Wavefunction Dynamics and Application to the Bose Glass Phase

Connor Behan

Department of Physics, Engineering Physics, and Astronomy,
Queen's University, Kingston, ON K7L 3N6, Canada[*]

(Dated: July 27, 2011)

Within the last twenty years, the technology for trapping and cooling atoms has advanced to the point where classically forbidden states of matter can be observed at measurable scales. Many experiments close to absolute zero use an optical lattice - a type of artificial crystal which can act as a quantum simulator and expose previously unseen phenomena. This thesis is a study of tunable parameters pertaining to the lattice (hopping frequency and disorder strength), which explores the wavefunctions of ultracold spinless bosons that begin on a single site and subsequently evolve in time. The main theoretical tools introduced are the Bose-Hubbard Hamiltonian, a model for optical lattices and the discrete non-linear Schrödinger equation, a lower-dimensional problem that has been proposed as an approximation to the dynamics of the Bose-Hubbard model. The methods chosen to simulate them are the Chebyshev expansion technique and the Cash-Karp adaptive Runge-Kutta method respectively. Parameters to test are chosen from the phase diagram for the superfluid to Mott-insulator phase transition. For perfectly periodic lattices, none of the results show evidence of this phase transition affecting the diffusion. However, the results do show transitions into the Bose glass phase. Based on comparisons between the two models, we conclude that the discrete non-linear Schrödinger equation does not reproduce the results of the Bose-Hubbard Hamiltonian under these conditions.
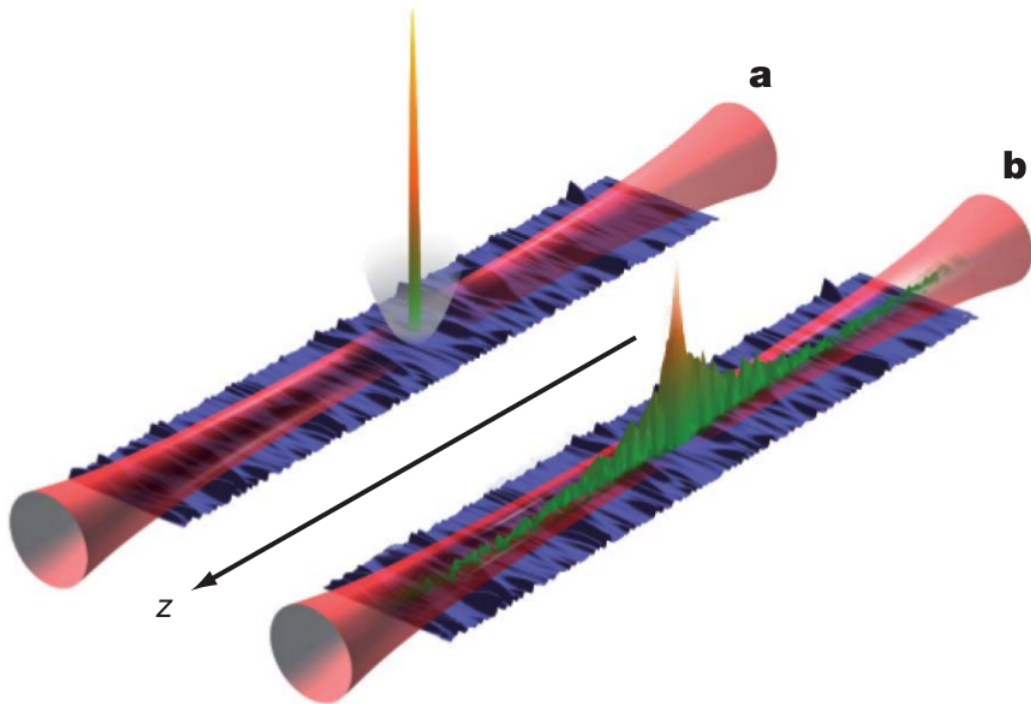
FIG. 1: A picture that sums up this thesis [1].

[*]Supervised by Robert J. Gooding

Contents

I. INTRODUCTION

Most people with an undergraduate physics education will be quick to say that the probabilistic nature of the universe, though a triumph of twentieth century physics, does not affect one's everyday experiences. A baseball may have a finite probability of impinging on a concrete wall and emerging on the other side but observing this would take on the order of 10^{100} throws. The probability density function of a marble in a box might be a superposition of trigonometric functions but it is a superposition of so many that the end result is very close to a delta function that allows us to see the marble in one place. And no one has ever played air hockey while worrying about the fact that the momentum of the puck is not a precisely known quantity. These examples illustrate the difficulty of studying the quantum mechanical properties of macroscopic materials, but by employing low temperatures, carefully designed experiments have succeeded in probing these limits of our understanding. They have also led to the development of a growing number of quantum devices.

The properties of ultracold atoms have been studied in the free continuum as well as on lattices. This thesis studies the dynamics of ultracold bosons on lattices. In practice, these are typically optical lattices that are created when one fires counter-propagating lasers at the trapped atoms. We will see that the periodic electric field creates an effective potential that is able to localize cold atoms. The dynamics of particles in an optical lattice have received a large amount of attention in the physics community, partly because the optical lattice is a candidate apparatus for building a quantum computer [2]. In the sections that follow, we will be interested in its ability to act as a quantum simulator - a setting where quantum statistics and non-classical phases of matter abound. Simulating the time evolution of a collection of bosons in an optical lattice requires a Hamiltonian, an initial condition and a method for integrating the equation of motion. The present study revolves around the Bose-Hubbard Hamiltonian where the initial condition is the one that places all of the bosons on the same site. We will see that this models a condition that is often produced experimentally [3] [4] [5] [6].

In the rest of this introduction, the details of optical lattice experiments are developed, leading into a section on the essential formalism. Here, two problems are introduced - one using the Bose-Hubbard Hamiltonian and one using an approximation to it. A large amount of effort has been expended finding numerical methods that are well suited to solving these problems. The numerical methods applicable to these two problems are presented separately. A summary of the results precedes the conclusion. To make this thesis self-contained, appendicies have been included that justify and extend many of the ideas that are mentioned in the main body of the text.

A. Low Temperature Experiments

A rule of thumb for forcing quantum effects to become important is making the de Broglie wavelength Λ very large compared to the inter-particle spacing $\left(\frac{V}{N_{\mathrm{p}}}\right)^{\frac{1}{3}}$ where N_{p} is the number of particles. Semi-classically, we can show that this is achieved by lowering the temperature. Harking back to de Broglie's original paper, the following relations are justified:

$$p = \frac{h}{\Lambda}$$

$$E = \frac{p^2}{2m}$$

The only additional input needed is the equipartition theorem, an understanding of which follows from the development in Appendix A. The expression for thermal energy that it gives is $E = \frac{3k_{\mathrm{B}}T}{2}$ where the 3 in the numerator is the number of dimensions of space. Putting these together we see that the desired condition is:

$$\left(\frac{V}{N_{\mathrm{p}}}\right)^{\frac{1}{3}} \ll \Lambda = \frac{h}{\sqrt{3mk_{\mathrm{B}}T}} \tag{1}$$

The left hand side of Eq. (1) will be very large for a temperature close to absolute zero.

In practice, the procedure designed to bring atoms to ultracold temperatures is the Doppler cooling technique where photons are repeatedly scattered off the atoms [7]. The frequency in the lab frame is set to a value just below one of the resonant frequencies of the bound electrons. Through the Doppler effect, this ensures that atoms moving away from the light source (into the trap) are transparent to the photons while atoms moving towards the light source (out of the trap) interact with the photons and slow down. It is hard to imagine that this would successfully contain free particles. Insetad, it is necessary to localize the particles in space while they are cooled. It has already been mentioned in the abstract, that an optical lattice may localize atoms in particular sites. This localization is not sufficient during the cooling stage for two reasons:

1. The energy difference between the potential maxima and minima in an optical lattice is on the order of the energy difference between electron orbitals within the atom, as we will see. This is negligible compared to the thermal energy of atoms at room temperature.

2. The laser arrangement used in Doppler cooling is in fact identical to the laser arrangement that creates an optical lattice so almost no improvement can be made by attempting to create *another* optical lattice.

Therefore, whenever an optical lattice is put to productive use, it resides inside a trap with a much larger confining potential well. The simplest device that is commonly used to provide the confining potential when cooling neutral atoms is a magnetic trap [8]. One-dimensional trapping can be understood if we think of an atom as a bar magnet being repelled by two strong permanent magnets on either side as in Fig. 2. A clever trap geometry makes it energetically prohibitive for the dipole to flip. Small displacements from the centre will be met with restoring forces that point back to the centre. The potential experienced by the atom is locally that of a simple harmonic oscillator.
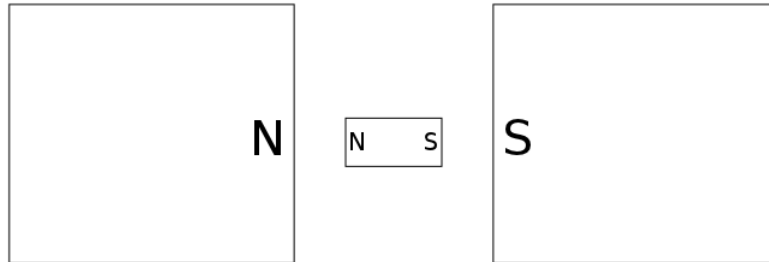

FIG. 2: A simplified picture of how magnets can aid in trapping atoms. Real traps are more complicated because they must stop the bar magnet in the middle from simply reversing its orientation.
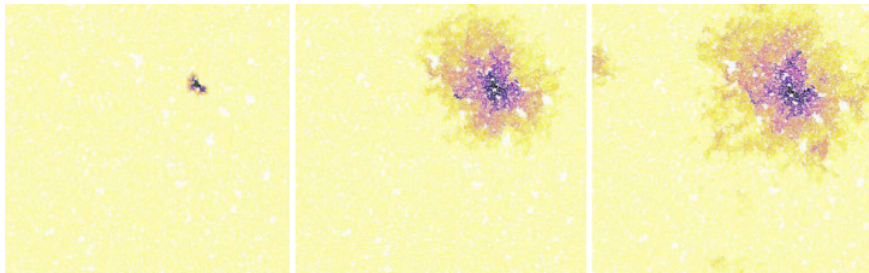

FIG. 3: An example of diffusion in a 2D lattice, of the type being studied in this thesis, with snapshots taken at three different times by [6].

When the strength of the magnetic trap is adjustable, the reason for using the "all on one site" initial condition later in this thesis becomes clear. In an optical lattice with many sites, the particles can be forced to all occupy the site closest to the centre of the magnetic trap by making the harmonic well very steep. When this trap is turned off, a diffusion of the bosons of the type simulated in this thesis is observed [3] [4] [5] [6]. In fact, the picture on the cover is an illustration of what a group doing precisely this type of 1D experiment saw [1]. Figure 3 shows a diffusion simulation that is similar in spirit.

It is worth discussing a technical point before we begin our discussion of optical lattices. The type of trap we are talking about will only affect an atom that has a magnetic moment. If that atom is a boson it must have at least spin-1 even though all our models for optical lattices that follow will apply only to spinless bosons. This is because only one alignment allows the atom to stay in the trap so no spin degree of freedom plays a role in the lattice [8]. Atoms that initially have the unfavourable spin are quickly ejected. Atoms that remain stay polarized so that their effective spins are 0. When used on a collection of bosons whose spins begin with random orientation, this technique will succeed in trapping at most half of the particles.

B. Optical Lattices

The periodic potentials that allow ultracold atoms to participate in an artificial crystal are formed by pairs of counter-propagating lasers. Since the magnetic trap overpowers the magnetic fields that come from the lasers, it is

primarily the Stark effect that causes the energy of a boson to shift. It shifts by a spatially varrying amount which we call the effective potential. Typical optical lattices are no more than a few millimetres wide. To see how the potential comes about, one must think about the wavefunctions of electrons in the bosonic atoms. They will be denoted $\left|\psi_0^{(0)}\right\rangle$, $\left|\psi_1^{(0)}\right\rangle$, $\left|\psi_2^{(0)}\right\rangle$, etc where $\left|\psi_0^{(0)}\right\rangle$ is the ground-state. Since the trapping potential varies very slowly over millimetre distances, these are well approximated by the electronic wavefunctions of free atoms; sometimes even by hydrogenic wavefunctions. In the dipole approximation, the Hamiltonian responsible for perturbing the ground-state energy will be:

$$\hat{\mathcal{H}}' = -\hat{\mathbf{p}} \cdot \mathbf{E}(\mathbf{r}, t)$$

where $\hat{\mathbf{p}}$ is the electric dipole moment operator. To first order, the ground-state energy does not change:

$$
\begin{aligned}
E_0^{(1)}(\mathbf{r}, t) &= \left\langle \psi_0^{(0)} \left| \hat{\mathcal{H}}' \right| \psi_0^{(0)} \right\rangle \\
&= -\left[\left\langle \psi_0^{(0)} \left| \hat{\mathbf{p}}_x E_x(\mathbf{r}, t) \right| \psi_0^{(0)} \right\rangle + \left\langle \psi_0^{(0)} \left| \hat{\mathbf{p}}_y E_y(\mathbf{r}, t) \right| \psi_0^{(0)} \right\rangle + \left\langle \psi_0^{0} \left| \hat{\mathbf{p}}_z E_z(\mathbf{r}, t) \right| \psi_0^{(0)} \right\rangle \right] \\
&= -\left[E_x(\mathbf{r}, t) \left\langle \hat{\mathbf{p}}_x \right\rangle + E_y(\mathbf{r}, t) \left\langle \hat{\mathbf{p}}_y \right\rangle + E_z(\mathbf{r}, t) \left\langle \hat{\mathbf{p}}_z \right\rangle \right] \\
&= 0
\end{aligned}
$$

because a neutral atom in its ground-state does not have a permanent dipole moment. To second order, we have:

$$
\begin{aligned}
E_0^{(2)}(\mathbf{r}, t) &= \sum_{j \neq 0} \frac{\left| \left\langle \psi_j^{(0)} \left| \hat{\mathcal{H}}' \right| \psi_0^{(0)} \right\rangle \right|^2}{E_j^0 - E_0^0} \\
&= \sum_{j \neq 0} \frac{\left| \left\langle \psi_j^{(0)} \left| \hat{\mathbf{p}}_{\boldsymbol{\eta}} \right| \psi_0^{(0)} \right\rangle \right|^2}{E_j^{(0)} - E_0^{(0)}} \left| E(\mathbf{r}, t) \right|^2
\end{aligned}
\qquad (2)
$$

where $\boldsymbol{\eta}$ is the direction of polarization in the electric field. We will not concern ourselves with what the number $\sum_{j \neq 0} \frac{\left| \left\langle \psi_j^{(0)} \left| \hat{\mathbf{p}}_{\boldsymbol{\eta}} \right| \psi_0^{(0)} \right\rangle \right|^2}{E_j^{(0)} - E_0^{(0)}}$ actually is. The important part is that it will not be zero [9] because of the position operators present in the expansion of $\hat{\mathbf{p}}$. If beams propagating in the $\pm x$ direction have amplitude $E_{0,x}$, wave-vector $k_{\mathrm{L},x}$, angular frequency $\omega_{\mathrm{L},x}$ and polarization $\boldsymbol{\eta}_x$ they can be written out as:

$$\mathbf{E}_{\pm x}(x, t) = E_{0,x} \Re \left(e^{i(\pm k_{\mathrm{L},x} x - \omega_{\mathrm{L},x} t)} \right) \boldsymbol{\eta}_x$$

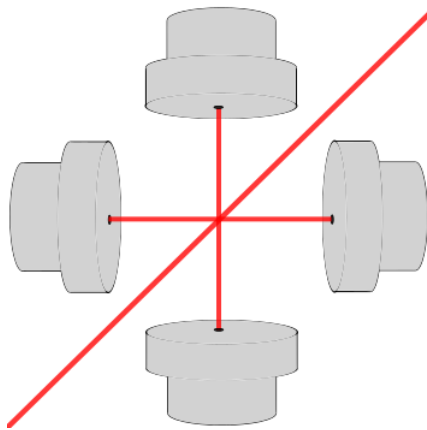The superposition of these lasers is a standing wave given by:

FIG. 4: An optical lattice no more than a few millimetres wide, is created where the coherent beams intersect. The beams are orthogonal so that the resulting lattice is cubic.

$$\begin{aligned}
\mathbf{E}_{+x}(x,t) + \mathbf{E}_{-x}(x,t) &= E_{0,x}\left[\Re\left(e^{i(k_{\mathrm{L},x}x-\omega_{\mathrm{L},x}t)}\right) + \Re\left(e^{i(-k_{\mathrm{L},x}x-\omega_{\mathrm{L},x}t)}\right)\right]\boldsymbol{\eta}_x \\
&= E_{0,x}\Re\left[e^{i(k_{\mathrm{L},x}x-\omega_{\mathrm{L},x}t)} + e^{i(-k_{\mathrm{L},x}x-\omega_{\mathrm{L},x}t)}\right]\boldsymbol{\eta}_x \\
&= E_{0,x}\Re\left[\left(e^{ik_{\mathrm{L},x}x} + e^{-ik_{\mathrm{L},x}x}\right)e^{i\omega_{\mathrm{L},x}t}\right]\boldsymbol{\eta}_x \\
&= 2E_{0,x}\cos\left(k_{\mathrm{L},x}x\right)\Re\left(e^{i\omega_{\mathrm{L},x}t}\right)\boldsymbol{\eta}_x \\
&= 2E_{0,x}\cos\left(k_{\mathrm{L},x}x\right)\cos\left(\omega_{\mathrm{L},x}t\right)\boldsymbol{\eta}_x
\end{aligned}$$

It is clear that if similarly polarized beams in the other two orthogonal directions are included, the net electric field
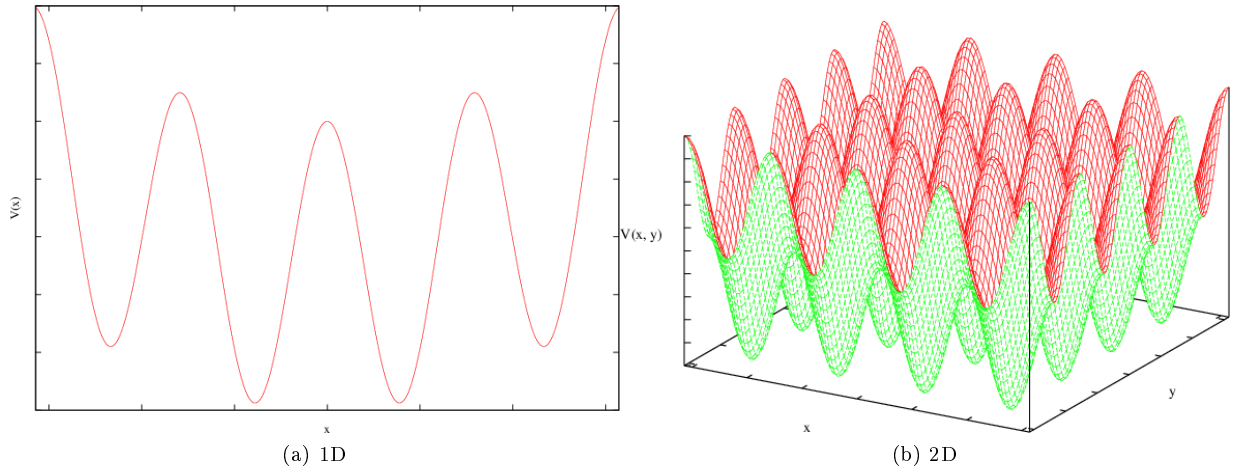

(a) 1D

(b) 2D

FIG. 5: Two plots of the optical lattice potential. The one-dimensional potential is exaggerated to the point where it has a parabolic offset. This effect of the surrounding trap would be observed for lasers with a large wavelength.

will be:

$$\begin{aligned}
\mathbf{E}(\mathbf{r},t) &= \mathbf{E}_{+x}(x,t) + \mathbf{E}_{-x}(x,t) + \mathbf{E}_{+y}(y,t) + \mathbf{E}_{-y}(y,t) + \mathbf{E}_{+z}(z,t) + \mathbf{E}_{-z}(z,t) \\
&= 2\left[E_{0,x}\cos\left(k_{\mathrm{L},x}x\right)\cos\left(\omega_{\mathrm{L},x}t\right)\boldsymbol{\eta}_x + E_{0,y}\cos\left(k_{\mathrm{L},y}y\right)\cos\left(\omega_{\mathrm{L},y}t\right)\boldsymbol{\eta}_y + E_{0,z}\cos\left(k_{\mathrm{L},z}z\right)\cos\left(\omega_{\mathrm{L},z}t\right)\boldsymbol{\eta}_z\right] \quad (3)
\end{aligned}$$

Dotting Eq. (3) with itself, we get:

$$E^2(\mathbf{r},t) = 4\left[E_{0,x}^2\cos^2\left(k_{\mathrm{L},x}x\right)\cos^2\left(\omega_{\mathrm{L},x}t\right) + E_{0,y}^2\cos^2\left(k_{\mathrm{L},y}y\right)\cos^2\left(\omega_{\mathrm{L},y}t\right) + E_{0,z}^2\cos^2\left(k_{\mathrm{L},z}z\right)\cos^2\left(\omega_{\mathrm{L},z}t\right)\right] \quad (4)$$

The optical lattice potential can now be recovered by substituting Eq. (4) into Eq. (2) and time averaging [10].

$$\begin{aligned}
V_{\mathrm{lat}}(\mathbf{r}) &= \overline{E_0^{(2)}(\mathbf{r},t)} \\
&= 2\sum_{j\neq 0}\frac{\left|\left\langle\psi_j^{(0)}\left|\hat{\mathfrak{p}}_{\boldsymbol{\eta}}\right|\psi_0^{(0)}\right\rangle\right|^2}{E_j^{(0)} - E_0^{(0)}}\left[E_{0,x}^2\cos^2\left(k_{\mathrm{L},x}x\right) + E_{0,y}^2\cos^2\left(k_{\mathrm{L},y}y\right) + E_{0,z}^2\cos^2\left(k_{\mathrm{L},z}z\right)\right] \\
&\equiv V_{0,x}\cos^2\left(k_{\mathrm{L},x}x\right) + V_{0,y}\cos^2\left(k_{\mathrm{L},y}y\right) + V_{0,z}\cos^2\left(k_{\mathrm{L},z}z\right) \quad (5)
\end{aligned}$$

The hills and vallies of an optical lattice can be modified by adjusting the intensities and the frequencies of the three laser pairs independently. Additionally, "superlattice" phenomena may be studied by making one optical lattice the envelope function for another. The simplest optical lattice creates an array of identical quantum dots at the point where the lasers intersect and is referred to as "ordered". Disordered systems are the ones that are harder to analyze. When we develop a fully quantized description of optical lattice phenomena (the Bose-Hubbard Hamiltonian), we will see an array of on-site energies that makes it easy to tell whether a given lattice is ordered or disordered.

C. Phase Transitions

Low-temperature experiments have attracted a large interest in the physics community for their potential to drastically alter properties of matter such as viscosity. Zero viscosity flow was first observed in Helium-4 in 1937 [11] and

since then a number of models have been proposed to explain this phenomenon. One of the most elegant for our purposes is due to Feynmann [12]. If a liquid of mass M slows from speed v to speed v' by creating an excitation, the following conservation laws must hold:

$$\frac{1}{2}Mv^2 = \frac{1}{2}Mv'^2 + E_{\text{ex}} \tag{6}$$

$$M\mathbf{v} = M\mathbf{v}' + \mathbf{p}_{\text{ex}} \tag{7}$$

Rearranging Eq. (7),

$$M\mathbf{v}' = M\mathbf{v} - \mathbf{p}_{\text{ex}}$$
$$M^2 v'^2 = M^2 v^2 - 2M\mathbf{v}\cdot\mathbf{p}_{\text{ex}} + p_{\text{ex}}^2$$
$$\frac{1}{2}Mv'^2 = \frac{1}{2}Mv^2 - \mathbf{v}\cdot\mathbf{p}_{\text{ex}} + \frac{p_{\text{ex}}^2}{2M}$$
$$\frac{1}{2}Mv^2 - E_{\text{ex}} = \frac{1}{2}Mv^2 - \mathbf{v}\cdot\mathbf{p}_{\text{ex}} + \frac{p_{\text{ex}}^2}{2M}$$
$$E_{\text{ex}} = \mathbf{v}\cdot\mathbf{p}_{\text{ex}} - \frac{p_{\text{ex}}^2}{2M}$$

Even in the case of a very large mass M and an excitation momentum parallel to the velocity, there is a minimum speed below which the fluid cannot slow down:

$$v \geq \frac{E_{\text{ex}}}{p_{\text{ex}}} \tag{8}$$

For a liquid whose dispersion relation is free of a band gap, this minimum is zero and therefore inconsequential. An
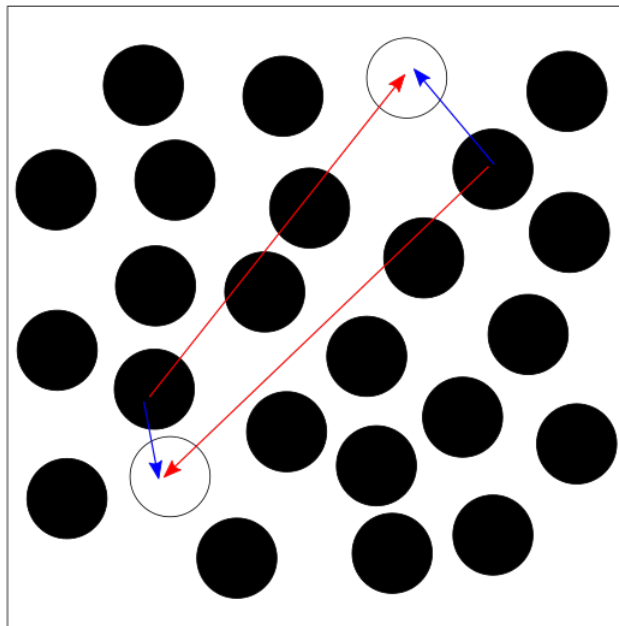

FIG. 6: If the circles represent indistinguishable bosonic atoms, the motion along the red arrows is completely equivalent to the motion along the blue arrows.

interesting argument using quantum statistics is that it is impossible to get rid of this energy gap between the two lowest energy states for a fluid of interacting bosons [12]. One can draw a picture of bosons in the first multi-particle excited state. If the particles udergo repulsive interactions and the energy of this state is to be arbitrarily close to the ground-state energy, it is a reasonable guess to draw the atoms so that they are uniformly spread out. A key property of excited states is that they possess one node. Therefore, one can flip the sign of the probability amplitude by rearranging the atoms over long distances. If the sign were to change following a rearrangement over short distances, this would give the wavefunction a steep gradient and thus an energy much higher than that of the ground-state. To

prevent a phonon from forming, the void we are left with when we move an atom must be filled by another atom that is initially far away. However, moving this pair of particles through a long path is the same as moving a different pair of particles through a short path since they are all indistinguishable. Moreover, interchanging the roles of the particles this way cannot change the wavefunction for a collection of bosons so we still have a wavefunction that changes sign rapidly enough to be far from the ground-state.

Feynmann's slick argument applies to any bosonic substance that does not freeze at the temperatures where quantum statistics become important. Helium-4 just happened to be the only such substance people knew about in the 1930s [11]. Flow without friction is exotic compared to our everyday experience. This gives credence to the argument that this behaviour constitutes a unique state of matter. Bosons placed in an optical lattice certainly have an energy gap so if the potential wells allow them to be approximated by fluid mechanics, the optical lattice experiment appears well poised to produce this "superfluid" phase. However, one can also imagine an insulating phase in which the potential wells are so deep that bosons stay localized in their own sites with a vanishingly small probability of escaping. It is unclear at this point whether nature distinguishes between these phases - they could simply represent two extremes in the way people choose to think about optical lattices.

Appendix A provides a precise definition of a phase transition: when the boundary between two phases is reached, an order parameter that affects the free energy becomes nonzero [13]. In this regard, the appearance of a superfluid is a genuine phase transition. After introducing the Bose-Hubbard Hamiltonian, we will make frequent reference to its phase diagram which involves several order parameters - one for each site on the lattice [14]. In an ordered lattice the two possible phases described above are the only ones available to the system. It may either exist as a superfluid with delocalized bosons, in which all of the order parameters are nonzero, or as a Mott-insulator with an integer number of bosons localized on each site, in which all of the order parameters are zero [15] [16]. A disordered lattice has a third phase called the Bose glass in which only some of the order parameters vanish [17]. The Bose glass phase displays complicated behaviour including islands of localized particles in various shapes.

II. FORMAL BACKGROUND

Any sufficiently thorough introduction to quantum mechanics will mention multi-particle wavefunctions and how they must be symmetric for bosons and anti-symmetric for fermions. After fixing the number of particles N_{p} and defining these states, it is natural to attempt to solve the Schrödinger equation where the Hamiltonian is given by:

$$\hat{\mathcal{H}} = \sum_{j=1}^{N_{\mathrm{p}}} \left[\frac{-\hbar^2}{2m} \nabla_j^2 + V(\mathbf{r}_j) \right] + \frac{1}{2} \sum_{j \neq l} V_{\mathrm{int}}(\mathbf{r}_j - \mathbf{r}_l) \tag{9}$$

Our goal is to use the Bose-Hubbard Hamiltonian to study many-body problems instead. The Bose-Hubbard Hamiltonian is not equal to Eq. (9); it involves simplifying assumptions valid for low-temperature lattice systems that serve to make the equations more tractable and the results more illuminating. In order to write it down, we will need to dispense with the elementary notation of Eq. (9) and invoke a formalism that is much more powerful.

A. The Bose-Hubbard Hamiltonian

This formalism is second quantization which involves operators that create and anhilate particles in certain states. The development in Appendix B reveals that instead of using Eq. (9), one should use the more useful form [18]:

$$\hat{\mathcal{H}} = \int_{\mathbb{R}^3} \hat{\Psi}^\dagger(\mathbf{r}) \left[\frac{-\hbar^2}{2m} \nabla_{\mathbf{r}}^2 + V(\mathbf{r}) \right] \hat{\Psi}(\mathbf{r}) \mathrm{d}\mathbf{r} + \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \hat{\Psi}^\dagger(\mathbf{r}) \hat{\Psi}^\dagger(\mathbf{r}') V_{\mathrm{int}}(\mathbf{r} - \mathbf{r}') \hat{\Psi}(\mathbf{r}') \hat{\Psi}(\mathbf{r}) \mathrm{d}\mathbf{r}' \mathrm{d}\mathbf{r} \tag{10}$$

Here, $\hat{\Psi}^\dagger(\mathbf{r})$ and $\hat{\Psi}(\mathbf{r})$ are the position space creation and anhilation operators respectively where the first adds a particle to the location $\mathbf{r}$ in space and the second removes it. Since these operators allow us to change the number of particles, they must already account for the spin statistics of bosons. This bosonic quality is present in the commutation relations that they satisfy:

$$\left[\hat{\Psi}^\dagger(\mathbf{r}), \hat{\Psi}^\dagger(\mathbf{r}') \right] = 0, \left[\hat{\Psi}(\mathbf{r}), \hat{\Psi}(\mathbf{r}') \right] = 0, \left[\hat{\Psi}(\mathbf{r}), \hat{\Psi}^\dagger(\mathbf{r}') \right] = \delta(\mathbf{r} - \mathbf{r}')$$

Field operators is another name given to these important objects. All calculations that have been carried out have focused on the Bose-Hubbard Hamiltonian which is well suited to describing optical lattices. It is based on discretized

versions of the field operators: $\hat{a}_j^\dagger$ creates a particle on site j and $\hat{a}_j$ anhilates one. Not surprisingly, these operators satisfy:

$$[\hat{a}_j, \hat{a}_l] = 0, \left[\hat{a}_j{}^\dagger, \hat{a}_l{}^\dagger\right] = 0, \left[\hat{a}_j, \hat{a}_l{}^\dagger\right] = \delta_{j,l}$$

Though more general forms can be written down, the standard form of the Bose-Hubbard Hamiltonian assumes three approximations:

1. Contact interactions of strength g.

2. Tight binding - tunelling only occurs between nearest-neighbour sites.

3. While on site j, bosons are described by the wavefunction $\Omega_j^{(0)}(\mathbf{r})$, the lowest band of that site's potential well.

The Bose-Hubbard Hamiltonian for spinless bosons is:

$$\begin{aligned}
\hat{\mathcal{H}} &= \sum_j \epsilon_j \hat{n}_j + \frac{U}{2} \sum_j \hat{n}_j \left(\hat{n}_j - 1\right) - J \sum_{\langle j,l \rangle \bullet} \hat{a}_j^\dagger \hat{a}_l \\
&= \sum_j \epsilon_j \hat{n}_j + \frac{U}{2} \sum_j \hat{n}_j \left(\hat{n}_j - 1\right) - J \sum_{\langle j,l \rangle \circ} \left(\hat{a}_j^\dagger \hat{a}_l + \hat{a}_j \hat{a}_l^\dagger\right)
\end{aligned} \tag{11}$$

where $\langle i,j \rangle \bullet$ denotes a sum over nearest-neighbours and $\langle i,j \rangle \circ$ denotes a sum over *distinct* nearest-neighbours. When deriving the Bose-Hubbard Hamiltonian from (10), the trick is to split up the potential as one term coming from a regular lattice V_{lat} and one from an external potential V_{ext} [10]. Disorder is thus realized as a non-uniform external potential. Appendix B shows that the parameters J, U and ϵ_j are defined as:

$$\begin{aligned}
J &= -\int_{\mathbb{R}^3} \Omega_j^{(0)*}(\mathbf{r}) \left[\frac{-\hbar^2}{2m}\nabla^2 + V_{\text{lat}}(\mathbf{r})\right] \Omega_l^{(0)}(\mathbf{r}) d\mathbf{r} \\
U &= g \int_{\mathbb{R}^3} \left|\Omega_0^{(0)*}(\mathbf{r})\right|^4 d\mathbf{r} \\
\epsilon_j &= \int_{\mathbb{R}^3} V_{\text{ext}}(\mathbf{r}) \left|\Omega_j^{(0)}(\mathbf{r})\right|^2 d\mathbf{r} \approx V_{\text{ext}}(\mathbf{r}_j)
\end{aligned}$$

where j and l are nearest-neighbours. We do not need to refer to these formal definitions however since the physical
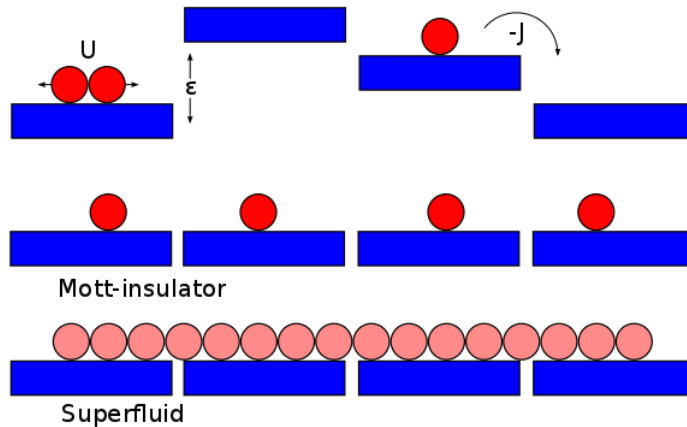
FIG. 7: The top of the diagram shows the meaning of each parameter in the Bose-Hubbard Hamiltonian. Based on their values, we either observer a Mott-insulator, which has an integer number of particles on each site or a superfluid, which has particles that are probabilistically spread out over the lattice.

roles played by J, U and ϵ_j are clear from (11). The term containing U will serve to increase the energy whenever two or more particles are on site j. This increase grows with the number of particles so we can associate U with the interactions. J is the only parameter involving a Hermitian operator defined not in terms of sites but in terms of

paris of sites. It is also represents the only way a state that gives definite positions to all the particles can be mixed with other such states. J is therefore called the hopping parameter and is an energy related to the typical hopping or tunnelling frequency of bosons in the lattice [10]. Figure 7 illustrates the physical effects that these parameters have. Already, the parameters show striking connections to the phases of interest described qualitatively. A high J to U ratio favours the appearance of a superfluid while a low one favours the appearance of a Mott-insulator. Without altering the material, this ratio can be changed by virtue of the fact that J depends on the potential. When the potential well on a lattice site is very deep, it becomes improbable for a particle within to tunnel to a neighbouring site. The on-site energies ϵ_j will quantify the disorder if they vary from site to site. Otherwise they can be set to zero without loss of generality. Something that is derived in Appendix B is that the order parameters for the transition from a Mott-insulator to a superfluid are the expectations $\langle \hat{a}_j \rangle$.

Even though the dynamics of this Hamiltonian can be solved exactly in principle, it is often necessary to approximate it. A standard technique for this uses mean-field theory. Mean-field theory is a general term that deals with approximating operators that have only a small deviation from their expected value. Appendicies A and B lead the reader through some applications of mean-field theory that are historically important.

B.   Approximations to the Dynamics

Our main tool for deriving the time-dependent behaviour of the mean-field Bose-Hubbard model will be the so-called Heisenberg picture of quantum mechanics. In Heisenberg's matrix mechanics, the state of a system is an immutable eigenvector of the Hamiltonian and the time-dependent properties of a system, if they exist, are carried by the operators. For an operator $\hat{A}$, the Heisenberg picture gives an equation of motion for the time derivative of $\hat{A}$ [18]. In all of the cases we will consider, the operator will depend on time in the Heisenberg picture but not in the Schrödinger picture. For this reason, the equation of motion is normally written:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{A}(t) = f\left(\hat{\mathcal{H}}, \hat{A}(t), \hat{A}\right)$$

However, for clarity, we will write:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{A}_{\mathscr{T}} = f\left(\hat{\mathcal{H}}, \hat{A}_{\mathscr{T}}, \hat{A}\right)$$

where $\mathscr{T}$ denotes the basis in which pure states of the system are independent of time. In other words, $_{\mathscr{T}}\langle\Psi|\,\hat{A}_{\mathscr{T}}(t)\,|\Psi\rangle_{\mathscr{T}}$ can only depend on time through the operator, whereas $\langle\Psi(t)|\,\hat{A}(t)\,|\Psi(t)\rangle$ generally depends on time through the state and the operator. First, we seek a unitary transformation $\hat{U}$ that makes this the case.

$$|\Psi(t)\rangle = \hat{U}\,|\Psi\rangle_{\mathscr{T}} \tag{12}$$

The operator $\hat{U}$ can be found if we substitute Eq. (12) into the time-dependent Schrödinger equation:

$$\hat{\mathcal{H}}\,|\Psi(t)\rangle \;=\; i\hbar\frac{\partial}{\partial t}\,|\Psi(t)\rangle$$
$$\hat{\mathcal{H}}\hat{U}\,|\Psi\rangle_{\mathscr{T}} \;=\; i\hbar\frac{\partial}{\partial t}\hat{U}\,|\Psi\rangle_{\mathscr{T}}$$
$$\frac{\partial\hat{U}}{\partial t} \;=\; \frac{-i\hat{\mathcal{H}}}{\hbar}\hat{U}$$
$$\hat{U} \;=\; e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}}$$

Now the time evolution of an arbitrary operator follows from interpreting $\hat{U}$ as a change-of-basis matrix.

$$\hat{A}\hat{U}\,|\Psi\rangle_{\mathscr{T}} \;= \hat{A}\,|\Psi(t)\rangle = \;\hat{U}\hat{A}_{\mathscr{T}}\,|\Psi\rangle_{\mathscr{T}}$$
$$\hat{A}_{\mathscr{T}} \;= \hat{U}^{\dagger}\hat{A}\hat{U} = \;e^{\frac{i\hat{\mathcal{H}}t}{\hbar}}\hat{A}e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}}$$

The rest is a simple matter of taking the derivative. Note that $\frac{\mathrm{d}(\hat{A}\hat{B})}{\mathrm{d}t} = \frac{\mathrm{d}\hat{A}}{\mathrm{d}t}\hat{B} + \hat{A}\frac{\mathrm{d}\hat{B}}{\mathrm{d}t}$ is the product rule for operators.

$$
\begin{aligned}
\frac{\mathrm{d}\hat{A}_{\mathscr{T}}}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{\frac{i\hat{\mathcal{H}}t}{\hbar}}\hat{A}e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}}\right) \\
&= e^{\frac{i\hat{\mathcal{H}}t}{\hbar}}\left(\frac{i}{\hbar}\hat{\mathcal{H}}\hat{A}e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\hat{A}e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}}\right)\right) \\
&= e^{\frac{i\hat{\mathcal{H}}t}{\hbar}}\left(\frac{i}{\hbar}\hat{\mathcal{H}}\hat{A}e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}} - \frac{i}{\hbar}\hat{A}\hat{\mathcal{H}}e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}} + \frac{\partial\hat{A}}{\partial t}e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}}\right) \\
&= \frac{i}{\hbar}\left(\hat{\mathcal{H}}\hat{A}_{\mathscr{T}} - \hat{A}_{\mathscr{T}}\hat{\mathcal{H}}\right) + e^{\frac{i\hat{\mathcal{H}}t}{\hbar}}\frac{\partial\hat{A}}{\partial t}e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}} \\
&= \frac{i}{\hbar}\left[\hat{\mathcal{H}},\hat{A}_{\mathscr{T}}\right] + e^{\frac{i\hat{\mathcal{H}}t}{\hbar}}\frac{\partial\hat{A}}{\partial t}e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}} \\
&= \frac{i}{\hbar}\left[\hat{\mathcal{H}},\hat{A}_{\mathscr{T}}\right] + \left(\frac{\partial\hat{A}}{\partial t}\right)_{\mathscr{T}}
\end{aligned}
\tag{13}
$$

We have used the fact that the Hamiltonian must commute with a Taylor series in itself. For an operator with no explicit time-dependence, Eq. (13) gives:

$$
i\hbar\frac{\mathrm{d}\hat{A}(t)}{\mathrm{d}t} = \left[\hat{A}(t),\hat{\mathcal{H}}\right]
\tag{14}
$$

We are interested in how this affects the anhilation operator at a given point. First, we will evaluate the commutator with respect to the most general Hamiltonian we have written down.

$$
\begin{aligned}
\left[\hat{\Psi}(\mathbf{r}), \int_{\mathbb{R}^3}\hat{\Psi}^\dagger(\mathbf{r}')\left(\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}')\right)\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}'\right] &= \int_{\mathbb{R}^3}\hat{\Psi}(\mathbf{r})\hat{\Psi}^\dagger(\mathbf{r}')\left(\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}')\right)\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}' \\
&\quad - \int_{\mathbb{R}^3}\hat{\Psi}^\dagger(\mathbf{r}')\left(\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}')\right)\hat{\Psi}(\mathbf{r}')\hat{\Psi}(\mathbf{r})\mathrm{d}\mathbf{r}' \\
&= \int_{\mathbb{R}^3}\left[\hat{\Psi}(\mathbf{r}),\hat{\Psi}^\dagger(\mathbf{r}')\right]\left(\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}')\right)\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}' \\
&= \int_{\mathbb{R}^3}\delta(\mathbf{r}-\mathbf{r}')\left(\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}')\right)\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}' \\
&= \left(\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\hat{\Psi}(\mathbf{r})
\end{aligned}
\tag{15}
$$

The commutator involving the interaction term in Eq. (10) can be evaulated with the help of the Jacobi identity: $[\hat{A},\hat{B}\hat{C}] = [\hat{A},\hat{B}]\hat{C} + \hat{B}[\hat{A},\hat{C}]$.

$$
\begin{aligned}
&\left[\hat{\Psi}(\mathbf{r}), \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\hat{\Psi}^\dagger(\mathbf{r}')\hat{\Psi}^\dagger(\mathbf{r}'')V_{\text{int}}(\mathbf{r}'-\mathbf{r}'')\hat{\Psi}(\mathbf{r}'')\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}''\mathrm{d}\mathbf{r}'\right] \\
&= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\hat{\Psi}(\mathbf{r})\hat{\Psi}^\dagger(\mathbf{r}')\hat{\Psi}^\dagger(\mathbf{r}'')V_{\text{int}}(\mathbf{r}'-\mathbf{r}'')\hat{\Psi}(\mathbf{r}'')\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}''\mathrm{d}\mathbf{r}' \\
&\quad - \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\hat{\Psi}^\dagger(\mathbf{r}')\hat{\Psi}^\dagger(\mathbf{r}'')V_{\text{int}}(\mathbf{r}'-\mathbf{r}'')\hat{\Psi}(\mathbf{r}'')\hat{\Psi}(\mathbf{r}')\hat{\Psi}(\mathbf{r})\mathrm{d}\mathbf{r}''\mathrm{d}\mathbf{r}' \\
&= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\left[\hat{\Psi}(\mathbf{r}),\hat{\Psi}^\dagger(\mathbf{r}')\hat{\Psi}^\dagger(\mathbf{r}'')\right]V_{\text{int}}\hat{\Psi}(\mathbf{r}'')\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}''\mathrm{d}\mathbf{r}' \\
&= \int_{\mathbb{R}^3}\delta(\mathbf{r}-\mathbf{r}')\int_{\mathbb{R}^3}\hat{\Psi}^\dagger(\mathbf{r}'')V_{\text{int}}(\mathbf{r}'-\mathbf{r}'')\hat{\Psi}(\mathbf{r}'')\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}''\mathrm{d}\mathbf{r}' \\
&\quad + \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\hat{\Psi}^\dagger(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}'')V_{\text{int}}(\mathbf{r}'-\mathbf{r}'')\hat{\Psi}(\mathbf{r}'')\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}''\mathrm{d}\mathbf{r}' \\
&= 2\int_{\mathbb{R}^3}\hat{\Psi}^\dagger(\mathbf{r}')V_{\text{int}}(\mathbf{r}-\mathbf{r}')\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}'\hat{\Psi}(\mathbf{r})
\end{aligned}
\tag{16}
$$

Note that in the last step we have used the fact that the interaction potential is symmetric in its two arguments. Equations (15) and (16) combine to give:

$$\left[\hat{\Psi}(\mathbf{r}), \hat{\mathcal{H}}\right] = \left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) + \int_{\mathbb{R}^3}\hat{\Psi}^\dagger(\mathbf{r}')V_{\text{int}}(\mathbf{r} - \mathbf{r}')\hat{\Psi}(\mathbf{r}')\mathrm{d}\mathbf{r}'\right]\hat{\Psi}(\mathbf{r})$$

Or (writing the time-dependence explicitly),

$$\left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) + \int_{\mathbb{R}^3}\hat{\Psi}^\dagger(\mathbf{r}', t)V_{\text{int}}(\mathbf{r} - \mathbf{r}')\hat{\Psi}(\mathbf{r}', t)\mathrm{d}\mathbf{r}'\right]\hat{\Psi}(\mathbf{r}, t) \tag{17}$$

If we choose a contact potential for the interactions, as we did when deriving the Bose-Hubbard Hamiltonian, Eq. (17) becomes:

$$\left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) + g|\hat{\Psi}(\mathbf{r}, t)|^2\right]\hat{\Psi}(\mathbf{r}, t)$$

and by Eq. (14) this is equal to $i\hbar\frac{\partial\hat{\Psi}(\mathbf{r}, t)}{\partial t}$. What this gives us is a non-linear partial differential equation for $\Phi(\mathbf{r}, t)$ which we define to be the expectation of $\hat{\Psi}(\mathrm{r}, t)$:

$$i\hbar\frac{\partial\Phi(\mathbf{r}, t)}{\partial t} = \left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\Phi(\mathbf{r}, t) + g\left\langle|\hat{\Psi}(\mathbf{r}, t)|^2\hat{\Psi}(\mathbf{r}, t)\right\rangle$$

Assuming that deviations from the expectation are small, zeroth order mean-field theory becomes a decoupling approximation that lets us rewrite the last term:

$$i\hbar\frac{\partial\Phi(\mathbf{r}, t)}{\partial t} = \left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) + g|\Phi(\mathbf{r}, t)|^2\right]\Phi(\mathbf{r}, t) \tag{18}$$

This is called the Gross-Pitaevskii equation and its solutions represent condensate wavefunctions of a dilute Bose gas [10]. This is because $|\Phi(\mathbf{r}, t)|^2 = \langle\hat{\rho}(\mathbf{r}, t)\rangle$, the expected number density as a function of position and time. This will coincide with the ground-state probability density when all particles participate in BEC. It is clear that without any interactions, Eq. (18) would be identical to the Schrödinger equation and for this reason, it is often called the non-linear Schrödinger equation. One can imagine a discretized version of it in which particles may only occupy a finite number of sites. This will transform the PDE into a set of coupled ODEs. It turns out that such a discretization follows immediately from using the Bose-Hubbard Hamiltonian in place of the full many-boson Hamiltonian in Eq. (14) [10]. For the external potentials:

$$\begin{aligned}\left[\hat{a}_m, \sum_j \epsilon_j\hat{a}_j^\dagger\hat{a}_j\right] &= \sum_{j\neq m}\epsilon_j\hat{a}_m\hat{a}_j^\dagger\hat{a}_j + \epsilon_m\hat{a}_m\hat{a}_m^\dagger\hat{a}_m \\ &\quad - \sum_{j\neq m}\epsilon_j\hat{a}_j^\dagger\hat{a}_j\hat{a}_m - \epsilon_m\hat{a}_m^\dagger\hat{a}_m\hat{a}_m \\ &= \epsilon_m\left(\hat{a}_m\hat{a}_m^\dagger - \hat{a}_m^\dagger\hat{a}_m\right)\hat{a}_m \\ &= \epsilon_m\hat{a}_m \end{aligned} \tag{19}$$

For the interaction term:

$$\begin{aligned}\left[\hat{a}_m, \sum_j \hat{a}_j^\dagger\hat{a}_j^\dagger\hat{a}_j\hat{a}_j\right] &= \sum_{j\neq m}\hat{a}_m\hat{a}_j^\dagger\hat{a}_j^\dagger\hat{a}_j\hat{a}_j + \hat{a}_m\hat{a}_m^\dagger\hat{a}_m^\dagger\hat{a}_m\hat{a}_m - \sum_{j\neq m}\hat{a}_j^\dagger\hat{a}_j^\dagger\hat{a}_j\hat{a}_j\hat{a}_m - \hat{a}_m^\dagger\hat{a}_m^\dagger\hat{a}_m\hat{a}_m\hat{a}_m \\ &= \hat{a}_m\hat{a}_m^\dagger\hat{a}_m^\dagger\hat{a}_m\hat{a}_m - \hat{a}_m^\dagger\hat{a}_m^\dagger\hat{a}_m\hat{a}_m\hat{a}_m \\ &= \hat{a}_m\hat{a}_m^\dagger\hat{a}_m^\dagger\hat{a}_m\hat{a}_m - \hat{a}_m^\dagger\left(\hat{a}_m\hat{a}_m^\dagger - 1\right)\hat{a}_m\hat{a}_m \\ &= \hat{a}_m\hat{a}_m^\dagger\hat{a}_m^\dagger\hat{a}_m\hat{a}_m - \hat{a}_m^\dagger\hat{a}_m\hat{a}_m^\dagger\hat{a}_m\hat{a}_m + \hat{a}_m^\dagger\hat{a}_m\hat{a}_m \\ &= \hat{a}_m\hat{a}_m^\dagger\hat{a}_m^\dagger\hat{a}_m\hat{a}_m - \left(\hat{a}_m\hat{a}_m^\dagger - 1\right)\hat{a}_m^\dagger\hat{a}_m\hat{a}_m + \hat{a}_m^\dagger\hat{a}_m\hat{a}_m \\ &= 2\hat{a}_m^\dagger\hat{a}_m\hat{a}_m \end{aligned} \tag{20}$$

And for the hopping term:

$$
\begin{aligned}
\left[\hat{a}_m, \sum_{\langle j,l \rangle \bullet} \hat{a}_j^\dagger \hat{a}_l \right] &= \sum_{\substack{\langle j,l \rangle \bullet \\ j,l \neq m}} \hat{a}_m \hat{a}_j^\dagger \hat{a}_l + \sum_{\langle j,m \rangle} \hat{a}_m \hat{a}_j^\dagger \hat{a}_m + \sum_{\langle l,m \rangle} \hat{a}_m \hat{a}_m^\dagger \hat{a}_l \\
&\quad - \sum_{\substack{\langle j,l \rangle \bullet \\ j,l \neq m}} \hat{a}_j^\dagger \hat{a}_l \hat{a}_m - \sum_{\langle j,m \rangle} \hat{a}_j^\dagger \hat{a}_m \hat{a}_m - \sum_{\langle l,m \rangle} \hat{a}_m^\dagger \hat{a}_l \hat{a}_m \\
&= \sum_{\langle j,m \rangle} \left(\hat{a}_m \hat{a}_j^\dagger \hat{a}_m + \hat{a}_m \hat{a}_m^\dagger \hat{a}_j \right) - \sum_{\langle j,m \rangle} \left(\hat{a}_j^\dagger \hat{a}_m \hat{a}_m + \hat{a}_m^\dagger \hat{a}_j \hat{a}_m \right) \\
&= \sum_{\langle j,m \rangle} \left(\hat{a}_m \hat{a}_j^\dagger \hat{a}_m + \hat{a}_m \hat{a}_m^\dagger \hat{a}_j \right) - \sum_{\langle j,m \rangle} \left(\hat{a}_m \hat{a}_j^\dagger \hat{a}_m + \left(\hat{a}_m \hat{a}_m^\dagger - 1 \right) \hat{a}_j \right) \\
&= \sum_{\langle j,m \rangle} \hat{a}_j
\end{aligned}
\tag{21}
$$

Putting these together with Eq. (14), we have:

$$
i\hbar \frac{\mathrm{d}\hat{a}_m(t)}{\mathrm{d}t} = \epsilon_m \hat{a}_m(t) + U |\hat{a}_m(t)|^2 \hat{a}_m(t) - J \sum_{\langle j,m \rangle} \hat{a}_j(t)
$$

Defining $\psi_m(t) = \langle \hat{a}_m(t) \rangle$ and using the zeroth order mean-field approximation again, we get the discrete non-linear Schrödinger equation:

$$
\frac{\mathrm{d}\psi_m(t)}{\mathrm{d}t} = \frac{-i}{\hbar} \left[\epsilon_m \psi_m(t) + U |\psi_m(t)|^2 \psi_m(t) - J \sum_{\langle j,m \rangle} \psi_j(t) \right]
\tag{22}
$$

It is instructive to think of the DNLS in the context of energy conservation. It can easily be checked that Eq. (22) is equal to $-\frac{\partial \mathcal{H}_{\mathrm{CL}}}{\partial \psi_m^*}$ where $\mathcal{H}_{\mathrm{CL}}$ is given by:

$$
\mathcal{H}_{\mathrm{CL}} = \frac{i}{\hbar} \left[\sum_j \epsilon_j |\psi_j|^2 + \frac{U}{2} \sum_j |\psi_j|^4 - J \sum_{\langle j,l \rangle \bullet} \psi_j^* \psi_l \right]
\tag{23}
$$

Moreover $\mathcal{H}_{\mathrm{CL}}$ is a valid Hamiltonian because $\frac{\partial \mathcal{H}_{\mathrm{CL}}}{\partial \psi_m}$ is indeed the complex conjugate of Eq. (22) and Hamilton's equations are satisfied. This illustrates a connection between the quantum and classical versions of Hamiltonian mechanics. In this sense Eq. (23) can be considered the classical limit of the Bose-Hubbard Hamiltonian even though it is not the expression we would get if we naïvely replaced $\hat{a}$ with ψ in Eq. (11) [19]. We will be interested in using Eq. (23) to verify the stability of a numerical method chosen to integrate the DNLS.

III.   SOLVING THE EXACT PROBLEM

The dynamics of the Bose-Hubbard Hamiltonian can be found by solving an eigenvalue problem. To write down the Schrödinger equation for this Hamiltonian, one must find the matrix representation of it. In the number eigenstate basis, this is straightforward since we know exactly what the creation and anhilation operators do to number eigenstates. For a system with two sites and two bosons, the number eigenstate basis is the set $\{ |2,0\rangle, |1,1\rangle, |0,2\rangle \}$ and the corresponding matrix is:

$$
\left[\hat{\mathcal{H}} \right] = \begin{bmatrix} \langle 2,0| \hat{\mathcal{H}} |2,0\rangle & \langle 2,0| \hat{\mathcal{H}} |1,1\rangle & \langle 2,0| \hat{\mathcal{H}} |0,2\rangle \\ \langle 1,1| \hat{\mathcal{H}} |2,0\rangle & \langle 1,1| \hat{\mathcal{H}} |1,1\rangle & \langle 1,1| \hat{\mathcal{H}} |0,2\rangle \\ \langle 0,2| \hat{\mathcal{H}} |2,0\rangle & \langle 0,2| \hat{\mathcal{H}} |1,1\rangle & \langle 0,2| \hat{\mathcal{H}} |0,2\rangle \end{bmatrix}
$$

This matrix is always finite-dimensional so it can easily be diagonalized when it is small. The question is, *how small do the experimental numbers need to be for this matrix to be small?* The answer is unfortunately pessimistic. For an

optical lattice with N_s sites and N_b bosons, the dimension of the Hilbert space (the number of number eigenstates having the bosons partitioned among the sites in all possible ways) is $\binom{N_b+N_s-1}{N_b}$. In other words, the size of the matrix we have to work with grows super-exponentially with the number of sites and the number of bosons - the systems become too large very quickly. It is therefore worth spending some time thinking about efficient methods for diagonalizing the Bose-Hubbard Hamiltonian that exploit some of its structure.

A.   Analytical Results

It should first be noted that without a coupling constant, the dynamics are trivial. If one inserts $J = 0$ into the Bose-Hubbard Hamiltonian, the resulting operator can be written solely in terms of number operators. Therefore, the matrix is already diagonal in the number eigenstate basis: if one sets up the system so that it resides in a number eigenstate, it will stay in that number eigenstate for the rest of time and the number of bosons on each site will not change. When $J \neq 0$, finding the eigenvalues of $\left[\hat{\mathcal{H}}\right]$ by hand for a general set of parameters is unrealistic. Instead we will restrict our attention to optical lattices in which $U = 0$ and all ϵ_j are the same. In this case, the Hamiltonian we are solving becomes:

$$-J \sum_{\langle j,l \rangle \circ} \left(\hat{a}_j^\dagger \hat{a}_l + \hat{a}_j \hat{a}_l^\dagger \right) \tag{24}$$

An elegant way to solve it is to first consider a system with only one boson. Our formula says that the dimension of this Hilbert space should be N_s. Indeed the number eigenstates are:

$$\{ |1, 0, \ldots, 0, 0\rangle, \ldots, |0, 0, \ldots, 0, 1\rangle \} \equiv \{ |1\rangle_1, \ldots, |1\rangle_{N_s} \}$$

Single particle eigenstates are now labelled by N_s wave vectors. In one dimension they are $0, \frac{2\pi}{N_s}, \frac{4\pi}{N_s}, \ldots, \frac{2(N_s-1)\pi}{N_s}$. For a given wave vector k, the corresponding single-particle eigenstate is:

$$|\psi_k\rangle = \frac{1}{\sqrt{N_s}} \sum_{j=0}^{N_s-1} e^{ijk} |1\rangle_j \tag{25}$$

A direct computation can prove that Eq. (25) is an eigenstate and indeed find the eigenvalue.

$$
\begin{aligned}
\hat{\mathcal{H}} |\psi_k\rangle &= -J \sum_{\langle l,m \rangle \circ} \left(\hat{a}_l^\dagger \hat{a}_m + \hat{a}_l \hat{a}_m^\dagger \right) \frac{1}{\sqrt{N_s}} \sum_{j=0}^{N_s-1} e^{ijk} |1\rangle_j \\
&= -\frac{J}{\sqrt{N_s}} \sum_{j=0}^{N_s-1} e^{ijk} \sum_{\langle l,m \rangle \circ} \hat{a}_l^\dagger \hat{a}_m |1\rangle_j - \frac{J}{\sqrt{N_s}} \sum_{j=0}^{N_s-1} e^{ijk} \sum_{\langle l,m \rangle \circ} \hat{a}_l \hat{a}_m^\dagger |1\rangle_j \\
&= -\frac{J}{\sqrt{N_s}} \sum_{j=0}^{N_s-1} e^{ijk} |1\rangle_{j-1} - \frac{J}{\sqrt{N_s}} \sum_{j=0}^{N_s-1} e^{ijk} |1\rangle_{j+1} \\
&= -\frac{J}{\sqrt{N_s}} \sum_{j=0}^{N_s-1} e^{i(j-1)k} e^{ik} |1\rangle_{j-1} - -\frac{J}{\sqrt{N_s}} \sum_{j=0}^{N_s-1} e^{i(j+1)k} e^{-ik} |1\rangle_{j+1} \\
&= -J e^{ik} |\psi_k\rangle - J e^{-ik} |\psi_k\rangle \\
&= -2J \cos(k) |\psi_k\rangle \tag{26}
\end{aligned}
$$

Now because we assumed that the bosons do not interact, we simply assign a wave vector to each boson. Each boson with wave vector k has energy $-2J \cos(k)$ and we add the energies. Not surprisingly, there are $\binom{N_b+N_s-1}{N_b}$ ways to assign one of the N_s wave vectors to the N_b bosons. If the bosons were distinguishable, this number would be $N_s^{N_b}$ but they are of course indistinguishable. The process of finding multi-particle wavefunctions and converting them to expected number densities is now completely algorithmic. However, for many sites it is usually tedious and leads to expressions that fill multiple pages. A special case that leads to a nice expression occurs when there are two sites. In these circumstances, we may write:

$$k_0 = 0, E_0 = -2J, k_1 = 1, E_1 = 2J$$

Substituting these values into Eq. (25) and using superposition, we have:

$$\begin{aligned} |\psi_0\rangle &= \frac{1}{\sqrt{2}}\left(|1,0\rangle + |0,1\rangle\right) \\ |\psi_1\rangle &= \frac{1}{\sqrt{2}}\left(|1,0\rangle - |0,1\rangle\right) \\ |\Psi(t)\rangle &= A_0 e^{\frac{2iJt}{\hbar}}|\psi_0\rangle + A_1 e^{\frac{-2iJt}{\hbar}}|\psi_1\rangle \end{aligned}$$

If all bosons begin on the first site, the initial conditions $\langle 1,0|\Psi(0)\rangle = 1$ and $\langle 0,1|\Psi(0)\rangle = 0$, tell us that $A_0 = \frac{1}{\sqrt{2}} = A_1$. Expanding the wavefunction in terms of the number eigenstates, we have:

$$\begin{aligned} |\Psi(t)\rangle &= \frac{1}{2}\left(e^{\frac{2iJt}{\hbar}} + e^{\frac{-2iJt}{\hbar}}\right)|1,0\rangle + \frac{1}{2}\left(e^{\frac{2iJt}{\hbar}} - e^{\frac{-2iJt}{\hbar}}\right)|0,1\rangle \\ &= \cos\left(\frac{2Jt}{\hbar}\right)|1,0\rangle + \sin\left(\frac{2Jt}{\hbar}\right)|0,1\rangle \\ \langle\hat{n}_0(t)\rangle &= |\langle 1,0|\Psi(t)\rangle|^2 \\ &= \cos^2\left(\frac{2Jt}{\hbar}\right) \end{aligned}$$

This is valid for one boson, but when an ensemble has no interactions, all bosons move through the lattice identically. This gives us the desired expressions with no approximations:

$$\langle\hat{n}_0(t)\rangle = N_{\mathrm{b}}\cos^2\left(\frac{2Jt}{\hbar}\right), \quad \langle\hat{n}_1(t)\rangle = N_{\mathrm{b}}\sin^2\left(\frac{2Jt}{\hbar}\right) \tag{27}$$

Even though these only apply to an uninteresting system, they are helpful in verifying the correctness of code that simulates more interesting systems. As we will see, they also allow one to compare different numerical methods in terms of accuracy.

B. Numerical Diagonalization

The most obvious way to improve upon the incredibly slow process of solving for the number densities by hand, is to diagonalize the matrix for the Bose-Hubbard Hamiltonian numerically. This is necessary to study large systems with interactions and especially large systems with disorder. When the Hamiltonian is a $d \times d$-matrix, the time-dependent wavefunction is:

$$|\Psi(t)\rangle = \sum_{j=1}^{d}\langle\psi_j|\Psi(0)\rangle\, e^{\frac{-iE_j t}{\hbar}}|\psi_j\rangle \tag{28}$$

which contains all the information about the probability of detecting each number eigenstate. The jth expected number density, $\langle\hat{n}_j(t)\rangle$ is the probability of finding the system in a state where one particle is on site j plus twice the probability of finding the system in a state where two particles are on site j plus three times the next probability and so on. While it is simple to keep evaluating Eq. (28) at different times, this method requires the eigenvalues $E_1, \ldots, E_d$ and eigenvectors $|\psi_1\rangle, \ldots, |\psi_d\rangle$. One of the most widely used codes for finding the eigenvalues and eigenvectors of an arbitrary matrix is the linear algebra package LAPACK. There are special functions in LAPACK for dealing with symmetric matrices but the Bose-Hubbard Hamiltonian is more than just a symmetric matrix. It is a symmetric matrix in which most of the entries are zero. These matrices have been termed "sparse". It is plain to see that the only contribution to off-diagonal matrix elements comes from the hopping term. Specifically, the (i, j) matrix element only survives if state i and state j differ by a nearest-neighbour hop. Starting with a number eigenstate, one can construct all such neighbouring states by going through all the sites. If there is at least one boson at a site, moving it to the left gives you one neighbour and moving it to the right gives you another. In a two-dimensional crystal, one can also move the boson backwards and forwards. If one is using periodic boundary conditions, like those used in this thesis, the number of nonzero entries in row i of $\left[\hat{\mathcal{H}}\right]$ is z times the number of occupied sites in state i. Normally z is equal to twice the number of dimensions, but in the one-dimensional crystal having two sites, $z = 1$ because each site only has one nearest-neighbour. Large amounts of memory can be saved with a sparse
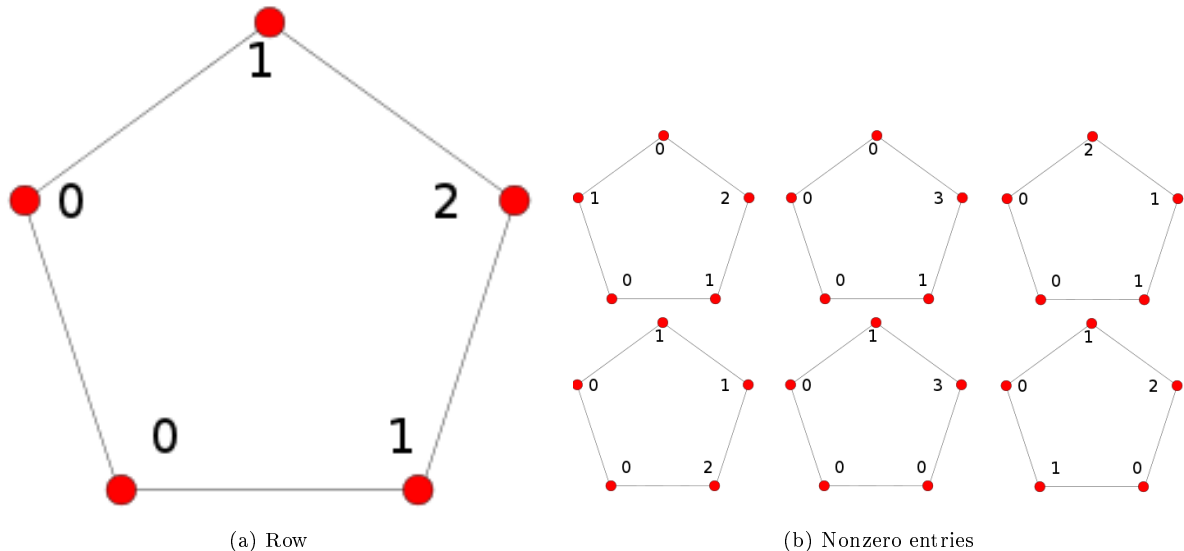
(a) Row

(b) Nonzero entries

FIG. 8: On the left is a configuration of the crystal with three occupied sites. The six states on the right are the neighbouring states differing from it by a nearest-neighbour hop. In the matrix representation of the Bose-Hubbard Hamiltonian, the row corresponding to the state on the left will only have nonzero entries in the positions corresponding to the six states on the right.

matrix by storing only the nonzero entries and their positions. Such storage schemes also greatly improve the speed of matrix multiplication by removing the need to multiply by zero and add zero repeatedly. These properties suggest that traditional diagonalization followed by superposition is not the best method for finding solutions to the Bose-Hubbard Hamiltonian because for a sparse matrix, the eigenvector matrix is almost guaranteed not to be sparse.

C. Propagator Expansions

An alternative route to diagonalization that we have already seen in Eq. (12), uses the formal solution to the Schrödinger equation known as the propagator:

$$|\Psi(t)\rangle = \hat{U}(t)|\Psi(0)\rangle$$
$$\hat{U}(t) = e^{\frac{-i\hat{\mathcal{H}}t}{\hbar}}$$

In order to work with the propagator, it is natural to write it as a series, truncate the series after the most significant terms and multiply the vector $|\Psi(0)\rangle$ by the matrix $\left[\hat{U}(t)\right]$ for increasingly large values of time. As a Taylor series, the propagator is:

$$\left[\hat{U}(t)\right] = \sum_{m=0}^{\infty} \frac{1}{m!}\left(\frac{-it}{\hbar}\right)^m \left[\hat{\mathcal{H}}\right]^m \tag{29}$$

For a managable number of terms, calculating Eq. (29) is faster than diagonalizing a matrix. The calculation, however, is slow enough that one should avoid carrying it out more than once. Therefore, instead of using the algorithm on the left, we use the algorithm on the right:

$$|\Psi(dt)\rangle = \hat{U}(dt)|\Psi(0)\rangle \qquad |\Psi(dt)\rangle = \hat{U}(dt)|\Psi(0)\rangle$$
$$|\Psi(2dt)\rangle = \hat{U}(2dt)|\Psi(0)\rangle \qquad |\Psi(2dt)\rangle = \hat{U}(dt)|\Psi(dt)\rangle$$
$$|\Psi(3dt)\rangle = \hat{U}(3dt)|\Psi(0)\rangle \qquad |\Psi(3dt)\rangle = \hat{U}(dt)|\Psi(2dt)\rangle$$
$$\cdots$$

The steps of this algorithm can be carried out about as quickly as those in the standard method, especially when the matrix $\left[\hat{U}(t)\right]$ is sparse. The main difference is that having to diagonalize a matrix introduces a large overhead, while calculating the propagator introduces an overhead that is comparatively small. In this sense, using the propagator can be seen as settling for approximate diagonalization. The two methods are equivalent when we keep an infinite number of terms but truncating the series is more justifiable than loosening the threshold after which a matrix is considered diagonal. While the Taylor series Eq. (29) is a familiar expansion, a more effective one for our purposes uses Chebyshev polynomials, a family of orthogonal polynomials important in approximation theory [20]. Chebyshev polynomials are defined so that $T_n(\cos(\theta)) = \cos(n\theta)$ and can be generated with the recurrence relation:

$$T_0(x) = 1, \; T_1(x) = x, \; T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)$$

With this definition, the Chebyshev polynomials form an orthonormal basis for a large class of functions, if one uses a slightly modified inner product:

$$\frac{1}{\pi} \int_{-1}^{1} \frac{T_0(x)T_0(x)}{\sqrt{1-x^2}} \mathrm{d}x = 1, \; \frac{2}{\pi} \int_{-1}^{1} \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}} \mathrm{d}x = \delta_{m,n}$$

This is the property that allows us to write the propagator as a sum of Chebyshev polynomials in the Hamiltonian rather than a sum of monomials in the Hamiltonian. To do this, we must rescale the Hamiltonian so that all of its eigenvalues lie in the interval $[-1, 1]$. In other words, $\left[\hat{\mathcal{H}}\right] = a\tilde{\mathcal{H}} + bI$ where $a = \frac{E_{\max} - E_{\min}}{2}$, $b = \frac{E_{\max} + E_{\min}}{2}$ and I is the identity matrix of appropriate size. Now,

$$\left[\hat{U}(t)\right] = e^{\frac{-it}{\hbar}\left(a\bar{\mathcal{H}} + b\right)} = e^{\frac{-ibt}{\hbar}}\left(c_0 + 2\sum_{m=1}^{\infty} c_m T_m\left(\tilde{\mathcal{H}}\right)\right) \tag{30}$$

To find the coefficients, we project the function we are trying to approximate $(e^{\frac{-iaxt}{\hbar}})$ onto each Chebyshev polynomial using the inner product under which they are orthogonal. This means:

$$c_m = \frac{1}{\pi} \int_{-1}^{1} \frac{T_m(x)e^{\frac{-iaxt}{\hbar}}}{\sqrt{1-x^2}} \mathrm{d}x = (-i)^m J_m\left(\frac{at}{\hbar}\right)$$

where J_m is the mth Bessel function of the first kind. These functions are ubiquitous enough to be implemented in the GNU C Library and their extremely fast decay $\left(J_m(x) = O\left(\frac{1}{m^{m+\frac{1}{2}}}\right)\right)$ is what allows our series in Chebyshev polynomials to converge quickly. To show that Eq. (30) converges more quickly than Eq. (29), these two expansions were used to simulate a two-site system. The solutions yielded were compared to the exact solutions Eq. (27). Error was calculated using sum of squared differences:

$$\left|\langle\hat{n}_0(dt)\rangle - N_{\mathrm{b}}\cos^2\left(\frac{2Jdt}{\hbar}\right)\right|^2 + \left|\langle\hat{n}_0(2dt)\rangle - N_{\mathrm{b}}\cos^2\left(\frac{4Jdt}{\hbar}\right)\right|^2 + \left|\langle\hat{n}_0(3dt)\rangle - N_{\mathrm{b}}\cos^2\left(\frac{6Jdt}{\hbar}\right)\right|^2 + \ldots$$

This resulted in a convincing table. In all cases of Table 1, the timestep dt is equal to 0.01 units of $\frac{\hbar}{J}$ and the

Number of terms	Error from Taylor series	Error from Chebyshev series
7	1973.022	0.619
8	59.926	0.116
9	0.356	0.017
10	0.025	0.016

TABLE I: An example of a Chebyshev series converging much more quickly than a Taylor series.

number of bosons N_{b} is equal to 100. To see that some of these errors really are noticable, the deviation between the two solutions is shown in Fig. 9. Note that this figure refers to an expectation of the number operator with the axis label "Number" - a convention used throughout this thesis. This is not the end of the story because of a potentially expensive step in computing the propagator with Chebyshev methods. The scale and shift factors, a and b, both require eigenvalues of the Bose-Hubbard Hamiltonian. It would certainly be unfortunate if our effort to avoid the costly diagonalization step required a full diagonalization anyway. It turns out that extremal eigenvalues
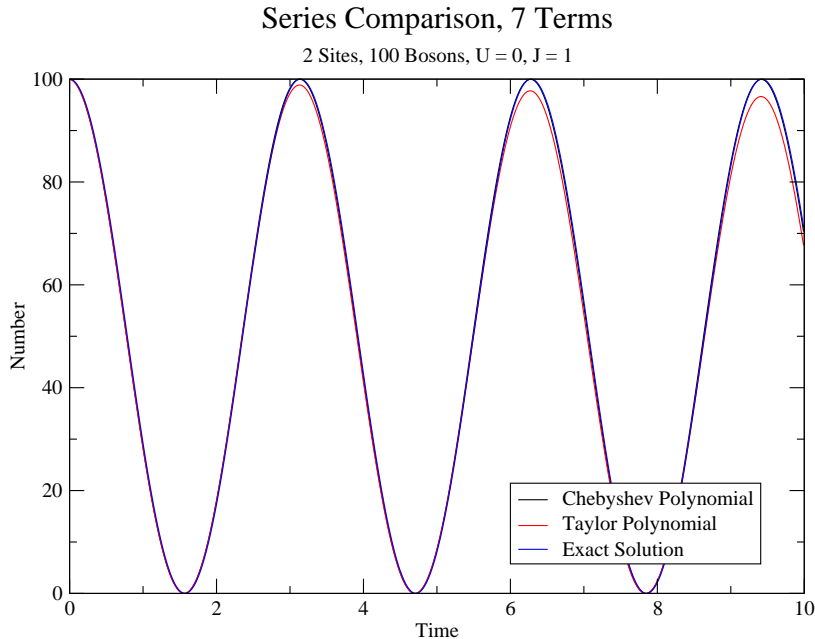
Series Comparison, 7 Terms

FIG. 9: This figure allows one to visualize the first row of Table 1. This means that the error that can be seen between the expected number versus time curve produced by the Chebyshev polynomial and the one produced by the Taylor polynomial is 1973.022.

can be found much more quickly than the entire spectrum. When using LAPACK with a Hamiltonian describing 6 sites and 5 bosons, finding the two extremal eigenvalues is six times faster than finding the whole eigenvector matrix. The savings, however, do not stop there. Various packages have been written to accomodate sparse matrices, such as TRLAN which implements the thickly-restarting Lanczos algorithm [21]. In the following table, its performance in finding the extremal eigenvalues has been compared to that of LAPACK. Also tested was "Customized TRLAN". This

System	Dimension	LAPACK	TRLAN	Customized TRLAN
6 sites, 5 bosons	252	0.09s	0.54s	0.04s
6 sites, 6 bosons	462	0.75s	0.82s	0.03s
7 sites, 6 bosons	924	5.77s	3.21s	0.07s
7 sites, 7 bosons	1716	44.38s	10.42s	0.17s
8 sites, 7 bosons	3432	422.36s	43.99s	0.57s

TABLE II: Comparing the speed with which three algorithms find the largest and smallest energies, necessary for Chebyshev expansion. All times were measured on a Pentium III.

used the exact same TRLAN function except with one additional argument. The additional argument was a pointer to a matrix multiplication function. This function was designed to take advantage of sparsity and was therefore much faster than standard matrix multiplication. When one wishes to know the short time dynamics of a large sparse Hamiltonian such as the Bose-Hubbard Hamiltonian, we see that the Chebyshev expansion provides a much faster route towards solving the Schrödinger equation than standard diagonalization.

IV.   SOLVING THE APPROXIMATE PROBLEM

In contrast to the exact Hamiltonian which gives us a problem where the difficulty of solving it grows exponentially with the size of the lattice, we have presented an approximate form of this problem, Eq. (22), where the difficulty of solving it only grows linearly. This is the discrete non-linear Schrödinger equation, which is a system of N_s coupled

differential equations - one for each site on the lattice. The initial conditions for this system mimic the initial condition for solutions to the exact eigenvalue problem. Instead of saying that $|\Psi(0)\rangle$ is the number eigenstate that has all $N_{\rm b}$ bosons on the first site, we say that $\psi(0)$ has $\sqrt{N_{\rm b}}$ in the first component and zeroes in all other components. An emphasis has been placed on seeing how such a state evolves in time using Runge-Kutta methods.

A. Explicit Runge-Kutta Methods

Runge-Kutta methods were designed to solve equations of the form:

$$\frac{\rm d}{{\rm d}t}\psi(t) = f(\psi(t), t)$$

like the DNLS equation. The Euler method:

$$
\begin{aligned}
\psi(dt) &\approx \psi(0) + f(\psi(0), 0)\, dt \\
\psi(2dt) &\approx \psi(dt) + f(\psi(dt), dt)\, dt \\
\psi(3dt) &\approx \psi(2dt) + f(\psi(2dt), 2dt)\, dt \\
&\cdots
\end{aligned}
$$

is an example of a first-order Runge-Kutta method which means the error between the exact solution and the solution found by performing the Euler method should be asymptotic to dt or $O(dt)$. Instead of giving the general definition right away, we will discuss the reasoning behind the simplest non-trivial example; the midpoint method. The finite difference approximation to the derivative can be written:

$$\frac{\rm d}{{\rm d}t}\psi(t) \approx \frac{\psi(t + dt) - \psi(t)}{dt}$$

This becomes more accurate if we use a centred finite diference:

$$
\begin{aligned}
\frac{\rm d}{{\rm d}t}\psi\left(t + \frac{dt}{2}\right) &\approx \frac{\psi(t + dt) - \psi(t)}{dt} \\
\psi(t + dt) &\approx \psi(t) + \frac{\rm d}{{\rm d}t}\psi\left(t + \frac{dt}{2}\right) dt \\
\psi(t + dt) &\approx \psi(t) + f\left(\psi\left(t + \frac{dt}{2}\right), t + \frac{dt}{2}\right) dt
\end{aligned}
$$

The solution evaluated at $t + \frac{dt}{2}$ is still unknown so it must be estimated further. If we estimate it using a Taylor series (the Euler method again), we have:

$$
\begin{aligned}
\psi\left(t + \frac{dt}{2}\right) &\approx \psi(t) + \frac{\rm d}{{\rm d}t}\psi(t)\frac{dt}{2} \\
&= \psi(t) + f(\psi(t), t)\frac{dt}{2}
\end{aligned}
$$

Substituting this, we arrive at a method that is $O(dt^2)$:

$$\psi(t + dt) \approx \psi(t) + f\left(\psi(t) + f(\psi(t), t)\frac{dt}{2}, t + \frac{dt}{2}\right) dt$$

This is the midpoint method, and it can also be written in the form:

$$
\begin{aligned}
\psi(t + dt) &\approx \psi(t) + dt\sum_{j=1}^{2} b_j d_j \\
d_1 &= f(\psi(t), t) \\
d_2 &= f(\psi(t) + a_{2,1}d_1 dt, t + c_2 dt)
\end{aligned}
$$

where the table

$$
\begin{array}{c|cc}
c_1 & a_{1,1} & a_{1,2} \\
c_2 & a_{2,1} & a_{2,2} \\
\hline
 & b_1 & b_2
\end{array}
$$

is given by

$$
\begin{array}{c|cc}
0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 \\
\hline
 & 0 & 1
\end{array}
$$

More generally, all Runge-Kutta methods are described by the equations:

$$
\boldsymbol{\psi}(t + dt) \;\approx\; \boldsymbol{\psi}(t) + dt \sum_{j=1}^{s} b_j d_j
$$

$$
d_j \;=\; f\left(\boldsymbol{\psi}(t) + dt \sum_{l=1}^{s} a_{j,l} d_l, t + c_j dt \right)
$$

and an $s \times s$ table of coefficients:

$$
\begin{array}{c|c}
\mathbf{c} & A \\
\hline
 & \mathbf{b}^{\mathrm{T}}
\end{array}
$$

called the Butcher tableau. The midpoint method is very easy to implement in code for solving differential equations. Another method which is easy to implement is the classic fourth-order Runge-Kutta method, whose Butcher tableau is:

$$
\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\
1 & 0 & 0 & 1 & 0 \\
\hline
 & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6}
\end{array}
$$

These methods can be readily implemented precisely because the matrix A in the tableau is lower triangular. This allows the intermediate slopes d_j to be explicitly calculated and for this reason, such methods are called explicit Runge-Kutta methods. Before we move on to implicit Runge-Kutta methods, there is an important modification one can make for explicit methods with a carefully chosen Butcher tableau. In all of the previous examples, the timestep dt was assumed to be constant. It makes more sense to choose large timesteps for intervals in which the Runge-Kutta method closely approximates the exact solution and to choose small timesteps for intervals in which the approximation is further off. We cannot actually measure the error between the approximate solution and the exact solution in practice, so we measure the error between one approximate solution and another approximate solution that is similar but more accurate. This can be made clear by looking at the coefficients discovered by Cash and Karp [22]:

$$
\begin{array}{c|cccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{5} & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 \\
\frac{3}{10} & \frac{3}{40} & \frac{9}{40} & 0 & 0 & 0 & 0 \\
\frac{3}{5} & \frac{3}{10} & \frac{-9}{10} & \frac{6}{5} & 0 & 0 & 0 \\
1 & \frac{-11}{54} & \frac{5}{2} & \frac{-70}{27} & \frac{35}{27} & 0 & 0 \\
\frac{7}{8} & \frac{1631}{55296} & \frac{175}{512} & \frac{575}{13824} & \frac{44275}{110592} & \frac{253}{4096} & 0 \\
\hline
 & \frac{2825}{27648} & 0 & \frac{18575}{48384} & \frac{13525}{55296} & \frac{277}{14336} & \frac{1}{4} \\
 & \frac{37}{378} & 0 & \frac{250}{621} & \frac{125}{594} & 0 & \frac{512}{1771}
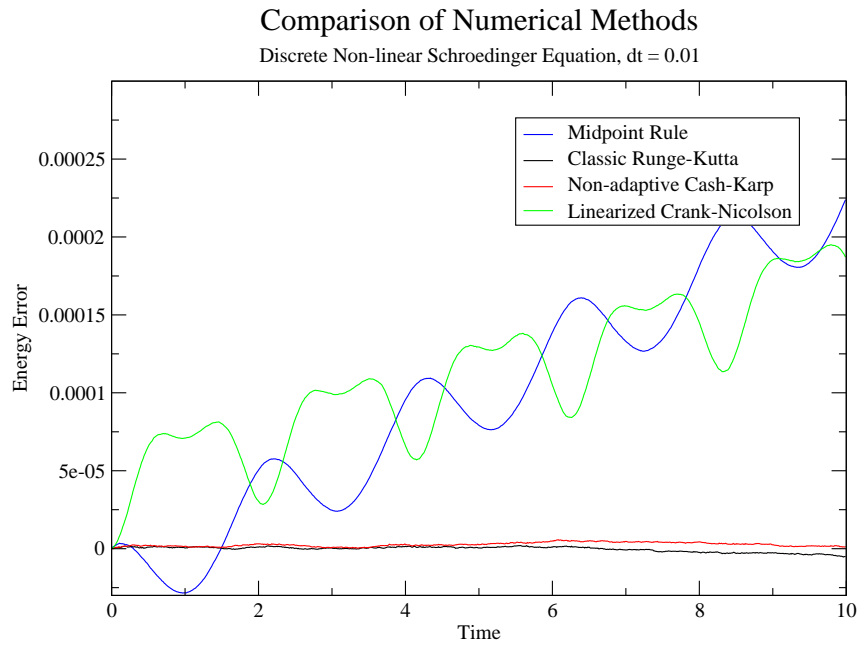\end{array}
$$

FIG. 10: The discrete non-linear Schrödinger equation is a Hamiltonian system which means that its energy drift over time is zero. This plot shows us how well some Runge-Kutta methods preserve this property.
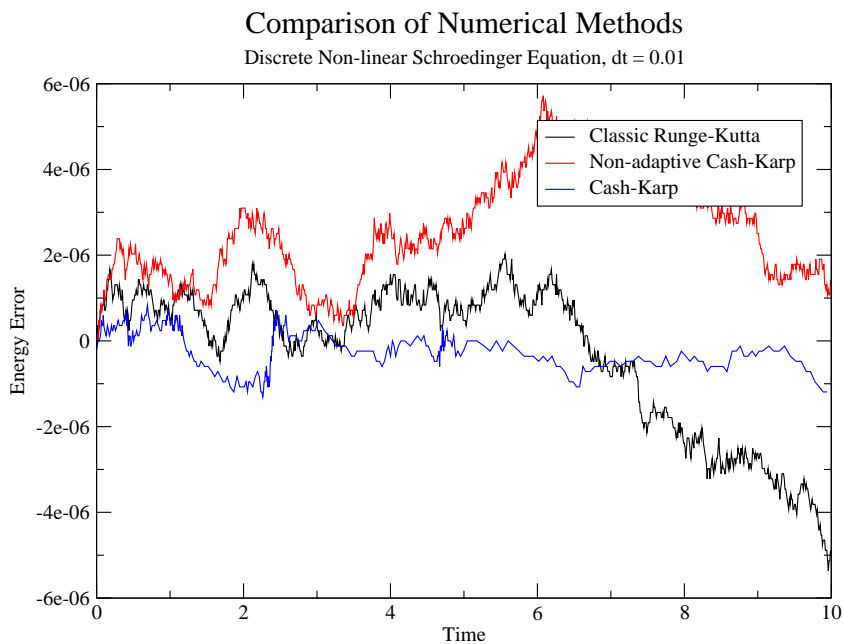

FIG. 11: A closeup showing some energy drifts that are too small to notice in Fig. 10. The method showing the smallest energy drift (which also took the smallest amount of CPU time) was Cash-Karp - the method of choice for this thesis.

Using the top row as the **b** vector yields a Runge-Kutta method that is $O(dt^4)$ while using the bottom row yields one that is $O(dt^5)$ [22]. The stepsize can therefore be adapted when the two methods are run concurrently. The function that chooses a stepsize dt using an error ε is not always obvious. Herein, all work with the discrete non-linear Schrödinger equation has used the following rules which may or may not be in use by other people. Since the less accurate method is $O(dt^4)$, it is approximately true that $\varepsilon = K dt^4$ for some K. We do not know K going in but its value becomes increasingly accurate if we assume that $K = 1$ initially and then measure the $\frac{\varepsilon}{dt^4}$ ratio at each step. If this ratio at a given step is greater than the average of all previous ratios, the error is too large and we decrease the stepsize by this factor. If this ratio is less than the average of all previous ratios, the stepsize is too small and we increase the stepsize by this factor. A sanity check is inserted so that the timestep can change by at most one order of magnitude at a time. Proponents of adaptive methods say that the integrator is *sprinting over flat ground and treading carefully over rocky terrian.*

Figures 10 and 11 rank the Runge-Kutta methods on how well they conserve the energy integral we found in Eq. (23). Here, "Non-adaptive Cash-Karp" refers to the fact that the fourth-order set of coefficients has been ignored and the Cash-Karp method has been used as if it were a regular fifth-order method. In the close-up, it is clear that adapting the timestep is beneficial. Even more impressive is the fact that the Cash-Karp method finished in 375 steps, whereas the others all took 1000.

B. An Implicit Runge-Kutta Method

One of the methods compared in Fig. 10 is based on an implicit Runge-Kutta method. The Crank-Nicolson Butcher tableau is:

$$
\begin{array}{c|cc}
1 & \frac{1}{2} & \frac{1}{2} \\
0 & 0 & 0 \\
\hline
 & \frac{1}{2} & \frac{1}{2}
\end{array}
$$

which means it is defined implicitly by:

$$\frac{\boldsymbol{\psi}(t+dt) - \boldsymbol{\psi}(t)}{dt} \approx \frac{1}{2}\left(f\left(\boldsymbol{\psi}(t+dt), t+dt\right) + f\left(\boldsymbol{\psi}(t), t\right)\right)$$

Solving this for $\boldsymbol{\psi}(t+dt)$ in terms of $\boldsymbol{\psi}(t)$ is not an easy task when the differential equation is non-linear. However, if our equation were linear and the function f amounted to multiplication by some matrix F, the Crank-Nicolson method would be easy to deal with.

$$
\begin{aligned}
\frac{\boldsymbol{\psi}(t+dt) - \boldsymbol{\psi}(t)}{dt} &\approx \frac{1}{2}\left(F\boldsymbol{\psi}(t+dt) + F\boldsymbol{\psi}(t)\right) \\
\boldsymbol{\psi}(t+dt) - \boldsymbol{\psi}(t) &\approx F\boldsymbol{\psi}(t+dt)\frac{dt}{2} + F\boldsymbol{\psi}(t)\frac{dt}{2} \\
\left(I - F\frac{dt}{2}\right)\boldsymbol{\psi}(t+dt) &\approx \left(I + F\frac{dt}{2}\right)\boldsymbol{\psi}(t) \\
\boldsymbol{\psi}(t+dt) &\approx \left(I - F\frac{dt}{2}\right)^{-1}\left(I + F\frac{dt}{2}\right)\boldsymbol{\psi}(t)
\end{aligned}
$$

The matrix F is not a constant though. For the DNLS it is:

$$
\begin{bmatrix}
U|\psi_1|^2 & -J & 0 & \ldots & -J \\
-J & U|\psi_2|^2 & -J & \ldots & 0 \\
0 & -J & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & -J \\
-J & \ldots & 0 & -J & U|\psi_{N_s}|^2
\end{bmatrix}
$$

This depends on the vector $\boldsymbol{\psi}$ already so the approach that has been taken in Fig. 10 has been substituting $\boldsymbol{\psi}\left(t + \frac{dt}{2}\right)$. The linearized word in the legend reflects the fact that this part does not follow the true Crank-Nicolson method [10]. It is possible that higher order implicit methods surpass the performance of Cash-Karp but this has been left as a topic for another project.

V. OPTICAL LATTICE SIMULATIONS

We are now at the point where we can observe some interesting physics in the solutions to the Bose-Hubbard model and its zeroth order mean-field approximation. As a start, understanding the dynamics when there is no disorder requires adjusting the number of bosons and the ratio J/U. A few cases uninteresting to us can be eliminated right away. After analyzing the requirements for the discrete non-linear Schrödinger equation to agree with the dynamics of the Bose-Hubbard Hamiltonian, differences between them and connections to well known phenomena will be presented. There is an infinite number of ways to introduce disorder so one has to be careful not to make conclusions that are biased to one particular "complexion". The approach that has been taken involves choosing a disorder strength W and allowing the ϵ_j values to fall randomly in the interval $\left[\frac{W}{2}, -\frac{W}{2}\right]$ and shifting them so that $\epsilon_0 = 0$. Turning up the disorder is accomplished by raising W while changing the complexion is accomplished by choosing a different seed for the random number generator.

A. Exploring the Parameter Space

Recall that when we derived the discrete non-linear Schrödinger equation, the only approximation we made beyond those of the Bose-Hubbard model itself, concerned the interaction term. If there are no interactions, the DNLS is exact so we need not consider the cases $U = 0$ and/or $N_b = 1$. To decide what parts of the parameter space we should consider when integrating the DNLS, it helps to think about the Bose-Hubbard phase diagram. The functions $\psi_j(t)$ that are returned when one solves the discrete non-linear Schrödinger equation are really the superfluid order parameters $\langle \hat{a}_j(t) \rangle$ for each site. This means that near the Mott-insulating phase, where the order parameters vanish, solutions to the discrete non-linear Schrödinger equation that are not identically zero begin to lose their physical meaning. We plan to use $|\psi_j(t)|^2$ to get the number of bosons on each site so the solutions must be normalizable and therefore nonzero by definition. It follows that if there is any Mott-insulating phase in our parameter space, we should not expect the DNLS to work there.
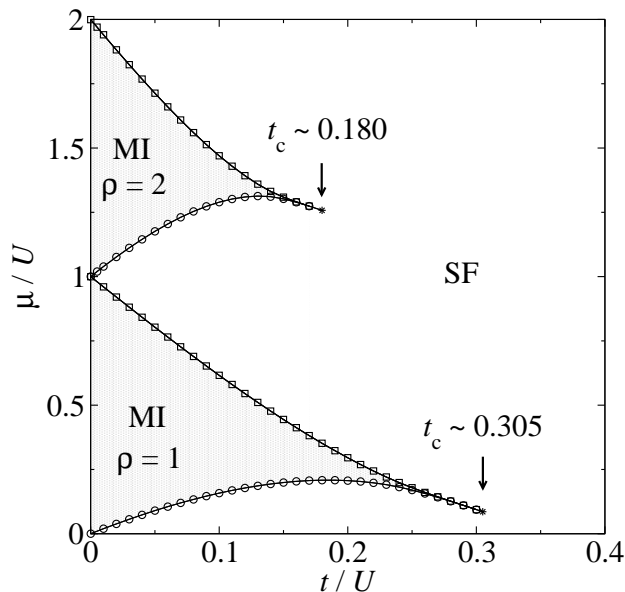

FIG. 12: An accurate phase diagram for the ordered Bose-Hubbard model by [14] using the symbol t where we have used J - another common convention.

The boundary between the Mott-insulator and the superfluid is divided into several lobes and each lobe has a critical value of J/U above which the system always exists as a superfluid. The largest such critical value is the vertex of the first lobe. From the mean-field phase diagram derived in Appendix B, it looks like the critical value of $\frac{Jz}{U}$ should be $3 - 2\sqrt{2}$ meaning that the critical J/U value in one dimension is $\frac{3}{2} - \sqrt{2} \approx 0.086$. This is far from the truth because the calculation there used first order mean field theory. This approximation suffers in low dimensions particularly around the critical points. Figure 12 shows the results of a group who did a more accurate calculation. It shows that we can be sure that our system is away from a Mott lobe when $J/U > 0.305$.

The chemical potential on the vertical axis assumes that one is working in the grand canonical ensemble where the number of particles (per site) is a dependent variable. All of our simulations take place in the canonical ensemble where the number of particles is the independent variable which causes the chemical potential to adjust itself accordingly. If we were to quantify the filling fraction to which the chemical potential values correspond on the diagram, we would see that the area around the first Mott lobe always has between 0 and 1 bosons per site. For argument's sake, if we were to simulate an optical lattice with 50 sites, we would have to go up to 50 bosons as well to fully explore this area. This is much larger than the systems we can simulate so a modest lattice with 8 sites has been chosen, with the 8 site, 8 boson problem being just at the threshold of the Pentium III available for this project.

Due to small size effects, it is important not to make the lattice much smaller. A particularly misleading phenomenon occurs in a lattice with six sites and non-interacting bosons. We already saw in Eq. (26) that the single-particle energies of non-interacting bosons were given by $E_k = -2J\cos(k)$. If we try all possible values for the wave vector, $k = \frac{2\pi r}{N_s}, r \in \{0, 1, ..., N_s - 1\}$, we get:

$$\cos(0) = 1, \cos\left(\frac{\pi}{3}\right) = \frac{1}{2}, \cos\left(\frac{2\pi}{3}\right) = -\frac{1}{2}, \cos(\pi) = -1, \cos\left(\frac{4\pi}{3}\right) = -\frac{1}{2}, \cos\left(\frac{5\pi}{3}\right) = \frac{1}{2}$$

These are all rational numbers and since they are energies, they appear as frequencies of the harmonics added together to produce $|\Psi(t)\rangle$. The ratios of the frequencies are rational purely because we chose a small lattice. The dynamics therefore show artificial periodicity, something that almost never happens in a general lattice. Two of our results for the exact Bose-Hubbard dynamics can give us some information about whether the discrete non-linear Schrödinger equation serves as a good approximation. The plots of Fig. 13 the expected number versus time are qualitatively
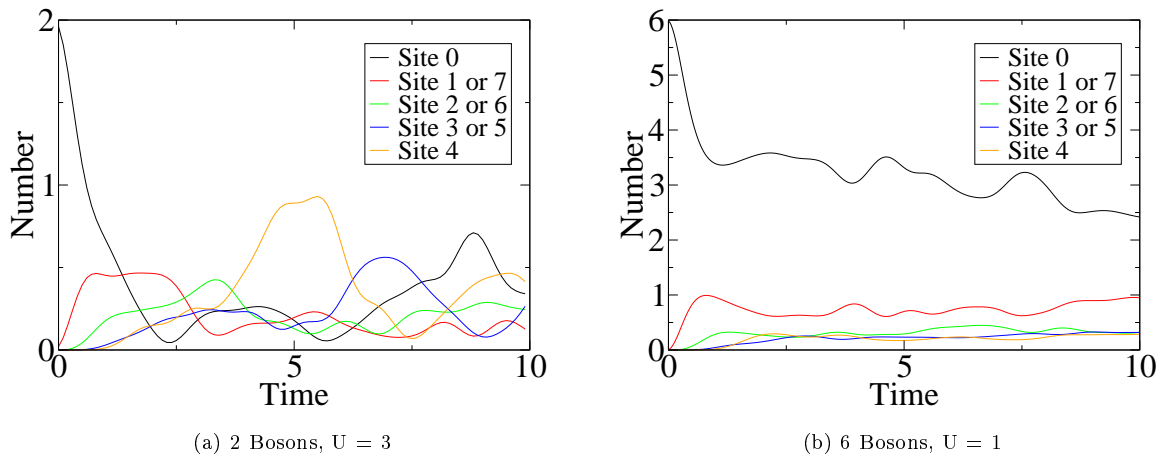

(a) 2 Bosons, U = 3   (b) 6 Bosons, U = 1

FIG. 13: Two solutions to the Bose-Hubbard Hamiltonian with 8 sites and J = 1 to which the discrete non-linear Schrödinger results will be compared.

different. In the left plot, site 4 is for a short time, the most highly populated site, while in the right plot, no site is ever more populated that site 0. If we find a function ψ that solves the discrete non-linear Schrödinger equation, it is not true that $x\psi$ solves it as well for a real number x. This property only holds for linear differential equations. However, what we can say is that if ψ solves the DNLS with interaction potential U, then $\sqrt{x}\psi$ solves the DNLS with interaction potential $\frac{U}{x}$:

$$\frac{d\psi_m(t)}{dt} = \frac{-i}{\hbar}\left[\epsilon_m\psi_m(t) + U|\psi_m(t)|^2\psi_m(t) - J\sum_{\langle j,m\rangle}\psi_j(t)\right]$$

$$\frac{d\sqrt{x}\psi_m(t)}{dt} = \frac{-i}{\hbar}\left[\epsilon_m\sqrt{x}\psi_m(t) + \frac{U}{x}|\sqrt{x}\psi_m(t)|^2\sqrt{x}\psi_m(t) - J\sum_{\langle j,m\rangle}\sqrt{x}\psi_j(t)\right]$$

This means that if we increase the number of bosons by a certain factor and decrease the interaction strength by that same factor, solutions to the discrete non-linear Schrödinger equation will not change. The DNLS can therefore

only apply in regions of the parameter space where exact solutions are close to being invariant under this scaling. Both plots in Fig. 13 were in the regime, previously thought to be safe for the DNLS. For the two boson plot, $J/U = 0.333 > 0.305$ and for the six boson plot $J/U = 1 > 0.305$. The parameters were chosen so that they exactly correspond to the scaling just discussed. Going from the left plot to the right plot is nothing but setting $x = 3$. The differences between these two plots suggest that the region of applicability for the DNLS must have J/U values well above 0.305. The results of the DNLS for 8 sites, 6 bosons and $J/U = 1$ still do not agree with those for the Bose-Hubbard Hamiltonian as can be seen in Fig 14.
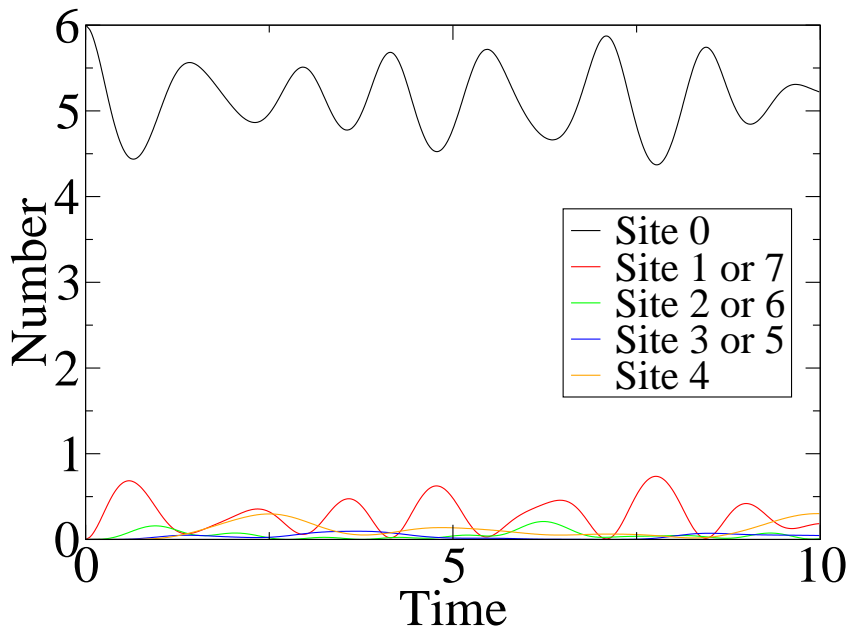

FIG. 14: Discrete non-linear Schrödinger Equation results for 8 Sites, 6 Bosons, J = U = 1. They are significantly different from the the results of the Bose-Hubbard Hamiltonain with the same choice of parameters: Fig. 13 (right).

It is beginning to look like the DNLS which we know to be valid in the non-interacting limit is *only* valid in the non-interacting limit. Note that Fig. 14 still displays signs of being a valid diffusion process that could conceivably occur on a lattice. The $\langle \hat{n}_1(t) \rangle$ and $\langle \hat{n}_7(t) \rangle$ curves are the same because there is nothing to distinguish the site that is "one unit left" of the site where bosons start from the site that is "one unit right" of the site where bosons start. The same is true for $\langle \hat{n}_2(t) \rangle$ and $\langle \hat{n}_6(t) \rangle$ and so on. When we consider disordered lattices, we will see that this symmetry is broken and all sites generally have their own distinct dynamics. It is also intuitive that the total number of bosons is conserved. At all times the sum of the expected numbers will be N_b. The curves do not look like they add up to the constant function 6 but we must remember that the curves belonging to pairs of indistinguishable sites must be given a weight of 2. Finally the progression of bosons off the initial site 0 follows a natural order. At the start, the most quickly increasing numbers are those on sites 1 and 7 - the nearest-neighbours. The second most quickly increasing numbers are those on sites 2 and 6 - the next nearest-neighbours. The same would be observed if particles were moving from place to place classically. The situation can be summarized well by a sequence of bar graphs.

A valid concern is that the finite size of the lattice introduces a discrepancy between the simulation and an experiment. This is true when the probability of finding bosons maximally far away from the initial site becomes high. Further spreading of the wave packet from this configuration essentially results in bosons "wrapping around" and returning to the initial site if periodic boundary conditions are used. It is usually not hard to discern when this effect is dominating and the dynamics are no longer relevant to real optical lattices which typically have thousands of sites. For the purposes of comparing the discrete non-linear Schrödinger equation to the Bose-Hubbard Hamiltonian, this is a non-issue. Periodic boundary conditions have been implemented with both of them so it still makes sense to ask whether these two models agree with eachother even if they share the same discrepancy with the actual underlying physics. There are also many lattice shapes that can be made in the lab so it is possible that an experiment may one day be done that truly does have periodic boundary conditions in the underlying physics.
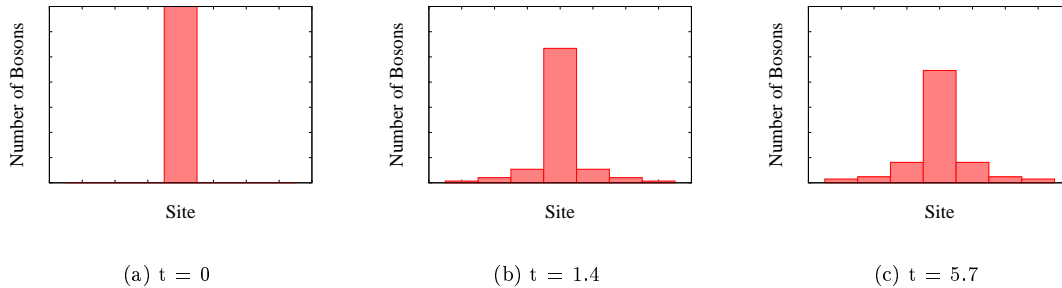
(a) t = 0   (b) t = 1.4   (c) t = 5.7

FIG. 15: The spreading of 7 bosons that begin at the central site of a 7 site lattice. Each frame is a snapshot of the expected number of bosons on each site. An odd number of sites has been chosen for aesthetic reasons.

B. Localization of Bosons

A case where it is quite clear that the "wrap around" effect occurs is that of 8 sites, 6 bosons and $J/U = 6$. For this case that has a very high value of J/U, the exact solution and the DNLS solution appear to agree over short times, but the Bose-Hubbard Hamiltonian appears driven toward a steady-state in which all sites are similarly occupied whereas the DNLS continues to oscillate. We say the dynamics of the Bose-Hubbard Hamiltonian have become localized. It is also possible to observe localization at early times, in which bosons largely do not leave the initial site in the first place.
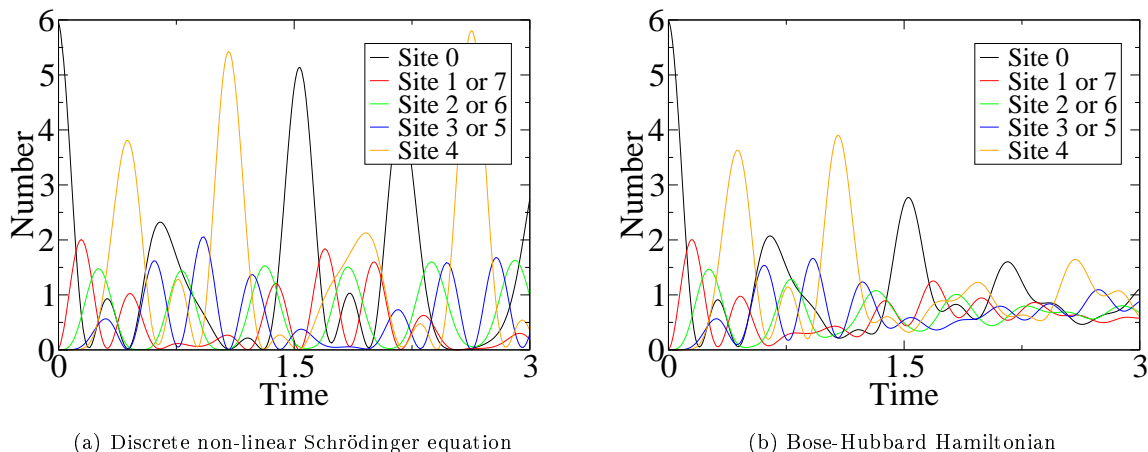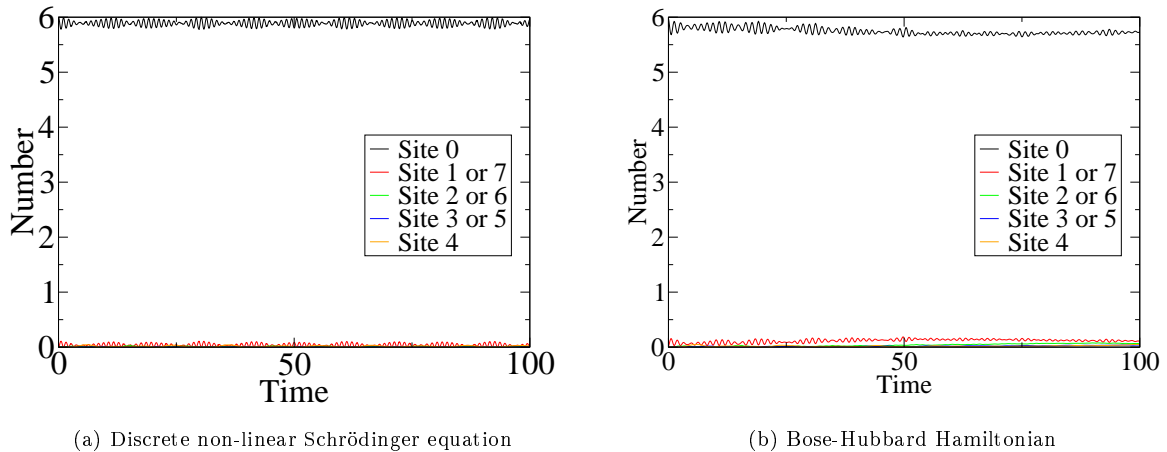

(a) Discrete non-linear Schrödinger equation   (b) Bose-Hubbard Hamiltonian

FIG. 16: Expected number versus time curves for 8 sites, 6 bosons, U = 1 and J = 6. One of them shows localization at late times.

From Fig. 17 it appears as though the DNLS and the exact Hamiltonian both admit this early time localization but on closer inspection we see that only the DNLS does. For 8 sites, 6 bosons and $J/U = 0.4 > 0.305$, the bosons do leave the initial site with a significant probability, it just takes thousands of timesteps before they do.

At this point, it is worth discussing why the bosons take so long to leave the initial site in this case. Compared to the hopping term, the repulsive interactions are stronger when $J/U = 0.4$ than when $J/U = 6$. Our experience suggests that this should lead to a "Coulomb explosion" and hence a faster diffusion, not a slower one. If the initial state of the system has all of the strongly repelling bosons on one site, the initial state has a large (expected) energy and the energy must be conserved. By this logic, the system should at later times, still be found in energetically unfavourable states. If our dynamical equations allowed energy to be *minimized* we would almost certainly see the spreading speed increase with J/U but we have not allowed the bosons to couple to their environment this way. They hold on to all of their energy, just like an electron placed in an f-orbital of a Hydrogen atom would stay in this

(a) Discrete non-linear Schrödinger equation

(b) Bose-Hubbard Hamiltonian

FIG. 17: Expected number versus time curves for 8 sites, 6 bosons, U = 1 and J = 0.4. Naïvely one might say that the bosons are localized on the initial site in both cases.
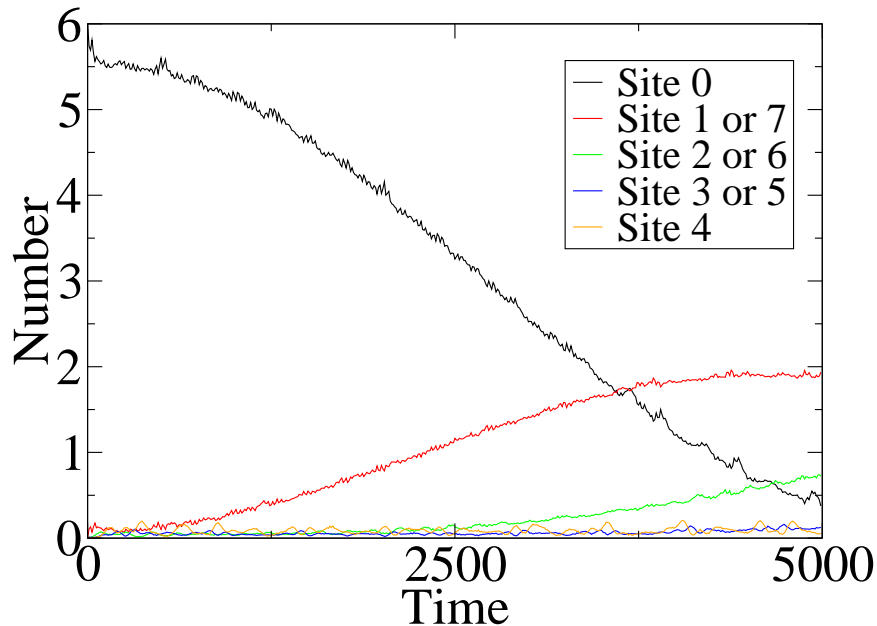

FIG. 18: Longer time dynamics for the Bose-Hubbard Hamiltonian with 8 sites, 6 bosons, U = 1 and J = 0.4. The same as the right of Fig. 17 except for the time scale.

eigenstate for all of time if not for decoherence. Classically it is still possible for charged particles to repel eachother and conserve energy. The repulsive potential energy is simply converted to kinetic energy when they fly away from eachohter at a constant speed without ever slowing down or changing direction. This also cannot happen with the Hamiltonian we are considering because our choice of using the canonical ensemble implicitly prevents bosons from leaving the lattice.

The discrete non-linear Schrödinger equation undergoes a qualitative change in localization as its parameters are adjusted. When there are 8 sites and 8 bosons, we can find a fairly narrow range of J/U values within which the dynamics change from that of delocalized particles to that of self-trapping, like we saw in Fig. 17. Again, the true

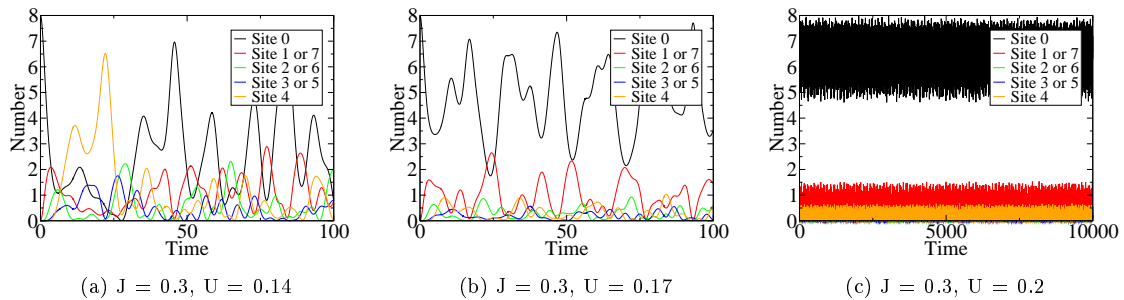(a) J = 0.3, U = 0.14   (b) J = 0.3, U = 0.17   (c) J = 0.3, U = 0.2

FIG. 19: Three solutions to the DNLS when modelling 8 sites and 8 bosons. The leftmost solution appears to show no localization while the rightmost solution is localized to the point where no other curve intersects the black curve. With this one, the time has been lengthened for emphasis.
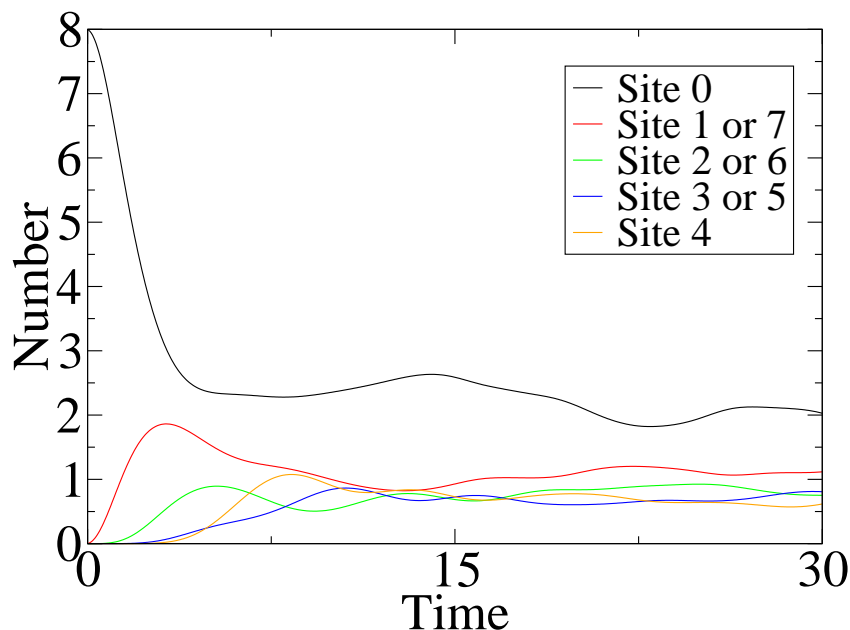

FIG. 20: The Bose-Hubbard Hamiltonian shows different behaviour from that of the DNLS when the parameters are 8 sites, 8 bosons, J = 0.3 and U = 0.17.

Bose-Hubbard model does not replicate this behaviour. The dynamics at the intermediate stage $J = 0.3, U = 0.17$ look different and it does not admit a solution where the bosons stay on site 0 indefinitely. This shift in the qualitative features of the DNLS is reminiscent of a phase transition, something we know to be consistent with the Bose-Hubbard model. One may ask questions like *do phase transitions allowed by the exact Hamiltonian have the same effect?* And *why haven't we seen a phase transition in the exact solutions yet?* We have not seen anything like a phase transition allowed by the phase diagram because the phase diagram assumes that the ensemble is in its ground state. In Appendix B, Mott lobes are defined in terms of ground state numbers of particles. The number eigenstate having all of the bosons on one site is certainly not the ground state and by conservation of energy, the system's state will not look like the ground state at a future time either. The DNLS is displaying transitions that are akin to phase transitions when we really should not be seeing them at all. Avoiding these "spurious transitions" is another reason not to use this equation. In a real experiment where energy exchange does occur (due to the nonzero $-TS$ term), the ground state and the well defined phases that come with it are eventually reached.

C. The Effects of Disorder

Whether disorder increases or decreases the rate at which wave packets spread is a question that has received much attention in the literature [5] [23] [24]. Before we answer it, we will see if the performance of the discrete non-linear Schrödinger equation improves when simulating a disordered lattice. A situation where the the two models disagreed previously was the 8 site, 6 boson $J/U = 6$ case that showed a steady-state.   If we simulate this again but with

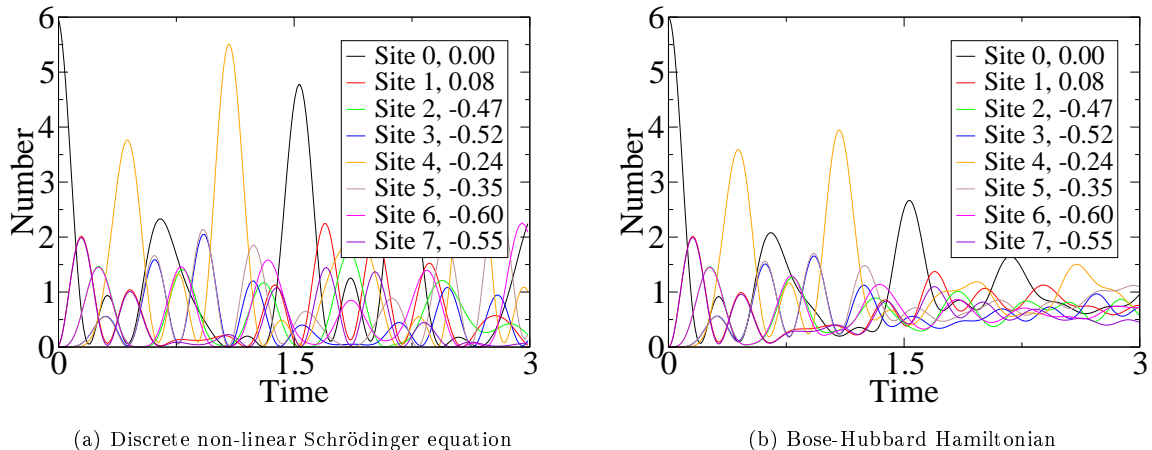(a) Discrete non-linear Schrödinger equation

(b) Bose-Hubbard Hamiltonian

FIG. 21: Expected number versus time curves for 8 sites, 6 bosons, U = 1, J = 6 and W = 1. A steady-state still appears for the exact solution only.
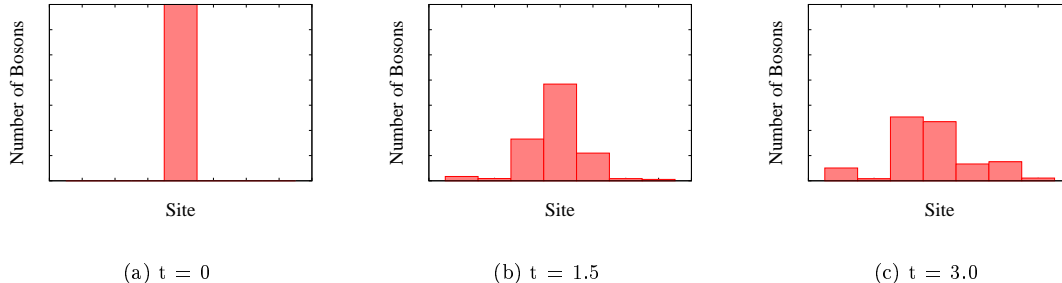

(a) t = 0

(b) t = 1.5

(c) t = 3.0

FIG. 22: Snapshots showing the spreading of 7 bosons that begin at the central site of a strongly disordered 7 site lattice.

$W = 1$ we see that a steady-state still appears for the Bose-Hubbard Hamiltonian but not the discrete non-linear Schrödinger equation. In fact if we compare Fig. 21 to Fig. 16, it is hard to notice major differences. All we can say is that as the potentials become "more random", the wavefunctions do as well. Indeed we see eight distinct curves in these plots because sites that are equidistant from the initial site are still fundamentally different because they have potential wells of different depths. This splitting manifests in Fig. 22 as well. If we remember that bosons in this situation had to circle the lattice several times before becoming evenly distributed, this is perhaps not the best choice of parameters for seeking measurable effects of the disorder strength. It would be more desirable to look for the effects of disorder during the initial stage of the expansion, before the boundary conditions become important.

Figure 23 is an example of this where after 100 timesteps, the system is still in the early stage of its expansion with no disorder. Something interesting happens when we set $W = 1$ and leave the other parameters unchanged. For two disorder complexions, one where the initial site has the highest energy and one where the initial site has the lowest energy, the two expansion rates are slower than the expansion rate in the absense of disorder. Also, the expansion rates for the two plots in Fig. 24 are very different from eachother. The plot that has bosons starting on the highest
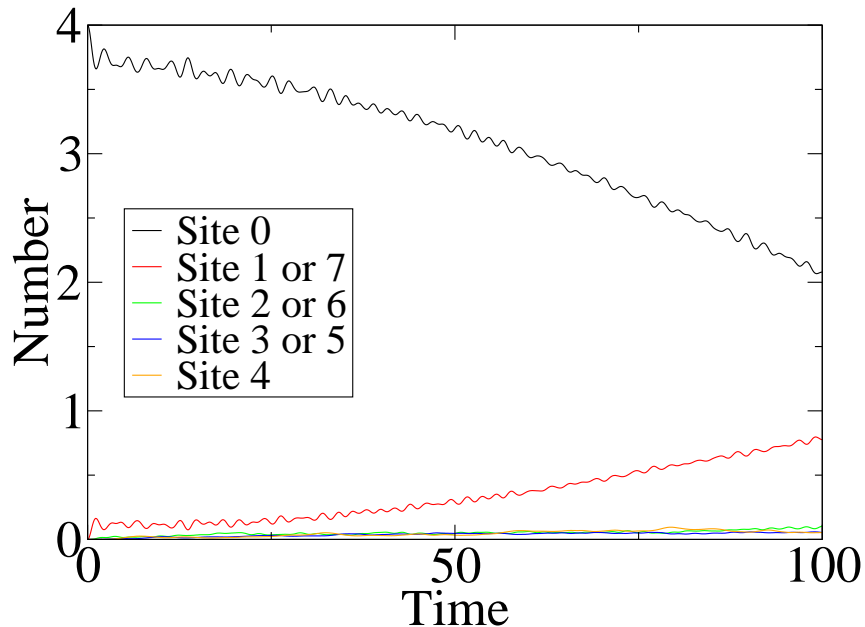
FIG. 23: Results of the Bose-Hubbard Hamiltonian with 8 sites, 4 bosons, U = 1, J = 0.3 and W = 0 showing an expansion that is slow but unimpeded.
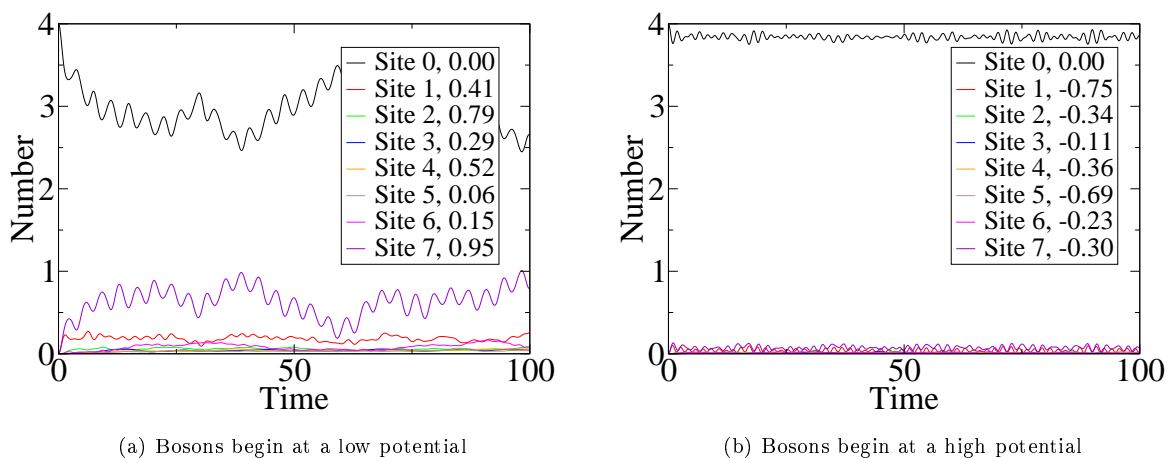
(a) Bosons begin at a low potential

(b) Bosons begin at a high potential

FIG. 24: Results of the Bose-Hubbard Hamiltonian with 8 sites, 4 bosons, U = 1, J = 0.3 and W = 1 for two complexions. These and other results show that the effect of disorder is to slow down diffusion because they all slow slower diffusion than the case that is ordered but otherwise identical - that of Fig. 23.

energy site shows a slower diffusion than the plot that has bosons starting on the lowest energy site. This is another example of energy conserving dynamics. Giving the bosons a high initial energy as in the right plot suppresses their probability of moving to a site that would give them a lower energy.

Something remarkable appears when we plot the rightmost figure over a much longer time interval. Figure 25 shows that over an interval thousands of times longer, the initial site still dominates the occupation of the bosons. This is the self-trapping that we failed to see in any of the ordered lattice plots. The system is in an excited state but this qualitative shift in the dynamics indicates that there may still be a phase transition taking place. Incidentally
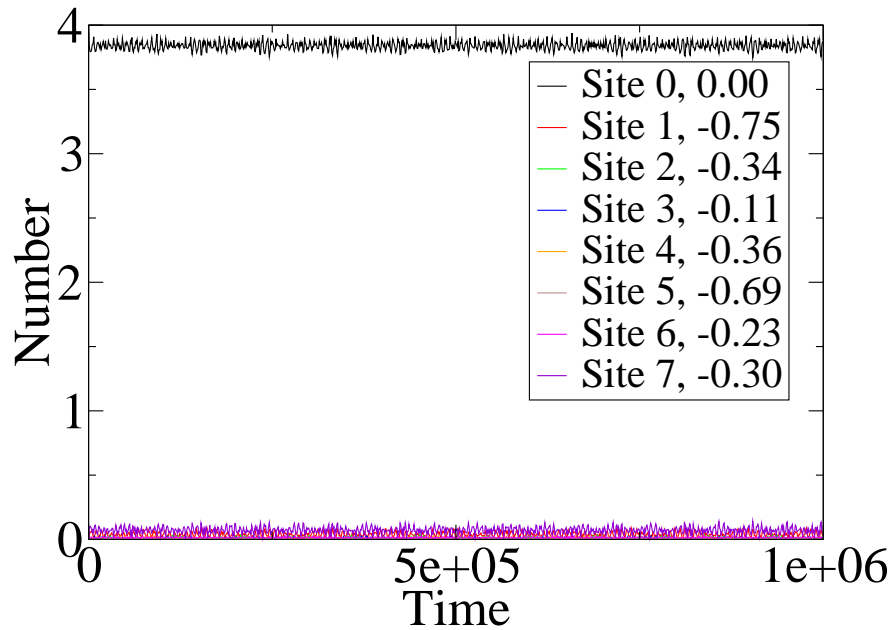
FIG. 25: A disorder complexion with 8 sites, 4 bosons, U = 1, J = 0.3 and W = 1 where to a good approximation, bosons never leave the initial site: the complexion in the right of Fig. 24 run over a longer time interval.
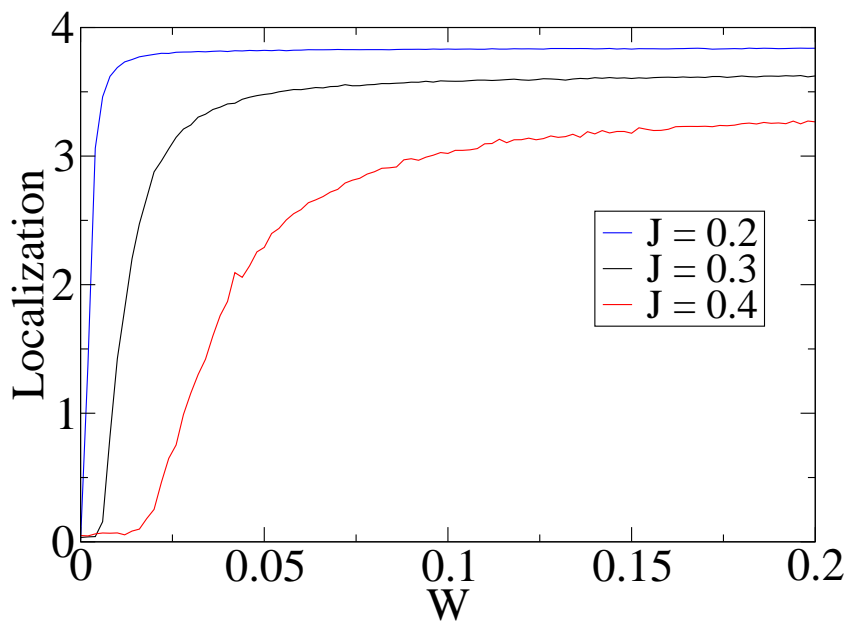
FIG. 26: A plot showing the disorder strength needed to achieve localization for 8 sites, 4 bosons, and U = 1. The vector containing the ϵ_j disorder values from ϵ_0 to ϵ_7 is $\boldsymbol{\epsilon} = (0, -0.75, -0.34, -0.11, -0.36, -0.69, -0.23, -0.3)W$. The black curve corresponds to J = 0.3 whose self-trapping was observed in Fig. 25. These transitions were also plotted for J = 0.2 and J = 0.4.

this suggests that the ordered version of the DNLS has features only present in the disordered version of the exact Hamiltonian [25].

It has been found in an experiment, a quasi-1D experiment even [5], that disordered lattices exhibit "subdiffusion", a slower diffusion process than would be observed without disorder. With a quantum mechanical system such as this, the act of performing the experiment may have an impact on the rate of diffusion. Using the Hamiltonian, we can predict the number of bosons that will be found on a given site but this is only an expectation. Measurements yielding different occupation numbers are still allowed and how probable these other measurements are is related to the variance. In a large experiment one may expect that at time t, most of the bosons are still on the initial site. If the variance is large, it is perfectly possible for one to find bosons on other sites in a measurement. However, if the variance is small, most measurements will cause the system to collapse back into the state $|1, 0, ..., 0\rangle$ and the diffusion will have to start again.

Something which may be able to tell us more about the self-trapping observed is a plot showing the disorder strength required to achieve it. To make this plot we quantify "how localized" a solution is by how far $\langle \hat{n}_0(t) \rangle$ drops below 4 in a span of one million timesteps. Occurring only in the presence of disorder, this phase is either a Bose glass or something very similar. Bose glasses are characterized by some of the on-site order parameters being zero and some being nonzero [17]. This gives rise to regions of localization interspersed among other regions of fluidity. This Bose glass that we see is one of the simplest cases. The region of localization is the initial site and it houses all of the bosons. It would be more interesting if there were a solution to the Bose-Hubbard model in which bosons diffused off of the initial site and subsequently regrouped at another site or family of sites and stayed there. These results have not been found, but no claim is being made that they do not exist.

VI. CONCLUSION

This study began with the open-ended goal of finding illuminating results in the expansion of a collection of bosons through an optical lattice. This led us to investigate many of the algorithms commonly used to simulate physical phenomena and incorporate some optimizations specific to the Bose-Hubbard Hamiltonian. By adjusting the parameters of the simulation carefully, in a way that was motivated by the physics of optical lattices, we were able to achieve our goal.

While the simulations of the Bose-Hubbard Hamiltonian we performed were very limited in size, we saw the beginnings of experimental phenomena such as the spreading of a wave packet and the transition to a self-trapping regime in the presence of disorder. Our efforts largely discredited the predictive power of the discrete non-linear Schrödinger equation for this problem even though it is widely used to find approximate dynamics for lattices governed by the Bose-Hubbard Hamiltonian [24] [26] [27] [28]. Our numerical results, having seen the best performance with the Runge-Kutta method due to Cash and Karp, challenge the popularity of using the Crank-Nicolson method to integrate the time-dependent Schrödinger equation. Indeed, articles have been written on why the Chebyshev expansion is a better choice as well [20] [6]. Even with the Chebyshev expansion at our disposal, the difficulty of solving large systems still grew exponentially, putting a simulation of realistic size beyond our reach.

There is a way to approximate the dynamical system with a small number of differential equations without using the discrete non-linear Schrödinger equation. It begins with a special form of the wavefunction called the Gutzwiller ansatz,

$$|\Psi(t)\rangle = \bigotimes_{m=1}^{N_s} |\Psi_m(t)\rangle \equiv \bigotimes_{m=1}^{N_s} \sum_{n=0}^{\infty} c_n^{(m)}(t) |n\rangle_m$$

which accommodates the same initial condition we have been using. Time-dependent perturbation theory is used to estimate the coefficients. Unlike the discrete non-linear Schrödinger equation which reduces the problem to N_s differential equations, this approach yields $N_s N_b$ differential equations - a number which still grows linearly as either the number of sites or the number of bosons is varied. The resulting equations are:

$$\frac{dc_p^{(m)}(t)}{dt} = \frac{i}{\hbar} \left[(\epsilon_m - \mu) p c_p^{(m)}(t) + \frac{U}{2} p(p-1) c_p^{(m)}(t) - J \sum_{\langle j,m \rangle} \left(\phi_j^* c_{p+1}^{(m)}(t) \sqrt{p+1} + \phi_j c_{p-1}^{(m)}(t) \sqrt{p} \right) + J \sum_{\langle j,l \rangle} \phi_j^* \phi_l c_p^{(m)} \right]$$

where ϕ_j is the order parameter defined by $\phi_j = \langle \Psi(t) | \hat{a}_j | \Psi(t) \rangle = \sum_{n=0}^{\infty} \sqrt{n} c_{n-1}^{(j)}(t) c_n^{(j)}(t)$. This system of equations, called the mean-field dynamical system is widely used to study the same type of expansion in an optical lattice and it has the advantage of working in the grand canonical ensemble [29]. Some authors have used it to study the initial

perturbation that drives the expansion by allowing the energies ϵ_j to depend on time [3]. Testing and using this dynamical system is left as an avenue for future work.

Condensed matter physics on optical lattices is a rich topic of which this thesis has only treated a small part. The results are consistent with accepted knowledge on the dynamics and phases of bosons in optical lattices [17] and the path through which one may try to extend them is clear. Lastly, the numerical methods that were repeatedly employed seem to have an affinity for this topic which we hope will be helpful to the physics community.

[1] J. Billy; V. Josse; Z. Zuo; A. Bernard; B. Hambrecht; P. Lugan; D. Clement; L. Sanchez-Palencia; P. Bouyer; A. Aspect. Direct observation of anderson localization of matter waves in a controlled disorder. *Nature*, 453:891–894, 2008.

[2] G. Pupillo; A. M. Rey; G. Brennen; C. J. Williams; C. W. Clark. Scalable quantum computation in systems with bose-hubbard dynamics. *Journal of Modern Optics*, 51(16):2395–2404, 2004.

[3] M. Snoek; W. Hofstetter. Two-dimensional dynamics of ultracold atoms in optical lattices. *Physical Review A*, 76(5), 2007.

[4] M. Larcher; F. Dalfovo; M. Modugno. Effects of interaction on the diffusion of atomic matter waves in one-dimensional quasiperiodic potentials. *Physical Review A*, 80(5), 2009.

[5] E. Lucioni; B. Deissler; L. Tanzi; G. Roati; M. Modugno; M. Zaccanti; M. Inguscio; G. Modugno. Observation of subdiffusion of a disordered interacting system. arXiv:1011.2362v1, November 2010.

[6] G. Schubert; H. Fehske. Dynamical aspects of two-dimensional quantum percolation. *Physical Review B*, 77(24), 2008.

[7] C. J. Foot. *Atomic Physics*, volume 7 of *Oxford Master Series in Physics*. Oxford University Press, 2004.

[8] J. J. Tollett. *A permanent magnet trap for cold atoms*. PhD thesis, Rice University, 1995.

[9] H. Pu. Hydrogen atom in external fields. http://www.owlnet.rice.edu/, 2010.

[10] R. V. Mishmash. Quantum many-body dynamics of ultracold bosons in one-dimensional optical lattices. Master's thesis, Colorado School of Mines, 2008.

[11] J. F. Annett; H. H. Wills. *Superconductivity, Superfluids, and Condensates*. Oxford University Press, 2002.

[12] A. J. M. Schmets; W. Montfrooij. Teaching superfluidity at the introductory level. arXiv:0804.3086v1, April 2008.

[13] J. M. Yeomans. *Statistical Mechanics of Phase Transitions*. Oxford University Press, 1992.

[14] S. Ejima; H. Fehske; F. Gebhard. Dynamic properties of the one-dimensional bose-hubbard model. *Europhysics Letters*, 93(3), 2011.

[15] D. Jaksch; C. Bruder; J. I. Cirac; C. W. Gardiner; P. Zoller. Cold bosonic atoms in optical lattices. *Physical Review Letters*, 81:3108–3111, 1998.

[16] J. Zakrzewski. Mean-field dynamics of the superfluid-insulator phase transition in a gas of ultracold atoms. *Physical Review A*, 71(4), 2005.

[17] V. I. Yukalov. Cold bosons in optical lattices. *Laser Physics*, 19(1):1–110, 2000.

[18] R. D. Mattuck. *A Guide to Feynmann Daigrams in the Many-Body Problem*. McGraw-Hill, 1976.

[19] M. Hiller. *Parametric Bose-Hubbard Hamiltonians: Quantum Dissipation, Irreversibility, and Pumping*. PhD thesis, University of Goettingen, 2007.

[20] H. Fehske; R. Schneider; A. Weisse. *Computational Many-Particle Physics*, volume 739 of *Lecture Notes in Physics*. Springer, 2008.

[21] K. Wu; H. Simon. Thick-restart lanczos method for large symmetric eigenvalue problems. *SIAM Journal on Matrix Analysis and Applications*, 22(2), 2000.

[22] W. H. Press; S. A. Teukolsky; W. T. Vetterling; B. P. Flannery. *Numerical Recipies in C*. Cambridge University Press, 2002.

[23] B. Deissler; M. Zaccanti; G. Roati; C. Errico; M. Fattori; M. Modugno; G. Modugno; M. Inguscio. Delocalization of a disordered bosonic system by repulsive interactions. *Nature Physics*, 6:354–358, 2010.

[24] G. Kopidakis; S. Komineas; S. Flach; S. Aubry. Absence of wave packet diffusion in disordered nonlinear systems. *Physical Review Letters*, 100(8), 2008.

[25] J. D. Bodyfelt; T. V. Laptyeva; G. Gligoric; D. O. Krimer; C. Skokos; S. Flach. Wave interactions in localizing media - a coin with many faces. arXiv:1102.4604v1, February 2011.

[26] S. Flach; D. O. Krimer; C. Skokos. Universal spreading of wave packets in disordered nonlinear systems. *Physical Review Letters*, 102(2), 2009.

[27] D. O. Krimer; S. Flach. Statistics of wave interactions in nonlinear disordered systems. *Physical Review E*, 82(4), 2010.

[28] C. Skokos; D. O. Krimer; S. Komineas; S. Flach. Delocalization of wave packets in disordered nonlinear chains. *Physical Review E*, 79(5), 2009.

[29] D. Jaksch; V. Venturi; J. I. Cirac; C. J. Williams; P. Zoller. Creation of a molecular condensate by dynamically melting a mott insulator. *Physical Review Letters*, 89, 2002.

[30] M. Plischke; B. Bergersen. *Equilibrium Statistical Physics*. World Scientific, 2006.

[31] R. K. P. Zia; E. F. Redish; S. R. McKay. Making sense of the legendre transform. *American Journal of Physics*, 77(7):614–622, 2008.

[32] H. Miller. Legendre transform in two-dimensional thermodynamics. http://math.mit.edu/ hrm/papers/thermo3.pdf, 2008.

[33] J. M. Grebowsky; J. P. McKelvey. Density-of-states function for the harmonic oscillator: A simple and direct approach. *American Journal of Physics*, 35(4):352–353, 1967.

[34] M. Le Bellac; F. Mortessagne; G. G. Batrouni. *Equilibrium and non-equilibrium statistical thermodynamics*. Cambridge University Press, 2004.

[35] P. F. Borges; H. Boschi-Filho; C. Farina. Quasiperiodic fields and bose-einstein condensation. arXiv:hep-th/9812045v1, December 1998.

[36] C. Hamaguchi. *Basic Semiconductor Physics*. Springer, 2010.

[37] B. Chen; X. Huang; S. Kou; Y. Zhang. Mott-hubbard transition of bosons in optical lattices with three-body interactions. *Physical Review A*, 78(4), 2008.

[38] J. K. Freericks; H. Monien. Phase diagram of the bose-hubbard model. *Europhysics Letters*, 26(7), 1994.

VII.   APPENDIX A: ENSEMBLES AND PHASES

One of the theoretical tools most frequently used to study quantum statistics is an ensemble of particles where certain parameters are fixed and others are allowed to vary. The constraints give rise to a free-energy whose minimization allows the direct calculation of the probabilities for an experimental outcome. The model we will discuss, called the grand canonical ensemble, offers a rich understanding of the differences between Fermi-Dirac statistics and Bose-Einstein statistics and its conditions are general enough to be satisfied in most experiments. We will begin by considering a system of fixed temperature T and chemical potential μ where the particles are indistinguishable. It will be in thermal equilibrium with a reservoir with which it may exchange particles and quanta of energy.

Of prime importance to us will be the set of single-particle energy states, denoted $\mathcal{S}$ and the set of allowed microstates for the entire system, $\mathcal{M}$. A microstate is defined by the number of particles in the first single-particle state, the number of particles in the second single-particle state *etc.* Formally, a microstate can be written $|n_1, n_2, \dots\rangle = \left|(n_r)_{r \in \mathcal{S}}\right\rangle$ where the number of particles is equal to:

$$N\left(\left|(n_r)_{r \in \mathcal{S}}\right\rangle\right) = \sum_{r \in \mathcal{S}} n_r \tag{31}$$

A.   Entropy Considerations

To gain a full understanding of the grand canonical ensemble, it will be necessary to develop a broad definition of entropy that can be used in an open system. This was done by Shannon in the 1940s and is discussed from a physics standpoint in [30]. The argument revolves around a large number of draws from a probability distribution $P(\lambda)$ that gives the probability of finding the ensemble in microstate λ. If we draw N times, we can expect state λ to occur $NP(\lambda)$ times in the limit of N being extremely large. It is now intuitive that the "amount of randomness" is directly related to the number of permutations or sequences that this experiment could probabilistically give us:

$$Z = \frac{N!}{\prod_{\lambda \in \mathcal{M}} (NP(\lambda))!} \tag{32}$$

An action that decreases this number can be viewed as a driving force pushing the system away from equilibrium. Any monotone function of Z will suffice as a measure of randomness. If we require the entropy S to be an extensive quantity we can determine the form of this monotone function. If we divide our system into two pieces as in the Gibbs paradox, we can take measurements from the first subsystem N_1 times and take measurements from the second subsystem N_2 times. For all of the Z_1 sequences coming from the first subsystem, there will be Z_2 sequences coming from the second, making the total number of sequences $Z = Z_1 Z_2$. The identity allowing us to solve for the entropy is now:

$$S(Z_1 Z_2) = S(Z_1) + S(Z_2) \tag{33}$$

It is first useful to write:

$$Z_1 \frac{dS(Z_1 Z_2)}{dZ_1} = Z_1 \frac{dS(Z_1 Z_2)}{d(Z_1 Z_2)} \frac{d(Z_1 Z_2)}{dZ_1} = Z_1 Z_2 \frac{dS(Z_1 Z_2)}{d(Z_1 Z_2)} \tag{34}$$

$$Z_2 \frac{dS(Z_1 Z_2)}{dZ_2} = Z_2 \frac{dS(Z_1 Z_2)}{d(Z_1 Z_2)} \frac{d(Z_1 Z_2)}{dZ_2} = Z_1 Z_2 \frac{dS(Z_1 Z_2)}{d(Z_1 Z_2)} \tag{35}$$

Then invoking Eq. (33), the left hand side of Eq. (34) becomes $Z_1 \frac{dS(Z_1)}{dZ_1}$ and the left hand side of Eq. (35) becomes $Z_2 \frac{dS(Z_2)}{dZ_2}$. Having established that

$$Z_1 \frac{dS(Z_1)}{dZ_1} = Z_2 \frac{dS(Z_2)}{dZ_2}$$

we can use separation of variables to set both sides equal to a constant K_1 and drop the subscripts.

$$
\begin{aligned}
Z \frac{dS(Z)}{dZ} &= K_1 \\
S(Z) &= K_1 \ln Z + K_2 \\
&= K_1 \ln Z \\
&= k_B \ln Z
\end{aligned}
$$

To make sure we express entropy in the right units, K_1 becomes Boltzmann's constant k_B. The boundary condition $S(1) = 0$ comes from the fact that there is no entropy if the system has only one sequence that it returns every time. What we have found is the entropy of the entire N-draw experiment. We can use extensivity again to find the entropy intrinsic to the probability distribution. We simply multiply the above by $\frac{1}{N}$ and take the limit as N approaches infinity.

$$
\begin{aligned}
S &= k_B \lim_{N\to\infty} \frac{1}{N} \ln \frac{N!}{\prod_{\lambda \in \mathcal{M}} (NP(\lambda))!} \\
&= k_B \lim_{N\to\infty} \frac{1}{N} \left[\ln N! - \sum_{\lambda \in \mathcal{M}} \ln (NP(\lambda))! \right]
\end{aligned}
$$

Stirling's approximation will not affect the limit so we can substitute $\ln N! = N \ln N - N$ to obtain:

$$
\begin{aligned}
S &= k_B \lim_{N\to\infty} \frac{1}{N} \left[N \ln N - N - N \sum_{\lambda \in \mathcal{M}} P(\lambda) \ln NP(\lambda) + N \sum_{\lambda \in \mathcal{M}} P(\lambda) \right] \\
&= k_B \lim_{N\to\infty} \left[\ln N - 1 - \sum_{\lambda \in \mathcal{M}} P(\lambda) \ln N - \sum_{\lambda \in \mathcal{M}} P(\lambda) \ln P(\lambda) + 1 \right] \\
&= -k_B \lim_{N\to\infty} \sum_{\lambda \in \mathcal{M}} P(\lambda) \ln P(\lambda) \\
&= -k_B \sum_{\lambda \in \mathcal{M}} P(\lambda) \ln P(\lambda)
\end{aligned}
\tag{36}
$$

B. Microstate and Macrostate Distributions

If we imagine an isolated system called the microcanonical ensemble, the ergodic hypothesis states that in a set of $\mathcal{W}$ microstates, each one has a probability of $\frac{1}{\mathcal{W}}$. Using this, Eq. (36) becomes:

$$
\begin{aligned}
S &= -k_B \frac{1}{\mathcal{W}} \sum_{\lambda=1}^{\mathcal{W}} \ln \frac{1}{\mathcal{W}} \\
&= k_B \ln \mathcal{W}
\end{aligned}
\tag{37}
$$

This is Boltzmann's formula and it is the same result we would get if we maximized S subject to the constraint that $\sum_{\lambda=1}^{\mathcal{W}} P(\lambda) = 1$ using Lagrange multipliers. To use entropy as a state-function, the natural independent variables are the number of particles N_p, the volume V and the internal energy U. If we consider a closed system rather than an isolated system, this amounts to allowing the internal energy to change. The system obtained by allowing the energy of a microcanonical ensemble to change is called a canonical ensemble. When U is changing it is helpful to transform the entropy into a state-function that no longer depends on U, instead depending on some conjugate variable called β which does not change. We will begin by defining a dimensionless entropy:

$$
\mathscr{S}(U, V, N_p) \equiv \frac{S(U, V, N_p)}{k_B}
$$

With this simplification made, the appropriate choice for β becomes $\frac{1}{k_B T}$ which does not change at constant temperature. The definition of temperature ensures that:

$$
\begin{aligned}
\frac{1}{T} &= \left(\frac{\partial S}{\partial U} \right)_{V, N_p} \\
\beta &= \left(\frac{\partial \mathscr{S}}{\partial U} \right)_{V, N_p}
\end{aligned}
\tag{38}
$$

This means our goal is come up with a function $\mathscr{A}(\beta, V, N_p)$ encoding the same information as $\mathscr{S}$ that doesn't depend on U but on the derivative of $\mathscr{S}$ with respect to U. This procedure, called a Legendre transformations, underlies some of the deepest symmetries in physics [31]. One must be careful when trying to accomplish this. It is certainly

possible to argue that the internal energy is a function of temperature and indeed this often matches our everyday experience. It is then simple to observe that $\mathscr{S}$ is already a function of β and we need only define our new function $\mathscr{A}$ by:

$$\mathscr{A}(\beta, V, N_{\mathrm{p}}) = \mathscr{S}(U(\beta), V, N_{\mathrm{p}})$$

This is not the Legendre transformation for a number of reasons, most notably a poor choice of independent variable for U. Saying that U is a function of β is as non-sensical as saying that the function:

$$f(x) = 2x$$

is a function of "2". While technically true, this leads us nowhere. The correct Legendre transformation is:

$$\mathscr{A}(\beta, V, N_{\mathrm{p}}) = \beta U - \mathscr{S}(U, V, N_{\mathrm{p}})$$
$$\mathscr{A} + \mathscr{S} = \beta U \tag{39}$$

The reason behind this definition lies in a symmetry that is the reverse of Eq. (38).

$$\begin{aligned}
\left(\frac{\partial \mathscr{A}}{\partial \beta}\right)_{V, N_{\mathrm{p}}} &= \beta \left(\frac{\partial U}{\partial \beta}\right)_{V, N_{\mathrm{p}}} + U - \left(\frac{\partial \mathscr{S}}{\partial \beta}\right)_{V, N_{\mathrm{p}}} \\
&= \left(\frac{\partial \mathscr{S}}{\partial U}\right)_{V, N_{\mathrm{p}}} \left(\frac{\partial U}{\partial \beta}\right)_{V, N_{\mathrm{p}}} + U - \left(\frac{\partial \mathscr{S}}{\partial \beta}\right)_{V, N_{\mathrm{p}}} \\
&= U \tag{40}
\end{aligned}$$

Dividing both sides of Eq. (39) by β tells us that the equation for a free-energy $A \equiv \frac{\mathscr{A}}{\beta}$ is:

$$A = U - TS$$

A is called the Helmholtz free-energy. It inherits an important property from the Legendre transformation, which can be derived using Eq. (38) and Eq. (40).

$$\begin{aligned}
\left(\frac{\partial \mathscr{A}}{\partial V}\right)_{\beta, N_{\mathrm{p}}} &= \beta \left(\frac{\partial U}{\partial V}\right)_{N_{\mathrm{p}}} - \left[\left(\frac{\partial \mathscr{S}}{\partial V}\right)_{U, N_{\mathrm{p}}} + \left(\frac{\partial \mathscr{S}}{\partial U}\right)_{V, N_{\mathrm{p}}} \left(\frac{\partial U}{\partial V}\right)_{N_{\mathrm{p}}}\right] \\
&= \beta \left(\frac{\partial U}{\partial V}\right)_{N_{\mathrm{p}}} - \left(\frac{\partial \mathscr{S}}{\partial V}\right)_{U, N_{\mathrm{p}}} - \beta \left(\frac{\partial U}{\partial V}\right)_{N_{\mathrm{p}}} \\
&= - \left(\frac{\partial \mathscr{S}}{\partial V}\right)_{U, N_{\mathrm{p}}} \tag{41}
\end{aligned}$$

Equation (41) makes it clear why we went to this kind of trouble. Minimizing the Hemlholtz free-energy A at constant temperature is the same as maximizing the entropy S at constant energy [32]. Differentiating A is all we need to do to predict the behaviour of the canonical ensemble. Allowing the number of particles to change, moves us from a closed system to an open one and this is the grand canonical ensemble. To find a suitable thermodynamic potential in this case, the hard work has already been done. We will take the dimensionless Helmholtz free-energy $\mathscr{A}(\beta, V, N_{\mathrm{p}})$ and perform a Legendre transformation one more time to transfer its dependence from the variable N_{p} to the fixed chemical potential μ.

$$\mathscr{B}(\beta, V, \mu') = \mu' N_{\mathrm{p}} - \mathscr{A}(\beta, V, N_{\mathrm{p}})$$
$$\mathscr{B} + \mathscr{A} = \mu' N_{\mathrm{p}}$$

where we have defined $\mu' \equiv \beta\mu$. The relation analagous to Eq. (41) holds:

$$\left(\frac{\partial \mathscr{B}}{\partial V}\right)_{\beta, \mu'} = - \left(\frac{\partial \mathscr{A}}{\partial V}\right)_{\beta, N_{\mathrm{p}}}$$

We define the grand potential $B \equiv -\frac{\mathscr{B}}{\beta}$ so that:

$$B = A - \mu N_{\mathrm{p}} = U - TS - \mu N_{\mathrm{p}} \tag{42}$$

and minimization of B with a fixed chemical potential corresponds to minimizing A with a fixed number of particles. When considering the statistical mechanics of the grand canonical ensemble it is important to recognize that the U term in Eq. (42) is really an ensemble average of different energies and similarly for N_p.

$$U = \sum_{\lambda \in \mathcal{M}} P(\lambda) E(\lambda) \tag{43}$$

$$N_\mathrm{p} = \sum_{\lambda \in \mathcal{M}} P(\lambda) N(\lambda) \tag{44}$$

Equations (43), (44) and (36) make it possible to expand the grand potential into a sum that depends on the microstates.

$$B = \sum_{\lambda \in \mathcal{M}} P(\lambda) E(\lambda) + \frac{1}{\beta} \sum_{\lambda \in \mathcal{M}} P(\lambda) \ln P(\lambda) - \mu \sum_{\lambda \in \mathcal{M}} P(\lambda) N(\lambda)$$

Differentiating with respect to all of the microstate energies and particle numbers, we get:

$$\frac{\partial B}{\partial E(\lambda)} = 0$$

$$\frac{\partial}{\partial E(\lambda)} \left(P(\lambda) E(\lambda) \right) + \frac{1}{\beta} \frac{\partial}{\partial E(\lambda)} \left(P(\lambda) \ln P(\lambda) \right) - \mu N(\lambda) \frac{\partial P(\lambda)}{\partial E(\lambda)} = 0$$

$$\left(\frac{\partial P(\lambda)}{\partial E(\lambda)} + P(\lambda) \right) + \frac{1}{\beta} \left(\frac{\partial P(\lambda)}{\partial E(\lambda)} + \frac{\partial P(\lambda)}{\partial E(\lambda)} \ln P(\lambda) \right) - \mu N(\lambda) \frac{\partial P(\lambda)}{\partial E(\lambda)} = 0$$

$$P(\lambda) + \frac{\partial P(\lambda)}{\partial E(\lambda)} \left(1 + \frac{1}{\beta} + \frac{1}{\beta} \ln P(\lambda) - \mu N(\lambda) \right) = 0 \tag{45}$$

and

$$\frac{\partial B}{\partial N(\lambda)} = 0$$

$$E(\lambda) \frac{\partial P(\lambda)}{\partial N(\lambda)} + \frac{1}{\beta} \frac{\partial}{\partial N(\lambda)} \left(P(\lambda) \ln P(\lambda) \right) - \mu \frac{\partial}{\partial N(\lambda)} \left(P(\lambda) N(\lambda) \right) = 0$$

$$E(\lambda) \frac{\partial P(\lambda)}{\partial N(\lambda)} + \frac{1}{\beta} \left(\frac{\partial P(\lambda)}{\partial N(\lambda)} + \frac{\partial P(\lambda)}{\partial N(\lambda)} \ln P(\lambda) \right) - \mu \left(\frac{\partial P(\lambda)}{\partial N(\lambda)} + P(\lambda) \right) = 0$$

$$P(\lambda) + \frac{\partial P(\lambda)}{\partial N(\lambda)} \left(1 - \frac{1}{\mu \beta} - \frac{1}{\mu \beta} \ln P(\lambda) - \frac{E(\lambda)}{\mu} \right) = 0 \tag{46}$$

We are now on the verge of finding the optimal probability for the grand canonical ensemble as we did for the microcanonical ensemble (and could have easily done for the canonical ensemble). It is a matter of substitution to show that

$$P(\lambda) = \frac{1}{\mathcal{Z}} e^{-\beta(E(\lambda) - \mu N(\lambda))} \tag{47}$$

is the simultaneous solution to the coupled equations Eq. (45) and Eq. (46). The exponential $e^{-\beta(E(\lambda) - \mu N(\lambda))}$ is called the Gibbs factor and $\mathcal{Z}$ is called the grand partition function. Normalization demands that the grand partition function be the sum of all possible Gibbs factors:

$$\mathcal{Z} = \sum_{\lambda \in \mathcal{M}} e^{-\beta(E(\lambda) - \mu N(\lambda))} \tag{48}$$

While Eq. (47) gives the probability distribution of microstates, the energy and the number of particles do not usually form a complete set of quantum numbers. To make comparisons with real experiments, we must take the probability that the system is in *a particular* microstate with energy E and particle number N_p and turn it into a probability that the system is in *any* microstate with energy E and particle number N_p. Since all of these states are equally likely, we multiply by how many there are.

$$P(E, N_\mathrm{p}) = \frac{\mathcal{W}(E, N_\mathrm{p})}{\mathcal{Z}} e^{-\beta(E - \mu N_\mathrm{p})} \tag{49}$$

The multiplicity $\mathcal{W}(E, N_\mathrm{p})$ is the number of microstates that belong to the (E, N_p) macrostate. For an isolated system it is equal to the total number $\mathcal{W}$. A related quantity is the density of states which has units of inverse energy. It is defined such that $g(E, N_\mathrm{p})\mathrm{d}E\mathrm{d}N_\mathrm{p}$ gives the number of microstates with energies and particle numbers lying in the intervals $[E, E + \mathrm{d}E]$ and $[N_\mathrm{p}, N_\mathrm{p} + \mathrm{d}N_\mathrm{p}]$ respectively. In the limit of very closely spaced energy levels, g is simply the continuous version of $\mathcal{W}$. However, the way in which it is typically used is different. For systems with a fixed particle number, the N_p dependence is redundant and the notation for the density of states becomes $g(E)$. When discussing single-particle energy states as opposed to ensemble microstates, the number of particles is 1 and the density of states is written $g(\epsilon)$. In this case, the density of states tells us whether or not the Hamiltonian gives rise to degeneracy.

C.   Finding the Density of States

Experiments done on condensed matter, in particular the optical lattice experiments mentioned in this thesis, trap atoms in a confining potential. The prototypical example of a confining potential is the infinite square well. If the gas of atoms to be cooled is placed inside a perfect $L_x \times L_y \times L_z$ box like this, the energy states will depend on three positive integers n_x, n_y and n_z according to:

$$\epsilon_{\infty\mathrm{SW}} = \frac{\hbar^2 \pi^2}{2m} \left[\left(\frac{n_x}{L_x} \right)^2 + \left(\frac{n_y}{L_y} \right)^2 + \left(\frac{n_z}{L_z} \right)^2 \right] \tag{50}$$

If L_x, L_y and L_z are all different, there is no reason for any of the energies in Eq. (50) to be degenerate. However, there are several combinations of the three quantum numbers whose corresponding energies are *approximately* degenerate and the number of such combinations grows with increasing energy. This statement can be made rigorous by calculating the density of states for this potential. When determining how many states exist at energy ϵ, we calculate $\#(\epsilon)$, the cumulative number of states *up to* energy ϵ and differentiate. Consider for a moment what would happen if we had $L_x = L_y = L_z \equiv L$. We would be able to factor $\frac{1}{L^2}$ out of Eq. (50), leaving ϵ proportional to n^2; the squared magnitude of the (n_x, n_y, n_z) vector. The number of states, $\#(n)$ would be the number of unit cubes inside the positive octant of a sphere of radius n (approximately the volume of such a sphere) and we would have a simple relationship between n and ϵ. The well we are dealing with is anisotropic so instead we define $K_x \equiv \frac{n_x}{L_x}$, $K_y \equiv \frac{n_y}{L_y}$ and

(a) Isotropic   (b) Anisotropic
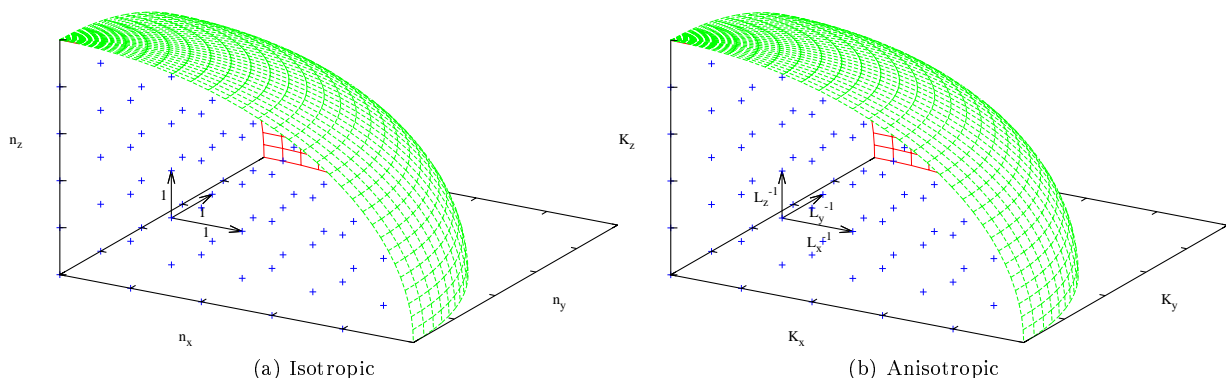
FIG. 27: The left figure shows the standard diagram used to find the density of states for particles in a box. The right figure shows a straightforward generalization applicable to boxes with different side-lengths.

$K_z \equiv \frac{n_z}{L_z}$. Now a similar relationship exists between K and ϵ. We just need to remember that we count the number of cubes not by taking the volume of the sphere, but by taking the volume of the sphere and dividing by the volume

of each cube which is $\frac{1}{L_x L_y L_z}$.

$$
\begin{aligned}
g_{\infty \text{SW}}(\epsilon) &= \frac{\mathrm{d}\#}{\mathrm{d}\epsilon} \\
&= \frac{\mathrm{d}}{\mathrm{d}\epsilon} \frac{\frac{1}{8}\frac{4}{3}\pi K^3}{\frac{1}{L_x L_y L_z}} \\
&= \frac{\mathrm{d}}{\mathrm{d}\epsilon} \frac{\pi L_x L_y L_z}{6} \left(\sqrt{\frac{2m\epsilon}{\hbar^2 \pi^2}} \right)^3 \\
&= \frac{L_x L_y L_z (2m)^{\frac{3}{2}}}{4\hbar^3 \pi^2} \sqrt{\epsilon}
\end{aligned}
\tag{51}
$$

As discussed in the introduction, a more realistic confining potential is that of a simple harmonic oscillator. The energy levels induced by this potential such as the one appearing in a magnetic trap are:

$$
\epsilon_{\text{SHO}} = \hbar \left[\omega_x \left(q_x + \frac{1}{2} \right) + \omega_y \left(q_y + \frac{1}{2} \right) + \omega_z \left(q_z + \frac{1}{2} \right) \right]
\tag{52}
$$

where q_x, q_y and q_z are non-negative integers. Shifting the energy will not affect the number of states so for our purposes we may use $\hbar \left(\omega_x q_x + \omega_y q_y + \omega_z q_z \right)$. If we define $Q_x \equiv \omega_x q_x$, $Q_y \equiv \omega_y q_y$ and $Q_z \equiv \omega_z q_z$, we can find the density of states for the simple harmonic oscillator using a formula similar to the one employed for the infinite square well. Our shape however, needs to change because Q is equal to $Q_x + Q_y + Q_z$ whereas K was equal to $\sqrt{K_x^2 + K_y^2 + K_z^2}$. The locus of points in three-space with a constant "sum of co-ordinates" norm is an octahedron where Q is equal to the side-length times $\frac{1}{\sqrt{2}}$. Using the pyramid formula:

$$
\begin{aligned}
g_{\text{SHO}}(\epsilon) &= \frac{\mathrm{d}\#}{\mathrm{d}\epsilon} \\
&= \frac{\mathrm{d}}{\mathrm{d}\epsilon} \frac{\frac{1}{8}\frac{2}{3} \left(\sqrt{2}Q \right)^2 Q}{\omega_x \omega_y \omega_z} \\
&= \frac{\mathrm{d}}{\mathrm{d}\epsilon} \frac{\epsilon^3}{6\hbar^3 \omega_x \omega_y \omega_z} \\
&= \frac{\epsilon^2}{2\hbar^3 \omega_x \omega_y \omega_z}
\end{aligned}
\tag{53}
$$

The frequencies, of course, depend on the mass of the atom.
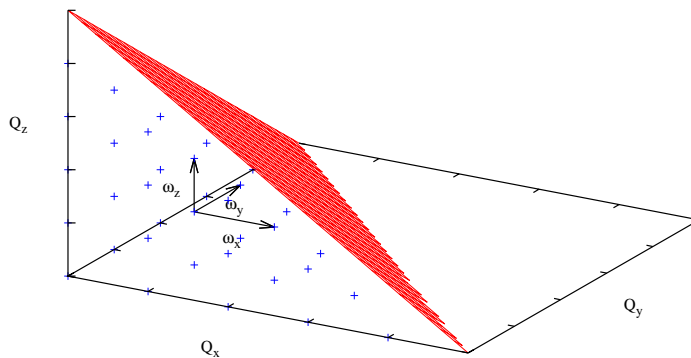
FIG. 28: State counting diagram for a simple harmonic oscillator. Amazingly, an entire paper has been written on how to use an octahedron to count simple harmonic oscillator states [33].

D. Non-interacting Distribution Functions

For this section, it is worth defining a toy set C which contains allowed occupation numbers. All this means is that $C = \{0, 1\}$ when we are dealing with fermions and $C = \{0, 1, 2, \dots\}$ when we are dealing with bosons. We already

have an expression for the number of particles in a given microstate, Eq. (31). If we further suppose that none of the particles interact, we can write down an expression for the total energy of the microstate in terms of the single particle energies, ϵ_r.

$$E\left(\left|(n_r)_{r \in \mathcal{S}}\right\rangle\right) = \sum_{r \in \mathcal{S}} n_r \epsilon_r \tag{54}$$

A single Gibbs factor can now be written as a product:

$$
\begin{aligned}
e^{-\beta\left(E\left(\left|(n_r)_{r \in \mathcal{S}}\right\rangle\right) - \mu N\left(\left|(n_r)_{r \in \mathcal{S}}\right\rangle\right)\right)} &= e^{-\beta\left(\sum_{r \in \mathcal{S}} n_r \epsilon_r - \mu \sum_{r \in \mathcal{S}} n_r\right)} \\
&= \prod_{r \in \mathcal{S}} e^{-\beta n_r(\epsilon_r - \mu)}
\end{aligned}
$$

which means the partition function looks like:

$$\mathcal{Z} = \sum_{\left|(n_r)_{r \in \mathcal{S}}\right\rangle \in C \times \cdots \times C} \prod_{r \in \mathcal{S}} e^{-\beta n_r(\epsilon_r - \mu)}$$

A useful fact, sometimes presented as being obvious [34], is that the sum and the product in this expression can be interchanged. We will prove this by inducting on the number of single-particle states $|\mathcal{S}|$. If there is only one single-particle state and $|\mathcal{S}| = 1$, the following is clearly true:

$$\prod_{r=1}^{|\mathcal{S}|} \sum_{n_r \in C} e^{-\beta n_r(\epsilon_r - \mu)} = \sum_{\left|(n_r)_{r=1}^{|\mathcal{S}|}\right\rangle \in C \times \cdots \times C} \prod_{r=1}^{|\mathcal{S}|} e^{-\beta n_r(\epsilon_r - \mu)}$$

If we assume that it holds for a general $|\mathcal{S}|$,

$$
\begin{aligned}
\prod_{r=1}^{|\mathcal{S}|+1} \sum_{n_r \in C} e^{-\beta n_r(\epsilon_r - \mu)} &= \left(\prod_{r=1}^{|\mathcal{S}|} \sum_{n_r \in C} e^{-\beta n_r(\epsilon_r - \mu)}\right) \sum_{n_{|\mathcal{S}|+1} \in C} e^{-\beta n_{|\mathcal{S}|+1}\left(\epsilon_{|\mathcal{S}|+1} - \mu\right)} \\
&= \left[\sum_{\left|(n_r)_{r=1}^{|\mathcal{S}|}\right\rangle \in C \times \cdots \times C} \prod_{r=1}^{|\mathcal{S}|} e^{-\beta n_r(\epsilon_r - \mu)}\right] \sum_{n_{|\mathcal{S}|+1} \in C} e^{-\beta n_{|\mathcal{S}|+1}\left(\epsilon_{|\mathcal{S}|+1} - \mu\right)} \\
&= \sum_{\left|(n_r)_{r=1}^{|\mathcal{S}|}\right\rangle \in C \times \cdots \times C} \left[\left(\prod_{r=1}^{|\mathcal{S}|} e^{-\beta n_r(\epsilon_r - \mu)}\right)\left(\sum_{n_{|\mathcal{S}|+1} \in C} e^{-\beta n_{|\mathcal{S}|+1}\left(\epsilon_{|\mathcal{S}|+1} - \mu\right)}\right)\right] \\
&= \sum_{\left|(n_r)_{r=1}^{|\mathcal{S}|}\right\rangle \in C \times \cdots \times C} \sum_{n_{|\mathcal{S}|+1} \in C} \prod_{r=1}^{|\mathcal{S}|+1} e^{-\beta n_r(\epsilon_r - \mu)} \\
&= \sum_{\left|(n_r)_{r=1}^{|\mathcal{S}|+1}\right\rangle \in C \times \cdots \times C} \prod_{r=1}^{|\mathcal{S}|+1} e^{-\beta n_r(\epsilon_r - \mu)}
\end{aligned}
$$

This factorization is often written $\mathcal{Z} = \prod_{r \in \mathcal{S}} \mathcal{Z}_r$. An interesting property of this is that each factor $\mathcal{Z}_r$ would in fact be the grand partition function of the entire system if the only single-particle state it had were state r. In this type of grand canonical ensemble, the occupation number of the unique single-particle state would define the microstate. The probability mass function Eq. (47) derived earlier, immediately tells us how this occupation number would be distributed.

$$P(n_r) = \frac{1}{\mathcal{Z}_r} e^{-\beta(E(n_r) - \mu N(n_r))} = \frac{1}{\mathcal{Z}_r} e^{-\beta n_r(\epsilon_r - \mu)}$$

To check our logic, one may ask the following question: *when an individual state is part of a multi-state grand canonical ensemble, can it be treated as its own "sub-grand canonical ensemble"?* We will show that the answer is *yes* by using

the probability above to reconstruct the microstate distribution for the multi-state ensemble.

$$
\begin{aligned}
P\left(\left|(n_r)_{r\in\mathcal{S}}\right\rangle\right) &= \prod_{r\in\mathcal{S}} P(n_r) \\
&= \prod_{r\in\mathcal{S}} \frac{1}{\mathcal{Z}_r} e^{-\beta n_r(\epsilon_r-\mu)} \\
&= \frac{1}{\mathcal{Z}} \prod_{r\in\mathcal{S}} e^{-\beta n_r(\epsilon_r-\mu)} \\
&= \frac{1}{\mathcal{Z}} e^{-\beta\left(\sum_{r\in\mathcal{S}} n_r\epsilon_r - \mu\sum_{r\in\mathcal{S}} n_r\right)} \\
&= \frac{1}{\mathcal{Z}} e^{-\beta\left(E\left(\left|(n_r)_{r\in\mathcal{S}}\right\rangle\right) - \mu N\left(\left|(n_r)_{r\in\mathcal{S}}\right\rangle\right)\right)}
\end{aligned}
$$

With this sanity check out of the way, we can write down the sub-grand partition function for Fermi-Dirac statistics.

$$
\mathcal{Z}_{r,\mathrm{FD}} = \sum_{n_r\in C} e^{-\beta n_r(\epsilon_r-\mu)} = \sum_{n_r=0}^{1} e^{-\beta n_r(\epsilon_r-\mu)} = 1 + e^{-\beta(\epsilon_r-\mu)} \tag{55}
$$

Using the formula for an infinite geometric series, we may do the same for Bose-Einstein statistics.

$$
\mathcal{Z}_{r,\mathrm{BE}} = \sum_{n_r\in C} e^{-\beta n_r(\epsilon_r-\mu)} = \sum_{n_r=0}^{\infty} e^{-\beta n_r(\epsilon_r-\mu)} = \frac{1}{1 - e^{-\beta(\epsilon_r-\mu)}} \tag{56}
$$

These formulae allow us to derive the celebrated statistical occupation functions for non-interacting fermions and bosons. The average number of particles observed in single-particle state r will be: $\sum_{n_r\in C} n_r P(n_r)$. Therefore, Eq. (55) tells us that the Fermi-function is:

$$
\begin{aligned}
f_{\mathrm{FD}}(\epsilon_r) &= \frac{1}{1 + e^{-\beta(\epsilon_r-\mu)}} \sum_{n_r=0}^{1} n_r e^{-\beta n_r(\epsilon_r-\mu)} \\
&= \frac{1}{e^{\beta(\epsilon_r-\mu)} + 1}
\end{aligned} \tag{57}
$$

With a little more work, Eq. (56) gives us the Bose-function:

$$
\begin{aligned}
f_{\mathrm{BE}}(\epsilon_r) &= \left(1 - e^{-\beta(\epsilon_r-\mu)}\right) \sum_{n_r=0}^{\infty} n_r e^{-\beta n_r(\epsilon_r-\mu)} \\
&= \left(1 - e^{-\beta(\epsilon_r-\mu)}\right) \sum_{n_r=1}^{\infty} n_r e^{-\beta n_r(\epsilon_r-\mu)} \\
&= \left(1 - e^{-\beta(\epsilon_r-\mu)}\right) e^{-\beta(\epsilon_r-\mu)} \frac{\partial}{\partial e^{-\beta(\epsilon_r-\mu)}} \sum_{n_r=0}^{\infty} e^{-\beta n_r(\epsilon_r-\mu)} \\
&= \left(1 - e^{-\beta(\epsilon_r-\mu)}\right) \frac{e^{-\beta(\epsilon_r-\mu)}}{\left(1 - e^{-\beta(\epsilon_r-\mu)}\right)^2} \\
&= \frac{1}{e^{\beta(\epsilon_r-\mu)} - 1}
\end{aligned} \tag{58}
$$

An understanding of quantum statistics allows us to derive an important phase transition that was predicted seventy years before it was observed [10]. We now discuss Bose-Einstein condensation - a phase transition that is ubiquitous in the study of low-temperature bosons.

E.   Bose-Einstein Condensation

At high temperatures in the thermodynamic limit, non-interacting bosons are spread out over a large range of energy states given by the distribution function Eq. (58). However, the form of this function guarantees the existence

of a temperature below which all particles added to the system condense into the ground-state. Assuming that the zero-point energy for single particles is very small, we can write the number of bosons in excited states as:

$$N_{\text{b, ex}} = \int_0^\infty f_{\text{BE}}(\epsilon) g(\epsilon) \mathrm{d}\epsilon$$

where $g(\epsilon)$ is the density of states. The phenomenon of BEC can be most clearly illustrated when the density of states is given by $g(\epsilon) = C\epsilon^{\alpha-1}$ where C and α are constants. Knowing that μ is always negative or zero, it is clear that the area under f_{BE} is maximized for $\mu = 0$. Therefore the highest number of bosons that can possibly reside in excited states is:

$$\begin{aligned}
N_{\text{b, ex}} &= \int_0^\infty \frac{C\epsilon^{\alpha-1}}{e^{\beta\epsilon}-1}\mathrm{d}\epsilon \\
&= C\beta^{-\alpha}\int_0^\infty \frac{(\beta\epsilon)^{\alpha-1}}{e^{\beta\epsilon}-1}\mathrm{d}(\beta\epsilon) \\
&= C\,(k_{\text{B}}T)^\alpha \int_0^\infty \frac{x^{\alpha-1}}{e^x-1}\mathrm{d}x
\end{aligned} \tag{59}$$

This number is not infinite as it would be for fermions that admit a positive chemical potential. Instead we see that for an arbitrary number of bosons N_{b}, it is not always possible to make $N_{\text{b, ex}} = N_{\text{b}}$ and we are left with a macroscopic fraction of bosons that must be in the ground-state. A certain temperature T_{c} will put us exactly on the cusp of this phase transition where $N_{\text{b, ex}} = N_{\text{b}}$ and $\mu = 0$. For this reason it is called the critical temperature. Substituting this into Eq. (59) gives us:

$$T_{\text{c}} = \frac{N_{\text{b}}^{\frac{1}{\alpha}}}{k_{\text{B}}\left[C\int_0^\infty \frac{x^{\alpha-1}}{e^x-1}\mathrm{d}x\right]^{\frac{1}{\alpha}}} \tag{60}$$

It is difficult to guess a simpler form for the integral in Eq. (60) but we will show that there is one by starting with the answer. The gamma function is recognizable as:

$$\begin{aligned}
\Gamma(\alpha) &= \int_0^\infty t^{\alpha-1}e^{-t}\mathrm{d}t \\
\Gamma(\alpha)\frac{1}{n^\alpha} &= \int_0^\infty x^{\alpha-1}e^{-nx}\mathrm{d}x
\end{aligned} \tag{61}$$

where we have let $t = nx$. Now if we sum both sides of Eq. (61) from $n = 1$ to ∞ we can express this integral in terms of the Riemann zeta function.

$$\begin{aligned}
\Gamma(\alpha)\sum_{n=1}^\infty \frac{1}{n^\alpha} &= \int_0^\infty x^{\alpha-1}\left(\sum_{n=0}^\infty e^{-nx} - 1\right)\mathrm{d}x \\
\Gamma(\alpha)\zeta(\alpha) &= \int_0^\infty x^{\alpha-1}\left(\frac{1}{1-e^{-x}} - 1\right)\mathrm{d}x \\
&= \int_0^\infty x^{\alpha-1}\left(\frac{e^x}{e^x-1} - \frac{e^x-1}{e^x-1}\right)\mathrm{d}x \\
&= \int_0^\infty \frac{x^{\alpha-1}}{e^x-1}\mathrm{d}x
\end{aligned}$$

It now follows that:

$$T_{\text{c}} = \frac{N_{\text{b}}^{\frac{1}{\alpha}}}{k_{\text{B}}\left[C\Gamma(\alpha)\zeta(\alpha)\right]^{\frac{1}{\alpha}}} \tag{62}$$

Recalling our expression Eq. (53) for $g_{\text{SHO}}(\epsilon)$, the density of states for a simple harmonic oscillator, we can substitute $\alpha = 3$ and $C = \frac{1}{2\hbar^3\omega_x\omega_y\omega_z}$ to obtain:

$$k_{\text{B}}T_{\text{c, SHO}} = \hbar\left[\frac{N_{\text{b}}\omega_x\omega_y\omega_z}{\zeta(3)}\right]^{\frac{1}{3}}$$

Onset of Bose-Einstein Condensation
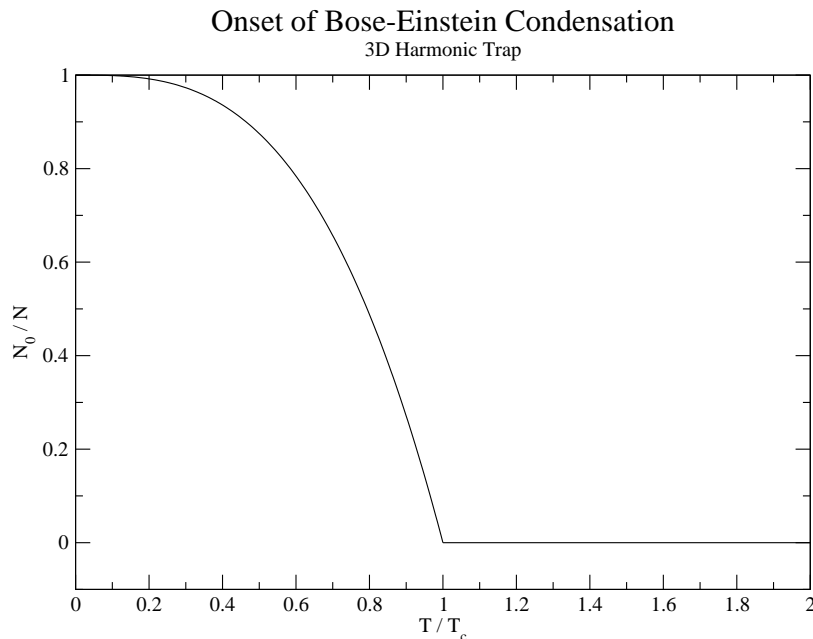### 3D Harmonic Trap

FIG. 29: The condensate fraction of a non-interacting Bose gas.

Another common result is the critical temperature for a Bose-Einstein condensate trapped in a box. In this case we use Eq. (51) for $g_{\text{SW}}(\epsilon)$ in which $\alpha = \frac{3}{2}$ and $C = \frac{L_x L_y L_z (2m)^{\frac{3}{2}}}{4\hbar^3 \pi^2}$.

$$
\begin{aligned}
k_{\text{B}} T_{\text{c, SW}} &= \left[\frac{4 N_{\text{b}} \hbar^3 \pi^2}{(2m)^{\frac{3}{2}} L_x L_y L_z \Gamma(\frac{3}{2}) \zeta(\frac{3}{2})} \right]^{\frac{2}{3}} \\
&= \frac{\hbar^2}{2m} \left[\frac{N_{\text{b}}}{L_x L_y L_z} \frac{4\pi^2}{\frac{1}{2}\sqrt{\pi}\zeta(\frac{3}{2})} \right]^{\frac{2}{3}} \\
&= \frac{2\pi\hbar^2}{m} \left[\frac{N_{\text{b}}/L_x L_y L_z}{\zeta(\frac{3}{2})} \right]^{\frac{2}{3}} \\
&= \frac{2\pi\hbar^2}{m} \left[\frac{\rho}{\zeta(\frac{3}{2})} \right]^{\frac{2}{3}}
\end{aligned}
$$

It is important to note that we were only able to derive the critical temperatures this way because the values of α we considered were all greater than one. For values less than or equal to one, the Gamma and Riemann zeta functions diverge and we see that Bose-Einstein condensation does not take place. This happens when we trap atoms in a 1D harmonic trap ($\alpha = 1$) or a 1D or 2D box ($\alpha = \frac{1}{2}$ and $\alpha = 1$ respectively). This has led to apparent confusion in the literature over the number of dimensions required for BEC [10] [35].

One can see the phase transition when plotting the fraction of particles in the ground-state. This fraction is zero when $T \geq T_{\text{c}}$. The formula Eq. (59) yields:

$$
N_{\text{b, ex}} = C \Gamma(\alpha) \zeta(\alpha) (k_{\text{B}} T)^{\alpha} = N_{\text{b}} \left(\frac{T}{T_{\text{c}}} \right)^{\alpha}
$$

Rearranging this, the condensate fraction is simply $1 - \left(\frac{T}{T_{\text{c}}} \right)^{\alpha}$.

F.   Landau Theory

It is unmistakable that the critical temperature is a point where the behaviour of the Bose gas under consideration changes qualitatively. Having seen this, we are now poised to answer the question of what a phase transition really is and what features of BEC we can expect to see in other phase transitions. The condensate fraction in Fig. 29 is clearly not differentiable at the critical temperature. Bose-Einstein condensation thus arises from a singularity which is part of the definition of a phase transition [13]. Also, the appearance of a full condensate was not sudden - the condensate appeared gradually after the critical temperature was passed. Transitions that behave this way are called "second-order phase transitions" or "continuous phase transitions". The more sudden ones are "first-order phase transitions". Landau developed a general theory of phase transitions that unifies many of these ideas. The main object that the theory deals with is a free-energy F that determines the statistical mechanics of a given state of matter. This could be the Helmholtz free-energy of the canonical ensemble or the grand potential of the grand canonical ensemble. The challenge is parametrizing F so that its Taylor series, $F(\phi) = \sum_{n=0}^{\infty} \frac{\mathrm{d}F}{\mathrm{d}\phi}(0)\phi^n$ has only even terms. This is essentially the condition that F be an even function of ϕ. Once this is done, we may write:

$$F(\phi) = b_0 + b_2\phi^2 + b_4\phi^4 + O(\phi^6)$$

If the terms in this series decay quickly enough, we can neglect the possibility of b_4 being negative. Afterall, a function that approaches $-\infty$ is not a physical energy. With this assumption, the value of ϕ that minimizes the free-energy depends heavily on the sign of b_2. Figure 30 shows that the critical conditions of a phase transition are those that
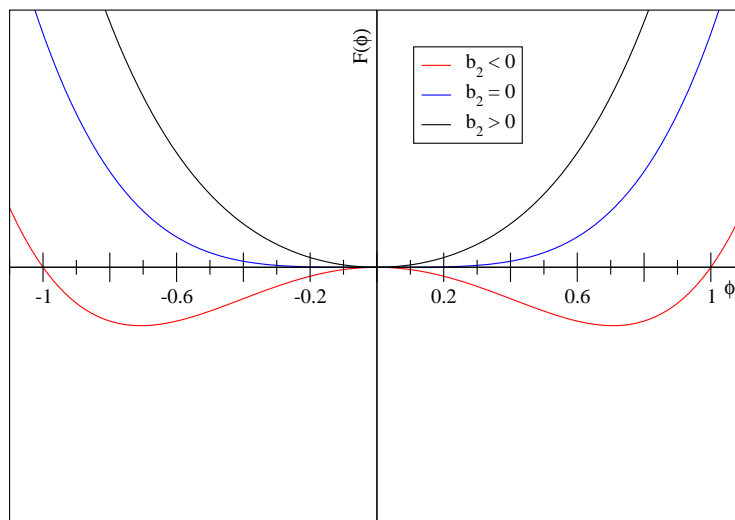
FIG. 30: Three example Landau free-energies. One of them is minimized for nonzero values of ϕ.

set b_2 equal to zero. Above this, $\phi = 0$ minimizes the free-energy but below this, the value of ϕ preferred by nature begins to take on a nonzero value. The red curve for example is minimized for $\phi \approx \pm 0.7$. This is exactly the behaviour of the condensate fraction in Fig. 29. When analyzing a phase transition this way, we call ϕ the order parameter. The condensate fraction happens to be the order parameter of a Bose-Einstein condensate. Other examples of order parameters are the polarization of a ferroelectric material or the difference between the atmospheric pressure and the vapour pressure for a liquid evaporating to a gas [13]. In the next section, we will derive another phase transition using a tool that is indispensable in the analysis of the Bose-Hubbard Hamiltonian - mean-field theory.

G. Magnetic Systems

Magnetism is well understood in materials composed of particles that all have a spin degree of freedom, *e.g.* iron atoms. In a ferromagnetic material, these spins align spontaneously and begin to exert the macroscopic forces we are familiar with. A paramagnetic material, on the other hand, only displays this kind of order in the presence of an external magnetic field due to thermal fluctuations that are strong enough to randomize the spins otherwise. Spin-$\frac{1}{2}$ electrons are described by a quantum number m_{s} that can take on the values $\pm\frac{1}{2}$. The energies of these "spin-up" and "spin-down" electrons are given by:

$$\epsilon = -g_{\mathrm{s}} m_{\mathrm{s}} \mu_{\mathrm{B}} B$$

where $g_{\mathrm{s}} \approx 2$ is the gyromagnetic ratio, μ_{B} is the Bohr magneton and B is the strength of the magnetic field. This allows us to define the partition function of a paramagnetic system:

$$\begin{aligned} Z &= e^{-\beta \epsilon_\uparrow} + e^{-\beta \epsilon_\downarrow} \\ &= e^{\beta \mu_{\mathrm{B}} B} + e^{-\beta \mu_{\mathrm{B}} B} \end{aligned} \tag{63}$$

The magnetization of the system is simply the net spin times the magnetization for which each spin is responsible. This can be expressed using Eq. (63) and probabilities:

$$\begin{aligned} M &= \mu_{\mathrm{B}} \left(N_\uparrow - N_\downarrow \right) \\ &= N_{\mathrm{p}} \mu_{\mathrm{B}} \left(P_\uparrow - P_\downarrow \right) \\ &= N_{\mathrm{p}} \mu_{\mathrm{B}} \left(\frac{e^{\beta \mu_{\mathrm{B}} B}}{Z} - \frac{e^{-\beta \mu_{\mathrm{B}} B}}{Z} \right) \\ &= N_{\mathrm{p}} \mu_{\mathrm{B}} \frac{e^{\beta \mu_{\mathrm{B}} B} - e^{-\beta \mu_{\mathrm{B}} B}}{e^{\beta \mu_{\mathrm{B}} B} + e^{-\beta \mu_{\mathrm{B}} B}} \\ &= N_{\mathrm{p}} \mu_{\mathrm{B}} \tanh \left(\beta \mu_{\mathrm{B}} B \right) \end{aligned} \tag{64}$$

Now we are left with the question of how a material actually becomes ferromagnetic. This process always involves spins interacting with eachother. A successful model describing this assumes that the spins only have two orientations relative to eachother and that they only interact with their nearest-neighbours. If two neighbouring spins s_j and s_l are anti-aligned, the system incurs an energy cost taken to be $\frac{J}{2}$ so that the underlying Hamiltonian becomes:

$$\mathcal{H}_{\mathrm{I}} = -J \sum_{\langle j,l \rangle_\bullet} s_j s_l \tag{65}$$

This model, called the Ising model, has been solved exactly in one and two dimensions but we will use mean-field theoy to obtain an approximate solution valid in any number of dimensions. Instead of saying that spin l always feels spin j where j changes as we change l, the mean-field approximation states that for any l, spin l feels a spin that represents the averaged contribution of all other spins. It is felt once for each nearest-neighbour that the spin has. In any reasonable crystal, this number is always the same. We call it the co-ordination number, z. The force on spin l which we demand to be $\mu_{\mathrm{B}} B_{\mathrm{I}}$ is:

$$-\frac{\partial \mathcal{H}_{\mathrm{I}}}{\partial s_l} = J \sum_{\langle j,l \rangle} s_j \equiv \mu_{\mathrm{B}} B_{\mathrm{I}} \tag{66}$$

Since Eq. (66) is now linear in the s_j terms, we can average it to get:

$$\mu_{\mathrm{B}} B_{\mathrm{I}} = J z m$$

were m is the average alignment of the spins. Inserting this into Eq. (64) it appears as though we still need more information to find a phase transition. However, we can again use the fact that m is the average alignment to rewrite the left hand side of the equation.

$$\begin{aligned} M &= N_{\mathrm{p}} \mu_{\mathrm{B}} \tanh \left(\beta J z m \right) \\ N_{\mathrm{p}} \mu_{\mathrm{B}} \left(P_\uparrow - P_\downarrow \right) &= N_{\mathrm{p}} \mu_{\mathrm{B}} \tanh \left(\beta J z m \right) \\ m &= \tanh \left(\beta J z m \right) \end{aligned}$$
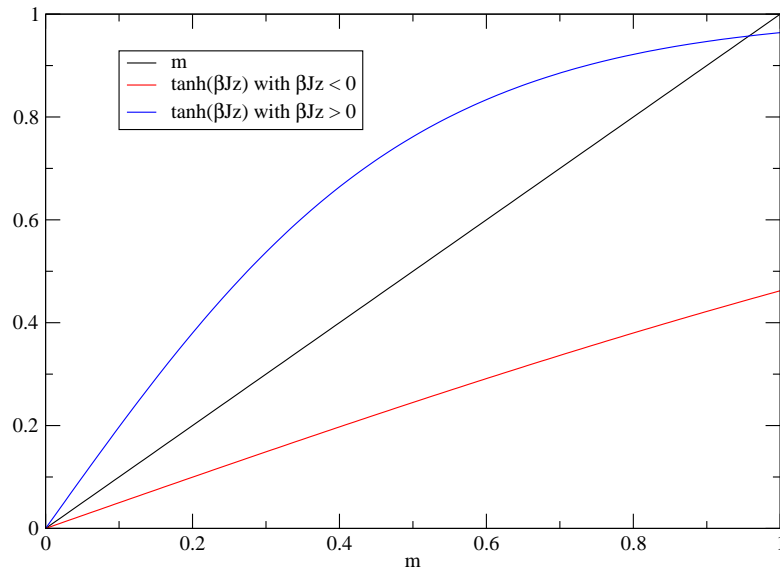
FIG. 31: A simple procedure to check that solutions to a transcendental equation exist.

This is a transcendental equation for m. Using the fact that the slope of the tanh function is less than one, nonzero solutions exist precisely when $\beta Jz > 1$. This gives two nice results. One is that the average alignment m is an order parameter for the paramagnetic to ferromagnetic phase transition. The other is that the critical temperature is $T_{\mathrm{c}} = \frac{Jz}{k_{\mathrm{B}}}$. Since this is a transcendental equation, we cannot plot m as a function of T directly but we will do the next best thing - plot T as a function of m and switch the axes.

$$
\begin{aligned}
m &= \tanh\left(\beta Jzm\right) \\
m &= \frac{e^{\beta Jzm} - e^{-\beta Jzm}}{e^{\beta Jzm} + e^{-\beta Jzm}} \\
m &= \frac{e^{2\beta Jzm} - 1}{e^{2\beta Jzm} + 1} \\
me^{2\beta Jzm} + m &= e^{2\beta Jzm} - 1 \\
e^{2\beta Jzm}(1 - m) &= 1 + m \\
\beta &= \frac{1}{2Jzm} \ln\left(\frac{1+m}{1-m}\right) \\
T &= \frac{2Jzm}{k_{\mathrm{B}} \ln\left(\frac{1+m}{1-m}\right)} \\
&= \frac{2T_{\mathrm{c}} m}{\ln\left(\frac{1+m}{1-m}\right)}
\end{aligned}
\tag{67}
$$

In the zero temperature limit, we see that m tends toward one of two values: $+1$ (all spin-up) and -1 (all spin-down). This chaotically chosen value corresponds to one of the two possible order parameters minimizing the free-energy in Fig. 30.
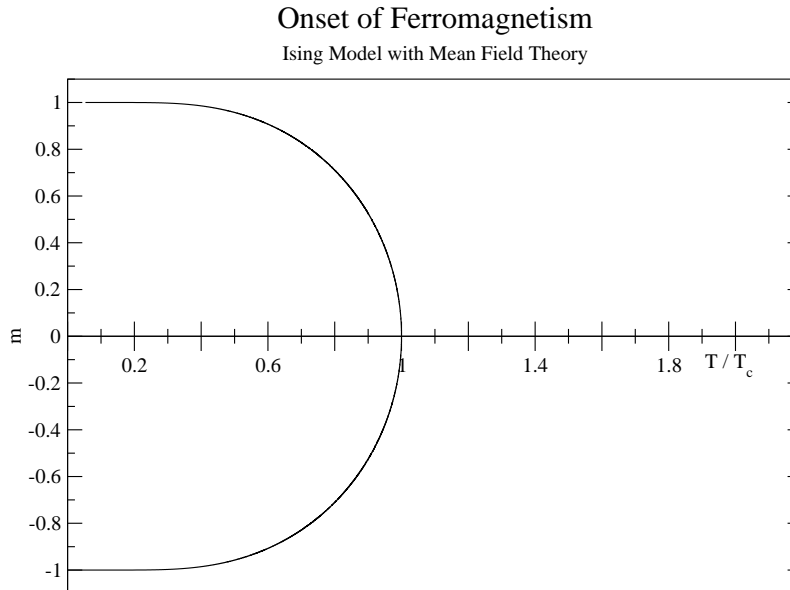
Onset of Ferromagnetism

Ising Model with Mean Field Theory

FIG. 32: Spontaneous alignment of spins in the paramagnetic to ferromagnetic phase transition.

VIII. APPENDIX B: THE MANY BOSON PROBLEM

Our goal in this section is to derive the Bose-Hubbard Hamiltonian and some of its interesting features starting with Eq. (9). The derivation that offers the richest understanding will develop second quantization along the way - a powerful tool that is repeatedly used in many-body physics.

A. Second Quantization

For a Hamiltonian that has a (possibly infinite) set of energy eigenstates $\mathcal{S}$, we have already used the ket $\left|(n_r)_{r \in \mathcal{S}}\right\rangle$ to say that there are n_1 particles in state 1, n_2 particles in state 2 and so on. Just as the set of energy eigenstates forms an orthonormal basis for the Hilbert space that can possibly describe one particle, these occupation number eigenstates form an orthonormal basis for the extended Hilbert space that can accomodate any number of particles. Without the occupaction number formalism (in what [18] calls *stone-age notation*), the state having one particle in the first energy eigenstate and two in the second looks like:

$$\sqrt{\frac{1}{3}} \left[\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\psi_2(\mathbf{r}_3) + \psi_1(\mathbf{r}_2)\psi_2(\mathbf{r}_1)\psi_2(\mathbf{r}_3) + \psi_1(\mathbf{r}_3)\psi_2(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\right]$$

With our new notation, the same state is simply $|1, 2, 0, 0, \ldots\rangle$ or better yet, $\frac{1}{2}\hat{c}_1^\dagger \hat{c}_2^\dagger \hat{c}_2^\dagger |0\rangle$, three creation operators acting on the vacuum state. Some explanation is in order for what these operators actually do. When anhilation and creation operators with index j act on an occupation number eigenstate, we require the resulting state to be proportional to the occupation number eigenstate whose j^th entry differs by one.

$$\hat{c}_j \left|\ldots, n_{j-1}, n_j, n_{j+1}, \ldots\right\rangle = \sqrt{n_j} \left|\ldots, n_{j-1}, n_j - 1, n_{j+1}, \ldots\right\rangle$$

$$\hat{c}_j^\dagger \left|\ldots, n_{j-1}, n_j, n_{j+1}, \ldots\right\rangle = \sqrt{n_j + 1} \left|\ldots, n_{j-1}, n_j + 1, n_{j+1}, \ldots\right\rangle$$

With the prefactors chosen this way, the operators obey the canonical commutation relations:

$$[\hat{c}_j, \hat{c}_l] = 0, \left[\hat{c}_j^\dagger, \hat{c}_l^\dagger\right] = 0, \left[\hat{c}_j, \hat{c}_l^\dagger\right] = \delta_{j,l} \tag{68}$$

The first two are clear because the operators act on the different indicies independently. To prove the third one, we write:

$$
\begin{aligned}
\left[\hat{c}_j, \hat{c}_j^{\dagger}\right] |\ldots, n_{j-1}, n_j, n_{j+1}, \ldots\rangle &= \hat{c}_j, \hat{c}_j^{\dagger} |\ldots, n_{j-1}, n_j, n_{j+1}, \ldots\rangle - \hat{c}_j^{\dagger}\hat{c}_j |\ldots, n_{j-1}, n_j, n_{j+1}, \ldots\rangle \\
&= (n_j + 1) |\ldots, n_{j-1}, n_j, n_{j+1}, \ldots\rangle - n_j |\ldots, n_{j-1}, n_j, n_{j+1}, \ldots\rangle \\
&= |\ldots, n_{j-1}, n_j, n_{j+1}, \ldots\rangle
\end{aligned}
\tag{69}
$$

Something we see from this is that $\hat{c}_j^{\dagger}, \hat{c}_j$ gives the number of particles in energy eigenstate j. It is the jth "number operator" while $\hat{c}_j, \hat{c}_j^{\dagger}$ is the jth "number plus one operator". This allows us to express a non-interacting Hamiltonian using creation and anhilation operators:

$$
\hat{\mathcal{H}}^{\circ} = \sum_j \epsilon_j \hat{c}_j^{\dagger}, \hat{c}_j
\tag{70}
$$

A Hamiltonian is, by definition, diagonal in the basis of energy eigenstates. However, this particular Hamiltonian must be diagonal in a basis where the individual particles are all in various single-particle energy eigenstates. Indeed it is diagonal in the occupation number basis because

$$
\hat{\mathcal{H}}^{\circ} |n_1, n_2, n_3, \ldots\rangle = \sum_j \epsilon_j n_j |n_1, n_2, n_3, \ldots\rangle
$$

returns the correct energy eigenvalue for the state $|n_1, n_2, n_3, \ldots\rangle$ made up of non-interacting particles. A more challenging task is writing a two-body operator, such as an interaction potential, in the algebra of creation and anhilation operators [18]. Consider such an operator $\hat{\Theta}$. What $\hat{\Theta}$ does to the state $|n_1, n_2, n_3, \ldots\rangle$ can be very general:

$$
\hat{\Theta} |n_1, n_2, n_3, \ldots\rangle = a_1 |n_1', n_2', n_3', \ldots\rangle + a_2 |n_1'', n_2'', n_3'', \ldots\rangle + \ldots
$$

However, in the primed and double-primed, *etc* states, we expect two of the entries to be one less than the corresponding entries in $|n_1, n_2, n_3, \ldots\rangle$ and two of the entries to be one more. If three numbers decreased and three other numbers increased, for example, this would tell us that we really had a three-body operator not a two-body one. This heuristic tells us that every occupation number eigenstate in this potentially very large expansion is just $\hat{c}_j^{\dagger}\hat{c}_l^{\dagger}\hat{c}_m\hat{c}_p |n_1, n_2, n_3, \ldots\rangle$ for some j, l, m, p. With:

$$
\hat{\Theta} = \frac{1}{4} \sum_{j,l,m,p} A_{j,l,m,p} \hat{c}_j^{\dagger}\hat{c}_l^{\dagger}\hat{c}_m\hat{c}_p
$$

established, we will show that the $A_{j,l,m,p}$ are really the matrix elements of the operator $\hat{\Theta}$ in the basis of doubly occupied states. Let $|1,1\rangle_{j',l'}$ denote the occupation number state with a 1 in the j' position, a 1 in the l' position and zeroes elsewhere. If $j' \neq l'$ and $m' \neq p'$ then:

$$
\begin{aligned}
{}_{j',l'}\langle 1,1| \hat{\Theta} |1,1\rangle_{m',p'} &= \frac{1}{4} \sum_{j,l,m,p} A_{j,l,m,p} \, {}_{j',l'}\langle 1,1| \hat{c}_j^{\dagger}\hat{c}_l^{\dagger}\hat{c}_m\hat{c}_p |1,1\rangle_{m',p'} \\
&= \frac{1}{4} \sum_{j,l,m,p} A_{j,l,m,p} \, {}_{j',l'}\langle 1,1| \hat{c}_j^{\dagger}\hat{c}_l^{\dagger}\delta_{\{m,p\},\{m',p'\}} |0\rangle \\
&= \frac{1}{4} \sum_{j,l,m,p} A_{j,l,m,p}\delta_{\{j,l\},\{j',l'\}}\delta_{\{m,p\},\{m',p'\}}
\end{aligned}
$$

The first Kronecker delta will evaulate to 1 if $(j,l) = (j',l')$ or if $(j,l) = (l',j')$. The second one will evaulate to 1 if $(m,p) = (m',p')$ or if $(m,p) = (p',m')$. Therefore the four surviving terms tell us that:

$$
\begin{aligned}
{}_{j',l'}\langle 1,1| \hat{\Theta} |1,1\rangle_{m',p'} &= \frac{1}{4} \left[A_{j',l',m',p'} + A_{j',l',p',m'} + A_{l',j',m',p'} + A_{l',j',p',m'}\right] \\
&= A_{j',l',m',p'}
\end{aligned}
$$

where we have used the fact that the indices can be freely permuted for bosons. Matrix elements for two-body operators are more commonly expressed in the basis of product states that have not been symmetrized. For our

operator $\hat{\Theta}$, we will denote such a matrix element by $\Theta_{j,l,m,p}$. We will relate these two types of matrix elements in position space:

$$
\begin{aligned}
{}_{j,l}\langle 1,1|\hat{\Theta}|1,1\rangle_{m,p} &= \iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{1}{\sqrt{2}}\left(\psi_j(\mathbf{r})\psi_l(\mathbf{r}') + \psi_j(\mathbf{r}')\psi_l(\mathbf{r})\right)^* \hat{\Theta}\left(\mathbf{r},\mathbf{r}'\right) \frac{1}{\sqrt{2}}\left(\psi_m(\mathbf{r})\psi_p(\mathbf{r}') + \psi_m(\mathbf{r}')\psi_p(\mathbf{r})\right) \mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r} \\
&= \frac{1}{2}\left[\iint_{\mathbb{R}^3\times\mathbb{R}^3} \psi_j^*(\mathbf{r})\psi_l^*(\mathbf{r}')\hat{\Theta}\left(\mathbf{r},\mathbf{r}'\right)\psi_m(\mathbf{r})\psi_p(\mathbf{r}')\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r} + \iint_{\mathbb{R}^3\times\mathbb{R}^3} \psi_j^*(\mathbf{r})\psi_l^*(\mathbf{r}')\hat{\Theta}\left(\mathbf{r},\mathbf{r}'\right)\psi_m(\mathbf{r}')\psi_p(\mathbf{r})\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r} \right. \\
&\quad \left. + \iint_{\mathbb{R}^3\times\mathbb{R}^3} \psi_j^*(\mathbf{r}')\psi_l^*(\mathbf{r})\hat{\Theta}\left(\mathbf{r},\mathbf{r}'\right)\psi_m(\mathbf{r})\psi_p(\mathbf{r}')\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r} + \iint_{\mathbb{R}^3\times\mathbb{R}^3} \psi_j^*(\mathbf{r}')\psi_l^*(\mathbf{r})\hat{\Theta}\left(\mathbf{r},\mathbf{r}'\right)\psi_m(\mathbf{r}')\psi_p(\mathbf{r})\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r} \right] \\
&= \frac{1}{2}\left[\Theta_{l,j,m,p} + \Theta_{l,j,p,m} + \Theta_{j,l,m,p} + \Theta_{j,l,p,m}\right] \\
&= 2\Theta_{j,l,m,p}
\end{aligned}
$$

Now that we know the two types of matrix elements are related by a factor of 2, a widely used expression for the general two-body operator follows:

$$
\hat{\Theta} = \frac{1}{2}\sum_{j,l,m,p} \Theta_{j,l,m,p}\hat{c}_j^{\dagger}\hat{c}_l^{\dagger}\hat{c}_m\hat{c}_p \tag{71}
$$

Two things about Eq. (71) are short-sighted. The first is that the derivations above only treats matrix elements in which $j \neq l$ and $m \neq p$. The three remaining cases add no new ideas to the proof and they do not change the result. The second is that we have carelessly permuted indicies relying on the fact that we only wish to describe bosons. If we were to derive Eq. (71) with fermions in mind as well, the result actually would change [18]. It would be:

$$
\hat{\Theta} = \frac{1}{2}\sum_{j,l,m,p} \Theta_{l,j,m,p}\hat{c}_j^{\dagger}\hat{c}_l^{\dagger}\hat{c}_m\hat{c}_p \tag{72}
$$

where the j and the l have been switched.

B. Changing Operators

Second quantization is not limited to energy eigenstates. We can define creation and anhilation operators for the eigenstates of the position operator which can be thought of as Dirac delta functions. An abuse of notation will be the occupation number eigenstate:

$$
|n_{\mathbf{r}_1}, n_{\mathbf{r}_2}, n_{\mathbf{r}_3}, \dots\rangle
$$

where we refer to the number of particles at the "first position", the number of particles in the "second position" and so on. Since this is continuously indexed, we should really write

$$
\left|(n_{\mathbf{r}})_{\mathbf{r}\in\mathbb{R}^3}\right\rangle
$$

We have already done most of the work that is needed to define "field operators" that destroy and create a particle at the point $\mathbf{r}$: $\hat{\Psi}(\mathbf{r})$ and $\hat{\Psi}^{\dagger}(\mathbf{r})$. The only rule we need to make is that $\hat{\Psi}^{\dagger}(\mathbf{r})$ acting on the vacuum produces a position eigenstate with a coefficient of 1 [18]. All other properties of these operators will be inherited from the formalism developed earlier. Let us write a position eigenstate in the basis of single-particle energy eigenstates:

$$
\delta(\mathbf{r} - \mathbf{r}_j) = \sum_l A_{j,l}\psi_l(\mathbf{r})
$$

Transforming both sides of the equation into the occupation number formalism, we have:

$$
\begin{aligned}
|1\rangle_{\mathbf{r}_j} &= \sum_l A_{j,l}|1\rangle_l \\
\hat{\Psi}^{\dagger}(\mathbf{r}_j)|0\rangle &= \sum_l A_{j,l}\hat{c}_l^{\dagger}|0\rangle \\
\hat{\Psi}^{\dagger}(\mathbf{r}_j) &= \sum_l A_{j,l}\hat{c}_l^{\dagger}
\end{aligned}
$$

The constants can now be found using the orthonormality of bound states.

$$\delta(\mathbf{r} - \mathbf{r}_j) = \sum_l A_{j,l} \psi_l(\mathbf{r})$$

$$\int_{\mathbb{R}^3} \delta(\mathbf{r} - \mathbf{r}_j) \psi_m^*(\mathbf{r}) \mathrm{d}\mathbf{r} = \sum_l A_{j,l} \int_{\mathbb{R}^3} \psi_l(\mathbf{r}) \psi_m^*(\mathbf{r}) \mathrm{d}\mathbf{r}$$

$$= \sum_l A_{j,l} \delta_{l,m}$$

$$= A_{j,m}$$

Therefore,

$$A_{j,l} = \int_{\mathbb{R}^3} \delta(\mathbf{r} - \mathbf{r}_j) \psi_l^*(\mathbf{r}) \mathrm{d}\mathbf{r} = \psi_l^*(\mathbf{r}_j)$$

Now we can go back and write concrete definitions for the field operators:

$$\hat{\Psi}^\dagger(\mathbf{r}) = \sum_j \psi_j^*(\mathbf{r}) \hat{c}_j^\dagger \tag{73}$$

$$\hat{\Psi}(\mathbf{r}) = \sum_j \psi_j(\mathbf{r}) \hat{c}_j \tag{74}$$

The canonical commutation relations obeyed by them are:

$$\left[\hat{\Psi}^\dagger(\mathbf{r}), \hat{\Psi}^\dagger(\mathbf{r}')\right] = 0, \left[\hat{\Psi}(\mathbf{r}), \hat{\Psi}(\mathbf{r}')\right] = 0, \left[\hat{\Psi}(\mathbf{r}), \hat{\Psi}^\dagger(\mathbf{r}')\right] = \delta(\mathbf{r} - \mathbf{r}') \tag{75}$$

Again, only the third one is non-trivial.

$$\left[\hat{\Psi}(\mathbf{r}), \hat{\Psi}^\dagger(\mathbf{r}')\right] = \left[\sum_j \psi_j(\mathbf{r}) \hat{c}_j, \sum_l \psi_l^*(\mathbf{r}') \hat{c}_l^\dagger\right]$$

$$= \sum_{j,l} \psi_j(\mathbf{r}) \psi_l^*(\mathbf{r}') \left[\hat{c}_j, \hat{c}_l^\dagger\right]$$

$$= \sum_{j,l} \psi_j(\mathbf{r}) \psi_l^*(\mathbf{r}') \delta_{j,l}$$

$$= \sum_j \langle \mathbf{r} | \psi_j \rangle \left(\langle \mathbf{r}' | \psi_j \rangle\right)^*$$

$$= \sum_j \langle \mathbf{r} | \psi_j \rangle \langle \psi_j | \mathbf{r}' \rangle$$

$$= \langle \mathbf{r} | \mathbf{r}' \rangle$$

$$= \delta(\mathbf{r} - \mathbf{r}')$$

We are now in a position to rewrite our first quantized Hamiltonian Eq. (9). We claim that the equivalent Hamiltonian in second quantization is Eq. (10). We will verify that the first piece is equal to the non-interacting Hamiltonian Eq. (70) by expanding the field operators:

$$\int_{\mathbb{R}^3} \sum_j \psi_j^*(\mathbf{r}) \hat{c}_j^\dagger(\mathbf{r}) \left[\frac{-\hbar^2}{2m} \nabla^2 + V(\mathbf{r})\right] \sum_l \psi_l(\mathbf{r}) \hat{c}_l \mathrm{d}\mathbf{r} = \sum_{j,l} \hat{c}_j^\dagger \hat{c}_l \int_{\mathbb{R}^3} \psi_j^*(\mathbf{r}) \epsilon_l \psi_l(\mathbf{r}) \mathrm{d}\mathbf{r}$$

$$= \sum_l \epsilon_l \hat{c}_l^\dagger \hat{c}_l$$

Similarly,

$$\iint_{\mathbb{R}^3 \times \mathbb{R}^3} \sum_j \psi_j^*(\mathbf{r})\hat{c}_j^\dagger \sum_l \psi_l^*(\mathbf{r}')\hat{c}_l^\dagger V_{\text{int}}(\mathbf{r}-\mathbf{r}') \sum_m \psi_m(\mathbf{r})\hat{c}_m \sum_p \psi_p(\mathbf{r}')\hat{c}_p \, \mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}$$

$$= \sum_{j,l,m,p} \hat{c}_j^\dagger \hat{c}_l^\dagger \hat{c}_m \hat{c}_p \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \psi_j^*(\mathbf{r})\psi_l^*(\mathbf{r}') V_{\text{int}}(\mathbf{r}-\mathbf{r}') \psi_m(\mathbf{r})\psi_p(\mathbf{r}') \, \mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}$$

$$= \sum_{j,l,m,p} V_{l,j,m,p}^{\text{int}} \hat{c}_j^\dagger \hat{c}_l^\dagger \hat{c}_m \hat{c}_p$$

which is exactly the result from Eq. (72).

C. Origin of the Bose-Hubbard Hamiltonian

When deriving a Hamiltonian that is meant to model bosons in a crystal, we must know a form for the energy eigenstates. By Bloch's theorem, these are labelled by two quantities, the wave vector $\mathbf{k}$ and the band index, n. A transition in a crystal is observed when the measured values of $\mathbf{k}$ and n change. Alternatively, one can construct position eigenstates using a linear combination that runs over all $\mathbf{k}$ and n. Then transitions are observed when one measures in the position basis and sees a change in position. There is a third way of thinking about the problem due to Wannier. One can instead restrict this linear combination to only the wave vector $\mathbf{k}$ and keep the band fixed. We usually cannot construct a perfectly localized delta function this way but by choosing the coefficients appropriately, we can construct functions that are approximately localized [36]. These Wannier functions will be denoted $\Omega_j^{(n)}(\mathbf{r})$ where n is the fixed band to which it belongs and $\mathbf{r}_j$ is the centre of the jth well in the periodic potential. Wannier functions are useful when we allow particles to "move from site to site" but not to change bands. For low-temperature bosonic systems, it is a good approximation that all particles occupy the lowest Bloch band. The extent to which bosons can hop between sites in this configuration is explained by the Bose-Hubbard Hamiltonian. Consider our expansions of the field operators Eq. (73) and Eq. (74). We did not use any special property of energy eigenstates when deriving these. The only relation between the functions and the operators in the terms of Eq. (74) that had to hold was that we had a state times the operator that removes a particle from that state. Similarly for Eq. (73) our terms had a state times the operator that adds a particle to the complex conjugate of that state. With out new functions defined, Wannier functions for the first band, we can define new anhilation and creation operators $\hat{a}_j$ and $\hat{a}_j^\dagger$ [10]. Now,

$$\hat{\Psi}^\dagger(\mathbf{r}) = \sum_j \Omega_j^{(0)*}(\mathbf{r})\hat{a}_j^\dagger$$

$$\hat{\Psi}(\mathbf{r}) = \sum_j \Omega_j^{(0)}(\mathbf{r})\hat{a}_j$$

The task is substitute these expansions into the second quantized Hamiltonian Eq. (10) and make sense of what we get. For the interaction piece, we already know this is:

$$\frac{1}{2} \sum_{j,l,m,p} U_{l,j,m,p} \hat{a}_j^\dagger \hat{a}_l^\dagger \hat{a}_m \hat{a}_p$$

$$U_{l,j,m,p} = \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \Omega_j^{(0)*}(\mathbf{r})\Omega_l^{(0)*}(\mathbf{r}') V_{\text{int}}(\mathbf{r}-\mathbf{r}') \Omega_m^{(0)}(\mathbf{r})\Omega_p^{(0)}(\mathbf{r}') \, \mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}$$

If we make the approximation that the interaction potential is a "billiard-ball" type interaction with strength g, an approximation valid for dilute Bose gases, $V_{\text{int}}(\mathbf{r}-\mathbf{r}') = g\delta(\mathbf{r}-\mathbf{r}')$ and we may write:

$$U_{l,j,m,p} = g \int_{\mathbb{R}^3} \Omega_j^{(0)*}(\mathbf{r})\Omega_l^{(0)*}(\mathbf{r})\Omega_m^{(0)}(\mathbf{r})\Omega_p^{(0)}(\mathbf{r}) \, \mathrm{d}\mathbf{r}$$

$$U = g \int_{\mathbb{R}^3} \left| \Omega_0^{(0)*}(\mathbf{r}) \right|^4 \mathrm{d}\mathbf{r} \tag{76}$$

where in the last step we have allowed only on-site interactions to take place and assumed that such interactions behave the same way on all sites. The corresponding term in our Hamiltonian becomes:

$$\frac{U}{2} \sum_j \hat{a}_j^\dagger \hat{a}_j^\dagger \hat{a}_j \hat{a}_j = \frac{U}{2} \sum_j \hat{n}_j(\hat{n}_j - 1) \tag{77}$$

Making sense of the rest of the Hamiltonian involves splitting up the potential, $V(\mathbf{r}) = V_{\text{lat}}(\mathbf{r}) + V_{\text{ext}}(\mathbf{r})$. The lattice potential is assumed to be periodic. The external potential is caused by a magnetic trap, an overlapping optical lattice or some other source of disorder that may cause the depths of the potential wells to differ. After substituting the external potential term into Eq. (10), we get:

$$\sum_{j,l} \epsilon_{j,l} \hat{a}_j^\dagger \hat{a}_l$$

$$\epsilon_{j,l} = \int_{\mathbb{R}^3} \Omega_j^{(0)*}(\mathbf{r}) V_{\text{ext}}(\mathbf{r}) \Omega_l^{(0)}(\mathbf{r}) d\mathbf{r}$$

If we assume that the external potential varies slowly over the optical lattice, we can take it out of the integral. The off-diagonal matrix elements then vanish due to the orthonormality of Wannier functions, leaving us with:

$$\epsilon_j = \int_{\mathbb{R}^3} V_{\text{ext}}(\mathbf{r}) \left| \Omega_j^{(0)}(\mathbf{r}) \right|^2 d\mathbf{r} \approx V_{\text{ext}}(\mathbf{r}_j) \tag{78}$$

The approximation gives us an intuitive interpretation of the ϵ_j values as simply being the on-site energies of the optical lattice. The corresponding term in the Hamiltonian is now:

$$\sum_j \epsilon_j \hat{a}_j^\dagger \hat{a}_j = \sum_j \epsilon_j \hat{n}_j \tag{79}$$

The last term we introduce will provide coupling between sites. As before, we will be left with a sum of operators and multiply this by some constant defined by an integral. The integral, a measure of the kinetic energy of the bosons, will be the hopping term and will quantify how often the bosons hop between nearest-neighbours, next-nearest-neighbours and so on. The remaining substitution we must perform gives us:

$$\sum_{j,l} -J_{j,l} \hat{a}_j^\dagger \hat{a}_l$$

$$J_{j,l} = -\int_{\mathbb{R}^3} \Omega_j^{(0)*}(\mathbf{r}) \left[\frac{-\hbar^2}{2m} \nabla^2 + V_{\text{lat}}(\mathbf{r}) \right] \Omega_l^{(0)}(\mathbf{r}) d\mathbf{r}$$

In the Bose-Hubbard Hamiltonian we are about to construct, we can freely shift the energy so that all $J_{j,j}$ will be zero [10]. The remaining values $J_{j,l}$ will describe bosons hopping between sites. We typically approximate this as nearest-neighbour hopping (the tight-binding approximation) and write:

$$J = -\int_{\mathbb{R}^3} \Omega_j^{(0)*}(\mathbf{r}) \left[\frac{-\hbar^2}{2m} \nabla^2 + V_{\text{lat}}(\mathbf{r}) \right] \Omega_l^{(0)}(\mathbf{r}) d\mathbf{r}, \langle j, l \rangle \tag{80}$$

We now write out the hopping term of the Bose-Hubbard Hamiltonian:

$$-J \sum_{\langle j,l \rangle_\bullet} \hat{a}_j^\dagger \hat{a}_l \tag{81}$$

Equations (77), (79) and (81), when put together, give the standard form of the Bose-Hubbard Hamiltonian as a truncation of the general many-body Hamiltonian:

$$\hat{\mathcal{H}} = \sum_j \epsilon_j \hat{n}_j + \frac{U}{2} \sum_j \hat{n}_j (\hat{n}_j - 1) - J \sum_{\langle j,l \rangle_\bullet} \hat{a}_j^\dagger \hat{a}_l \tag{82}$$

An important property of the Bose-Hubbard Hamiltonian is that it does not change if we shift all the field operators by a constant phase factor - *i.e.* replace $\hat{a}_j$ with $\hat{a}_j e^{i\theta}$ for all j. This means that the wavefunctions in an undisturbed system will not be driven to adopt any particular phase angle over any other.

D. The Bose-Hubbard Phase Diagram

In the absence of hopping, Eq. (82) can be diagonalized with nothing but the eigenstates of the number operators. An understanding of the phase transitions that occur in a system described by the Bose-Hubbard model can be gained

by approximating the hopping Hamiltonian with mean-field theory. First we define a family of parameters ϕ_j such that:

$$\hat{a}_j = \phi_j + \delta\hat{\phi}_j$$

Here, ϕ_j is the expectation of the jth anhilation operator and $\delta\hat{\phi}_j$ is an operator representing the deviation of the jth anhilation operator's value away from its mean. Nearest-neighbour products can now be expanded as:

$$
\begin{aligned}
\hat{a}_j^\dagger \hat{a}_l &= \left(\phi_j^* + \delta\hat{\phi}_j^\dagger\right)\left(\phi_l + \delta\hat{\phi}_l\right) \\
&= \phi_j^* \phi_l + \phi_j^* \delta\hat{\phi}_l + \phi_l^* \delta\hat{\phi}_j^\dagger + \delta\hat{\phi}_l \delta\hat{\phi}_j^\dagger \\
&\approx \phi_j^* \phi_l + \phi_j^* \delta\hat{\phi}_l + \phi_l^* \delta\hat{\phi}_j^\dagger \\
&= \phi_j^* \phi_l + \phi_j^* \left(\hat{a}_l - \phi_l\right) + \phi_l^* \left(\hat{a}_l^\dagger - \phi_l^*\right) \\
&= \phi_l \hat{a}_j^\dagger + \phi_j^* \hat{a}_l - \phi_j^* \phi_l
\end{aligned}
\tag{83}
$$

where we have used first order mean-field theory to eliminate terms where the deviation appears to second order. Our goal is to find an order parameter that vanishes at the boundary between two phases but right now, we have one for every site on the lattice. If we restrict our interest to a particular phase transition (Mott-insulator to superfluid) that occurs in a perfectly periodic optical lattice and drop the subscripts on the energies ϵ_j, we will be justified in dropping the subscripts on the ϕ_j as well since all the sites will be indistinguishable. It is important to note that the phase transitions of interest will take place in the grand canonical ensemble - a regime in which the number of bosons is not fixed. Since the relevant free-energy in this case is the grand potential, the mean-field Hamiltonian we use will take the form:

$$\hat{\mathcal{H}}_{\text{MF}} = \hat{\mathcal{H}} - \mu\hat{N} = \hat{\mathcal{H}} - \mu\sum_j \hat{n}_j$$

where μ is the chemical potential. Equation (83) can be inserted into the Bose-Hubbard Hamiltonian to give:

$$
\begin{aligned}
\hat{\mathcal{H}}_{\text{MF}} &= \sum_j (\epsilon_j - \mu)\hat{n}_j + \frac{U}{2}\sum_j \hat{n}_j(\hat{n}_j - 1) - J\sum_{\langle j,\, l\rangle_\bullet}\left(\phi_l\hat{a}_j^\dagger + \phi_j^*\hat{a}_l - \phi_j^*\phi_l\right) \\
&= (\epsilon - \mu)\sum_j \hat{n}_j + \frac{U}{2}\sum_j \hat{n}_j(\hat{n}_j - 1) - J\sum_{\langle j,\, l\rangle_\bullet}\left(\phi\hat{a}_j^\dagger + \phi^*\hat{a}_l - |\phi|^2\right)
\end{aligned}
\tag{84}
$$

The form of Eq. (82) showed us that we can shift the phase of the entire system with impunity. Now, the mean-field approach allows us to answer an important question: what happens when we shift the phase of *just part of the system*? When shifting the phase angle of a single site's field operator, we see that the Hamiltonian remains invariant precisely when $\phi = 0$. The state of matter that the system occupies when $\phi = 0$ will therefore be insulating, whereas the state on the other side of the boundary will exhibit some degree of stiffness in the phase. When a disturbance propagates through the lattice of bosons and causes phases to varry over long ranges and form a gradient, we find that there is a nonzero probability current:

$$
\begin{aligned}
\mathbf{J} &= \frac{\hbar}{2mi}\left(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*\right) \\
&= \frac{\hbar}{2mi}\left(|\psi|e^{-i\theta}\boldsymbol{\nabla}|\psi|e^{i\theta} - |\psi|e^{i\theta}\boldsymbol{\nabla}|\psi|e^{-i\theta}\right) \\
&= \frac{\hbar|\psi|}{2mi}\left(e^{-i\theta}\left(e^{i\theta}\boldsymbol{\nabla}|\psi| + ie^{i\theta}|\psi|\boldsymbol{\nabla}\theta\right) - e^{i\theta}\left(e^{-i\theta}\boldsymbol{\nabla}|\psi| - ie^{-i\theta}|\psi|\boldsymbol{\nabla}\theta\right)\right) \\
&= \frac{\hbar|\psi|}{2mi}\left(2i|\psi|\boldsymbol{\nabla}\theta\right) \\
&= \frac{\hbar}{m}|\psi|^2\boldsymbol{\nabla}\theta
\end{aligned}
$$

This complements the semi-classical argument for superfluidity discussed previously. We now see that it is phase coherence that gives rise to this zero viscosity flow [11]. Continuing with Eq. (84), we can write the sums over

nearest-neighbour pairs in terms of the number of nearest-neighbours per site, z.

$$
\begin{aligned}
\hat{\mathcal{H}}_{\mathrm{MF}} &= \frac{U}{2} \sum_j \hat{n}_j(\hat{n}_j - 1) - (\mu - \epsilon) \sum_j \hat{n}_j - J\phi z \sum_j \hat{a}_j^\dagger - J\phi^* z \sum_l \hat{a}_l + J N_{\mathrm{s}} z |\phi|^2 \\
&= \sum_j \left[\frac{U}{2} \hat{n}_j(\hat{n}_j - 1) - (\mu - \epsilon)\hat{n}_j + Jz|\phi|^2 - Jz\left(\phi \hat{a}_j^\dagger + \phi^* \hat{a}_j \right) \right] \\
&\equiv \sum_j \hat{\mathcal{H}}_j
\end{aligned}
$$

Since the mean-field Hamiltonian Eq. (84) can be written as a sum where each term in the sum has the same dependence on ϕ, we need only consider one Hamiltonian in the sum which we will write as:

$$
\hat{\mathcal{H}}_j = \hat{\mathcal{H}}^\circ + \hat{\mathcal{H}}' \tag{85}
$$

where the solved Hamiltonian is diagonal in the number eigenstate basis and the coupling term is considered the perturbation:

$$
\begin{aligned}
\hat{\mathcal{H}}^\circ &= \frac{U}{2} \hat{n}(\hat{n} - 1) - (\mu - \epsilon)\hat{n} + Jz|\phi|^2 \\
\hat{\mathcal{H}}' &= -Jz\left(\phi \hat{a}^\dagger + \phi^* \hat{a}\right)
\end{aligned}
$$

The free energy is minimized when the ground-state number of particles occupies each site. For the site being considered in Eq. (85) we will refer to this state as $|n_{\mathrm{g}}\rangle$. The recipie for time-independent perturbation theory applies well in this situation [37].
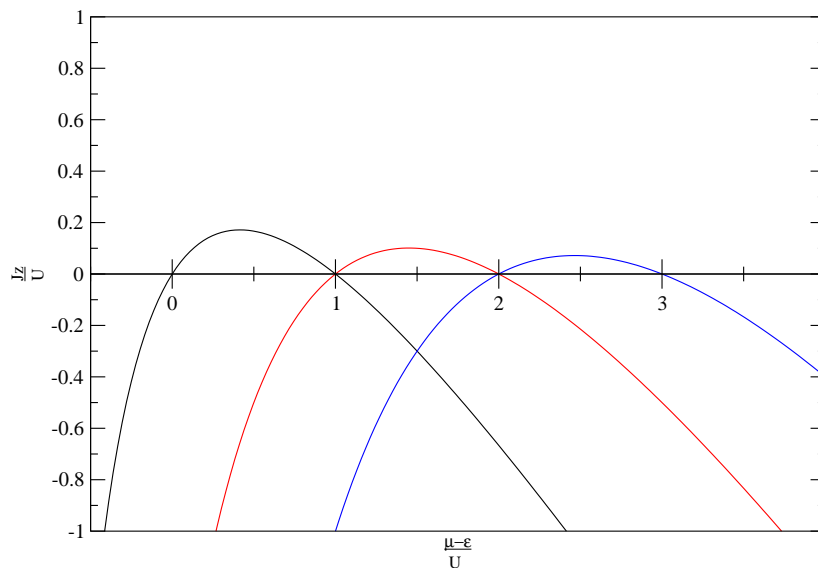

FIG. 33: The contribution of each ground-state particle number to the overall phase diagram. The black curve is for $n_{\mathrm{g}} = 1$, the red for $n_{\mathrm{g}} = 2$ and the blue for $n_{\mathrm{g}} = 3$.

$$
\begin{aligned}
E_{n_\mathrm{g}}^{(0)} &= \langle n_\mathrm{g}|\, \hat{\mathcal{H}}^\circ \,|n_\mathrm{g}\rangle \\
&= \frac{U}{2} n_\mathrm{g}(n_\mathrm{g}-1)(\mu-\epsilon)n_\mathrm{g} + Jz|\phi|^2 \\
&\equiv b_0 + Jz|\phi|^2 \\
E_{n_\mathrm{g}}^{(1)} &= \langle n_\mathrm{g}|\, \hat{\mathcal{H}}' \,|n_\mathrm{g}\rangle \\
&= 0 \\
E_{n_\mathrm{g}}^{(2)} &= \sum_{p \neq n_\mathrm{g}} \frac{\left| \langle n_\mathrm{g}|\, \hat{\mathcal{H}}' \,|p\rangle \right|^2}{E_{n_\mathrm{g}}^{(0)} - E_p^{(0)}} \\
&= \frac{\left| \langle n_\mathrm{g}|\, \hat{\mathcal{H}}' \,|n_\mathrm{g}+1\rangle \right|^2}{E_{n_\mathrm{g}}^{(0)} - E_{n_\mathrm{g}+1}^{(0)}} + \frac{\left| \langle n_\mathrm{g}|\, \hat{\mathcal{H}}' \,|n_\mathrm{g}-1\rangle \right|^2}{E_{n_\mathrm{g}}^{(0)} - E_{n_\mathrm{g}-1}^{(0)}} \\
&= (Jz)^2 \left(\frac{|\phi|^2(n_\mathrm{g}+1)}{(\mu-\epsilon) - Un_\mathrm{g}} + \frac{|\phi|^2 n_\mathrm{g}}{U(n_\mathrm{g}-1) - (\mu-\epsilon)} \right) \\
&\equiv b_2|\phi|^2 \\
E_{n_\mathrm{g}}^{(3)} &= \sum_{p \neq n_\mathrm{g}} \sum_{q \neq n_\mathrm{g}} \frac{\langle n_\mathrm{g}|\, \hat{\mathcal{H}}' \,|q\rangle \langle q|\, \hat{\mathcal{H}}' \,|p\rangle \langle p|\, \hat{\mathcal{H}}' \,|n_\mathrm{g}\rangle}{\left(E_p^{(0)} - E_{n_\mathrm{g}}^{(0)} \right)\left(E_q^{(0)} - E_{n_\mathrm{g}}^{(0)} \right)} - \langle n_\mathrm{g}|\, \hat{\mathcal{H}}' \,|n_\mathrm{g}\rangle \sum_{p \neq n_\mathrm{g}} \frac{\left| \langle n_\mathrm{g}|\, \hat{\mathcal{H}}' \,|p\rangle \right|^2}{\left(E_{n_\mathrm{g}}^{(0)} - E_p^{(0)} \right)^2} \\
&= 0
\end{aligned}
$$

As described in Appendix A, the vanishing odd terms fit the perscription of a Landau free-energy which undergoes a phase transition when it is minimized for nonzero values of the superfluid order parameter ϕ.

$$
E = b_0 + (b_2 + Jz)|\phi|^2 + b_4|\phi|^4 + O(|\phi|^6) \tag{86}
$$

The system crosses the phase boundary when the second coefficient in Eq. (86) is zero. Demanding this, we can solve for what the phase boundary actually is.
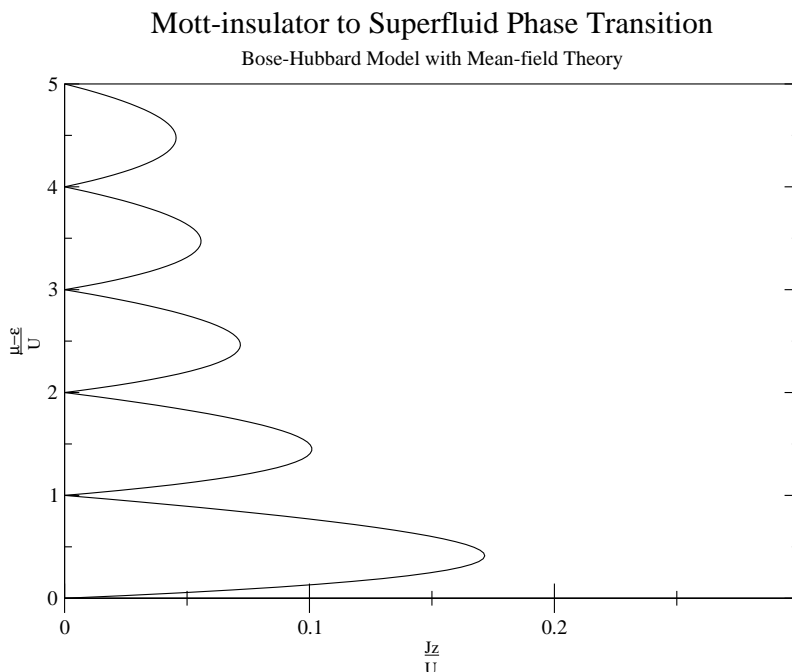
Mott-insulator to Superfluid Phase Transition
Bose-Hubbard Model with Mean-field Theory

FIG. 34: Mean-field phase diagram for the phase transition that affects ultracold bosons in an ordered optical lattice.

$$b_2 + Jz = 0$$

$$\frac{(Jz)^2(n_{\mathrm{g}}+1)}{(\mu-\epsilon)-Un_{\mathrm{g}}} + \frac{(Jz)^2 n_{\mathrm{g}}}{(\mu-\epsilon)-Un_{\mathrm{g}}+U} + Jz = 0$$

$$Jz((\mu-\epsilon)-Un_{\mathrm{g}}+U)(n_{\mathrm{g}}+1) - Jzn_{\mathrm{g}}((\mu-\epsilon)-Un_{\mathrm{g}}) + ((\mu-\epsilon)-Un_{\mathrm{g}})((\mu-\epsilon)-Un_{\mathrm{g}}+U) = 0$$

$$\frac{Jz}{U}\left(\frac{\mu-\epsilon}{U}-n_{\mathrm{g}}+1\right)(n_{\mathrm{g}}+1) - \frac{Jz}{U}n_{\mathrm{g}}\left(\frac{\mu-\epsilon}{U}-n_{\mathrm{g}}\right) + \left(\frac{\mu-\epsilon}{U}-n_{\mathrm{g}}\right)\left(\frac{\mu-\epsilon}{U}-n_{\mathrm{g}}+1\right) = 0$$

$$\frac{Jz}{U} = \frac{\left(n_{\mathrm{g}}-\frac{\mu-\epsilon}{U}\right)\left(1-\left(n_{\mathrm{g}}-\frac{\mu-\epsilon}{U}\right)\right)}{\frac{\mu-\epsilon}{U}+1} \tag{87}$$

As shown in Fig. 33, each integer value of n_{g} excludes a different region where the order parameter must be zero. Note that it is always possible to express energies in units of the interaction strength U because if there are no interactions, a perfect BEC forms and these phases do not exist. If we interchange the axes and plot only the first-quadrant, we can reproduce the phase diagram the way it is usually drawn [38] [14].

IX.   APPENDIX C: CHEBYSHEV EXPANSION CODE

```c
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#define chebyshev_coefficients(a, b) chebyshev_coefficients[(a) * terms + (b)]
/* The dimension of the "Krylov subspace" used in the algorithm */
#define MAXLAN  30

/* An implementation of sparse matrices */
struct _sparse_entry { int pos; float val; };
typedef struct _sparse_entry sparse_entry;
struct _sparse { int num; sparse_entry *array; };
typedef struct _sparse sparse;
sparse *hamiltonian;

/* Custom Matrix Vector Multiplication function, compatible with our sparse matrix storage */
void diag_op(int *nrow, int *ncol, double *xin, int *ldx, double *yout, int *ldy) {
  int i, j;

  for (i = 0; i < *ldy; i++) {
    *(yout + i) = 0.0;

    for (j = 0; j < hamiltonian[i].num; j++) {
      *(yout + i) += *(xin + hamiltonian[i].array[j].pos) * hamiltonian[i].array[j].val;
    }
  }
}

/* Diagonalization function built on TRLAN */
void diagonalize(int dim, int min_or_max, float *ev) {
  int nev = 1; /* We only want one eigenvalue */
  int ned = 1; /* It should either be the "first-largest" or "first-smallest" */
  int i, lwrk, ipar[32];
  double eval[nev], evec[dim * nev], wrk[MAXLAN * (MAXLAN + 10)];

  lwrk = MAXLAN * (MAXLAN + 10);
  ipar[0] = 0;
  ipar[1] = min_or_max; /* Whether we want the largest or smallest eigenvalue */
  ipar[2] = ned;
  ipar[3] = 0;
  ipar[4] = MAXLAN;
```

```c
    ipar[5] = 1;            /* Restarting scheme 1 */
    ipar[6] = 2000;         /* Maximum number of Matrix Vector Multiplications */
    ipar[7] = 0;            /* For people who want to use MPI */
    ipar[8] = 1;            /* Verbosity */
    ipar[9] = 99;           /* Fortran IO unit number used to write log messages */
    ipar[10] = 1;           /* When this is 1, starting eigenvalues are used as a guess */
    ipar[11] = -5;          /* A flag having to do with checkpoint files */
    ipar[12] = 98;          /* Fortran IO unit number used to write checkpoint files */
    ipar[13] = 3 * dim;     /* The number of floating-point operations per Matrix Vector Multiplication */
    wrk[0] = 1.4901e-8;     /* Relative tolerance on residual norms */
    /* Sets the eigenvalues and eigenvectors to their initial guesses */
    for (i=0; i < nev; ++i) {
      eval[i] = 0.0;
    }
    for (i=0; i < dim; ++i) {
      evec[i] = 1.0;
    }
    /* Call the Fortran function TRLAN77 */
    trlan77_(diag_op, ipar, &dim, &nev, eval, evec, &dim, wrk, &lwrk);
    /* Returns the desired eigenvalue as a side-effect */
    *ev = eval[0];
}

/* Multiplies two matrices. */
sparse *multiply_sparse(sparse *matrix1, sparse *matrix2, int dimension) {
  sparse *ret = malloc(dimension * sizeof(sparse));
  int i, j, k, l;
  int stay_in_loop, found;

  for (i = 0; i < dimension; i++) {
    /* In the worst case, the matrix is dense */
    ret[i].array = malloc(dimension * sizeof(sparse_entry));
  }

  for (i = 0; i < dimension; i++) {
    ret[i].num = 0;

    for (j = 0; j < matrix1[i].num; j++) {
      for (k = 0; k < matrix2[matrix1[i].array[j].pos].num; k++) {
        l = 0;
        found = 0;
        stay_in_loop = 1;
        if (ret[i].num > 0) {
          while (stay_in_loop) {
            if (ret[i].array[l].pos == matrix2[matrix1[i].array[j].pos].array[k].pos) {
              found = 1;
              stay_in_loop = 0;
            } else if (l == ret[i].num - 1) {
              stay_in_loop = 0;
            }
            l++;
          }
          l--;
        }

        if (found) {
          ret[i].array[l].val += matrix2[matrix1[i].array[j].pos].array[k].val
                                 * matrix1[i].array[j].val;
```

```c
      } else {
        l = ret[i].num;
        ret[i].array[l].val = matrix2[matrix1[i].array[j].pos].array[k].val
                            * matrix1[i].array[j].val;
        ret[i].array[l].pos = matrix2[matrix1[i].array[j].pos].array[k].pos;
        ret[i].num++;
      }
    }
  }

  ret[i].array = realloc(ret[i].array, ret[i].num * sizeof(sparse_entry));
}

  return ret;
}

/* Looksup where the (i, j) location of a sparse matrix is actually stored */
void get_indices(int terms, sparse **matrices, int i, int j, int *index_array) {
  int stay_in_bigger_loop = 1;
  int stay_in_loop = 1;
  int found = 0;
  int k = 0;
  int l;

  /* Finds where the entry is in the most dense matrix */
  while (stay_in_loop) {
    if (matrices[terms - 1][i].array[k].pos == j) {
      index_array[terms - 1] = k;
      found = 1;
      stay_in_loop = 0;
    } else if (k == matrices[terms - 1][i].num - 1) {
      stay_in_loop = 0;
    }

    k++;
  }

  if (!found) {
    for (k = 0; k < terms; k++) {
      index_array[k] = -1;
    }
    return;
  }

  k = terms - 2;

  /* Now searches the more sparse matrices */
  while (stay_in_bigger_loop) {
    stay_in_loop = 1;
    found = 0;
    l = 0;

    while (stay_in_loop) {
      if (matrices[k][i].array[l].pos == j) {
        index_array[k] = l;
        found = 1;
        stay_in_loop = 0;
      } else if (l == matrices[k][i].num - 1) {
```

```
          stay_in_loop = 0;
        }

        l++;
      }

      if (!found) {
        for (l = 0; l <= k; l++) {
          index_array[l] = -1;
        }
        stay_in_bigger_loop = 0;
      } else if (k == 0) {
        stay_in_bigger_loop = 0;
      }

      k--;
    }
}

/* Returns a copy of an integer array. */

int *int_copy_of(int *set, int size) {
  int *new_set = malloc(size * sizeof(int));
  int i;

  for (i = 0; i < size; i++) {
    new_set[i] = set[i];
  }

  return new_set;
}

/* We have a set of sets and we want to add a set to it in the next available spot. */

void put_in_matrix(int *set, int **matrix) {
  int i = 0;

  while (matrix[i] != 0) {
    i++;
  }

  matrix[i] = set;
}

/* Generates all distinguishable n_s, tuples of integers where the sum of all the integers is n_b. */

void populate_basis_set(int n_s, int free_bosons, int start, int *prefix_state, int **basis_set) {
  int i, j;

  if (free_bosons == 1) {
    for (i = start; i < n_s; i++) {
      int *state = int_copy_of(prefix_state, n_s);
      state[i] = 1;
      put_in_matrix(state, basis_set);
    }
    free(prefix_state);
  } else {
    int *state = int_copy_of(prefix_state, n_s);
```

```c
      state[start] = free_bosons;
      put_in_matrix(state, basis_set);

      if (start < (n_s - 1)) {
        for (i = free_bosons; i > 0; i--) {
          int *new_prefix_state = int_copy_of(prefix_state, n_s);
          new_prefix_state[start] = free_bosons - i;
          populate_basis_set(n_s, i, start + 1, new_prefix_state, basis_set);
        }
      }
    }
  }
}

/* Returns n choose k according to the binomial theorem without wasting cycles with factorials. */

int choose(int n, int k) {
  float result = 1.0;
  int i;

  for (i = 1; i <= k; i++) {
    result *= ((float) (n - (k - i))) / ((float) i);
  }

  return result;
}

void dynamics(int n_s, int n_b, float interaction_u, float hopping_j, float *epsilon, float max_time,
              float dt, int terms) {
  /* Variables for generating the number eigenstate basis */
  int dimension = choose(n_s + n_b - 1, n_b);
  int **basis_set = calloc(dimension, sizeof(int *));
  int *state = calloc(n_s, sizeof(int));

  /* Variables for generating the Hamiltonian matrix */
  float entry;
  int *how_far_we_got = malloc(dimension * sizeof(int));
  hamiltonian = malloc(dimension * sizeof(sparse));
  int current_spot;
  int difference_count;
  int first_differing_index;
  int second_differing_index;
  int i, j, k, l;
  int todo;

  /* Variables for timing the routines */
  struct timeval tv1;
  struct timeval tv2;
  int t1, t2;

  int *index_array = malloc(terms * sizeof(int));
  sparse **matrices = malloc(terms * sizeof(sparse *));
  float **real_matrix = malloc(dimension * sizeof(float *));
  float **imaginary_matrix = malloc(dimension * sizeof(float *));
  long int *chebyshev_coefficients = malloc(terms * terms * sizeof(long int));
  float max_energy, min_energy, scale_factor, shift_factor, t, error;

  /* Variables for writing the files */
  FILE **fp = malloc(n_s * sizeof(FILE *));
```

```c
char *filename = malloc(8 + log10((float) n_s) * sizeof(char));

gettimeofday(&tv1, NULL);
t1 = tv1.tv_usec + 1000000 * tv1.tv_sec;
populate_basis_set(n_s, n_b, 0, state, basis_set);
gettimeofday(&tv2, NULL);
t2 = tv2.tv_usec + 1000000 * tv2.tv_sec;
printf("Basis generation: %d us\n", t2 - t1);

gettimeofday(&tv1, NULL);
t1 = tv1.tv_usec + 1000000 * tv1.tv_sec;
/* The number of nonzero entries in a row is twice the number of occupied sites */
for (i = 0; i < dimension; i++) {
  hamiltonian[i].num = 1;
  for (j = 0; j < n_s; j++) {
    if (basis_set[i][j] != 0) {
      if (n_s == 2) {
        /* This part is tricky but we almost never used two sites */
        hamiltonian[i].num += 1;
      } else {
        hamiltonian[i].num += 2;
      }
    }
  }
  hamiltonian[i].array = malloc(hamiltonian[i].num * sizeof(sparse_entry));
}
for (i = 0; i < dimension; i++) {
  current_spot = 0;
  /* Terms coming from hopping */
  for (j = 0; j < i; j++) {
    k = 0;
    difference_count = 0;
    first_differing_index = 0;
    second_differing_index = 0;
    entry = 0.0;
    while (difference_count < 3 && k < n_s) {
      if (basis_set[i][k] != basis_set[j][k]) {
        difference_count++;
        second_differing_index = first_differing_index;
        first_differing_index = k;
      }
      k++;
    }
    /* For the inner product to survive, states need two places where they differ in boson number */
    if (difference_count == 2) {
      /* These places must be nearest neighbours in the periodic sense */
      if ((abs(second_differing_index - first_differing_index) == 1
          || (abs(second_differing_index - first_differing_index) == (n_s - 1)))) {
        /* The difference in boson number itself must also be 1 */
        if (abs(basis_set[i][first_differing_index] - basis_set[j][first_differing_index]) == 1) {
          if (basis_set[j][first_differing_index] > basis_set[i][first_differing_index]) {
            entry -= hopping_j * sqrt(basis_set[j][first_differing_index])
                               * sqrt(basis_set[i][second_differing_index]);
          } else {
            entry -= hopping_j * sqrt(basis_set[i][first_differing_index])
                               * sqrt(basis_set[j][second_differing_index]);
          }
          hamiltonian[i].array[current_spot].pos = j;
```

```
          hamiltonian[i].array[current_spot].val = entry;
          hamiltonian[j].array[how_far_we_got[j]].pos = i;
          hamiltonian[j].array[how_far_we_got[j]].val = entry;
          how_far_we_got[j]++;
          current_spot++;
        }
      }
    }
  }
  entry = 0.0;
  /* Terms coming from interactions */
  for (k = 0; k < n_s; k++) {
    entry += epsilon[k] * basis_set[i][k]
          + (interaction_u / 2.0) * basis_set[i][k] * (basis_set[i][k] - 1);
  }
  hamiltonian[i].array[current_spot].pos = i;
  hamiltonian[i].array[current_spot].val = entry;
  current_spot++;
  how_far_we_got[i] = current_spot;
}
free(how_far_we_got);
gettimeofday(&tv2, NULL);
t2 = tv2.tv_usec + 1000000 * tv2.tv_sec;
printf("Matrix calculation: %d us\n", t2 - t1);

/* Stores the Chebyshev polynomials in a matrix */
chebyshev_coefficients(0, 0) = 1;
chebyshev_coefficients(1, 0) = 0;
chebyshev_coefficients(1, 1) = 1;

for (i = 2; i < terms; i++) {
  if (i % 2 == 0) {
    if (i % 4 == 0) {
      chebyshev_coefficients(i, 0) = 1;
    } else {
      chebyshev_coefficients(i, 0) = -1;
    }
    for (j = 2; j <= i; j += 2) {
      chebyshev_coefficients(i, j) = 2 * chebyshev_coefficients(i - 1, j - 1)
                                   - chebyshev_coefficients(i - 2, j);
    }
  } else {
    for (j = 1; j <= i; j += 2) {
      chebyshev_coefficients(i, j) = 2 * chebyshev_coefficients(i - 1, j - 1)
                                   - chebyshev_coefficients(i - 2, j);
    }
  }
}

matrices[0] = malloc(dimension * sizeof(sparse));
for (i = 0; i < dimension; i++) {
  real_matrix[i] = calloc(dimension, sizeof(float));
  imaginary_matrix[i] = calloc(dimension, sizeof(float));

  matrices[0][i].num = hamiltonian[i].num;
  matrices[0][i].array = malloc(hamiltonian[i].num * sizeof(sparse_entry));
  for (j = 0; j < hamiltonian[i].num; j++) {
    matrices[0][i].array[j].val = hamiltonian[i].array[j].val;
```

```c
      matrices[0][i].array[j].pos = hamiltonian[i].array[j].pos;
  }
}

gettimeofday(&tv1, NULL);
t1 = tv1.tv_usec + 1000000 * tv1.tv_sec;
diagonalize(dimension, -1, &min_energy);
diagonalize(dimension, 1, &max_energy);
gettimeofday(&tv2, NULL);
t2 = tv2.tv_usec + 1000000 * tv2.tv_sec;
printf("Diagonalization: %d us\n", t2 - t1);

for (i = 0; i < n_s; i++) {
  sprintf(filename, "nd%d.dat", i);
  fp[i] = fopen(filename, "w");
}

gettimeofday(&tv1, NULL);
t1 = tv1.tv_usec + 1000000 * tv1.tv_sec;

scale_factor = (max_energy - min_energy) / (2.0 - 0.01);
shift_factor = (max_energy + min_energy) / 2.0;

if (terms < 1.5 * scale_factor * dt) {
  fprintf(stderr, "The number of terms is less than %f!\n", 1.5 * scale_factor * dt);
}

/* Rescales the Hamiltonian so that it's eigenvalues are between -1 and 1 */
for (i = 0; i < dimension; i++) {
  for (j = 0; j < matrices[0][i].num; j++) {
    if (matrices[0][i].array[j].pos == i) {
      matrices[0][i].array[j].val = (matrices[0][i].array[j].val - shift_factor) / scale_factor;
    } else {
      matrices[0][i].array[j].val /= scale_factor;
    }
  }
  free(hamiltonian[i].array);
}
free(hamiltonian);

for (i = 1; i < terms; i++) {
  matrices[i] = multiply_sparse(matrices[i - 1], matrices[0], dimension);
}

/* There is no more matrix multiplication to do so we use explicit storage here */
for (i = 0; i < dimension; i++) {
  real_matrix[i][i] = j0f(scale_factor * dt);
  for (k = 1; k < terms; k++) {
    if (k % 2 == 0) {
      real_matrix[i][i] += 2.0 * jnf(k, scale_factor * dt);
    }
    for (j = 0; j <= i; j++) {
      get_indices(terms, matrices, i, j, index_array);

      /* Be careful tweaking things as this part is very sensitive to signs */
      for (l = 1; l <= k; l += 2) {
        if ((k % 4 == 1) && (index_array[l - 1] != -1)) {
          imaginary_matrix[i][j] -= 2.0 * jnf(k, scale_factor * dt)
```

```c
                            * chebyshev_coefficients(k, l) * matrices[l - 1][i].array[index_array[l
        } else if ((k % 4 == 2) && (index_array[l] != -1)) {
          real_matrix[i][j] -= 2.0 * jnf(k, scale_factor * dt)
                            * chebyshev_coefficients(k, l + 1) * matrices[l][i].array[index_array[l]
        } else if ((k % 4 == 3) && (index_array[l - 1] != -1)) {
          imaginary_matrix[i][j] += 2.0 * jnf(k, scale_factor * dt)
                            * chebyshev_coefficients(k, l) * matrices[l - 1][i].array[index_array[l
        } else if (index_array[l] != -1) {
          real_matrix[i][j] += 2.0 * jnf(k, scale_factor * dt)
                            * chebyshev_coefficients(k, l + 1) * matrices[l][i].array[index_array[l]
        }
      }
    }
  }
}

for (i = 0; i < terms; i++) {
  for (j = 0; j < dimension; j++) {
    free(matrices[i][j].array);
    for (k = 0; k < j; k++) {
      real_matrix[k][j] = real_matrix[j][k];
      imaginary_matrix[k][j] = imaginary_matrix[j][k];
    }
  }
}
free(matrices);
gettimeofday(&tv2, NULL);
t2 = tv2.tv_usec + 1000000 * tv2.tv_sec;
printf("Matrix generation: %d us\n", t2 - t1);

gettimeofday(&tv1, NULL);
t1 = tv1.tv_usec + 1000000 * tv1.tv_sec;
/* Variables for using the expansion */
float *real_part = malloc(dimension * sizeof(float));
float *imaginary_part = malloc(dimension * sizeof(float));
float *real_initial_state = calloc(dimension, sizeof(float));
float *imaginary_initial_state = calloc(dimension, sizeof(float));
int index;
float *probability = malloc(dimension * sizeof(float));
float *number = malloc(n_s * sizeof(float));
real_initial_state[0] = 1.0;
t = 0.0;
error = 0.0;

while (t < max_time) {
  for (i = 0; i < dimension; i++) {
    real_part[i] = 0.0;
    imaginary_part[i] = 0.0;

    for (j = 0; j < dimension; j++) {
      real_part[i] += real_matrix[i][j] * real_initial_state[j]
                    - imaginary_matrix[i][j] * imaginary_initial_state[j];
      imaginary_part[i] += real_matrix[i][j] * imaginary_initial_state[j]
                    + imaginary_matrix[i][j] * real_initial_state[j];
    }

    probability[i] = real_part[i] * real_part[i] + imaginary_part[i] * imaginary_part[i];
  }
```

```c
  for (i = 0; i < dimension; i++) {
    real_initial_state[i] = real_part[i];
    imaginary_initial_state[i] = imaginary_part[i];
  }

  for (i = 0; i < n_s; i++) {
    number[i] = 0.0;
    for (j = 0; j < dimension; j++) {
      number[i] += probability[j] * basis_set[j][i];
    }
    fprintf(fp[i], "%lf\t%lf\n", t, number[i]);
  }

  t += dt;
  /* Error from the exact solution for two sites systems, meaningless otherwise */
  error += pow(number[0] - ((float) n_b / 2.0) * (1.0 + cos(2 * t)), 2);
  }
  gettimeofday(&tv2, NULL);
  t2 = tv2.tv_usec + 1000000 * tv2.tv_sec;
  printf("Expansion: %d us\n", t2 - t1);
  printf("Error: %0.10f\n", error);

  for (i = 0; i < n_s; i++) {
    fclose(fp[i]);
  }
}

int main() {
  int n_s, n_b, terms, seed, i;
  float interaction_u, hopping_j, disorder_w, max_time, dt;

  printf("Random seed (0 for no disorder): ");
  scanf("%d", &seed);

  printf("Number of sites: ");
  scanf("%d", &n_s);

  printf("Number of bosons: ");
  scanf("%d", &n_b);

  printf("Interaction energy: ");
  scanf("%f", &interaction_u);

  printf("Hopping energy: ");
  scanf("%f", &hopping_j);

  if (seed) {
    printf("Disorder Strength: ");
    scanf("%f", &disorder_w);
  } else {
    disorder_w = 0.0;
  }

  printf("Number of terms: ");
  scanf("%d", &terms);

  printf("Time over which to run: ");
```

```c
    scanf("%f", &max_time);

    printf("Timestep: ");
    scanf("%f", &dt);

    float *epsilon = malloc(n_s * sizeof(float));

    srand48(seed);
    for (i = 0; i < n_s; i++) {
      if (seed) {
        epsilon[i] = disorder_w * (drand48() - 0.5);
      } else {
        epsilon[i] = 0.0;
      }
    }

    dynamics(n_s, n_b, interaction_u, hopping_j, epsilon, max_time, dt, terms);
}
```