k\lambda = \pi n \Rightarrow k = \pi n/2\lambda$ for some integer n and the last term vanishes while the second term will contribute with a negative or positive sign:

$$O(\lambda) = \frac{\mathcal{N}^2}{2} \sqrt{\frac{\pi}{a}} e^{-a\lambda^2} \left\{ (a^2\lambda^2 - a/2) e^{-\left(\frac{n\pi}{2\lambda}\right)^2/a} + (-1)^n \left(a^2\lambda^2 - a/2 - \left(\frac{n\pi}{2\lambda}\right)^2 \right) \right\}. \quad (5.28)$$

The normalization is important:

$$\mathcal{N}^{-2} = \int_{-\infty}^{\infty} dx \cos^2(kx) e^{-ax^2} = \frac{1}{2} \int_{-\infty}^{\infty} dx (1 + \cos(2kx)) e^{-ax^2} = \frac{1}{2} \sqrt{\frac{\pi}{a}} \left(1 + e^{-\left(\frac{n\pi}{2\lambda}\right)^2/a} \right). \quad (5.29)$$

For n even I find:

$$O(\lambda) = e^{-a\lambda^2} \left\{ (a^2\lambda^2 - a/2) - \left(\frac{n\pi}{2\lambda}\right)^2 \frac{1}{1 + e^{-\left(\frac{n\pi}{2\lambda}\right)^2/a}} \right\}. \quad (5.30)$$

For a given a and large n , $O \sim n^2$. If O is seen as part of a Fourier series, it would be suppressed by n or n^2 but would still be important. Note that the part multiplied by n^2 is not multiplied by any power of a . As $a \rightarrow \infty$ corresponds to small lattice separation, n will have to be correspondingly larger to make that term significant. For a real Wannier function I would expect some connection between a and the amplitudes of the Wannier modes, which may or may not make oscillatory integrands relevant. It is precisely the non-uniform convergence of $f_k(x)$ as $a \rightarrow \infty$ that is the reason that the oscillatory character is important for the value of the integral. I conclude that at least in this model, the oscillatory character can be important even for small a .

5.1.4 Overlaps in the quadratic lattice

For $V_0 = 25E_r$, the “self-energies” were computed as $t_{ii}^{xx} = t_{ii}^{yy} \approx -49.8$, confirming that the p_x and p_y states are degenerate in the quadratic lattice. Furthermore, $t_{ii}^{xy} = t_{ii}^{yx} = 0$ because of the symmetries of the potential and states. In fact, $t_{ij}^{xy} = t_{ij}^{yx}$ for any pair of neighbor sites i, j . I found $t^{xx} \approx -7.1$, $t^{yy} \approx 2.4$ between horizontal pairs, while $t^{xx} \approx 2.4$, $t^{yy} \approx -7.1$ for vertical pairs. Hence p_x, p_y orbitals are more likely to tunnel in the direction of their nodes, which, of course, is a result of the orbitals being more extended in these directions. All orbital-changing tunneling is suppressed and thus only the onsite interaction converts pairs in one orbital states into the other.

In section 7.3, I discuss the solution of the quadratic lattice in terms of a mean-field approximation, starting from a coherent-state ansatz. The interaction Hamiltonian is

$$\hat{V} = \sum_{i\alpha} \frac{U_{\alpha\alpha}}{2} \hat{n}_{i\alpha} (\hat{n}_{i\alpha} - 1) + \sum_{i\alpha \neq \beta} U_{\alpha\beta} \hat{n}_{i\alpha} \hat{n}_{i\beta} + \frac{U_{\alpha\beta}}{2} (\hat{a}_{i\alpha}^\dagger \hat{a}_{i\alpha}^\dagger \hat{a}_{i\beta} \hat{a}_{i\beta} + h.c.). \quad (5.31)$$

The first two terms favor singly occupied and unoccupied sites while the third term mixes x and y states, so the solution can be expected to be some singly occupied mix of x and y states. In section 7.3, I show that the solution, in first-quantized notation, is a *vortex* state with a wavefunction

$$\psi_{j,\pm}(\mathbf{x}) = \psi_{jx}(\mathbf{x}) \pm i\psi_{jy}(\mathbf{x}), \quad (5.32)$$

where $\psi_{j\alpha}(\mathbf{x})$ are the $\alpha = x, y$ Wannier functions. These functions are eigenstates of the z component of $\hat{\mathbf{L}} = \hat{\mathbf{x}} \times \hat{\mathbf{p}}$, the angular momentum operator, with eigenvalues $\pm\hbar$. They therefore have a non-zero angular momentum and violate time-reversal symmetry! The wavefunction (5.32) seems to violate Feynman’s *no-node* theorem [C.09], which suggests that it should always be possible to write the ground state of a many-body boson system with a real wavefunction. But we must remember that the Wannier functions are not the energy eigenfunctions, and neither does the no-node theorem hold for excited bands. Figure 5.5 shows a plot of the phase and absolute value of a vortex state in “physical space”. The solution is referred to as a vortex solution because it has a *winding number* $n = \pm 1$. The winding number is defined by

$$\oint d\mathbf{x} \arg[\psi_{j,\pm}(\mathbf{x})] = 2\pi n, \quad (5.33)$$

where the integral is any closed path around the center of the wavefunction. Note that the vortex solution has a discontinuity in phase around the central minimum which precisely reflects the non-zero winding number: this is a feature of the plotting software which always chooses the first branch of the Riemann surface.

If the tunneling is small but non-zero, the Hamiltonian will couple these states into a pattern. This means that it is not permissible to treat individual sites as independent. The tunneling constrains how neighbouring sites are connected, creating coherence between sites. For the quadratic lattice, the bipartite nature of the lattice leads to a “checkerboard” pattern of vortex or “anti-vortex” states on each of the two sublattices [AS05] [JJ12]. Again, see chapter 7 for a full treatment of this problem.

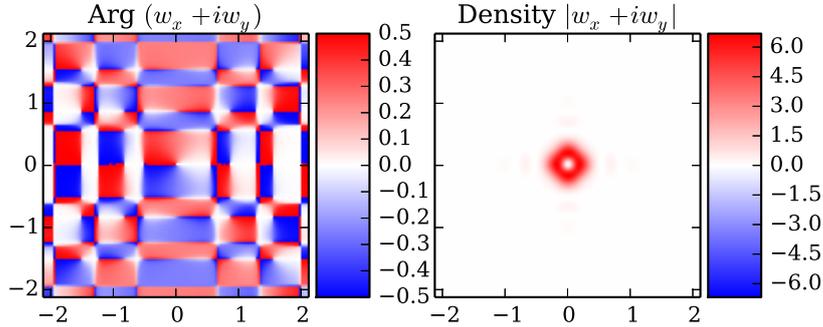


Figure 5.5: Vortex solution. The left plot shows the phase (in units of 2π), while the right plot shows the density $|w_x(\mathbf{x}) + w_y(\mathbf{x})|^2$. Note the central vortex structure, surrounded by smaller vortices in an intricate pattern. Notice the branch cuts, which are lines where the phase has turned a full 2π . All branch cuts are connected to centers, where the wavefunction has to be zero to avoid being singular (this effect is most clearly visible as the center hole of the density plot). A spatially varying phase implies that the superfluid host some persistent current - the vortex.

5.2 The zig-zag lattice

The zig-zag lattice potential consists of repeated ladders of potential wells. The ladders are separated by potential maxima so strong that the atoms can be said to be confined to each chain. In the third direction, the system is limited by another optical lattice with an amplitude so high that tunneling is also completely suppressed. Indeed, using large potential amplitudes in the transverse directions, experiments have demonstrated [BDZ08, p. 13] that the atoms will form regular tubular structures in the z -direction. The goal of choosing this particular lattice is to create a quasi-one-dimensional system where the interaction between sites on the “angles” of the chain cannot be ignored. The system can be effectively modelled as a one-dimensional system with interactions between nearest-neighbor (diagonal) and next-nearest-neighbor (horizontal) sites.

5.2.1 Constructing the zig-zag lattice

The zig-zag potential can be created by superimposing four standing waves:

$$V(\mathbf{x}) = V_0 \left[\sin^2(\mathbf{k}_1 \cdot \mathbf{x}/2) + \sin^2(\mathbf{k}_2 \cdot \mathbf{x}/2) + \sin^2((\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{x}/2) + g \sin^2(\mathbf{k}_s \cdot \mathbf{x}/2 - \pi/4) \right], \quad (5.34)$$

where V_0 , the strength of the lattice, and $g > 0$ are free parameters. The fourth term proportional to g is called the *superlattice*. As mentioned in the end of section

4.4, it is necessary to set $V_0 \leq 25E_r$ for the system to have time to delocalize. The reciprocal vectors are²

$$\begin{aligned} k_1 &= k_y \hat{\mathbf{y}}, \\ k_2 &= k_x \hat{\mathbf{x}} + \frac{1}{2} k_y \hat{\mathbf{y}}, \\ k_s &= \frac{1}{2} k_y \hat{\mathbf{y}}, \end{aligned} \tag{5.35}$$

where

$$\begin{aligned} k_x &= \frac{2\pi}{2 \cos(\theta)}, \\ k_y &= \frac{2\pi}{\sin(\theta)}. \end{aligned}$$

The third free parameter is θ , the angle between the horizontal rows and the diagonal of the lattice.

Setting $g = 0$ corresponds to a regular triangular lattice of angle θ . Depending on the angle, this lattice will have a number of symmetries: it will at the very least be invariant under translation by any integer-coefficient linear combination of lattice vectors $\mathbf{a}_1 = 2 \cos(\theta) \hat{\mathbf{x}}$, $\mathbf{a}_2 = \cos(\theta) \hat{\mathbf{x}} + \sin(\theta) \hat{\mathbf{y}}$. Thus the lattice is a Bravais lattice as described in section 3.1.2, and there is one site per primitive cell. There will be an additional reflection symmetry in the x and y plane and, for $\theta = \pi/3$, rotation symmetry by angles $\pi/3$.

As soon as the superlattice is turned on, a number of symmetries are inevitably broken. This will have profound consequences on the solutions. The first and most important consequence is that discrete translation invariance is broken. The primitive cell now contains two sites, which can be taken as a diagonal pair of sites in each row.

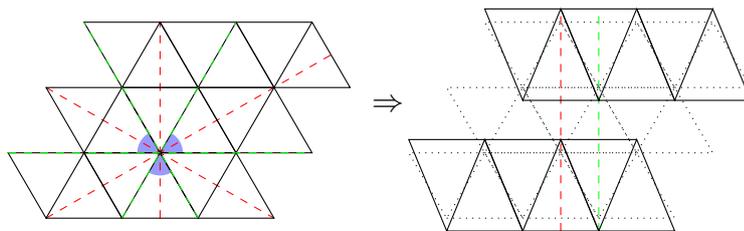


Figure 5.6: Symmetries of the triangular lattice and symmetry breaking in the transition to a zig-zag lattice. The nodes mark the minima of the zig-zag potential for $g = 0$ (left) and $g > 0$ (right). The triangular lattice has 60° rotation symmetry (blue angles), three reflection axes through any triangle (red, dotted) and reflection axes along all sides of triangles (green). Of course, there is also a translation symmetry by any lattice vector \mathbf{X} . The zig-zag lattice keeps axes of reflection symmetry through (red) and between (green) every triangle in the direction of the superlattice. There is still horizontal translation symmetry by one lattice spacing and a vertical translation symmetry by two lattice periods.

Another important effect of the superlattice is that it changes the positions of the minima of the potential. Qualitatively, each zig-zag strip has its minima squeezed

²I choose to work in rescaled coordinates, in which the lattice spacing ~ 1 .

towards the middle of the strip in the y -direction. Figure 5.6 illustrates the symmetry reduction and shift schematically. Since I will later assume the Wannier functions to be centered at the minima of the potential, it is convenient to have the distances and angles between the sites in closed-form expressions. I therefore introduce an affine transformation $y \rightarrow y' = ay + b$ which brings the first minimum back to $\mathbf{x} = \mathbf{0}$ and the second minimum to $\mathbf{x} = \mathbf{a}_2, \mathbf{a}_1, \mathbf{a}_1 + \mathbf{a}_2$ and so on. The coefficients a, b were found numerically by optimization of the position of the minima. Physically, the affine transformation could be implemented by a shifting of the phase of all lasers in the y -direction and a tuning of the parameter k_y as defined above. The shifted potential is shown in figure 5.7 in a heatmap plot and a 3D projection.

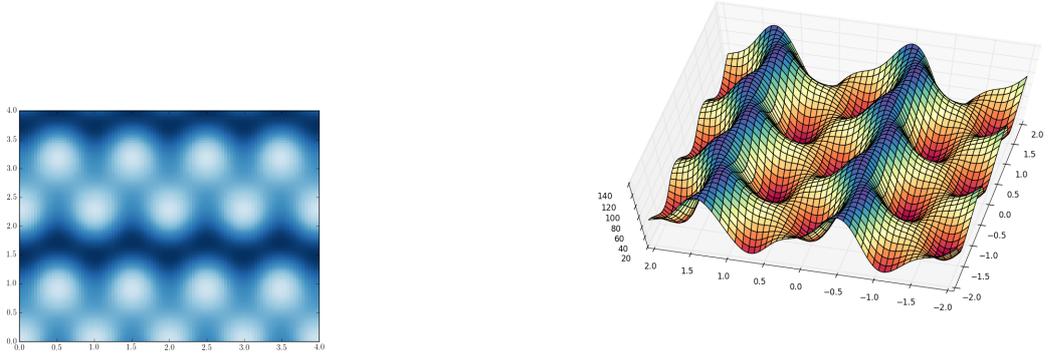


Figure 5.7: Two different views of the zig-zag potential with $\theta = \pi/3$ and $g = 2$. To the left: a heatmap plot where minima are light and maxima are blue. To the right: a color-coded 3D projection: minima are red, maxima are blue.

5.2.2 Wannier states in the zig-zag lattice

The potential (5.34) is non-separable. To obtain the full energy eigenfunctions would be numerically challenging, and finding the full Wannier functions would involve integrating over a hexagonal Brillouin zone. As is common, I chose to work with the two-dimensional Wannier functions for a square lattice, which are more easily obtained since the Schrödinger equation for this potential is separable. Whether the full Wannier states, the single-site separable states or their harmonic approximation are used, they are all in the end functional bases. There is no guarantee however, that this subspace of all functions are enough to describe the solution. At least for the triangular lattice, the true Wannier functions resemble the separable ones[WC03]. Recall equation (3.20) for the tunnelings:

$$t_{ij}^{\alpha\beta} = \int d\mathbf{x} \psi_{\mathbf{x}_i\alpha}^\dagger(\mathbf{x}) \left[-\nabla^2 + V(\mathbf{x}) \right] \psi_{\mathbf{x}_j\beta}(\mathbf{x}). \quad (5.36)$$

The values of the tunnelings will obviously depend on the functional basis. $t_{ii}^{\alpha\beta}$, the *onsite-energies*, are conventionally absorbed into the interaction, so that the values of certain couplings in the interaction, as well as tunnelings, will depend on the choice of basis. Even though the mathematics may be easier or harder for a particular single-site basis, they must in the end give the same physical predictions.

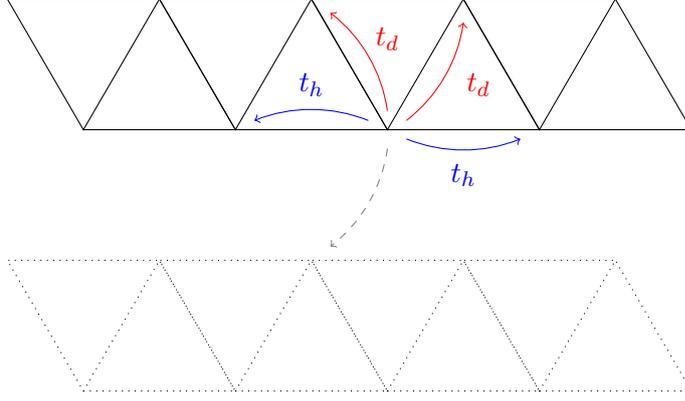


Figure 5.8: Illustration of what is meant by diagonal and horizontal hopping in the lattice, denoted t_d and t_h respectively. The probability for tunneling between different zig-zag ladders is so small that it may be ignored in calculations.

One interesting variation on the theme of p -band Wannier functions are rotated states. Imagine rotating the nodes of the Wannier functions relative to the lattice, mathematically described by

$$\psi_{i\alpha} = \sum_{\beta} U_{\alpha\beta} \tilde{\psi}_{i\beta}, \quad (5.37)$$

where U is an orthogonal matrix with determinant 1. Especially interesting from symmetry considerations are those Wannier bases which has nodes in the x/y direction (such as those obtained from the quadratic lattice) and those states rotated by 45° . While this has some interesting consequences for the form of the Hamiltonian, the rotated Wannier functions were not used for the results of this thesis.

5.2.3 Overlap integrals

The tunneling and interaction strength are couplings determined by the overlap of the single-particle states in the Fock basis. Figure 5.8 illustrates the types of tunnelings which are involved schematically.

Note that the Hermitian property of the Hamiltonian

$$\sum_{ij,\alpha\beta} t_{ij}^{\alpha\beta} \hat{a}_{i\alpha}^\dagger \hat{a}_{j\beta} = \left(\sum_{ij,\alpha\beta} t_{ij}^{\alpha\beta} \hat{a}_{i\alpha}^\dagger \hat{a}_{j\beta} \right)^\dagger = \sum_{ij,\alpha\beta} (t_{ij}^{\alpha\beta})^* \hat{a}_{j\beta}^\dagger \hat{a}_{i\alpha} \quad (5.38)$$

leads to the constraint $(t_{ji}^{\beta\alpha})^* = t_{ij}^{\alpha\beta}$. Moreover, in the absence of external gauge fields, it is possible to choose the tunnelings to be real [JFJO11] [Aue94, p. 23]. If $t_o^{xx} = t_o^{yy}$, the Hamiltonian will not favor any of the orbital states, while if they are different this will act as a kind of chemical potential, favouring one of the orbital states. t_o^{xy}, t_o^{yx} correspond to terms in the Hamiltonian which convert a single p_x state to a p_y state, “on-site”. The interaction does have a term converting $p_x \leftrightarrow p_y$ but only converts pairs (therefore this part of the interaction conserves particle number modulo 2). The Hamiltonian has a local \mathbb{Z}_2 parity symmetry provided that $t_o^{xy} = t_o^{yx} = 0$. The symmetry only holds locally because of the fact that tunneling can convert orbital states between sites.

Computation of the overlap integrals is not such a simple integration routine as it may seem. This problem stems from the fact that the integrand is often

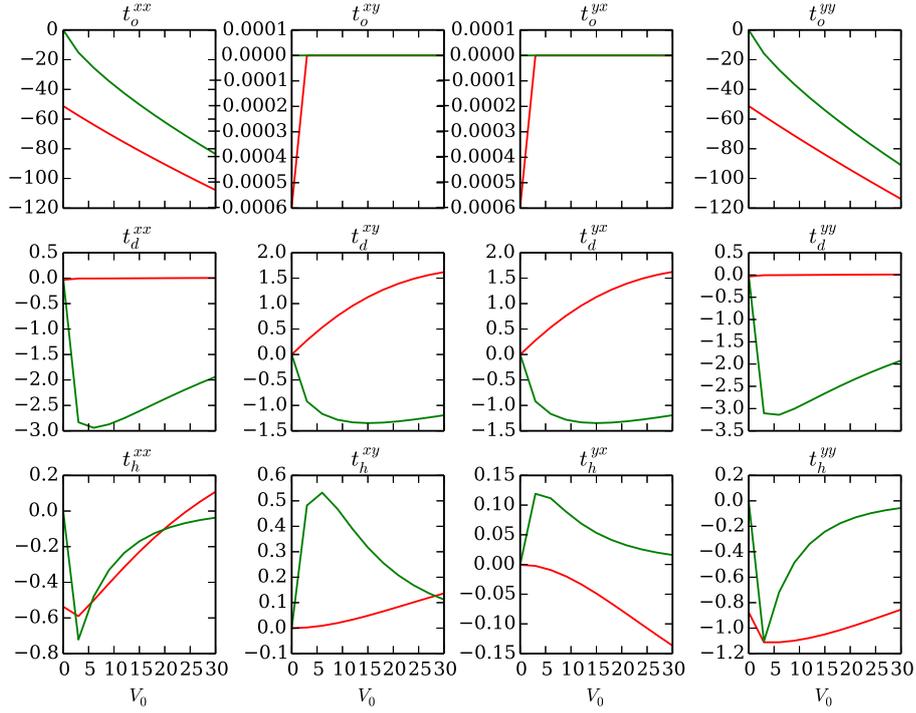


Figure 5.9: Comparison between the overlap integrals, computed with the full separable Wannier functions (green) and with the harmonic approximation (red), computed for a range of V_0 . What is found is that even for relatively large potential amplitudes, the harmonic approximation fails to qualitatively reproduce the overlap integrals.

- (almost) antisymmetric, leading to large performance losses for relatively small changes in integration grid size and integration area.
- not well localized because of the “wiggly” nature of the Wannier states, which is the very feature which make them an orthonormal basis.

Both these reasons motivated extensive convergence tests of the integrals in all numerical parameters: the truncation of the Fourier series for the Mathieu equation, the number of quasimomenta used for the integral approximating the Wannier functions and the region of integration. Just as for the quadratic lattice, the harmonic approximation gives the wrong results when calculating overlap integrals. Figure 5.9 compares the harmonic approximation with a calculation made with the “full” Wannier functions. At least for $V_0 < 30E_r$, the true Wannier states give strikingly different predictions.

It is interesting to understand how θ and g , the free parameters of the lattice, influence the physics. After second quantization, all physical information about the system should be contained in the choice of single-particle basis, and in the couplings. Figure 5.10 shows a plot of the overlap integrals computed for a range of θ with $g = 2$ held fixed. For each point in each subplot, a lattice potential was generated and optimized as described above. Depending on θ , the minima of the potential will be more or less elongated. It is reasonable to take this into account

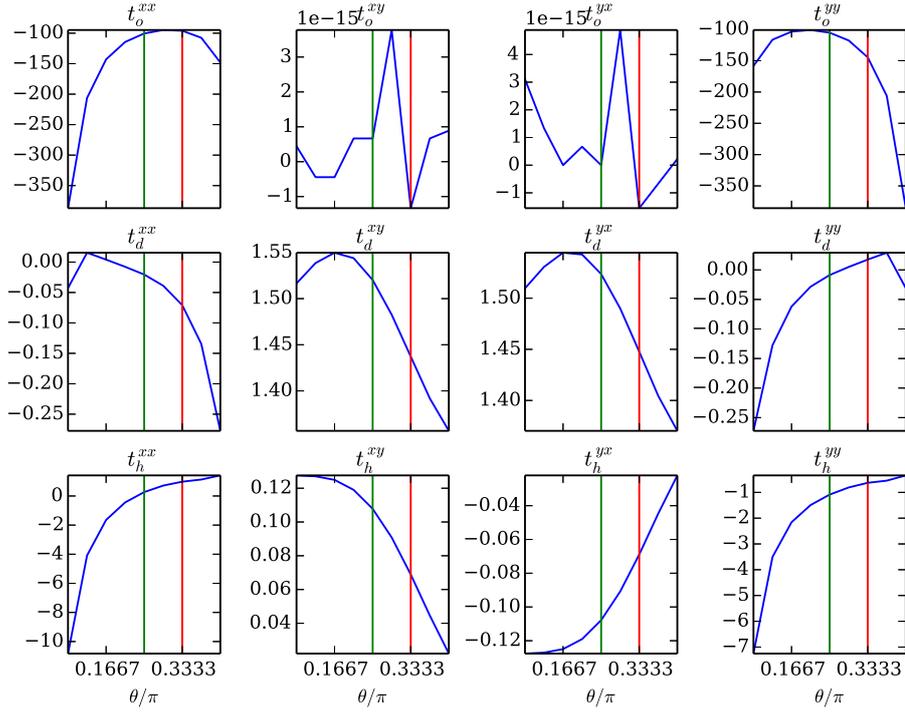


Figure 5.10: Self-energies and tunnelings computed for a range of values of θ , where $V_0 = 25E_r$. The first row shows the onsite energies t_o^{xx}, t_o^{yy} . Note that t_o^{xy}, t_o^{yx} are negligible. The second and third row show the diagonal and horizontal tunnelings respectively. The angles $\theta = \pi/4$ and $\theta = \pi/3$ have been marked with vertical lines.

when calculating the separable Wannier functions: this was done by setting the lattice constants used in the calculations equal to $\ell_x = \cos(\theta), \ell_y = \sin(\theta)$. $\theta = \pi/4$ then leads to Wannier functions such as in the quadratic lattice but with lattice constant $\ell_x = \ell_y = 1/\sqrt{2}$. Other angles squeeze the lattice in the x - or y -direction. Note that for $\theta = \pi/4$, t_h^{xx} is negligible. t_d^{xx} and t_d^{yy} are also unimportant compared to the large terms: t_h^{yy}, t_d^{xy} and t_d^{yx} are the significant couplings.

Fixing $g = 2$, $\theta = \pi/4$, I obtain the set of tunneling coefficients in table 5.1. Due to the superlattice, the difference between self-energies $2\mu = t_o^{xx} - t_o^{yy} \approx 20$, which breaks degeneracy of the p_x and p_y states. If μ is one order of magnitude larger than the tunnelings, it will be the dominant term, leading to a state with only one orbital populated. As will be seen in chapter 8, the chemical potential can be interpreted as an external magnetic field in the Mott phase: a strong chemical potential leads to a simple polarized spin model. Since the superlattice is also necessary to uncouple the different chains, it seems like the zig-zag potential would show no interesting dynamics.

However, there is a way out. By use of a Raman coupling as described in section 4.4, tunnelings and self-energies may be tuned at will. While this makes experimental realization harder, it enables us to tune μ as well as the other couplings. It is necessary to shift μ and all the horizontal tunnelings to generate interesting behavior in a spin model, as will be shown in chapter 8.

		Even	Odd
Diagonal (nearest-neighbor)	t_d^{xx}	0	0
	t_d^{xy}	1.5	-1.5
	t_d^{yx}	1.5	-1.5
	t_d^{yy}	0	0
Horizontal (next-nearest neighbor)	t_h^{xx}	0	0
	t_h^{xy}	0.1	-0.1
	t_h^{yx}	-0.1	0.1
	t_h^{yy}	-1	-1

Table 5.1: Tunneling coefficients calculated for a zig-zag lattice with $g = 2$, $V_0 = 25E_r$ and $\theta = \pi/4$, for even and odd sites. As was already shown, $t_{ij}^{xy} = t_{ji}^{yx}$. Note the pattern of alternating signs of the diagonal and horizontal xy tunnelings.

As argued, the phases of the tunneling coefficients play an important role for the properties of model. More precisely, if we imagine a particle hopping around in the lattice in such a way that it forms a loop (often referred to as *Wilson loop*), then we can define the *Peierls phase* [JFJO11]

$$\phi = \angle[t_{ij}^{\alpha\beta} t_{jk}^{\gamma\delta} \dots t_{li}^{\epsilon\alpha}], \quad (5.39)$$

where $t_{ij}^{\alpha\beta}$ the corresponding tunneling coefficients form a closed loop, and it is understood that the phase is defined modulo 2π . This phase is *gauge invariant*, i.e. no local unitary transformation will alter its value, and can be seen as a synthetic magnetic flux penetrating the loop. Thus, for a non-zero ϕ the model is equivalent to the one of a charged particle in a magnetic field. Noting that some of the tunneling coefficients are negative in the present model, the natural question arises whether it is possible to find loops such that $\phi = \pi$. It is, however, easy to convince oneself that the Peierls phase actually vanishes for any loop in the zig-zag lattice.

5.2.4 The second-quantized Hamiltonian

From the overlap integrals, I obtain a Hamiltonian of the form

$$\hat{H} = \hat{\mathcal{T}} + \hat{\mathcal{V}}, \quad (5.40)$$

where the kinetic part of the Hamiltonian is

$$\hat{\mathcal{T}} = - \sum_{j,\alpha,\beta} t_d^{\alpha\beta} (\hat{a}_{j\alpha}^\dagger \hat{a}_{(j+1)\beta} + h.c.) - \sum_{j,\alpha,\beta} t_h^{\alpha\beta} (\hat{a}_{j\alpha}^\dagger \hat{a}_{(j+2)\beta} + h.c.), \quad (5.41)$$

where the sums are over nearest- and next-nearest neighbors, and the interaction part of the Hamiltonian is

$$\hat{\mathcal{V}} = \hat{\mathcal{V}}_{nn} + \hat{\mathcal{V}}_{FD}, \quad (5.42)$$

$$\hat{\mathcal{V}}_{nn} = \sum_{j,\alpha} \frac{U_{\alpha\alpha}}{2} \hat{n}_{j\alpha} (\hat{n}_{j\alpha} - 1) + \sum_{j,\alpha \neq \beta} U_{\alpha\beta} \hat{n}_{j\alpha} \hat{n}_{j\beta}, \quad (5.43)$$

$$\hat{\mathcal{V}}_{FD} = \sum_{j,\alpha \neq \beta} \frac{U_{\alpha\beta}}{2} (\hat{a}_{j\alpha}^\dagger \hat{a}_{j\alpha}^\dagger \hat{a}_{j\beta} \hat{a}_{j\beta} + h.c.). \quad (5.44)$$

\hat{V}_{nn} couples density fluctuations to each other and favors a constant density. \hat{V}_{FD} represents scattering of pairs of p_α states into a pair of p_β states, and couples the two orbital states at each site in competition with the hopping and density terms. Note the similarity between (5.40) and (3.42). The zig-zag chain may be seen as an extension of the Bose-Hubbard model to two “flavors” or internal states and next-nearest neighbor interaction. The interaction term in (5.40) will also lead to a Mott-superfluid quantum phase transition. In the following chapters, this Hamiltonian and its phases will be investigated in the limits of strong and weak interactions.

Chapter 6

Exact diagonalization

Physics is actually too hard for
physicists.

David Hilbert

The Hamiltonian (5.40),

$$\hat{H} = \hat{\mathcal{T}} + \hat{\mathcal{V}}, \quad (6.1)$$

has terms which are quartic in creation and annihilation operators. It is these quartic terms which makes the Hamiltonian impossible to diagonalize exactly. If exact diagonalization was accessible, all expectations and correlation functions could be computed readily. Hamiltonians with interaction terms generally can only be diagonalized numerically, for a truncated number of sites. In fact, this can only be done for a few sites¹, which is not helpful since we are interested in the thermodynamic limit. To make progress, it is necessary to use approximations and study the problem in certain regimes of parameter space. In the limiting case of $U = 0$, corresponding to an extreme case of the superfluid phase, the Hamiltonian consists only of its *ideal* or *kinetic* part,

$$\hat{\mathcal{T}} = - \sum_{j,\alpha,\beta} t_d^{\alpha\beta} \left(\hat{a}_{j\alpha}^\dagger \hat{a}_{(j+1)\beta} + h.c. \right) - \sum_{j,\alpha,\beta} t_h^{\alpha\beta} \left(\hat{a}_{j\alpha}^\dagger \hat{a}_{(j+2)\beta} + h.c. \right) - \mu \sum_j (\hat{n}_{jx} - \hat{n}_{jy}), \quad (6.2)$$

which *can* be diagonalized analytically. In this chapter, the chemical potential has been included in the kinetic Hamiltonian since it has the same bilinear form. Diagonalizing the Hamiltonian makes it possible to compute expectation values and correlation function.

In the parameter regime considered, $t_d^{xx} = t_d^{yy} = t_h^{xx} = 0$ to a good approximation and $t_d^{xy} = t_d^{yx} = 1.5$, $t_h^{yy} = -1$ and $t_h^{xy} = -t_h^{yx} = 0.1$. The sum in $\hat{\mathcal{T}}$ repeats in a period of two sites, since this is the number of sites in a primitive cell. A Fourier transformation would not make the problem easier, since the alternating signs would

¹Since the Hilbert space grows exponentially with the system size, numerical diagonalization is intractable even for 10 sites.

couple modes via a mechanism like

$$\begin{aligned} \sum_j (-1)^j \hat{a}_j^\dagger \hat{a}_{j+1} &= \sum_{k,q} \left(\frac{1}{\mathcal{N}^2} \sum_j (-1)^j e^{-i\pi(k-q)} \right) \hat{a}_k^\dagger \hat{a}_q = \\ \sum_{k,q} \left(\frac{1}{\mathcal{N}^2} \sum_j e^{-i\pi((k+1)-q)} \right) \hat{a}_k^\dagger \hat{a}_q &= \sum_{k,q} \delta_{k+1,q} \hat{a}_k^\dagger \hat{a}_q = \sum_k \hat{a}_k^\dagger \hat{a}_{k+1} \end{aligned} \quad (6.3)$$

Note that the alternating signs repeat in a period of two sites, corresponding to the primitive cell of the zig-zag lattice. Define the *vector* annihilation operator as

$$\hat{v}_j = \begin{pmatrix} \hat{a}_{jx1} \\ \hat{a}_{jy1} \\ \hat{a}_{jx2} \\ \hat{a}_{jy2} \end{pmatrix}, \quad (6.4)$$

where j labels the cells and 1,2 denotes the two sites within each cell. This is just a convenient notation to rewrite the tunnelings within and between the cells as tunnelings between internal states of the cells. In terms of the vector operators, the Hamiltonian can be rewritten as a sum over a number of matrix quadratic forms:

$$\hat{\mathcal{T}} = - \sum_j \left(\hat{v}_j^\dagger A \hat{v}_{j+1} + \hat{v}_{j+1}^\dagger A^\dagger \hat{v}_j + \hat{v}_j^\dagger B \hat{v}_j \right), \quad (6.5)$$

where

$$A = \begin{pmatrix} 0 & -t_h^{xy} & 0 & 0 \\ t_h^{xy} & t_h^{yy} & 0 & 0 \\ 0 & -t_d^{xy} & 0 & t_h^{xy} \\ -t_d^{xy} & 0 & -t_h^{xy} & t_h^{yy} \end{pmatrix}, \quad (6.6)$$

$$B = \begin{pmatrix} \mu & 0 & 0 & t_d^{xy} \\ 0 & -\mu & t_d^{xy} & 0 \\ 0 & t_d^{xy} & \mu & 0 \\ t_d^{xy} & 0 & 0 & -\mu \end{pmatrix}, \quad (6.7)$$

Now, the form (6.5) can be put in a diagonal form by Fourier transform over the *cells*:

$$\hat{v}_j = \frac{1}{\mathcal{N}} \sum_k \hat{v}_k e^{i\pi k j}, \quad (6.8)$$

so that the Hamiltonian can be written as

$$\hat{\mathcal{T}} = - \sum_k \hat{v}_k^\dagger \left(A e^{ik} + A^\dagger e^{-ik} + B \right) \hat{v}_k = - \sum_k \hat{v}_k^\dagger T(k) \hat{v}_k. \quad (6.9)$$

The 4×4 matrix

$$-T(k) = - \begin{pmatrix} \mu & -2it_h^{xy} \sin(k) & 0 & (1 - e^{-ik})t_d^{xy} \\ 2it_h^{xy} \sin(k) & 2t_h^{yy} \cos(k) - \mu & (1 - e^{-ik})t_d^{xy} & 0 \\ 0 & (1 - e^{ik})t_d^{xy} & \mu & 2it_h^{xy} \sin(k) \\ (1 - e^{ik})t_d^{xy} & 0 & -2it_h^{xy} \sin(k) & 2t_h^{yy} \cos(k) - \mu \end{pmatrix} \quad (6.10)$$

was then diagonalized numerically for a set of $k \in [-\pi, \pi]$. The tunneling t_h^{xy} and the chemical potential μ were treated as free parameters and varied. The reason for

this is that these tunneling controls next-nearest neighbor couplings. Any physics related to the number of sites in the Wigner-Seitz cell should be controlled by these tunnelings.

Because the Fourier transformation is over the cells of the primitive lattice, the four elements of the eigenvectors may be interpreted as the linear combination of Fourier coefficients for each orbital state and site in each cell. The spectrum and solutions for the four bands, are plotted in figure 6.1 for $t_h^{xy} = 0$ and $\mu = 0$, in figure 6.2 for $t_h^{xy} = 0.1$ and $\mu = 0$. Figure 6.3 shows the spectrum and solutions for $t_h^{xy} = 0.1$, $\mu = 10$. The “solutions” schematically shown below the spectrum are convenient representations of the eigenvectors of $-T(k)$ for the minimal value of k . This eigenvector consists of four complex elements, related to the x, y components at the two sites in each cell. Each element is represented as an arrow where the magnitude and direction of the vector corresponds to the absolute value and phase. In section 7.3 it will be seen that there is a connection between the mean-field order parameter and the coefficients of the eigenvectors.

For $t_h^{xy} = 0$, figure 6.1 shows that the minimal-energy excitations in the first band have $k_{min} = \pm\pi$ but with a slight “density” imbalance in favor of p_y . Remembering the form of the Fourier transformation, this means that the solutions for each primitive cell repeat with alternating sign giving a pattern with periodicity 4. This is interesting since it is twice the size of the primitive cell - in a loose sense the solution is “anti-ferromagnetic” over primitive cells. For $t_h^{xy} \neq 0$, the solution changes abruptly. This indicates that there is a quantum phase transition to a vortex-like solution - but the phase is in fact not extended to any finite range of t_h^{xy} and it is therefore not correct to talk of a phase transition. For $t_h^{xy} \neq 0$, k_{min} is no longer a multiple of π , leading to a “wave-like” solution, still with a slight density imbalance in favor of p_y . Furthermore, four values in the “Brillouin zone” are allowed, suggesting that the solution can be any linear combination of these waves. In the limit of $t_h^{xy} \rightarrow \infty$, $k_{min} = (2n + 1)\pi/2$ for n integer are all solutions. For $\mu = 10$ and generally for $\mu \neq 0$, the wavelike pattern survives but populations between the orbitals are shifted. In figure 6.3, the ground state only populates the p_x orbital, which has nodes along the direction of the chain. Otherwise, the periodicity of the solution is not changed.

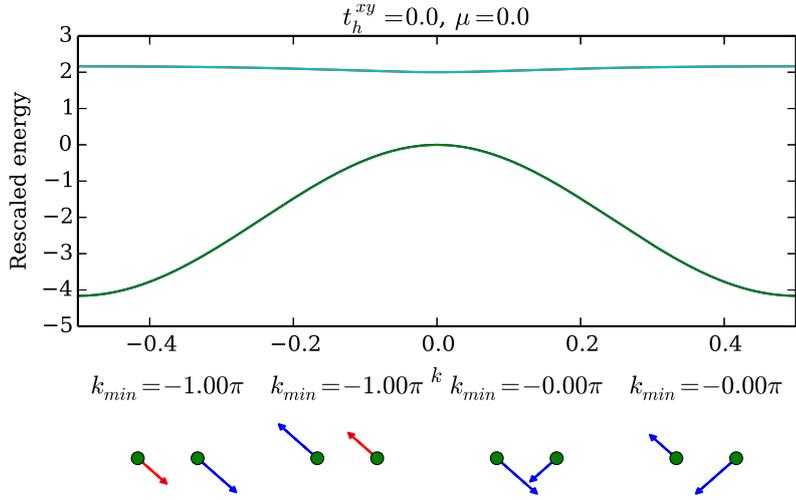


Figure 6.1: Spectrum and solutions for the ideal Hamiltonian for $t_h^{xy} = 0$, $\mu = 0$. The four bands are shown from lowest to highest energy from left to right. Note that the two lower and two upper bands are degenerate.

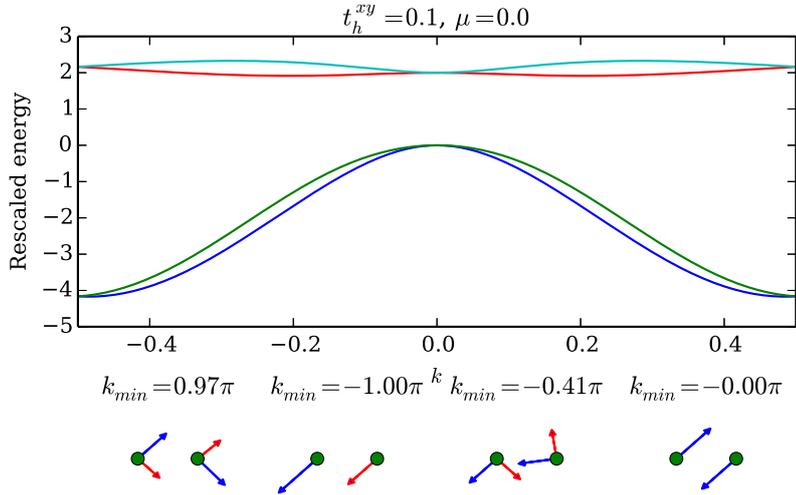


Figure 6.2: Same as figure 6.1, but with $t_h^{xy} = 0.1$. The four bands are shown from lowest to highest energy from left to right. Note that the second band now has density variation while the first band is actually a vortex-antivortex configuration with a quasimomentum not a multiple of π . In the limit $t_h^{xy} \rightarrow \infty$, $k = (2n + 1)\pi/2$.

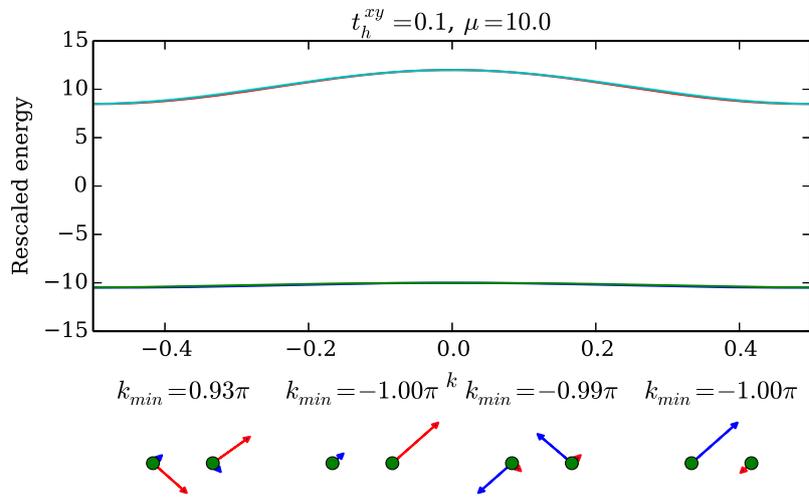


Figure 6.3: Spectrum and solutions for the ideal Hamiltonian for $t_h^{xy} = 0.1$, $\mu = 10$. The four bands are shown from lowest to highest energy from left to right. Note that the wavelike pattern survives as $\mu \rightarrow \infty$.

Chapter 7

Mean-field approximations

The shortest path between two truths in the real domain passes through the complex domain.

Jacques Hadamard

In this chapter, I will first demonstrate the use of a coherent state mean-field approximation for a Hubbard-like toy model. It is shown that the sign of the tunneling coefficients (referred to as the *couplings*, a term in statistical physics and quantum field theory) determine the phase of the system. The mean-field theory will then further be applied to investigate the phase transition in the quantum Ising chain and the vortex-antivortex order in the quadratic lattice, as described in section 5.1. Finally, the technique will be applied to the zig-zag lattice to further understanding of the order and phase transitions. This is a technique which is useful for understanding the system in the superfluid phase.

For any mean-field Hamiltonian which is a sum of terms, if it is possible to minimize each term independently, this is also a global minimum as, obviously,

$$\text{Min}[H] \geq \sum_j \text{Min}[H_j]. \quad (7.1)$$

While some mean-field approximations are soluble exactly, one often has to resort to numerical methods to find the global minimum configuration of a mean-field Hamiltonian. For this purpose, I wrote a software package for efficient mean-field minimization. Given a Hamiltonian represented as a function of the couplings and order parameters, the software returns the mean-field configuration and energy *as a function of the order parameters*. This higher-order function technique is useful for then finding the phases in coupling constant space. The algorithm uses the *BFGS* quasi-Newton method [JS06, p. 136] included in the `scipy` package for Python. Since the algorithm uses only first derivatives, it may get stuck in a local minimum. To reduce this risk, M random configurations $\{\psi_{i\alpha}\}$ are generated and H_{MF} is then minimized from each initial guess. As M becomes very big, it is reasonable to assume that some random configuration will be “close” to the global minimum configuration. While M will depend on the dimensionality of the system (the number of sites), it is possible to find a reasonable value by comparing the energy and minimum configuration for each iteration, as M becomes very big. The minimum energies for each configuration were saved and plotted in a histogram. One would expect that

the global minimum would appear as a cut in the lower end of the histogram for large enough M . For 16 sites, $M = 100 - 1000$, a significant fraction of the iterations found the ground state energy.

7.1 Toy model of next-nearest neighbor tunneling

I will now investigate the effects of next-nearest tunneling in a mean-field approximation. Consider a one-dimensional Hubbard-like toy model with non-negligible tunneling between nearest neighbors and next-nearest neighbors:

$$\hat{H} = -t_d \sum_i (\hat{a}_i^\dagger \hat{a}_{i+1} + \text{h.c.}) - t_h \sum_i (\hat{a}_i^\dagger \hat{a}_{i+2} + \text{h.c.}) + U \sum_i \hat{n}_i (\hat{n}_i - 1). \quad (7.2)$$

The actual values of t_d and t_h will depend on the lattice geometry. This model differs from the zig-zag Hamiltonian (5.40) in that it only has one “orbital”, or one type of particle. A coherent state variational ansatz, as described in section 2.5, $|\Psi\rangle(\{\alpha_j\}) = \prod_j |\alpha\rangle_j$ yields the mean-field Hamiltonian

$$\begin{aligned} H_{MF} &= \langle H \rangle = \\ &= -t_d \sum_i (\alpha_i^* \alpha_{i+1} + \text{transp.}) - t_h \sum_i (\alpha_i^* \alpha_{i+2} + \text{transp.}) + U \sum_i |\alpha_i|^4. \end{aligned} \quad (7.3)$$

Additionally, the particle number of the system is fixed, so that $N = \sum_j |\alpha_j|^2 = \text{constant}$. The interaction term is minimized by a constant density on each site $|\alpha_j|^2 = n$. Assuming that the interaction can be minimized independently of the kinetic part, I set $|\alpha_j|^2 = n$. The onsite order parameter can then be written as $\alpha_j = \sqrt{n} e^{i\theta_j}$ and the mean-field Hamiltonian as

$$H_{MF}/n = -2t_d \sum_i \cos(\theta_i - \theta_{i+1}) - 2t_h \sum_i \cos(\theta_i - \theta_{i+2}), \quad (7.4)$$

where I have dropped the last term since it only shifts the energy under the assumption of constant density. First, note that the system is invariant under multiplication by a pure phase, since it only depends on the differences of phases. This will show up as an overall phase in the numerical solutions.

In (t_d, t_h) -space, i.e. the plane of possible values of the coupling constants, there will be regions where different mean-field solutions dominate. These regions are the phases of the model, characterized by some order parameter containing information about the systems physical properties. For $t_h = 0$, the system is just a chain with nearest-neighbor interactions. Depending on the sign of t_d , the chain will be in the “ferromagnetic” or “anti-ferromagnetic” phase. Note that in the classical Ising model, the individual spins are ± 1 , while here the order parameters can assume any complex phase. The terminology here means that the order parameters all have the same phase (ferromagnetic) or an alternating phase (anti-ferromagnetic). For $t_d = 0$, the chain can be separated in *two* nearest-neighbor chains, since even sites only interact with even and odd sites only interact with odd. The sign of t_h determines if the chain is ferro- or anti-ferromagnetic. In the whole $t_d, t_h > 0$ quadrant the system is ferromagnetic, while it is anti-ferromagnetic in the quadrant $t_d < 0, t_h > 0$. This is because for $t_h > 0$, the second sum in (7.4) can be minimized on both the even and odd sublattice simultaneously. Only the sign of t_d determines

if the Hamiltonian is minimized by a ferro- or anti-ferromagnetic phase. For $t_h < 0$, the situation is different. Here there is necessarily some conflict between the terms, leading to some other phase difference between the sites.

Analytically, let me make the ansatz $\theta_j = kj$, where $kN = 2\pi m \Rightarrow k = 2\pi m/N$ for any integer m . Then:

$$H_{MF}/n = -2t_d \sum_i \cos(k) - 2t_h \sum_i \cos(2k) = -2N (t_d \cos(k) + t_h \cos(2k)) \quad (7.5)$$

$$\frac{dH_{MF}}{dk} = 0 \Rightarrow t_d \sin(k) = -2t_h \sin(2k) \Rightarrow t_d/t_h = -4 \cos(k), \quad (7.6)$$

with the solution $k = \arccos(-t_d/4t_h)$ which has the domain $-4t_h \leq t_d \leq 4t_h$. In the thermodynamic limit, a *spiral* order is expected to dominate in the wedge $-4t_h \leq t_d \leq 4t_h$ in the $t_h < 0$ half-plane. Note that on the $t_d = 0$ and $t_h < 0$ line, $k = \pi/2$ forming a texture of length 4. This texture can also be interpreted as an anti-symmetric configuration on the two sublattices. $t_d = 4t_h$ and $t_h < 0$ corresponds to $k = \pi$, which is the anti-ferromagnetic configuration already found in the lower right quadrant of the (t_h, t_d) -plane. I therefore expect the anti-ferromagnetic configuration to extend out to the wedge. $t_d = -4t_h, t_h < 0$ corresponds to $k = 0$, the ferromagnetic configuration. I expect the ferromagnetic phase to extend out from the upper right quadrant. Figure 7.1 shows the expected phase diagram. The pattern was verified by numerically minimizing the mean-field energy for a chain of 20 sites with periodic boundary conditions. The results are shown in nine selected plots of the order in figure 7.2.

The ferromagnetic phase has the energy per site $H_{MF}/Nn = -2(t_d + t_h)$, while the anti-ferromagnetic phase has energy $H_{MF}/Nn = -2(-t_d + t_h)$. The spiral phase has energy per site

$$H_{MF}/Nn = -2 \left[-t_d^2/4t_h + \left(2(t_d/4t_h)^2 - 1 \right) t_h \right] = -\frac{r}{4 \cos \theta} \left(8 \cos^2 \theta - \sin^2 \theta \right) \quad (7.7)$$

where r, θ defines the point (t_h, t_d) in polar coordinates. Figure 7.3 shows the energy and its first derivative, which depends linearly on the radius, as a function of angle θ in the (t_h, t_d) -plane. At $\theta = 0 = 2\pi$, corresponding to $t_d = 0, t_h > 0$, the energy has a cusp and its first derivative is discontinuous. This indicates a discontinuous or first-order phase transition as discussed in section 3.2. The transition to the spiral phase is marked with dashed, red lines. At these values of θ , both the energy and its first derivative is continuous, but the derivative has a cusp indicating a discontinuous second derivative. Hence the transition to the spiral phase is a continuous phase transition!

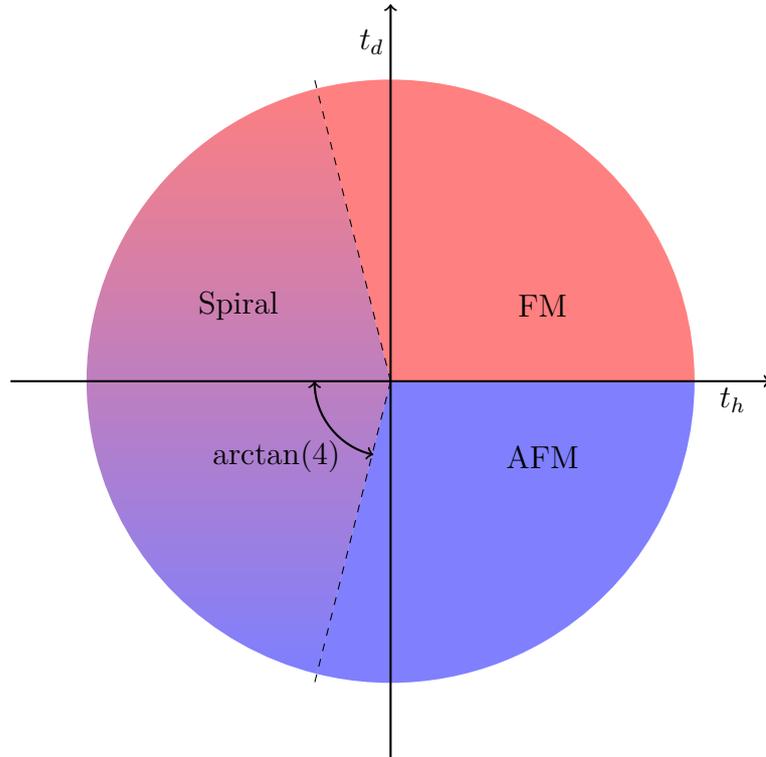


Figure 7.1: Phase diagram for the NNN toy model. For $t_d, t_h > 0$ and in a wedge around the positive t_d -axis, a ferromagnetic (FM) phase persists, while the system is anti-ferromagnetic (AFM) in a region mirrored in the t_h -axis. In a wedge of angle $2 \arctan(4)$ around the negative t_h -axis, the system has a spiral order parameter, with a period which is dependent on the angle to the origin. Note that the color is only drawn in a circle, while the full phase diagram fills the whole plane within the angles indicated!

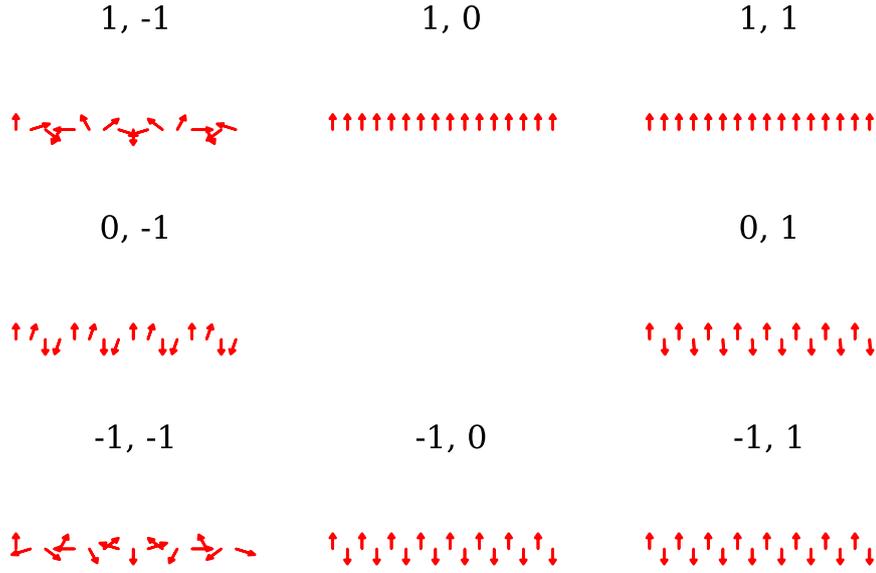


Figure 7.2: Minimum energy configurations for the NNN toy model with nearest and next-nearest neighbor couplings, for a range of couplings t_d and t_h . Note the spiral order parameter for $t_d < 0$

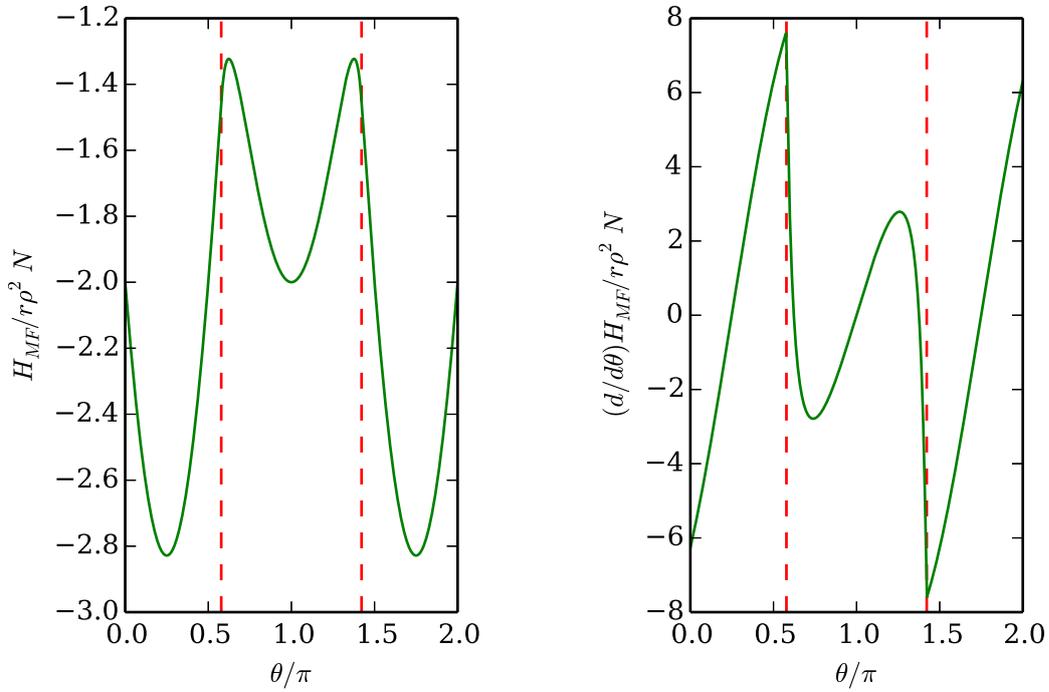


Figure 7.3: Energy divided by site, density and radius $r = \sqrt{t_h^2 + t_d^2}$ for a range of angles in the phase diagram for the NNN toy model. The dotted lines mark the border between the spiral phase and the ferro-/anti-ferromagnetic phases.

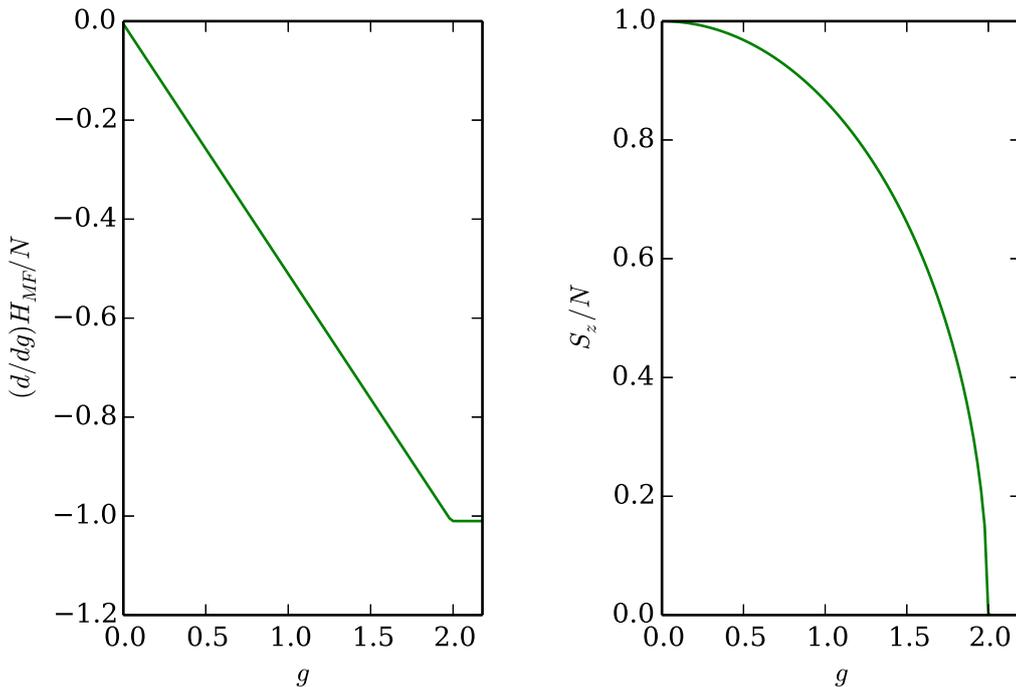


Figure 7.4: Energy per site for the mean-field approximation of the quantum Ising model and the first derivative, plotted as functions of g .

7.2 Mean-field treatment of the quantum Ising model

Using an ansatz which is a product state of spin coherent states [Aue94, p. 72], the Hamiltonian (3.27) can be semi-classically approximated as

$$\begin{aligned}
 H_{mf}/S &= -J \sum_j \cos(\theta_j) \cos(\theta_{j+1}) - Jg \sum_j \sin(\theta_j) \cos(\varphi_j) = \\
 &= \sum_j [-J \cos(\theta_j) \cos(\theta_{j+1}) + h \sin(\theta_j) \cos(\varphi_j)] \quad (7.8)
 \end{aligned}$$

where $S_j^z = S \cos(\theta_j)$, $S_j^x = S \sin(\theta_j) \cos(\varphi_j)$ and $S_j^y = S \sin(\theta_j) \sin(\varphi_j)$ are the mean-field order parameters. This Hamiltonian can be numerically minimized for different ratios of g . The minimum energy and its first derivative is plotted in figure 7.4. Since the first derivative of the energy is continuous, mean-field energy claims that the quantum phase transition is a continuous phase transition.

7.3 Mean-field derivation of vortex-antivortex structure

In section 5.1 I claimed that, in the limit where the hopping is weak, bosons in the p band loaded in a square lattice show a vortex-antivortex structure which is a product state of wavefunctions of the form (5.32). With mean-field theory, it is

possible to derive this solution. For the first excited orbital states, the appropriate coherent state ansatz is a product state $|\Psi\rangle = \prod_j |\psi_j^x, \psi_j^y\rangle_j$ where

$$|\alpha_j^x, \alpha_j^y\rangle_j = \exp\left(-\frac{|\alpha_j^x|^2 + |\alpha_j^y|^2}{2}\right) \sum_{n_x, n_y} \frac{(\alpha_j^x)^{n_x} (\alpha_j^y)^{n_y}}{\sqrt{n_x! n_y!}} |n_x n_y\rangle_j. \quad (7.9)$$

Writing the complex order parameters in polar form as $\psi_{jx} = \sqrt{n_{jx}} e^{i\theta_{jx}}$,

$$\begin{aligned} \hat{\mathcal{V}} &= \sum_{j\alpha} \frac{U_{\alpha\alpha}}{2} \hat{n}_{j\alpha} (\hat{n}_{j\alpha} - 1) + \sum_{j\alpha \neq \beta} U_{\alpha\beta} \hat{n}_{j\alpha} \hat{n}_{j\beta} + \frac{U_{\alpha\beta}}{2} (\hat{a}_{j\alpha}^\dagger \hat{a}_{j\alpha}^\dagger \hat{a}_{j\beta} \hat{a}_{j\beta} + \text{h.c.}) \\ \Rightarrow H_{\text{MF}} &= \sum_{j\alpha} \frac{U_{\alpha\alpha}}{2} n_{j\alpha} + \sum_{j\alpha \neq \beta} U_{\alpha\beta} \sqrt{n_{j\alpha}} \sqrt{n_{j\beta}} + \frac{U_{\alpha\beta}}{2} (\psi_{j\alpha}^* \psi_{j\alpha}^* \psi_{j\beta} \psi_{j\beta} + \text{h.c.}) \\ &= \sum_{j\alpha} \frac{U_{\alpha\alpha}}{2} n_{j\alpha} + \sum_{j\alpha \neq \beta} U_{\alpha\beta} \sqrt{n_{j\alpha}} \sqrt{n_{j\beta}} + \frac{U_{\alpha\beta}}{2} \sqrt{n_{j\alpha}} \sqrt{n_{j\beta}} \cos(2(\theta_{j\alpha} - \theta_{j\beta})). \end{aligned} \quad (7.10)$$

In the quadratic lattice, it is possible to minimize the interaction Hamiltonian and the kinetic part simultaneously [AS05]. Assuming constant density $n_j = n$

$$\hat{\mathcal{T}} = - \sum_{\langle ij \rangle \alpha \beta} t^{\alpha\beta} (\psi_{i\alpha}^* \psi_{j\beta} + \text{conj.}) = -2n \sum_{\langle ij \rangle, \alpha \beta} t^{\alpha\beta} \cos(\theta_{i\alpha} - \theta_{j\beta}). \quad (7.11)$$

Here the arbitrary site-indexing i has been replaced by a vector $\mathbf{i} = i_x, i_y$ meant to stand for the row and column of the site, and $t^{xy} = 0$ which is the reason that the phases in x and y can be set independently by the interaction term. For $t^{xx} > 0, t^{yy} < 0$ horizontally and $t^{xx} < 0, t^{yy} > 0$ vertically, as is the case in the quadratic lattice, the kinetic Hamiltonian is minimized term by term if the phase is locked in a *checkerboard* pattern, which is shown in figure 7.5, given by the single-site mean-field solution

$$\psi_{i_x} = (-1)^{i_y}, \quad (7.12)$$

$$\psi_{i_y} = (-1)^{i_x}. \quad (7.13)$$

This pattern ensures that all terms in the sum in the kinetic Hamiltonian have negative sign, since the product of two horizontal neighbors in the p_x state (in the p_y state) is positive (negative) and the opposite holds in the vertical direction. Hence the kinetic hamiltonian is minimized.

To find the spatial many-body wave function which the mean-field solution corresponds to, expand the expectation value of the position operator (2.10) in terms of Wannier functions as

$$\begin{aligned} \langle \psi_j(\mathbf{x}) \rangle &= \sum_j (w_{xj}(\mathbf{x}) \langle \hat{a}_{xj} \rangle + w_{yj}(\mathbf{x}) \langle \hat{a}_{yj} \rangle) = \\ &= \sum_j (\psi'_{xj} \psi_{xj}(\mathbf{x}) + \psi'_{yj} \psi_{yj}(\mathbf{x})), \end{aligned} \quad (7.14)$$

where $\psi'_{\alpha j}$ are here the mean-field order parameters, $\psi_{\alpha j}(\mathbf{x})$ are the Wannier functions and the dash has been added to avoid confusion. Hence the wavefunction at each site can be said to be in the *vortex-antivortex pattern* as was described in section 5.1.4.

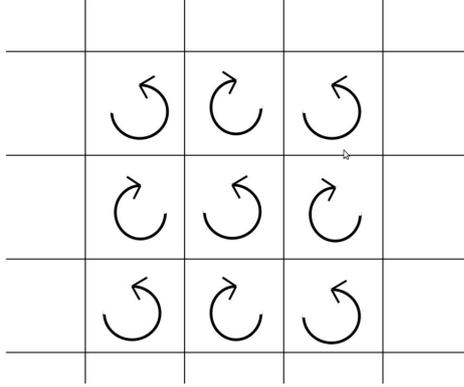


Figure 7.5: The checkerboard pattern which minimizes energy in a mean-field approximation of the quadratic lattice. The clockwise vortices indicate vortex solutions $|\psi_+\rangle$ and the anti-clockwise anti-vortex solutions $|\psi_-\rangle$. The phase on each site is not indicated in the figure.

7.4 Mean-field approximation of the zig-zag chain

The zig-zag Hamiltonian will now be solved in a mean-field approximation. Assuming a coherent state ansatz in the form (7.9), the normal ordered expectation value of the Hamiltonian (5.40) is:

$$\begin{aligned}
 H_{MF} = \langle : \hat{H} : \rangle = & - \sum_{i\alpha\beta} t_{i,d} \left(\psi_{i\alpha}^* \psi_{(i+1)\beta} + \psi_{(i+1)\alpha}^* \psi_{i\beta} \right) - \sum_{i\alpha\beta} t_{i,h} \left(\psi_{i\alpha}^* \psi_{(i+2)\beta} + \psi_{(i+2)\alpha}^* \psi_{i\beta} \right) + \\
 & + \sum_{i\alpha} \frac{U_{\alpha\alpha}}{2} |\psi_{i\alpha}|^4 + \sum_{i\alpha \neq \beta} U_{\alpha\beta} \left\{ |\psi_{i\alpha}|^2 |\psi_{i\beta}|^2 + \frac{1}{2} \left[(\psi_{i\alpha}^*)^2 (\psi_{i\beta})^2 + (\psi_{i\beta}^*)^2 (\psi_{i\alpha})^2 \right] \right\},
 \end{aligned} \tag{7.15}$$

where the index $i = 0, \dots, N - 1$. N has to be a multiple of 2 if periodic boundary conditions are applied ($i = 0$ identified with $i = N$), since there is no other way to fit the tunnelings in a consistent way. The mean-field Hamiltonian is a function of the $2N$ complex variables $\psi_{i\alpha}$. Conservation of particle number means that the order parameters additionally satisfy $N = \sum_i (|\psi_{ix}|^2 + |\psi_{iy}|^2)$. This constraint can be handled either by introducing a Lagrange multiplier to constrain the Hamiltonian or by finding coordinates where the constraint is “built-in”, so called *hyperspherical coordinates*. For this thesis, periodic boundary conditions and constrained coordinates were used.

Figure 7.6 shows the minimum energy configuration of $M = 1000$ iterations. The complex order parameters are shown as arrows located at each site, with the magnitude and direction proportional to the absolute value (density) and phase of the order parameter, as in chapter 6. For $U_0 = 0$ and $t_h^{xy} = 0$, the solution agrees with the exact solution found in chapter 6. For $U_0 = 0$ and $t_h^{xy} = 2$ however, the mean-field solution is a variation in densities, which was not the result found by exact diagonalization. However, recall that the exact solution for any value of t_h^{xy} has a spectrum with minima k_{min} symmetrically around $k = 0$. Thus, any linear combination of these states are solutions. If two “wave” solutions for $k_{min} = \pm\pi/2$ are added, the result is the density pattern observed for $U_0 = 0$, $t_h^{xy} = 2$.

For any $U_0 > 0$, there is a spontaneous symmetry breaking into one of the two values of k_{min} since the pattern needs to minimize the onsite interaction, which can

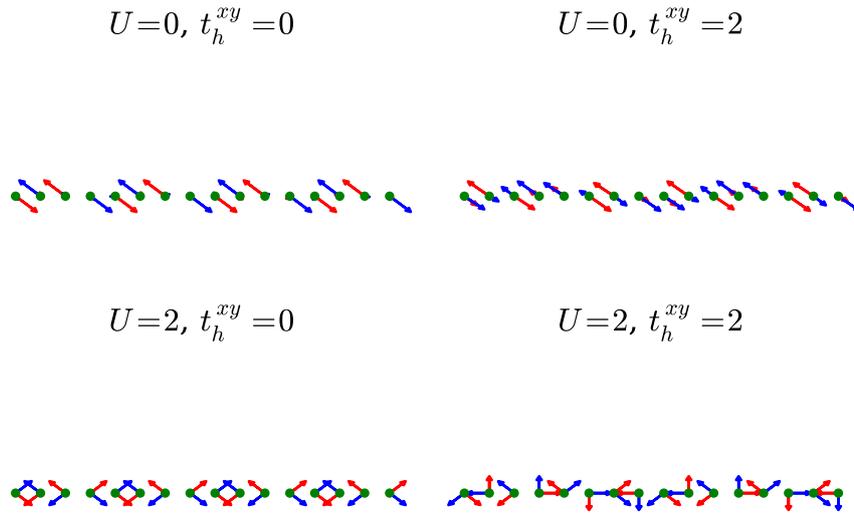


Figure 7.6: Minimum energy configuration for 16 sites, found by $M = 1000$ optimization iterations and periodic boundary conditions. The complex order parameters are visualized as arrows where ψ_j^x are red and ψ_j^y are blue. For $U_0 = 0$, the upper row, the solutions do not have to be in a vortex configuration, numerically choosing the minimum as a linear combination of two waves of opposite direction. For $U_0 > 0$, the lower row, there solution has to choose either direction of the wave. Note the periodicity of four for $t_h^{xy} = 0$ and eight for $t_h^{xy} = 2$.

be done independently as shown in section 7.3. This is visible for both $t_h^{xy} = 0$ and $t_h^{xy} = 2$, where the solution otherwise follows the results from exact diagonalization. Since the “phase” without vortex structure only appears for $U_0 = 0$ exactly, this is not a proper phase transition. The periodicity of the solutions is four for $t_h^{xy} = 0$ and sixteen for $t_h^{xy} = 2$. The latter is related to the limited size of the simulation - which was checked numerically by varying the number of sites in the simulation. With open boundary conditions or in the thermodynamic limit, the results should agree with the exact solution.

In section 8.4, a coherent state ansatz is tried for an effective Hamiltonian in the Mott phase. This should in principle give the same results as the full mean-field treatment in this section, but the results are hard to check for numerical reasons: as $U_0 \gg t$ the first derivative of the energy is much steeper in certain directions. The order between sites becomes impossible to find for the numerical algorithm and so far, I have not succeeded in finding a useful closed-form expression where the Mott phases are visible.

Chapter 8

The Mott insulator phase

The problems of the real world are primarily those you are left with when you refuse to apply their effective solutions.

Edsger W. Dijkstra

In the strongly interacting regime, the zig-zag lattice will undergo a quantum phase transition to the Mott insulator phase. Assuming that the number of particles matches the number of sites, the system will be in a state with one particle at each site. The sites are not completely isolated however: the atoms can interact quantum mechanically via “virtual” processes where two or more atoms interact. These processes are low-energy quantum fluctuations around the ground state which can be described by a simplified, *effective* Hamiltonian. In this chapter, an effective Hamiltonian for the Mott₁ phase will be developed by second-order perturbation theory. It will then be shown that this effective Hamiltonian can be interpreted as a one-dimensional spin problem with next-nearest neighbor coupling.

8.1 Perturbation theory

It is possible to partition the Fock space of the system in *singly* occupied states and states where *at least one* site is occupied by more than one boson. [Aue94, p. 25]. Define the *projection operator* \hat{P} as the operator projecting onto the subspace of states with only one atom per site¹. Define $\hat{Q} = 1 - \hat{P}$ as the complimentary projector onto the subspace of states with at least one double occupancy. The Hamiltonian for the full state $|\psi\rangle$ can then be written with block matrices as

$$\begin{pmatrix} \hat{P}\hat{H}\hat{P} & \hat{P}\hat{H}\hat{Q} \\ \hat{Q}\hat{H}\hat{P} & \hat{Q}\hat{H}\hat{Q} \end{pmatrix} \begin{pmatrix} \hat{P}|\psi\rangle \\ \hat{Q}|\psi\rangle \end{pmatrix} = E \begin{pmatrix} \hat{P}|\psi\rangle \\ \hat{Q}|\psi\rangle \end{pmatrix} \Leftrightarrow$$

$$(\hat{P}\hat{H}\hat{P} + \hat{P}\hat{H}\hat{Q})|\psi\rangle = E\hat{P}|\psi\rangle \tag{8.1}$$

$$(\hat{Q}\hat{H}\hat{P} + \hat{Q}\hat{H}\hat{Q})|\psi\rangle = E\hat{Q}|\psi\rangle, \tag{8.2}$$

¹ $\hat{P} = \sum_s |s\rangle \langle s|$, where the sum is over all singly occupied states.

where I have used $\hat{P}^2 = \hat{P}$, $\hat{Q}^2 = \hat{Q}$ which holds for any projection operator. Formally, it is always possible to eliminate $\hat{Q}|\psi\rangle$ from the first equation by rewriting the second as

$$\hat{Q}|\psi\rangle = (E - \hat{Q}\hat{H}\hat{Q})^{-1} \hat{Q}\hat{H}\hat{P}|\psi\rangle. \quad (8.3)$$

Inserting this expression into (8.1) yields

$$(\hat{P}\hat{H}\hat{P} + \hat{P}\hat{H}\hat{Q}(E - \hat{Q}\hat{H}\hat{Q})^{-1}\hat{Q}\hat{H}\hat{P})|\psi\rangle = E|\psi\rangle. \quad (8.4)$$

This is so far an exact expression. The first term describes self-interaction within the singly occupied subspace. In the form (5.40), $\hat{P}\hat{T}\hat{P} = 0$ since the chemical potential has been absorbed in the interaction part, but $\hat{P}\hat{V}\hat{P} \neq 0$. The second term describes interaction via virtual processes with the multiply occupied subspace. In the Mott phase, that strong interaction leads to a large energy gap between states with one atom per site and states with some double occupancy. Thus one would expect $\hat{H}_{22} \gg E$ which makes it possible to expand the second term to zeroth order in E/U

$$(E - \hat{Q}\hat{H}\hat{Q})^{-1} = -(\hat{Q}\hat{H}\hat{Q})^{-1} + \mathcal{O}(E/U), \quad (8.5)$$

so that the effective Hamiltonian may be written as

$$\hat{H}_{\text{eff}} = \hat{P}\hat{H}\hat{P} - \hat{P}\hat{H}\hat{Q}(\hat{Q}\hat{H}\hat{Q})^{-1}\hat{Q}\hat{H}\hat{P} + \mathcal{O}(E/U). \quad (8.6)$$

Furthermore, in our case, $\hat{P}\hat{V}\hat{Q} = 0$ since a state in the doubly occupied subspace cannot be brought to the singly occupied by the interaction, which is on-site. Also $\hat{Q}\hat{V}\hat{P} = 0$ by a similar argument. I find $\hat{P}\hat{H}\hat{Q} = \hat{P}\hat{T}\hat{Q}$ and $\hat{Q}\hat{H}\hat{P} = \hat{Q}\hat{T}\hat{P}$ so that the effective Hamiltonian can be simplified to

$$\hat{H}_{\text{eff}} = \hat{P}\hat{V}\hat{P} - \hat{P}\hat{T}\hat{Q}(\hat{Q}\hat{V}\hat{Q})^{-1}\hat{Q}\hat{T}\hat{P} + \mathcal{O}(E/U). \quad (8.7)$$

The first term describes the process of an atom interacting with itself. The second term describes the virtual process of an atom jumping from one site to another atom, interacting and jumping back. Note that the second term is smaller: it is proportional to t^2/U since it includes two kinetic terms and the inverse of the interaction. Higher terms in the expansion would be multiplied by higher powers of t/U since this is the only dimensionless small parameter that can be formed. This is the reason for calling this a second-order perturbation theory.

While this seems like an arbitrary trick, it is in fact the simplest case of *Kato-Takahashi perturbation theory* [D.10, p. 63], a type of perturbation theory which is suitable for dealing with Hamiltonians with degenerate spectra. One could expand the *resolvent operator* $\mathcal{G}(E) = (E - \hat{H})^{-1}$ in powers of E/U , a process which leads to the *t-J Hamiltonian* and the *Heisenberg* model in the Bose-Hubbard model [Aue94, p. 25]. This could in principle be used to find the energy and effective Hamiltonian to any order in E/U . Let me also note that the expansion treats the ideal part of the Hamiltonian as a perturbation, which is sometimes called a *strong coupling expansion*.

8.2 The effective Hamiltonian

Equation (8.7) only involves processes where one or two sites are involved at a time. Because of this, it is possible to treat single sites in the first term and pairs of sites for the second term. The first term of the effective Hamiltonian is easily found to be

$$\hat{P}\hat{\mathcal{V}}\hat{P} = \mu \sum_j \frac{1}{2} (\hat{n}_{xj} - \hat{n}_{yj}). \quad (8.8)$$

For each pair, the second term can be represented by a 4×4 matrix in a subspace of Fock space since there are four possible states for each pair of sites: $S = \{|xx\rangle, |xy\rangle, |yx\rangle, |yy\rangle\}$. The notation $|\alpha\beta\rangle$ simply means that there is a Wannier p_α -state at site A, and a Wannier p_β -state at site B. For doubly occupied states in the pair there are $D = \{|xx\rangle, |xy\rangle, |yy\rangle, |;yy\rangle, |;xy\rangle, |;xx\rangle\}$ where everything to the left of the semi-colon is at site A, and everything to the right at site B. Transitions between these states may be described in both matrix notation and by second-quantized operators. Since it is in this second-quantized form I want to be able to write the effective Hamiltonian, the correspondence between the representations has to be found. The matrix elements in the effective Fock Hamiltonian above must correspond to *c-number*² prefactors to these strings of creation and annihilation operators. The only allowed transitions in the second-order perturbation will be quadratic in creation and annihilation operators. The correspondence is:

$$\begin{pmatrix} A_{ij}^{xx} & E_{ij}^{xy} & E_{ji}^{xy} & F_{ij} \\ E_{ij}^{xy*} & A_{ij}^{xy} & S_{ij} & E_{ji}^{yx} \\ E_{ji}^{xy*} & S_{ij}^* & A_{ij}^{yx} & E_{ij}^{yx} \\ F_{ij}^* & E_{ji}^{yx*} & E_{ij}^{yx*} & A_{ij}^{yy} \end{pmatrix} \leftrightarrow \begin{pmatrix} \hat{n}_{ix}\hat{n}_{ix} & \hat{n}_{ix}\hat{f}_{jyx} & \hat{f}_{iyx}\hat{n}_{jx} & \hat{f}_{iyx}\hat{f}_{jyx} \\ \hat{n}_{ix}\hat{f}_{jxy} & \hat{n}_{ix}\hat{n}_{iy} & \hat{f}_{iyx}\hat{f}_{jxy} & \hat{f}_{ixy}\hat{n}_{jy} \\ \hat{f}_{ixy}\hat{n}_{jx} & \hat{f}_{ixy}\hat{f}_{jyx} & \hat{n}_{iy}\hat{n}_{jx} & \hat{n}_{iy}\hat{f}_{ixy} \\ \hat{f}_{ixy}\hat{f}_{jyx} & \hat{f}_{iyx}\hat{n}_{jy} & \hat{n}_{jy}\hat{f}_{jyx} & \hat{n}_{iy}\hat{n}_{iy} \end{pmatrix}. \quad (8.9)$$

where the “flip” operator $\hat{f}_{j\alpha\beta} = \hat{a}_{j\alpha}^\dagger \hat{a}_{j\beta}$ has been introduced for convenience. The site indices ij indicate that this gives the contribution to the effective Hamiltonian from one of the pairs. Observe that the self-adjoint property of the Hamiltonian means that all diagonal elements of the matrix to the left are real. Also note the $x \leftrightarrow y$ symmetry along the upper right-lower left diagonal. The latter (and of course the first) property holds for any number of single-atom states.

To solve this problem, I chose to write a software tool to find effective Hamiltonians such as (8.7) generally. The tool was written using the `sympy` library for Python and is capable of finding the effective Hamiltonian for any geometry and dimension, type of interaction, number of orbital states and set of tunneling coefficients. This could be useful for finding the effective Hamiltonians of systems in higher excited bands or higher Mott phases. However, for the moment the program is not yet capable of calculating some of the higher-order corrections which may be important for some lattice geometries. This will be implemented in later versions but could not be finished for this thesis. It could be used for searching a large set of tunneling coefficients and geometries for interesting effective Hamiltonians, making quantum engineering of spin models simpler and more efficient. Since second-order perturbations from and to Mott₁ cannot involve more than a pair of sites at a time,

²*c-number* is quantum field theory slang for anything that is not a second-quantized operator, usually a complex number.

the diagonal and horizontal pairs lead to contributions of similar form, except that t_d has to be replaced with t_h . For this thesis, the calculation was also done by hand. In matrix notation, the contribution from one pair to the second-order perturbation can be written as $TV^{-1}T^\dagger$ where

$$T = \begin{pmatrix} t_+^{xx} & t_+^{yx} & 0 & 0 & t_-^{yx} & t_-^{xx} \\ t_+^{xy} & t_+^{yy} & 0 & t_-^{yx} & t_-^{xx} & 0 \\ 0 & t_+^{xx} & t_+^{yx} & 0 & t_-^{yy} & t_-^{xy} \\ 0 & t_+^{xy} & t_+^{yy} & t_-^{yy} & t_-^{xy} & 0 \end{pmatrix} \quad (8.10)$$

where $+, -$ is shorthand for ij, ji denoting the *direction* (since tunneling may not be the same in both directions, as is the case for the orbital changing next-nearest neighbor tunneling) and

$$V = \begin{pmatrix} U_{xx} + \mu & 0 & U_{xy}/2 & 0 & 0 & 0 \\ 0 & U_{xy} & 0 & 0 & 0 & 0 \\ U_{xy}/2 & 0 & U_{yy} - \mu & 0 & 0 & 0 \\ 0 & 0 & 0 & U_{yy} - \mu & 0 & U_{xy}/2 \\ 0 & 0 & 0 & 0 & U_{xy} & 0 \\ 0 & 0 & 0 & U_{xy}/2 & 0 & U_{xx} + \mu \end{pmatrix}, \quad (8.11)$$

which is block-diagonal and has the inverse

$$V^{-1} = \begin{pmatrix} -4(U_{yy}-\mu)/U^2 & 0 & 2U_{xy}/U^2 & 0 & 0 & 0 \\ 0 & 1/U_{xy} & 0 & 0 & 0 & 0 \\ 2U_{xy}/U^2 & 0 & -4(U_{xx}+\mu)/U^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4(U_{xx}+\mu)/U^2 & 0 & 2U_{xy}/U^2 \\ 0 & 0 & 0 & 0 & 1/U_{xy} & 0 \\ 0 & 0 & 0 & 2U_{xy}/U^2 & 0 & -4(U_{yy}-\mu)/U^2 \end{pmatrix}. \quad (8.12)$$

where $U^2 = U_{xy}^2 - 4(U_{xx} + \mu)(U_{yy} - \mu)$. Note that the second-order correction will have two poles at $\mu = \pm\sqrt{U_{xx}^2 - U_{xy}^2/4}$ (assuming $U_{xx} = U_{yy}$ which is true for this lattice). From the perspective of Kato-Takahashi perturbation theory, the poles could be interpreted as indications that the assumption $E \ll \hat{H}_{22}$ breaks down so that the perturbation theory is no longer valid. This could be because large absolute values of μ increases the energy in the singly occupied state. The effective Hamiltonian can now be written as

$$\begin{aligned} \hat{H}_{\text{eff}} = & \mu \sum_j \frac{1}{2} (\hat{n}_{jx} - \hat{n}_{jy}) - \sum_{ij\alpha\beta} A_{ij}^{\alpha\beta} \hat{n}_{i\alpha} \hat{n}_{j\beta} + \\ & - \sum_{ij\alpha\neq\beta} \{ F_{ij} \hat{f}_{i\alpha\beta} \hat{f}_{j\alpha\beta} + S_{ij} \hat{f}_{i\alpha\beta} \hat{f}_{j\beta\alpha} + 2E_{ij}^{\alpha\beta} \hat{n}_{i\alpha} \hat{f}_{j\alpha\beta} + \text{h.c.} \}, \end{aligned} \quad (8.13)$$

where

$$A^{xx} = \frac{1}{U_{xy}U^2} \left(-4U_{xy} |t_-^{xx}|^2 (U_{yy} - \mu) - 4U_{xy} |t_+^{xx}|^2 (U_{yy} - \mu) + U^2 (|t_-^{yx}|^2 + |t_+^{yx}|^2) \right), \quad (8.14)$$

$$A^{yy} = \frac{1}{U_{xy}U^2} \left(-4U_{xy} |t_-^{yy}|^2 (U_{xx} + \mu) - 4U_{xy} |t_+^{yy}|^2 (U_{xx} + \mu) + U^2 (|t_-^{xy}|^2 + |t_+^{xy}|^2) \right), \quad (8.15)$$

$$A^{xy} = \frac{1}{U_{xy}U^2} \left(-4U_{xy} |t_+^{xy}|^2 (U_{yy} - \mu) - 4U_{xy} |t_-^{yx}|^2 (U_{xx} + \mu) + U^2 (|t_-^{xx}|^2 + |t_+^{yy}|^2) \right), \quad (8.16)$$

$$A^{yx} = \frac{1}{U_{xy}U^2} \left(-4U_{xy} |t_-^{xy}|^2 (U_{yy} - \mu) - 4U_{xy} |t_+^{yx}|^2 (U_{xx} + \mu) + U^2 (|t_-^{xx}|^2 + |t_+^{yy}|^2) \right), \quad (8.17)$$

$$E^{xy} = \frac{1}{U_{xy}U^2} \left(2U_{xy}^2 t_-^{xx} t_-^{yx} - 4U_{xy} t_+^{xx} t_+^{yx} (U_{yy} - \mu) + U^2 (t_-^{xx} t_-^{yx} + t_+^{yx} t_+^{yy}) \right), \quad (8.18)$$

$$E^{yx} = \frac{1}{U_{xy}U^2} \left(2U_{xy}^2 t_+^{xy} t_+^{yy} - 4U_{xy} t_-^{yx} t_-^{yy} (U_{xx} + \mu) + U^2 (t_-^{xx} t_-^{xy} + t_+^{xy} t_+^{yy}) \right), \quad (8.19)$$

$$F = \frac{1}{U_{xy}U^2} \left(2U_{xy}^2 (t_-^{xx} t_-^{yy} + t_+^{xx} t_+^{yy}) + U^2 (t_-^{xy} t_-^{yx} + t_+^{xy} t_+^{yx}) \right), \quad (8.20)$$

$$S = \frac{1}{U_{xy}U^2} \left(2U_{xy}^2 (t_-^{xy} t_-^{yx} + t_+^{xy} t_+^{yx}) + U^2 (t_-^{xx} t_-^{yy} + t_+^{xx} t_+^{yy}) \right), \quad (8.21)$$

As demonstrated in chapter 5, the orbital-changing tunnelings alternate in sign between sites, and the tunnelings are different between diagonal and horizontal pairs. From the above equations, it is then clear that only the terms E^{xy} , E^{yx} , F and S alternate along the chain.

8.3 Schwinger spin-boson mapping

Because of the fixed density, the orbital degrees of freedom may be interpreted as a spinor³ and the effective Hamiltonian as a spin Hamiltonian. This is known as a *Schwinger spin-boson mapping*. For a larger number of orbital states the system can be mapped to higher spin models. The map for spin 1/2 states is [AS10, p. 88]

$$\hat{S}_i^z = \frac{1}{2} (\hat{a}_{ix}^\dagger \hat{a}_{ix} - \hat{a}_{iy}^\dagger \hat{a}_{iy}) = \frac{1}{2} (\hat{n}_{ix} - \hat{n}_{iy}), \quad (8.22)$$

$$\hat{S}_i^+ = \hat{S}_i^x + i\hat{S}_i^y = \hat{f}_{xy}, \quad (8.23)$$

$$\hat{S}_i^- = \hat{S}_i^x - i\hat{S}_i^y = \hat{f}_{yx}, \quad (8.24)$$

$$(8.25)$$

³In the sense of non-relativistic 2-spinors here, not the full relativistic spinors.

and additionally, the constraint $\hat{n}_x + \hat{n}_y = 1$ holds, as the system is in the Mott₁ phase. Also note that the self-adjoint property of the Hamiltonian in the Fock basis means that any diagonal terms (only involving number operators) must be real, and any orbital-changing terms must appear with their hermitian conjugate (so that the sum is hermitian).

Using the spin-boson mapping, I found

$$\hat{S}_i^z = \frac{1}{2} (\hat{n}_{ix} + \hat{n}_{iy}) = \hat{n}_{ix} - \frac{1}{2} = \frac{1}{2} - \hat{n}_{iy} \Rightarrow \begin{cases} \hat{n}_{ix} = \frac{1}{2} (1 + 2\hat{S}_i^z) \\ \hat{n}_{iy} = \frac{1}{2} (1 - 2\hat{S}_i^z) \end{cases}, \quad (8.26)$$

which can be used to calculate

$$\hat{n}_{\alpha i} \hat{n}_{\alpha j} = \frac{1}{4} (1 \pm 2\hat{S}_i^z) (1 \pm 2\hat{S}_j^z) = \frac{1}{4} [1 \pm 2(\hat{S}_i^z + \hat{S}_j^z) + 4\hat{S}_i^z \hat{S}_j^z]. \quad (8.27)$$

The \pm depends on if α is x or y . Furthermore

$$\hat{n}_{\alpha i} \hat{n}_{\beta j} = \frac{1}{4} (1 \pm 2\hat{S}_i^z) (1 \mp 2\hat{S}_j^z) = \frac{1}{4} [1 \pm 2(\hat{S}_i^z - \hat{S}_j^z) - 4\hat{S}_i^z \hat{S}_j^z], \quad (8.28)$$

where again \pm depends on if α is x or y (remember that the sum over orbitals is only over $\alpha \neq \beta$). Also

$$\begin{aligned} \hat{n}_{\alpha i} \hat{f}_{\alpha \beta j} &= \frac{1}{2} (1 \pm 2\hat{S}_i^z) (\hat{S}_j^x \pm i\hat{S}_j^y) = \frac{1}{2} (\hat{S}_j^x \pm i\hat{S}_j^y) + i\hat{S}_i^z \hat{S}_j^y \pm \hat{S}_i^z \hat{S}_j^x \\ &\Rightarrow \hat{n}_{\alpha i} \hat{f}_{\alpha \beta j} + \text{h.c.} = \hat{S}_i^x \pm 2\hat{S}_i^z \hat{S}_j^x, \end{aligned} \quad (8.29)$$

$$\begin{aligned} \hat{f}_{\alpha \beta i} \hat{n}_{\alpha j} &= (\hat{S}_i^x \pm i\hat{S}_i^y) \frac{1}{2} (1 \pm 2\hat{S}_j^z) = \frac{1}{2} (\hat{S}_i^x \pm i\hat{S}_i^y) + i\hat{S}_i^y \hat{S}_j^z \pm \hat{S}_i^x \hat{S}_j^z \\ &\Rightarrow \hat{f}_{\alpha \beta i} \hat{n}_{\alpha j} + \text{h.c.} = \hat{S}_i^x \pm 2\hat{S}_i^z \hat{S}_j^z, \end{aligned} \quad (8.30)$$

$$\begin{aligned} \hat{f}_{\alpha \beta i} \hat{f}_{\alpha \beta j} &= \hat{S}^\pm \hat{S}^\pm = (\hat{S}_i^x \pm i\hat{S}_i^y) (\hat{S}_j^x \pm i\hat{S}_j^y) = \hat{S}_i^x \hat{S}_j^x \pm i(\hat{S}_i^x \hat{S}_j^y + \hat{S}_i^y \hat{S}_j^x) - \hat{S}_i^y \hat{S}_j^y \\ &\Rightarrow \hat{f}_{\alpha \beta i} \hat{f}_{\alpha \beta j} + \text{h.c.} = 2(\hat{S}_i^x \hat{S}_j^x - \hat{S}_i^y \hat{S}_j^y), \end{aligned} \quad (8.31)$$

$$\begin{aligned} \hat{f}_{\alpha \beta i} \hat{f}_{\beta \alpha j} &= \hat{S}^\pm \hat{S}^\mp = (\hat{S}_i^x \pm i\hat{S}_i^y) (\hat{S}_j^x \mp i\hat{S}_j^y) = \hat{S}_i^x \hat{S}_j^x \pm i(\hat{S}_i^y \hat{S}_j^x - \hat{S}_i^x \hat{S}_j^y) + \hat{S}_i^y \hat{S}_j^y \\ &\Rightarrow \hat{f}_{\alpha \beta i} \hat{f}_{\beta \alpha j} + \text{h.c.} = 2(\hat{S}_i^x \hat{S}_j^x + \hat{S}_i^y \hat{S}_j^y), \end{aligned} \quad (8.32)$$

where I have anticipated that $E_{ij}^{\alpha\beta}$, S_{ij} and F_{ij} are real, since all tunnelings can be chosen to be real in the absence of external gauge fields[JFJO11]. Also note that the first order perturbation from the chemical potential maps directly to a field in the z -direction.

8.4 The spin Hamiltonian

The spin Hamiltonian can be written as

$$\hat{H}_{\text{spin}} = - \sum_j \hat{\mathbf{S}}_j J_d \hat{\mathbf{S}}_{j+1} - \sum_j \hat{\mathbf{S}}_j J_h \hat{\mathbf{S}}_{j+2} + \sum_j \mathbf{D}_j \cdot \hat{\mathbf{S}}_j \times \hat{\mathbf{S}}_{j+2} + \mathbf{h} \cdot \sum_j \hat{\mathbf{S}}_j, \quad (8.33)$$

where J_d, J_h are diagonal 3×3 matrices. The first two terms are called the *spin-spin* interactions, respectively between nearest and next-nearest neighbors along the chain. The third term is called a *Dzyaloshinskii-Moriya interaction* or *anti-symmetric exchange*. Here $D_j = (-1)^j D \hat{\mathbf{y}}$ so that the term involves the anti-symmetrized product of the x and z components of the spin operator as

$$\sum_j \mathbf{D}_j \cdot \hat{\mathbf{S}}_j \times \hat{\mathbf{S}}_{j+2} = D \sum_j (-1)^j (\hat{S}_j^x \hat{S}_{j+2}^z - \hat{S}_j^z \hat{S}_{j+2}^x). \quad (8.34)$$

The fourth term is called an *external field* since it acts on the spins like an external magnetic field in the \mathbf{h} direction. It includes contributions from both the first-order and second-order terms in the effective Hamiltonian.

The coupling constants were found by summing over the orbitals for each site as

$$J_{ij}^x = 4(S_{ij} + F_{ij}), \quad (8.35)$$

$$J_{ij}^y = 4(S_{ij} - F_{ij}), \quad (8.36)$$

$$J_{ij}^z = 2(A_{ij}^{xx} + A_{ij}^{yy} - (A_{ij}^{xy} + A_{ij}^{yx})), \quad (8.37)$$

$$D_{ij}^{xz} = 2(E_{ij}^{xy} - E_{ij}^{yx}), \quad (8.38)$$

$$h_i^x = \sum_q (E_{iq}^{xy} + E_{iq}^{yx}), \quad (8.39)$$

$$h_i^z = \mu + \sum_q ((A_{iq}^{xx} - A_{iq}^{yy}) - (A_{qi}^{xx} - A_{qi}^{yy})). \quad (8.40)$$

So that, for example, the diagonal elements of the matrix J_d are found by summing over indices $i, j = i, (i + 1)$. The sum over q is over nearest- and next-nearest neighbors of site i . Its origin is the summation over pairs in the effective Hamiltonian. Since $t_d^{xx} = t_d^{yy} = 0$, $E_d^{xy} = E_d^{yx} = 0$ cancelling the Dzyaloshinskii-Moriya interaction between nearest neighbors. For horizontal pairs $E_{ij}^{\alpha\beta} = -E_{ji}^{\alpha\beta}$ so that $h_i^x = 0$ for all sites, but the anti-symmetric interaction does not vanish since $E^{xy} \neq E^{yx}$.

It is not possible to directly vary the value of the coupling constant, but it is possible to do so indirectly by tuning the interactions and tunnelings. Which coupling constants are possible to realize in the zig-zag lattice? First, notice that the x and z components of the Dzyaloshinskii-Moriya vector \mathbf{D} are zero because F, S and $E^{\alpha\beta}$ are *real*. In general, for example,

$$E_{ij}^{\alpha\beta} \hat{n}_{\alpha i} \hat{f}_{\alpha\beta j} + \text{h.c.} = \Re(E_{ij}^{\alpha\beta}) (\hat{S}_i^x \pm 2\hat{S}_i^z \hat{S}_j^x) \pm i\Im(E_{ij}^{\alpha\beta}) (\hat{S}_j^y + 2\hat{S}_j^z \hat{S}_j^z) \quad (8.41)$$

Since complex tunnelings correspond to *artificial gauge fields*[JFJO11], these components of the anti-symmetric interaction are generally related to such fields. *Transverse magnetic fields* given by $h^x \neq 0$, are possible when there are pairs of sites between which $t^{xx}, t^{xy}, t^{yy} \neq 0$ and $t_+^{xy} = t_+^{yx}$, which follows from the argument about h^x above.

The size of the spin-spin couplings can be calculated in the absence of a chemical potential and by assuming $t_+^{\alpha\alpha} = t_-^{\alpha\alpha}$. Then J^x, J^y are proportional to

$$S \pm F \propto t^{xx}t^{yy} \pm t_+^{xy}t_+^{yx}, \quad (8.42)$$

while

$$J^z \propto 2|t^{xy}|^2 - |t^{xx}|^2 - |t^{yy}|^2, \quad (8.43)$$

$$D \propto (t^{xx} - t^{yy})t_+^{xy}. \quad (8.44)$$

From these equations, one can see that it is possible to tune J^z to any positive or negative value. Setting $J^z = 0$ corresponds to (for real tunnelings) $2|t^{xy}|^2 = |t^{xx}|^2 + |t^{yy}|^2$ which leads to $J^x \propto (t^{xx} \pm t^{yy})^2 / 2$ and a similar positive-definite value for J^y . Hence either J^x or J^y is *as large as the anti-symmetric interaction*, for any tuned value of the tunnelings. Therefore, I conclude that the D-M term is never large enough to dominate the solution in itself. Note that the term may still be large enough to *influence* it, for example by breaking the parity symmetry.

From (8.42) and (8.43), it is clear that the relative strength of the nearest- and next-nearest spin-spin couplings can be set by tuning the tunnelings. Because of this, it should be possible to use the zig-zag model to simulate spin models with a varying strength of next-nearest neighbor couplings. For the zig-zag lattice, $t_h^{xx} = 0$ and $t_h^{yy} = 1.5$ in the positive direction. From equation (8.43) it is then clear that J_h^z can be tuned to negative values by setting $|t_h^{yy}| \leq 2$. Figure 8.1 shows a plot of the next-nearest neighbor couplings for $t_h^{yy} = -2, h_z = h_{z0}$ and $U_0 = 100$. The indicated phases will be discussed in the next paragraphs.

Note that spin Hamiltonians such as (8.33) are often simplified by the application of a *Jordan-Wigner transformation* [S.11, p. 46] [M.]. In this case, however, such a transformation is not an alternative. The Jordan-Wigner transformation rewrites the problem as a fermionic Hamiltonian by a non-local transformation

$$\hat{S}_j^z = \hat{a}_j \hat{a}_j^\dagger - \hat{a}_j^\dagger \hat{a}_j \quad (8.45)$$

$$\hat{S}_j^x = - \left(\bigotimes_{n=1}^{j-1} \hat{S}_n^z \right) (\hat{a}_j + \hat{a}_j^\dagger) \quad (8.46)$$

$$\hat{S}_j^y = i \left(\bigotimes_{n=1}^{j-1} \hat{S}_n^z \right) (\hat{a}_j^\dagger - \hat{a}_j) \quad (8.47)$$

Because of the fact that

$$\begin{aligned} \hat{S}_j^z \hat{S}_j^z &= (\hat{a}_j \hat{a}_j^\dagger - \hat{a}_j^\dagger \hat{a}_j)(\hat{a}_j \hat{a}_j^\dagger - \hat{a}_j^\dagger \hat{a}_j) = \hat{a}_j \hat{a}_j^\dagger \hat{a}_j \hat{a}_j^\dagger + \hat{a}_j^\dagger \hat{a}_j \hat{a}_j^\dagger \hat{a}_j = \\ &= (1 - \hat{a}_j^\dagger \hat{a}_j) \hat{a}_j \hat{a}_j^\dagger + (1 - \hat{a}_j \hat{a}_j^\dagger) \hat{a}_j^\dagger \hat{a}_j = \hat{a}_j \hat{a}_j^\dagger + \hat{a}_j^\dagger \hat{a}_j = 1, \end{aligned} \quad (8.48)$$

where the fermion anti-commutation rules $\{\hat{a}_j, \hat{a}_j\} = \{\hat{a}_j^\dagger, \hat{a}_j^\dagger\} = 0$ and $\{\hat{a}_j^\dagger, \hat{a}_j\} = 1$ were used,

$$\begin{aligned} \hat{S}_j^x \hat{S}_{j+1}^x &= \left(\bigotimes_{n=1}^{j-1} \hat{S}_n^z \hat{S}_n^z \right) (\hat{a}_j + \hat{a}_j^\dagger) \hat{S}_j^z (\hat{a}_{j+1} + \hat{a}_{j+1}^\dagger) = \\ &= (\hat{a}_j + \hat{a}_j^\dagger)(\hat{a}_j \hat{a}_j^\dagger - \hat{a}_j^\dagger \hat{a}_j)(\hat{a}_{j+1} + \hat{a}_{j+1}^\dagger). \end{aligned} \quad (8.49)$$

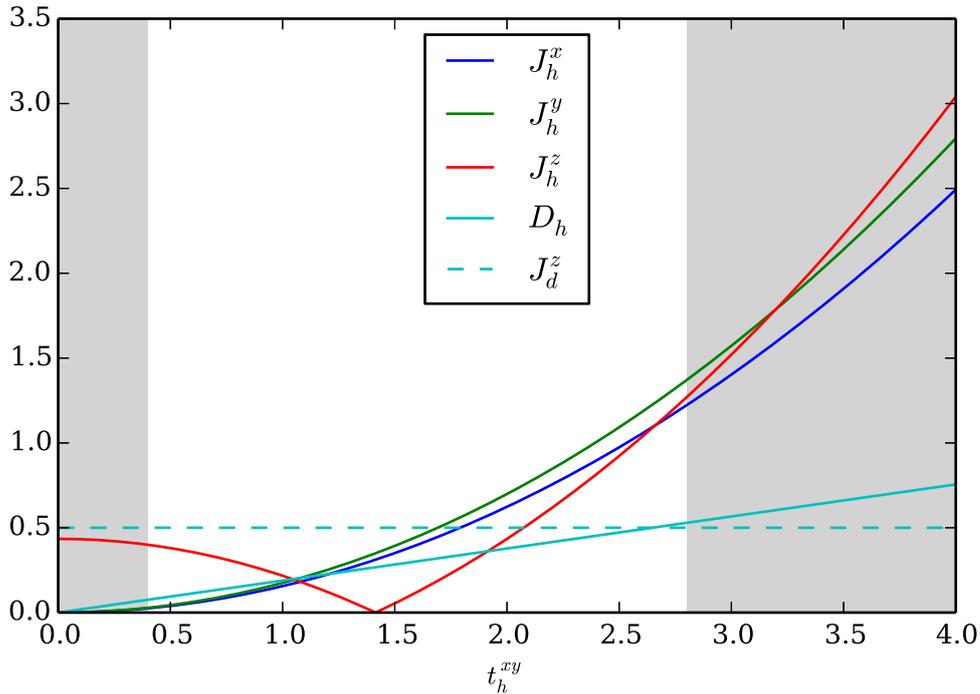


Figure 8.1: Magnitude of next-nearest neighbor coupling constants J_h^x , J_h^y , J_h^z and D for a range of values of t_h^{xy} while keeping $U_0 = 100$ and $t_h^{yy} = -2$ fixed. The dotted cyan line shows the size of J_d^z which is the largest nearest-neighbor coupling. The left area marked in light gray shows where a two up-two down-phase dominates with all spins pointing in the $\pm z$ direction. The right area marked in light gray shows where a phase which is polarized in z spin component dominates. In between is a phase which is anti-ferromagnetic in the y spin component.

Since $(\hat{a}_j + \hat{a}_j^\dagger)(\hat{a}_j \hat{a}_j^\dagger - \hat{a}_j^\dagger \hat{a}_j) = \hat{a}_j^\dagger \hat{a}_j \hat{a}_j^\dagger - \hat{a}_j \hat{a}_j^\dagger \hat{a}_j = \hat{a}_j^\dagger - \hat{a}_j$, the above expression reduces to

$$\hat{S}_j^x \hat{S}_{j+1}^x = (\hat{a}_j^\dagger - \hat{a}_j) (\hat{a}_{j+1} + \hat{a}_{j+1}^\dagger). \quad (8.50)$$

However, no such cancellation occurs for products $\hat{S}_j^x \hat{S}_{j+2}^z$ such as those in the anti-symmetric term. Only a Dzyaloshinskii-Moriya component in the $\hat{\mathbf{z}}$ component can be transformed to a local term of Fermi operators. Therefore, the Jordan-Wigner transformation is of no use for this particular spin Hamiltonian. Note that it is possible to “rotate” the Jordan-Wigner transformation as is done in [S.11, p. 47] so that the anti-symmetric term becomes local, but this comes to the price of a non-local term from the external field. Even in the special case $h_z = 0$, the fermionic Hamiltonian have terms containing operators acting on three different sites, which is a consequence of the next-nearest neighbor interaction.

8.5 Mean-field approximation of the spin model

The spin model was once again solved in a mean-field approximation. It is clear that for spin-1/2 particles quantum fluctuations are typically very important (especially in lower dimensions). Nevertheless, a mean-field analysis is sufficient to give insight into the phases of the spin system [JA10]. The phase diagram in the (t_h^{xy}, μ) plane was investigated for $t_h^{yy} = -2$, because of the interesting negative value of J_h^z discussed in the last section. Figure 8.2 shows a heatmap plot of the first derivative of the energy. For clarity, four plots along lines $t_h^{xy} = \text{constant}$ were also produced, showing the first derivative of the energy. These are shown in figure 8.3. Additionally, a plot of the first derivative along the $h_z = 0$ axis is included in figure 8.4.

On the line t_h^{xy} , the top left plot in figure 8.3 and the vertical axis in figure 8.2, there are a total of four phase transitions visible: two outer continuous phase transitions and two inner discontinuous phase transitions. The inner discontinuous phase transition is visible in the right plot in figure 8.2. The lines $t_h^{xy} = 1$, $t_h^{xy} = 2$ are characterized by two continuous phase transitions while for $t_h^{xy} > 2.8$ the phase transition is again discontinuous. A further plot along the line $h_z = 0$, shown in figure 8.4, indicate that there are two discontinuous phase transitions at $t_h^{xy} \approx 0.4$ and $t_h^{xy} \approx 2.8$.

In figure 8.2, it is also apparent that there are four phases: one in a smaller semi-circle around $t_h^{xy} = h_z = 0$, one inside a larger “gothic arch” shape and outside a third and fourth phase. Figure 8.6 shows three of these spin configurations. The first phase, here called the *4-pattern*, dominate in a semi-circle around $t_h^{xy} = h_z = 0$. This is the configuration expected from minimizing an anti-ferromagnetic next-neighbor coupling $-\sum_j J_h^z S_j^z S_{j+2}^z$ where $J_h^z < 0$. Then it holds that $S_j^z S_{j+2}^z = -1$ for every pair. Since this is not the only large term for these values of the couplings, it has to result from *competition* between terms. In the wider “arch”-shaped region, the spins are in the *y-AFM* phase. This is obviously a configuration resulting from a nearest-neighbor spin-spin coupling $J_d^y < 0$, but again has to result from competition between terms. Note that it is not possible to continuously deform the first phase into the second. The phase is polarized in the $\pm z$ direction for non-zero values of μ . For large enough values of μ and/or large enough values of t_h^{xy} the *y*

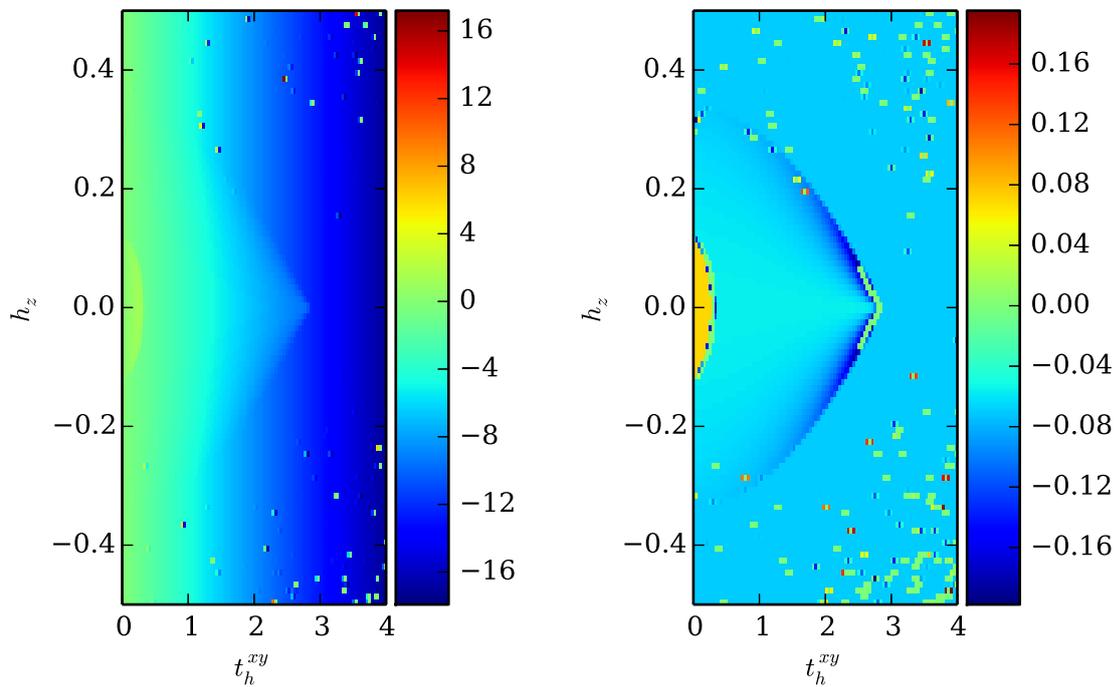


Figure 8.2: Heatmaps of the first derivative of the minimized mean-field energy H_{mf} for the spin Hamiltonian, in the t_h^{xy} and h_z direction ($h_z = \mu - h_{z0}$ where $h_{z0} \approx 0.1$). The left plot shows $\frac{\partial E}{\partial h_z}$ while the right plot shows $\frac{\partial E}{\partial t_h^{xy}}$. $t_h^{yy} = -2, t_d^{xy} = 1.5$ for both figures. The “artefacts” are due to numerical errors in the minimization.

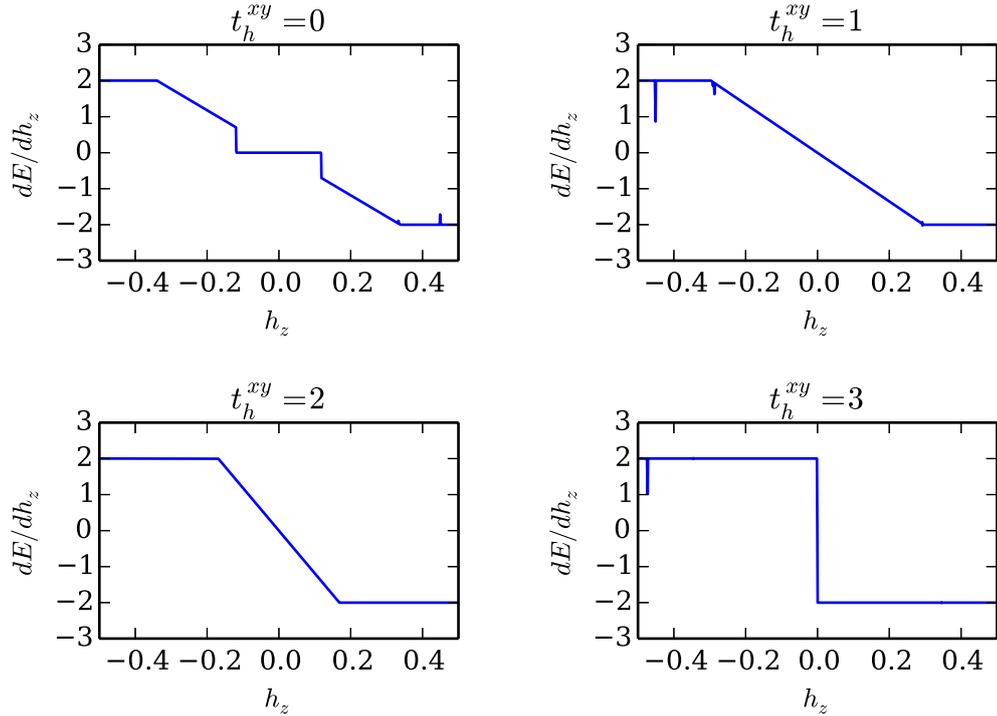


Figure 8.3: Sweeps along lines in figure 8.2. $\frac{\partial E}{\partial h_z}$ as a function of h_z at constant $t_h^{xy} = 0, 1, 2$ and 3 from left to right, top to bottom. Discontinuous transitions are appearing symmetrically in the first and fourth figure. Additionally, continuous phase transitions occur, also symmetrically, for $t_h^{xy} = 0, 1$ and 2.

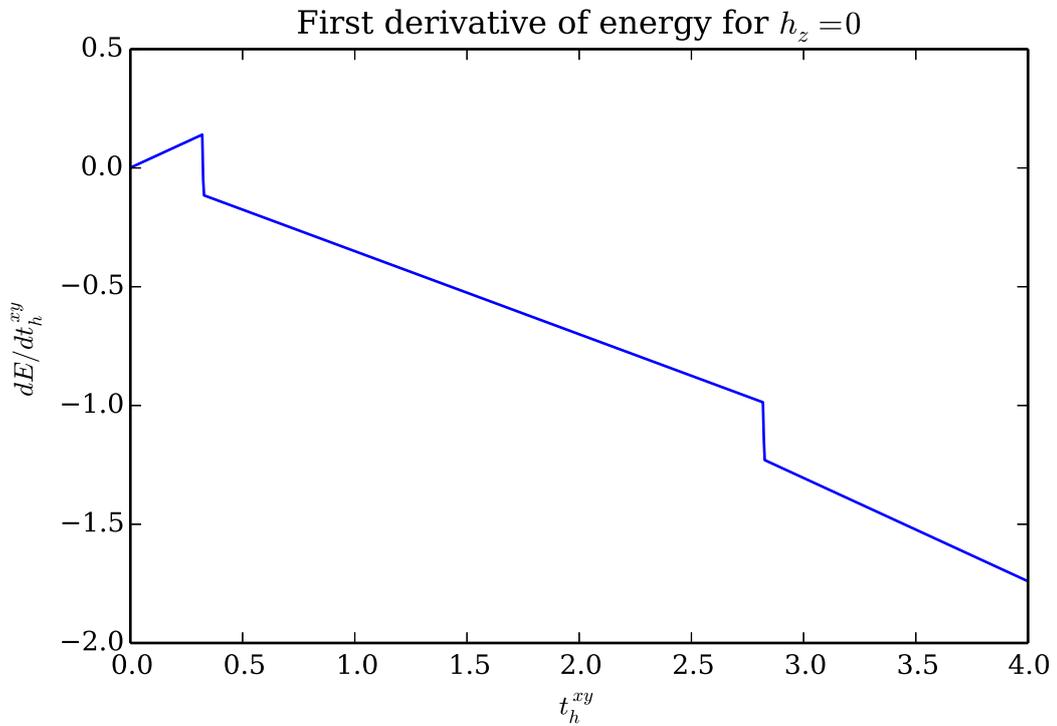


Figure 8.4: Sweep along the line $h = 0$ in figure 8.2, showing $\frac{\partial E}{\partial t_h^{xy}}$ as a function of t_h^{xy} with $h_z = 0$. Two first order phase transitions are visible.

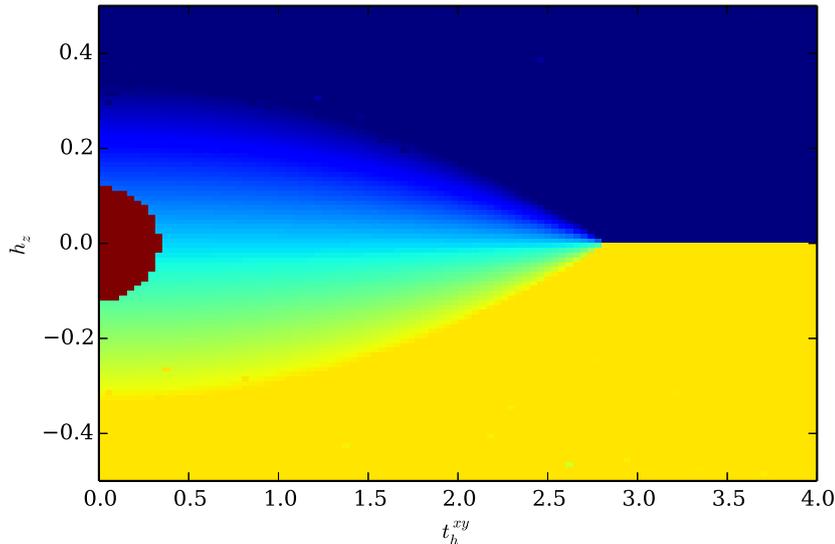


Figure 8.5: Heatmap plot showing the different configurations of the spin Hamiltonian in the (t_h^{xy}, h_z) plane. The 4-pattern (dark red) dominates in two “lobes”, around $h_z = 0$ and $t_h^{xy} < 3$. A phase which is anti-ferromagnetic in the y spin component and with a varying, polarized z component surrounds this phase (the z projection per site varies from -1 (yellow) to $+1$ (blue)). For large h the system is entirely polarized in the $\pm z$ direction, since the external field term dominates.

component of the spin is zero. The second phase has deformed into a third and fourth phase, which is polarized up or down in the z component, by a continuous phase transition. These phases are referred to as the z - P phases. Figure 8.5 shows the pattern of configurations, as found by comparing the order parameters at nearest- and next-nearest sites. I also note that for $|t_h^{xy}| > 2$, the phase diagram only has a discontinuous phase transition between the z - P phases.

When translating back to the original bosonic variables, the polarized phases correspond to a phase with only p_x or p_y occupied. The y - AFM phase for $h_z \approx 0$ corresponds to the wavelike vortex-antivortex solution - but in the Mott phase such long-range correlation is lost. The z - AFM phase corresponds to the zero-momentum phase found for $t_h^{xy} = 0$ and $U = 0$ in the exact diagonalization. It is interesting that this phase can never exist in practice within the superfluid phase, but becomes a stable phase “within” the Mott phase.

Figure 8.2 shows an interesting feature of the latter transition: as t_h^{xy} approaches ≈ 2.8 from smaller values, the first derivative of the energy is always smooth but approaches a sharp edge. At the edge, all first-order derivatives are discontinuous. On the line segment $h_z = 0, t_h^{xy} > 2.8$, the y - AFM phase will undergo a spontaneous symmetry breaking to either of the z - P phases. Such a spontaneous symmetry breaking accompanying a first-order phase transition is rare but possible.

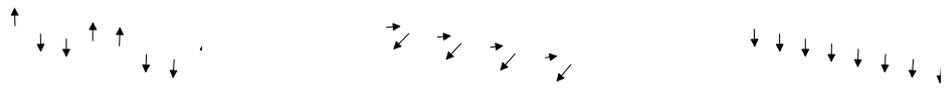


Figure 8.6: Mean-field spin configurations shown as a projected 3D plot. The order parameters are plotted as arrows, where the z component of the spin is up, the x component is along the direction of the chain and the y component is perpendicular to the chain. Note that this is only a method of plotting, and that the actual order parameters are not spatially extended but are pseudo-spin representations of populations of p_x and p_y atoms. Left: the first phase, anti-ferromagnetic in every other site. Middle: the second phase, polarized in the z component and anti-ferromagnetic in the y component. Right: the third and fourth phase is entirely polarized in the z component.

Chapter 9

Conclusions

In this thesis, physics of cold atoms loaded in the zig-zag lattice has been investigated in four steps. First, overlap integrals were computed from a separable basis of Wannier functions. I found that the tunnelings can be tuned by varying the angles in the lattice and the strength of the superlattice. Second, the system was diagonalized in the absence of interactions. This revealed wavelike solutions with a period longer than the primitive cell of the lattice and dependent on the tunnelings. Third, an approximation of the wavefunction in the superfluid phase was found by a variational calculation, showing that the periodicity survives in the presence of interactions. Fourth, in the Mott phase, perturbation theory was used to derive an effective Hamiltonian. By a Schwinger spin-boson mapping, this was shown to be equivalent to a spin Hamiltonian with next-nearest neighbor couplings. By tuning the coupling constants of the spin model, it was shown that, on the mean-field level, the system could be put in at least four different phases.

In conclusion, the zig-zag lattice has many interesting properties in both the superfluid and the Mott phase. In the superfluid phase, the mean-field solution has a periodicity larger than the primitive cell of the zig-zag lattice. This periodicity may be varied by tuning the value of the tunnelings. Within the Mott phase, it is possible to realize phases where nearest-neighbor and next-nearest neighbor tunnelings dominate. This realizes spin models with solutions which are periodic over two or four sites.

However, there remains work to be done on the zig-zag lattice. In the Mott phase, the coupling constants of the spin Hamiltonian are determined by the values of the tunnelings and the interaction parameters. I have suggested methods for creating anti-symmetric exchanges in other spin components and transverse magnetic fields, but there may be other phases which might be found by parameter sweeps in the tunnelings and interactions. The zig-zag lattice is in the end a potential quantum simulator: it is therefore important to understand the full set of possible spin Hamiltonians that can be realized in practice.

Another aspect that has not yet been investigated is the *topology* of the zig-zag lattice [MAV12, p. 429]. There are also many possible extensions of the model which have been tried for other lattice geometries: investigating Mott phases with higher occupancy or including the effect of atoms in the p_z state. It is also possible to consider optical lattices with atoms in the d band or higher, as was done in [FJPJ15].

The coherent state ansatz used throughout this thesis is one of many possible

variational trial states. It has the disadvantage of being a simple product state which cannot capture the effect of entanglement. The *Gutzwiller* trial states [DCI⁺98]

$$\begin{aligned}
 |\Psi_G\rangle &= \prod_i |\phi_i\rangle, \\
 |\phi_i\rangle &= \sum_{n=0}^{\infty} f_n^{(i)} |n\rangle_i
 \end{aligned}
 \tag{9.1}$$

where $|n\rangle_i$ is the Fock state with n bosons at site i and $\{f_n^{(i)}\}$ is the set of variational parameters. Expanding (9.1) yields a superposition where the size of every term depends on $f_i^{(n)}$. By letting $f_i^{(n)} \rightarrow 0$ for large n , it is possible to reduce the entanglement with states with high occupancy at some sites, for example. This could be useful for describing the order in the Mott phase.

The mean-field theoretical solutions may also be improved by considering perturbations around the mean-field solution, leading to what is sometimes called the *Bogoliubov equations* [PS08, p. 191]. This is equivalent to rewriting the annihilation operators in the form $\hat{a}_{j\alpha} = \psi_{j\alpha} + \delta\hat{a}_{j\alpha}$, where $\psi_{j\alpha}$ is the mean-field order parameter. The Hamiltonian is then expanded in powers of $\delta\hat{\psi}$, which is assumed to be a small perturbation. The Bogoliubov scheme would also make it possible to study the stability of the mean-field approximations systematically.

Exact diagonalization is not possible due to the exponential size of Hilbert space as mentioned in the introduction. However, there are several methods and numerical schemes which could and should be tried on this system. One possibility is using the *diffusion Monte Carlo (DMC)* method [JBK96], which propagates a first guess of the solution with a Green's function until the system has decayed to the ground state. Since the Hamiltonian is one-dimensional, it is also a suitable candidate for the *Density Matrix Renormalization Group (DMRG)* method. The idea of this method is to exclude the highly entangled subspace of Hilbert space from the calculation and handle short-range correlations between sites in a computationally efficient way [R.14]. Numerical diagonalization methods are left for future research, since there is not space or time to include it in this thesis.

Bibliography

- [AM76] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. Brooks/Cole, 1976.
- [AS05] Isacsson A. and Girvin S.M. Multi-flavor bosonic hubbard models in the first excited bloch band of an optical lattice. *Phys. Rev. A* *72*, 053604, 2005.
- [AS10] A. Altland and B. Simons. *Condensed Matter Field Theory*. Cambridge University Press, 2010.
- [Aue94] A. Auerbach. *Interacting Electrons and Quantum Magnetism*. Springer-Verlag, 1994.
- [Bal10] L.E. Ballentine. *Quantum Mechanics: A Modern Development*. World Scientific, 2010.
- [BDZ08] I. Bloch, J. Dalibard, and W. Zwerger. Many-body physics with ultracold gases. *Rev. Mod. Phys.* *80*, 885, 2008.
- [C.09] Wu C. Unconventional bose-einstein condensations beyond the "no-node" theorem. *Mod. Phys. Lett. B* *23*, 1, 2009.
- [CLM10] A. Collin, J. Larson, and J. P. Martikainen. Quantum states of p -band bosons in optical lattices. *Phys. Rev. A*, 81:023605, Feb 2010.
- [D.10] Ruhl D. *Strong-Coupling Solution of the Dynamical Mean-Field Equations for the Mott-Hubbard Insulator on a Bethe Lattice*. PhD thesis, Philips Universitat Marburg, 2010.
- [DCI+98] Jaksch D., Bruder C., Cirac J. I., Gardiner C. W., and Zoller P. Cold bosonic atoms in optical lattices. *Phys. Rev. Lett.* *81*, 3108, 1998.
- [Fey81] R.P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics* *21* (6-7): 467-488, 1981.
- [FG10] Mandl F. and Shaw G. *Quantum Field Theory*. Wiley, 2010.
- [FJPJ15] Pinheiro F., Martikainen J-P., and Larson J. Phases of d -orbital bosons in optical lattices. *New Journal of Physics*, *17* 053004, 2015.
- [GC14] Dvali G. and Gomez C. Black holes as critical point of quantum phase transition. *The European Physical Journal C*, 2014.

- [II15] Gradstheyn I.S. and Ryzhik I.M. *Table of Integrals, Series and Products*. Elsevier Press, 2015.
- [J.11] Sethna J. *Entropy, Order Parameters and Complexity*. Clarendon Press, 2011.
- [JA10] Lancaster J. and Mitra A. Quantum quenches in an xxz spin chain from a spatially inhomogeneous initial state. *Phys. Rev. E* 81: 6, 061134, 2010.
- [JBK96] Kosztin J, Faber B., and Schultzen K. Introduction to the diffusion monte carlo method. *Am. J. Phys.* 64, 633, 1996.
- [JFJO11] Dalibard J, Gerbier F, Gediminas Juzeliūnas, and Patrik Öhberg. *Colloquium* : Artificial gauge potentials for neutral atoms. *Rev. Mod. Phys.*, 83:1523–1543, Nov 2011.
- [JJ12] Larson J. and Martikainen J.P. Multiorbital bosons in bipartite optical lattices. *Phys. Rev. A* 86, 023611, 2012.
- [JMB10] Eisert J., Cramer M., and Plenio M. B. Colloquium: Area laws for the entanglement entropy. *Rev. Mod. Phys.* 82, 277, 2010.
- [JS06] Nocedal J. and Wright S.J. *Numerical Optimization*. Springer, 2006.
- [M.] Nielsen M. The fermionic canonical commutation relations and the jordan-wigner transform. (link).
- [Mah90] G.D. Mahan. *Many-Particle Physics*. Plenum Press, 1990.
- [MAV12] Lewenstein M., Sanpera A., and Ahufinger V. *Ultracold Atoms in Optical Lattices*. Oxford University Press, 2012.
- [MB06] Plischke M. and Bergersen B. *Equilibrium Statistical Physics*. World Scientific, 2006.
- [MD95] Peskin M. and Schroeder D. *An Introduction to Quantum Field Theory*. Westview Press, 1995.
- [Mey01] P. Meystre. *Atom optics*. Springer, 2001.
- [MOT⁺02] Greiner M., Mandel O., Esslinger T., Hänsch T.W., and Bloch I. Quantum phase transition from a superfluid to a mott insulator in a gas of ultracold atoms. *Nature* 415, 39-44, 2002.
- [N.92] Goldenfeld N. *Lectures on Phase Transitions and the Renormalization Group*. Perseus Books, 1992.
- [NAJ⁺12] Marzari N., Mostofi A.A., Yates J.R., Souza I., and Vanderbilt D. Maximally localized wannier functions: Theory and applications. *Rev. Mod. Phys.* 84, 1419, 2012.
- [PB11] R.K. Pathria and P.D. Beale. *Statistical mechanics*. Elsevier, 2011.

- [PBML13] F. Pinheiro, G.M. Bruun, J-P. Martikainen, and J. Larson. Xyz quantum heisenberg models with p-orbital bosons. *Phys. Rev. Lett.* *111* 205302, 2013.
- [PML12] F. Pinheiro, J-P. Martikainen, and J. Larson. Confined p-band bose-einstein condensates. *Phys. Rev. A* *85* 033638, 2012.
- [PS08] C. Pethick and H. Smith. *Bose-Einstein Condensation in Dilute Gases*. Cambridge University Press, 2008.
- [PSanB⁺13] Joanna Pietraszewicz, Tomasz Sowiński, Mirosław Brewczyk, Maciej Lewenstein, and Mariusz Gajda. Spin dynamics of two bosons in an optical lattice site: A role of anharmonicity and anisotropy of the trapping potential. *Phys. Rev. A*, 88:013608, Jul 2013.
- [R.14] Orus R. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics* *349* 117-158, 2014.
- [RPC⁺99] Matthews M. R., Anderson B. P., Haljan P. C., Hall D. S., Wieman C. E., and Cornell E. A. Vortices in a bose-einstein condensate. *Phys. Rev. Lett.* *83*, 2498, 1999.
- [S.11] Sachdev S. *Quantum Phase Transitions*. Cambridge University Press, 2011.
- [Sak94] J.J. Sakurai. *Modern Quantum Mechanics*. Addison-Wesley, 1994.
- [TSAI07] Müller T., Fölling S., Widera A., and Bloch I. State preparation and dynamics of ultracold atoms in higher lattice orbitals. *Phys. Rev. Lett.* *99*, 200405, 2007.
- [W.03] Zwerger W. Mott–hubbard transition of cold atoms in optical lattices. *Journal of Optics B: Quantum and Semiclassical Optics*, Volume 5, Number 2, 2003.
- [WAM⁺10] Bakr W.S., Peng A., Tai M.E., Ma R., Simon J., Gillen I., Foelling S., Pollet L., and Greiner M. Probing the superfluid to mott insulator transition at the single atom level. *Science* *329*. no. 5991, pp. 547 - 550, 2010.
- [WC03] D. M. Whittaker and M. P. Croucher. Maximally localized wannier functions for photonic lattices. *Phys. Rev. B*, 67:085204, Feb 2003.