Accelerated Tempering Dynamics in HMC Simulations of Lattice Field Theory

MPhys Project Report

Jack Frankland

March 26, 2018

Abstract

Supervisors: Dr Roger Horsley, Dr Brian Pendleton

Personal statement

Acknowledgments

Contents

1 Introduction

The Hybrid Monte Carlo (HMC) algorithm (also referred to as Hamiltonian Monte Carlo) was originally developed by Duane, Kennedy, Pendleton and Roweth in [1] for the purposes of numerical simulation of lattice field theory. HMC is an example of a Markov Chain Monte Carlo (MCMC) method where states are proposed via a Markov chain in order to produce samples distributed according to some probability distribution, these samples can then be used to estimate the expectation value of some function of the samples under the probability distribution. Although MCMC was original introduced via the *Metropolis algorithm* in [2], HMC makes improvements on this method by using Hamiltonian dynamics to propose new states in the Markov chain. The idea is that we consider the variables of interest as *position* and introduce auxiliary *momentum* variables which we can use to define a Hamiltonian and evolve the system in computer time via a leapfrog integration method to propose new states. We alternate between drawing the momentum variables from a multivariate Gaussian distribution and performing a Metropolis update where the new state state is proposed via the Hamiltonian dynamics. The advantage of the HMC method is that the proposed state can have a high probability of acceptance and be distant from the current state, it therefore avoids the problems of slow state exploration that results from using random-walk proposals such as in the original Metropolis algorithm [3]. Although popular amongst the lattice field theory community the HMC method has also been used in statistics, e.g. [4], [5] and [6] for applications such as neural network models. However, our application of HMC will be to quantum mechanical systems and in particular the case of quantum harmonic and anharmonic oscillators.

Standard HMC can have difficulty sampling from the different areas of a probability distribution if those areas are separated by a region of low probability; we will refer to these areas as *isolated modes* of the distribution. An example of this in quantum mechanics is that of a 1-dimensional double well potential, where the probability of finding a particle in either well is high, whilst finding it in the region between the wells is low. If we are interested in generating samples for say the position of the particle using the HMC algorithm, then for a deep enough well there will be an asymmetry in the number of samples recorded in each well. This is due to the energy cost associated with moving between the wells, and so the Hamiltonian dynamics in the HMC simulation will tend to propose an update state in the well the current state is in, leading to the asymmetry in the recorded samples. The aim of this project is to write an HMC simulation of the anharmonic oscillator and introduce a tempering parameter as proposed in [7] and [3] into the dynamics to investigate whether this increases the frequency at which samples move between the two wells. In terms of our simulation the benefit of successful tempering would be more accurate estimations of observed quantities such as position and energy expectation values. However, in terms of a broader goal, successful tempering results in HMC simulations of simple quantum mechanical systems, may suggest it has potential applications in lattice field theory where isolated modes are a more serious problem. In certain field theories, including QCD, the states we wish to construct via our Markov chain are in distinct topological sectors which are labelled by a topological charge. For small lattice spacings moving between these sectors can take a very long time, since they are separated by a region of high Euclidean action (i.e. low

probability in the distribution) and so the transitions are statistically suppressed (see [8], [9] and [10]). It is therefore possible that tempered dynamics will have applications to theories such as lattice QCD where isolated modes arise as regions of distinct topological charge. In general the *autocorrelation time* which measures how correlated subsequent samples in any Monte Carlo simulation are will be high if the system is trapped in an isolated mode. The purpose of tempering is to try and diminish the autocorrelation time without the need for extra computation.

In this report we follow the work of Creutz and Freedman in [11] who constructed a MCMC simulation of the harmonic and anharmonic oscillators, however we will use HMC where they used the Metropolis algorithm. Reproducing the results in [11] using the Metropolis algorithm is a popular choice for undergraduate projects e.g. [12], [13] and [14] which provide a useful comparison for our simulation which uses a different algorithm. We begin the report by following the steps taken in [11] where we introduce and define the path integral of quantum mechanics and show that for a discrete lattice in Euclidean time the path integral can be considered as a canonical partition function. This connection between quantum mechanics and statistical physics is what will enable us to apply Monte Carlo methods and we will see that the quantum expectation values in the ground state correspond to classical expectation values of the statistical system with a canonical distribution. We will then briefly review the topic of Monte Carlo methods and explain how they can be used to approximate expectation values for functions of random variables under some probability distribution via a Markov chain as well as defining detailed balance and ergodicity which are conditions required for the Monte Carlo simulation to be valid. Before introducing the HMC algorithm we very briefly review Hamiltonian dynamics, explain how Hamilton's equations can be integrated numerically via the leapfrog method and show that as well as being reversible, the leapfrog method preserves volume on the phase space of position and momentum coordinates, which we we will see is necessary for detailed balance to hold. Having provided the relevant background we are then able to define the steps of the HMC algorithm and explain how it can be used to sample from a canonical distribution by defining axillary momenta coordinates and a Hamiltonian. In this section we will also discuss the issue of ergodicity and show that HMC obeys detailed balance. Due to the stochastic nature of MCMC generated samples in the Markov chain can be correlated and we briefly discuss the effect of this on data analysis. The results section begins by applying the HMC algorithm to the canonical distribution function for the quantum system. We examine the results of our simulation for the harmonic oscillator where we are able to compare with numerical and exact results for the discrete theory in [11],[12], [13] and [14]. The simple system of the harmonic oscillator provides a good testing ground for our simulation where we can easily compare our results to theory. Moving on to the anharmonic oscillator we are able to apply our simulation to a system that does not have an analytic solution. For the case of the anharmonic oscillator we are able to compare our results to those acquired through alternative methods [15] and simulations [13] however we will observe in this section that the HMC simulation begins to fail in the case of the isolated modes in the double well potential. In particular the double well system can have difficulty tunnelling between the two modes. At this point we will examine the effects of introducing a tempering parameter into the simulation, with the hope that it will increase the tunnel rate. We will show that volume preservation still holds for tempered dynamics,

which is necessary if it is still to obey detailed balance.

2 Background

2.1 Quantum Mechanics

2.1.1 The Path Integral

In quantum mechanics we are often interested in calculating the path integral:

$$\langle x_b, t_b | x_a, t_a \rangle = \int_{x_a}^{x_b} \mathcal{D}x\,(t) \exp\left(\frac{i}{\hbar} S_M\,[x\,(t)]\right),\tag{1}$$

where we will use the notation:

$$Z_{ba} = \langle x_b, t_b | x_a, t_a \rangle \tag{2}$$

$$= \langle x_b | \, e^{-i(t_b - t_a)\hat{H}/\hbar}\, | x_a \rangle \,,\tag{3}$$

where $\hat{H}$ is the Hamiltonian operator of a spinless particle of mass m moving in a potential:

$$\hat{H}\,(\hat{p}, \hat{x}) = \frac{\hat{p}^2}{2m} + \hat{V}\,(\hat{x})\,.\tag{4}$$

Z_{ba} is the transition amplitude for a particle of mass m in position eigenstate (in the Heisenberg picture) $|x_a, t_a\rangle$ to move to position eigenstate $|x_b, t_b\rangle$; this gives us the probability amplitude of a particle at x_a at time t_a to move to position x_b at time t_b. The term on the right of equation 1 is known as the *Feynman path integral*. The measure $\int \mathcal{D}x\,(t)$ is an integral over all paths between x_a and x_b and $S_M\,[x\,(t)]$ is the Minkowski action of a particle of mass m on the path $x\,(t)$ which is defined by:

$$S_M\,[x\,(t)] = \int_{t_a}^{t_b} dt \left[\frac{1}{2} m \left(\frac{dx}{dt}\right)^2 - V(x)\right],\tag{5}$$

where $x\,(t_a) = x_a$ and $x\,(t_b) = x_b$ are the boundary conditions, and $V\,(x)$ (now a function) is the potential the particle is in.

Due to the oscillating integrand in equation 1, it is not clear the integral will converge, and the integral measure needs to be defined before we proceed. In order to do this we follow the steps taken by Creutz and Freedman in [11] to get equation 1 into a form we can work with.
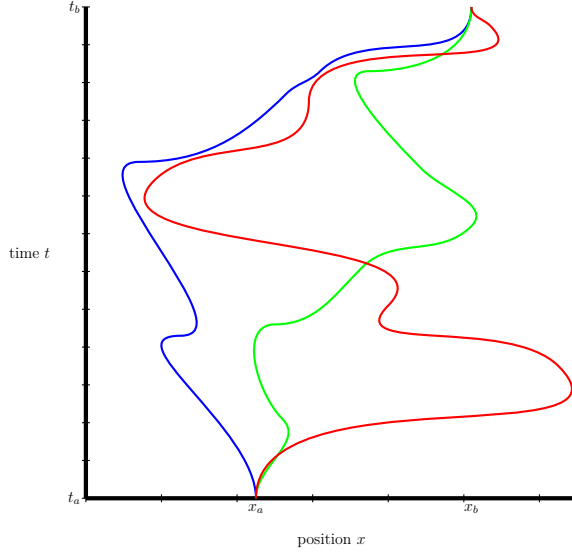
Figure 1. Three of the infinitely many paths from (x_a, t_a) to (x_b, t_b) that contribute to the path integral.
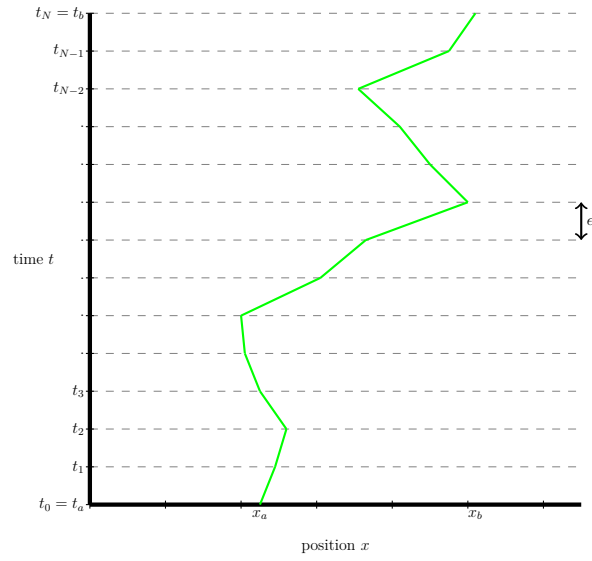

Figure 2. Discretising time and a path from (x_a, t_a) to (x_b, t_b) onto a lattice of time spacing ϵ.

2.1.2 Connecting to Statistical Physics

In the transition amplitude we integrate over an infinite number of paths, figure 1 shows three such paths that contribute to the path integral. The first step in calculating the path integral is to discretise time, this is shown in figure 2 where we have discretised the green path in figure 1 onto a time lattice and we assume the particle travels along a straight line between the time sites. For each time site on the lattice t_i we have a position $x_i = x(t_i)$ where $i = 0, 1, \ldots, N$. We introduce the notation:

$$\boldsymbol{x} = (x_0, x_1, \ldots, x_N) \tag{6}$$

to denote a particular path on the lattice where each position coordinate has been specified and we refer to equation 6 as a *configuration* on the lattice. In our notation for the labelling of the position eigenstates in equation 1 we also have $x_b = x_N = x(t_N)$ and $x_a = x_0 = x(t_0)$ in order to match with figure 2. ϵ is the spacing between lattice sites and so $\epsilon = \frac{t_b - t_a}{N} = t_{i+1} - t_i$ and for $k = 0, 1, \ldots, N$ we have $t_k = t_a + k\epsilon$. In order to discretise the action in equation 5 we approximate the derivative by a forward difference and the integral as a Riemann sum since we assume the particle is travelling on a straight line between lattice sites:

$$S_M(\boldsymbol{x}) = \sum_{i=0}^{N-1} \epsilon \left[\frac{1}{2} m \left(\frac{x_{i+1} - x_i}{\epsilon} \right)^2 - V(x_i) \right]. \tag{7}$$

Since for $i = 1, 2, \ldots, N-1, -\infty < x_i < \infty$ we may define the integral measure in equation 1 as:

$$\int_{x_a}^{x_b} \mathcal{D}x \sim \prod_{n=1}^{N-1} \int_{-\infty}^{\infty} dx_n, \tag{8}$$

which up to normalisation integrates over all possible routes through the lattice; so that for our discrete time lattice, the path integral is given by:

$$Z_{ba} \sim \int_{-\infty}^{+\infty} \prod_{i=1}^{N-1} dx_i \exp\left(\frac{i}{\hbar} S_M(\boldsymbol{x}) \right). \tag{9}$$

In the limit that $N \to \infty$ (or equivalently $\epsilon \to 0$) we recover (up to normalisation) equation 1 from equation 9 exactly. The normalisation in this expression is irrelevant since we will see that in the expectation values we wish to calculate any normalisation would cancel anyway.

In order to work with the discrete path integral we have one final step. We make a *Wick rotation* into imaginary time; this is done via the substitution:

$$\tau = it. \tag{10}$$

Applying this to the discretised theory developed above by defining $a = i\epsilon$, we now have a lattice in imaginary time, of lattice spacing a, substituting a into equation 7:

$$S_M(\boldsymbol{x}) = i \sum_{i=0}^{N-1} a \left[\frac{1}{2} m \left(\frac{x_{i+1} - x_i}{a} \right)^2 + V(x_i) \right] \tag{11}$$

$$= i S_E(\boldsymbol{x}), \tag{12}$$

where we have redefined the notation slightly so that $x_i = x(\tau_i)$ since we now have a lattice in imaginary time. The quantity

$$S_E(\boldsymbol{x}) \equiv \sum_{i=0}^{N-1} a \left[\frac{1}{2} m \left(\frac{x_{i+1} - x_i}{a} \right)^2 + V(x_i) \right] \tag{13}$$

is the discretised *Euclidean action*; it has this name because the effect of the Wick transformation is that it turns the Minkowski metric ds_M on the coordinates (x, y, z, t) into the Euclidean metric ds_E on the coordinates (x, y, z, τ) and vice-versa:

$$ds_M^2 = -dt^2 + dx^2 + dy^2 + dz^2 \tag{14}$$
$$= d\tau^2 + dx^2 + dy^2 + dz^2 \tag{15}$$
$$= ds_E^2. \tag{16}$$

Equation 12 is a very useful result since upon substitution into the discrete path integral we find:

$$Z_{ba} \sim \int_{-\infty}^{+\infty} \prod_{i=1}^{N-1} dx_i \exp \left(-\frac{1}{\hbar} S_E(\boldsymbol{x}) \right) \tag{17}$$

which we refer to as the *discrete euclidean path integral* and will converge since the integrand is now exponentially suppressed. We will impose periodic boundary conditions by identifying the first and last lattice sites (i.e. take $x_b = x_a$) and then integrate over that site, which can be quantified through:

$$Z = \text{Tr}(Z_{ba}) = \int dx_a \int dx_b \delta(x_b - x_a) Z_{ba} \tag{18}$$

$$= \int_{-\infty}^{+\infty} \prod_{i=0}^{N-1} dx_i \exp \left(-\frac{1}{\hbar} S_E(\boldsymbol{x}) \right), \tag{19}$$

where in equation 19 the identification of the first and last lattice sites is now implicit. We have the standard result from statistical physics that for a system with a fixed number N of continuous degrees of freedom labelled by x_i for $i = 0, 1, \ldots N - 1$, so that the vector $\boldsymbol{x} = (x_0, \ldots, x_{N-1})$ describes the system, then the canonical partition function is given by:

$$\mathcal{Z} \sim \int_{-\infty}^{+\infty} \prod_{i=0}^{N-1} dx_i \exp(-\beta H(\boldsymbol{x})), \tag{20}$$

with $\beta = \frac{1}{k_B T}$ where T is the system temperature and k_B is the Boltzmann constant. Comparing equation 19 to equation 20 we can see that the discretised Euclidean path integral is a classical canonical partition function on a system with N degrees of freedom, provided that we take:

$$\frac{1}{\hbar} S_E(\boldsymbol{x}) = \beta H(\boldsymbol{x}), \tag{21}$$

and impose periodic boundary conditions. We then have a Boltzmann factor given by $\exp \left(-\frac{1}{\hbar} S_E(\boldsymbol{x}) \right)$. So in summary we now have a classical interpretation of our quantum calculation of the path integral; our lattice is essentially a one dimensional crystal of size N

at temperature T with a continuous variable x_i at each crystal site, and its classical Hamiltonian (in the statistical system and units where $\hbar = \beta = 1$) is given by $S_E(\boldsymbol{x})$ which couples the nearest neighbours variables x_i and places each variable in its own potential.

Since we are interested in calculating properties of the quantum system, e.g. position expectation, ground state energy, the first excited state energy etc. we will explore how the quantum expectation values relate to those of the statistical system. Note that upon applying the Wick transformation and periodic boundary conditions to the continuum theory path integral in equation 3:

$$Z = \int_{-\infty}^{+\infty} dx \, \langle x | e^{-(\tau_b - \tau_a)\hat{H}/\hbar} | x \rangle = \mathrm{Tr}\left(e^{-(\tau_b - \tau_a)\hat{H}/\hbar} \right). \tag{22}$$

Following [11] we define for any operator $\hat{A}$ the quantity:

$$\langle \hat{A} \rangle = \frac{\mathrm{Tr}\left(e^{-(\tau_b - \tau_a)\hat{H}/\hbar} \hat{A} \right)}{\mathrm{Tr}\left(e^{-(\tau_b - \tau_a)\hat{H}/\hbar} \right)} \tag{23}$$

and after repeating the above discretisation theory for equation 23 can be written as [11]:

$$\langle \hat{A} \rangle = \frac{\int_{-\infty}^{+\infty} \prod_{i=0}^{N-1} dx_i A(x_0, \ldots, x_{N-1}) \exp\left(-\frac{1}{\hbar} S_E(\boldsymbol{x}) \right)}{\int_{-\infty}^{+\infty} \prod_{i=0}^{N-1} dx_i \exp\left(-\frac{1}{\hbar} S_E(\boldsymbol{x}) \right)} \tag{24}$$

$$= \langle A \rangle \tag{25}$$

where $A(x_0, \ldots x_{N-1})$ is a now a function on our lattice position variables. Equation 24 is precisely the statistical expectation value $\langle A \rangle$ of the function A on our system with the canonical distribution function:

$$P(\boldsymbol{x}) = \frac{1}{Z} \exp\left(-\frac{1}{\hbar} S_E(\boldsymbol{x}) \right) \tag{26}$$

where Z is the partition function for the discrete theory derived as above. We can see from equation 24 that the normalisation in Z is indeed not relevant for calculating such expressions as it would cancel between the numerator and the denominator in the discretisation. It can also be shown that in the limit that $\tau_b - \tau_a \to \infty$, that is, for a large enough lattice equation 23 takes the form [11]

$$\langle \hat{A} \rangle = \frac{\sum_{n=0}^{\infty} e^{-\frac{1}{\hbar} E_n (\tau_b - \tau_a)} \langle n | \hat{A} | n \rangle}{\sum_{n=0}^{\infty} e^{-\frac{1}{\hbar} E_n (\tau_b - \tau_a)}} \tag{27}$$

$$= \langle 0 | \hat{A} | 0 \rangle. \tag{28}$$

So, in order to calculate the expectation of some operator e.g. position, position squared or the energy in the the ground state, we can utilise equation 24 and calculate the expectation of the corresponding function on the statistical system under the canonical distribution $\langle A \rangle$[1].

---

[1]There is a clash of notation with regards to using the angle brackets here. However, since it is useful to know in what context the values are being calculated, we will always use hats to denote quantum expectations and include the state with respect to which the expectation is being calculated, i.e. $\langle 0 | \hat{A} | 0 \rangle$ and use $\langle A \rangle$ for the statistical expectation. This distinction is somewhat redundant since we have seen the two quantities are equivalent for a large enough lattice, but it is useful to determine the context in which we are working.

In practice the high dimensional integral in 24 is not possible to compute analytically so we use Monte Carlo methods which we develop in the next chapter.

Due to divergences in the kinetic term of the expectation of the action $\langle S_E \rangle$ of $\mathcal{O}\left(1/a\right)$[11] that occur in the calculation for the ground state energy given by:

$$E_0 = \langle 0 | \hat{H} | 0 \rangle \tag{29}$$

$$= \lim_{\tau_b - \tau_a \to \infty} \frac{\text{Tr}\left(e^{-(\tau_b - \tau_a)\hat{H}/\hbar}\hat{H}\right)}{\text{Tr}\left(e^{-(\tau_b - \tau_a)\hat{H}/\hbar}\right)} \tag{30}$$

$$= \lim_{\tau_b - \tau_a \to \infty} \frac{-1}{\tau_b - \tau_a} \frac{\partial}{\partial\left(\hbar^{-1}\right)} \ln Z \tag{31}$$

$$= \lim_{\tau_b - \tau_a \to \infty} \frac{-1}{\tau_b - \tau_a} \frac{\partial}{\partial\left(\hbar^{-1}\right)} \ln \int_{-\infty}^{+\infty} \prod_{i=0}^{N-1} dx_i \exp\left(-\frac{1}{\hbar}S_E\left(\boldsymbol{x}\right)\right), \tag{32}$$

in the discrete theory, we need an alternative method for calculating the lowest lying energy eigenvalues in our system as $a \to 0$. Rather than use the *point splitting* fix given in [16] we will follow the method used in [11], [13] and [14] and utilise the quantum *virial theorem* which relates the expectation value of kinetic energy to that of the potential for stationary states as:

$$2 \langle n | \hat{T} | n \rangle = \langle n | \hat{x}\hat{V}'\left(\hat{x}\right) | n \rangle \tag{33}$$

the derivation of which can be found in [17] and [18]. Using equation 24 we then have that the lowest lying energy eigenstates can be calculated as:

$$E_0 = \langle 0 | \hat{H} | 0 \rangle \tag{34}$$

$$= \langle 0 | \hat{T} + \hat{V} | 0 \rangle \tag{35}$$

$$= \langle 0 | \frac{1}{2}\hat{x}\hat{V}'\left(\hat{x}\right) + \hat{V}\left(\hat{x}\right) | 0 \rangle \tag{36}$$

$$= \langle \frac{1}{2}xV'\left(x\right) + V\left(x\right) \rangle \tag{37}$$

for a large enough lattice. It should be noted that the virial theorem as stated in 33 is a continuum results and its applicability to the discrete case may not be valid so correction terms may be needed. Indeed, we found that when calculating the first excited state energy via the naive method of computing $E_1 = \Delta E + E_0$ where the energy gap $\Delta E = E_1 - E_0$ was computed via the correlation function which is a result from discrete theory, and E_0 via the virial theorem, a continuum result, there was a discrepancy in the case of the harmonic oscillator with the analytic result for E_1 which can be computed exactly.

It is interesting to note that in quantum mechanics, due to the uncertainty principle $\Delta x \Delta p \geq \frac{\hbar}{2}$, $\hbar$ provides a measure of quantum fluctuations in our system. As $\hbar \to 0$ we recover classical physics and in this limit the only path in the path integral that contributes to the transition amplitude is the classical one. On the other hand, in statistical mechanics T provides a measure of statistical fluctuations in our system, and in the limit that $T \to 0$ these fluctuations go to zero. Hence taking the limit that $\hbar \to 0$ and $T \to 0$ we see statistical mechanics on a (real) crystal lattice is equivalent quantum mechanics in imaginary time.

3 Methods

3.1 Monte Carlo Methods

Monte Carlo methods provide a way of numerically evaluating integrals where the error is independent of the dimensionality of the integral. For the purposes of our work we will be interested in estimating expectation values of functions (which are the observables of our statistical system) $\phi(X)$ of some d-dimensional random variable X under some probability distribution with density function $f(\boldsymbol{x})$:

$$\langle \phi(X) \rangle = \int \prod_{i=1}^{d} dx_i f(\boldsymbol{x}) \phi(\boldsymbol{x}), \tag{38}$$

since this takes the form of the path integral we wish to calculate, where $\boldsymbol{x}$ correspond to the configurations on the lattice, $f(\boldsymbol{x})$ the canonical distribution and $\phi(\boldsymbol{x})$ some function on the lattice such as mean position. In order to estimate the expectation value/integral we approximate the expectation as a sample mean on some finite set of M samples of X as (we will refer to this as the *Monte Carlo estimate*):

$$\langle \phi(X) \rangle \approx \frac{1}{M} \sum_{n=1}^{M} \phi(\boldsymbol{x}_n). \tag{39}$$

It can be shown that the error in 39 is always $\mathcal{O}\left(1/\sqrt{M}\right)$ independent of the dimension of the integral, and that the Monte method wins over quadrature methods (traditional numerical integration schemes) in terms of computational cost vs. accuracy for $d > 3$ [19]. It is now just a question of how to choose the samples $\boldsymbol{x}_n$ to get a expectation value with a small error. From the form of equation 38 drawing uniform samples from the support of f is clearly unwise, since this would give equal weight to contributions to our approximation of the expectation value for samples that are very unlikely to be realised by the random variable X under the density function f as those which are very likely to be realised. We therefore choose to draw samples according to the distribution of X, this is known as *importance sampling*. In order to draw samples according to a probability distribution we use a *Markov chain* which for our purposes can be defined as a sequence of random variables (samples) for which the probability of the next random variable in the sequence $\boldsymbol{x}'$ taking on a particular value depends only on the value of the current random variable $\boldsymbol{x}$, we will denote this probability by $P(\boldsymbol{x} \to \boldsymbol{x}')$. By starting from some arbitrary state (which in our simulation is a configuration) through the stochastic sequence of configurations we will eventually end up sampling from the desired equilibrium distribution with density function $f(\boldsymbol{x})$. This idea of a *Markov Chain Monte Carlo* method was first introduced by Metropolis et al. in [2] via the Metropolis algorithm. Hybrid Monte Carlo is a method for constructing such a Markov chain and it is discussed in the next section.

A Markov chain will converge on a distribution with density function $P(\boldsymbol{x})$ provided the chain is ergodic and obeys detailed balance. Ergodicity is the condition that any sample

that can be realised under the probability distribution can be reached through the Markov chain and the detailed balance condition is given by:

$$P\left(\boldsymbol{x}\right)P\left(\boldsymbol{x}\to\boldsymbol{x}'\right)=P\left(\boldsymbol{x}'\right)P\left(\boldsymbol{x}'\to\boldsymbol{x}\right). \tag{40}$$

During an actual simulation, one only starts to calculate the values of observables via equation 39 after an equilibration period, this is due to the fact that initially samples will not be drawn according to the correct probability distribution and it is only once the Markov chain has converged that they will.

We may consider our Monte Carlo simulation as consisting of three main stages. First to start the simulation we provide an initial configuration to begin the chain. We then have to perform a sufficient number of updates/steps of the chain such that we are drawing from the correct distribution, which in our case is the canonical distribution for our statistical system. Once we are at equilibrium we may compute the values of any observables on any configuration and use them to calculate the Monte Carlo estimate via equation 39. In order to provide an initial lattice configuration we found that after some experimentation, a reasonable method was to provide a so called *hot start* by uniformly distributing the position values of the lattice sites on the interval $[-1, 1]$. To determine when and if samples are being drawn from the equilibrium distribution one can observe the evolution of some set of observables, for example position, and see when these values begin to converge. This method has potential issues in that there are the possibilities of *metastable states* [20] for which it may seem as if the system has reached equilibrium, when in fact it is sampling from region of the configuration space for which it is metastable, i.e. it will remain sampling from this region for a long period of time but will not remain indefinitely. This issue will arise for the case of the anharmonic oscillator, where the possibility of the particle getting "stuck" in one well can lead to Monte Carlo estimates on $\langle x \rangle$ being non-zero for a symmetric well about zero. However, in this case we are able to identify when this metastability has occurred, since by symmetry we know the true answer. Once we have achieved equilibrium we are able to evaluate observables via the configurations, however due to the possibility of correlations between samples in the Markov chain this is not entirely trivial; the analysis of recorded data is discussed at the end of this chapter.

3.2 Hybrid Monte Carlo

3.2.1 Hamiltonian Dynamics

Due to the fundamental importance of Hamiltonian dynamics in the HMC algorithm in this section we provide a quick review of Hamilton's equations and explain the numerical method used to integrate these equations so that they may be used in computational simulations.

Hamiltonian Dynamics enables us to solve the time evolution of a system of canonical coordinates $(\boldsymbol{q}, \boldsymbol{p})$ where the d-dimensional $\boldsymbol{q} = (q_1, q_2, \ldots, q_d)$ and $\boldsymbol{p} = (p_1, p_2, \ldots, p_d)$ vectors are

called *position* and *momentum* respectively. The system is then characterised by a Hamiltonian function $H(\boldsymbol{q}, \boldsymbol{p})$ on the $2d$-dimensional *phase space* occupied by the $(\boldsymbol{q}, \boldsymbol{p})$ vectors. **Hamilton's equations**:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \tag{41}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \tag{42}$$

for $i = 1, 2, \ldots d$, provide the time evolution of the vectors $\boldsymbol{q}$ and $\boldsymbol{p}$, so that if the state of the system is $(\boldsymbol{q}, \boldsymbol{p})$ at time t then Hamilton's equations give a mapping T_s to the state $(\boldsymbol{q}', \boldsymbol{p}')$ at time $t + s$. We will see in section 3.2.2 for the purposes of the implementing the HMC algorithm we may assume that the Hamiltonian is of the form:

$$H(\boldsymbol{q}, \boldsymbol{p}) = U(\boldsymbol{q}) + K(\boldsymbol{p}), \tag{43}$$

Where $U(\boldsymbol{q})$ and $K(\boldsymbol{p})$ are known as the *potential* and *kinetic* terms respectively. Equations 41 and 42 then become:

$$\frac{dq_i}{dt} = \frac{\partial K}{\partial p_i} \tag{44}$$

$$\frac{dp_i}{dt} = -\frac{\partial U}{\partial q_i}. \tag{45}$$

Since we intend to use the solutions to Hamilton's equations in the HMC simulation we need to find a way of integrating them numerically; in [3] it is argued that a successful method for this is *leapfrog integration*. In the leapfrog method to make the small time step ϵ from t to $t + \epsilon$ in position and momentum we use the equations:

$$p_i(t + \epsilon/2) = p_i(t) + \epsilon/2 \frac{dp_i}{dt}(t) \tag{46}$$

$$= p_i(t) - \epsilon/2 \frac{\partial U}{\partial q_i}(\boldsymbol{q}(t)), \tag{47}$$

$$q_i(t + \epsilon) = q_i(t) + \epsilon \frac{dq_i}{dt}(t + \epsilon/2) \tag{48}$$

$$= q_i(t) + \epsilon \frac{\partial K}{\partial p_i}(\boldsymbol{p}(t + \epsilon/2)), \tag{49}$$

$$p_i(t + \epsilon) = p_i(t + \epsilon/2) + \epsilon/2 \frac{dp_i}{dt}(t + \epsilon/2) \tag{50}$$

$$= p_i(t + \epsilon/2) - \epsilon/2 \frac{\partial U}{\partial q_i}(\boldsymbol{q}(t + \epsilon)), \tag{51}$$

where in the second equality in each step we have made the substitution for Hamilton's equations. In equation 46 we begin with a half step in each momentum component to go from t to $t+\epsilon/2$. Using the half stepped momenta we then do a full step from t to $t+\epsilon$ in each position component in equation 48. We end in equation 50 with a second half step in the

momentum components to go from $t + \epsilon/2$ to $t + \epsilon$ using the updated position components. Iterating this process we do as many steps of size ϵ as we like.

This generalises to making l steps in position and momentum and we can combine any two half steps in momentum that follow one another to simplify the algorithm which is often more convenient for computation purposes. Starting at $t = 0$:

1. We begin with a half step in momentum:

$$p_i \left(\epsilon/2\right) = p_i \left(0\right) - \epsilon/2 \frac{\partial U}{\partial q_i} \left(\boldsymbol{q} \left(0\right)\right), \tag{52}$$

and a full step in position:

$$q_i \left(\epsilon\right) = q_i \left(0\right) + \epsilon \frac{\partial K}{\partial p_i} \left(\boldsymbol{p} \left(\epsilon/2\right)\right). \tag{53}$$

2. Then make $l - 1$ alternating steps in momentum and position:

$$p_i \left(n\epsilon + \epsilon/2\right) = p_i \left(n\epsilon - \epsilon/2\right) - \epsilon \frac{\partial U}{\partial q_i} \left(\boldsymbol{q} \left(n\epsilon\right)\right), \tag{54}$$

$$q_i \left(n\epsilon + \epsilon\right) = q_i \left(n\epsilon\right) + \epsilon \frac{\partial K}{\partial p_i} \left(\boldsymbol{p} \left(n\epsilon + \epsilon/2\right)\right), \tag{55}$$

for $n = 1, \ldots, l - 1$.

3. Then a final half step in momentum:

$$p_i \left(l\epsilon\right) = p_i \left(l\epsilon - \epsilon/2\right) - \epsilon/2 \frac{\partial U}{\partial q_i} \left(\boldsymbol{q} \left(l\epsilon\right)\right). \tag{56}$$

The reasoning for the name *leapfrog* becomes apparent here since apart from the initial and final half steps in momentum, the momentum and position values are used alternatively to calculate the position and momentum respectively at the next step and the variables "leap" over one another at on offset of $\epsilon/2$.

An important feature of the leap frog integration scheme that makes it applicable to HMC is that it preserves volume in phase space exactly, which we will see leads to detailed balance. For Hamiltonian dynamics in continuous time, preservation of volume in phase space is known as Louville's theorem and its proof is given in [21]. For the discrete dynamics integrated via the leap frog scheme we can easily verify that volume preservation is still the case. First let us define the mappings $T_q \left(\epsilon\right) : \left(\boldsymbol{q}, \boldsymbol{p}\right) \to \left(\boldsymbol{q}', \boldsymbol{p}'\right)$ and $T_p \left(\epsilon\right) : \left(\boldsymbol{q}, \boldsymbol{p}\right) \to \left(\boldsymbol{q}', \boldsymbol{p}'\right)$ on phase space such that:

$$T_q \left(\epsilon\right) \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} = \begin{pmatrix} \boldsymbol{q} + \epsilon \nabla_p K \left(\boldsymbol{p}\right) \\ \boldsymbol{p} \end{pmatrix}, \tag{57}$$

$$T_p \left(\epsilon\right) \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} = \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} - \epsilon \nabla_q U \left(\boldsymbol{q}\right) \end{pmatrix}, \tag{58}$$

then (with a slight abuse of notation) the Jacobians of these mappings will be of the form:

$$J = \det \begin{pmatrix} \frac{\partial q'_i}{\partial q_j} & \frac{\partial q'_i}{\partial p_j} \\ \frac{\partial p'_i}{\partial q_j} & \frac{\partial p'_i}{\partial p_j} \end{pmatrix} \tag{59}$$

so that:

$$J_q = \det \begin{pmatrix} \delta_{ij} & \epsilon \partial_{p_i} \partial_{p_j} K(\boldsymbol{p}) \\ 0 & \delta_{ij} \end{pmatrix} = 1 \tag{60}$$

$$J_p = \det \begin{pmatrix} \delta_{ij} & 0 \\ -\epsilon \partial_{q_i} \partial_{q_j} U(\boldsymbol{q}) & \delta_{ij} \end{pmatrix} = 1. \tag{61}$$

For a single step of size ϵ the leap frog algorithm equations 46, 48 and 50 define a mapping on phase space $T(\epsilon) : (\boldsymbol{q}(t), \boldsymbol{p}(t)) \rightarrow (\boldsymbol{q}(t+\epsilon), \boldsymbol{p}(t+\epsilon))$ that can be written using the mappings given above as:

$$T(\epsilon) = T_p(\epsilon/2) T_q(\epsilon) T_p(\epsilon/2). \tag{62}$$

Since each map in the composition on the right hand side of this equation has a Jacobian of 1 we have that the Jacobian of $T(\epsilon)$ is 1. For the total trajectory which consists of l steps of size ϵ we may write the mapping that via leapfrog integration takes us from the start to the end of the trajectory as:

$$\text{traj}(\epsilon, l) = (T(\epsilon))^l \tag{63}$$

which of course also a Jacobian of 1 since each mapping in the composition has Jacobian 1 and hence the leapfrog integration preserves volume on the phase space since as a mapping its Jacobian is 1. We will return to this argument in section 4.1.2 where we will show that the leap frog integration scheme preserves volume even for the case of tempered dynamics.

For a Hamiltonian of the form in equation 43 with a symmetric kinetic term $K(\boldsymbol{p}) = K(-\boldsymbol{p})$ (which we will see is always the case for our work) then by negating the momenta at the end of a trajectory and applying the leapfrog equations a second time we will return to the initial point in phase space and hence the dynamics is reversible. Negating the momenta at the end of a leapfrog trajectory will also preserve the volume.

3.2.2   Sampling and the Hamiltonian

For a physical system in thermodynamic equilibrium with a fixed number of degrees of freedom and with energy function $E(\boldsymbol{x})$, where the vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)$ denotes the state/configuration of the system that depends on d variables, the canonical Boltzmann distribution (probability density function for the states of the system) in units where $k_b = 1$ is given by:

$$P(\boldsymbol{x}) = \frac{1}{\mathcal{Z}} \exp\left(-\frac{E(\boldsymbol{x})}{T}\right), \tag{64}$$

13

where $\mathcal{Z}$ is the canonical partition function for the system which provides the normalisation and is given by

$$\mathcal{Z} = \int \prod_{i=1}^{d} dx_i \exp\left(-\frac{E(\boldsymbol{x})}{T}\right). \tag{65}$$

Alternatively any probability distribution $P(\boldsymbol{x})$ can be constructed as the Boltzmann distribution for the energy function $E(\boldsymbol{x}) = -T(\log P(\boldsymbol{x}) + \log \mathcal{Z})$ for some choice of T and $\mathcal{Z}$.

A Hamiltonian of the form $H(\boldsymbol{q}, \boldsymbol{p}) = U(\boldsymbol{q}) + K(\boldsymbol{p})$ is an energy function on the $2d$-dimensional phase space given by position and momentum $(\boldsymbol{q}, \boldsymbol{p})$, with $\boldsymbol{q} = (q_1, q_2, \ldots, q_d)$ and $\boldsymbol{p} = (p_1, p_2, \ldots, p_d)$. We may define a Boltzmann distribution on those variables by:

$$P(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{\mathcal{Z}} \exp\left(-\frac{H(\boldsymbol{q}, \boldsymbol{p})}{T}\right), \tag{66}$$

with

$$\mathcal{Z} = \int \prod_{i=1}^{d} dx_i dp_i \exp\left(\frac{-H(\boldsymbol{q}, \boldsymbol{p})}{T}\right) \tag{67}$$

Then, if $H(\boldsymbol{q}, \boldsymbol{p}) = U(\boldsymbol{q}) + K(\boldsymbol{p})$ this factorises as:

$$P(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{\mathcal{Z}_{\boldsymbol{q}}} \exp\left(-\frac{U(\boldsymbol{q})}{T}\right) \frac{1}{\mathcal{Z}_{\boldsymbol{p}}} \exp\left(-\frac{K(\boldsymbol{p})}{T}\right) = P(\boldsymbol{q}) P(\boldsymbol{p}), \tag{68}$$

(where $\mathcal{Z}_{\boldsymbol{q}}$ and $\mathcal{Z}_{\boldsymbol{p}}$ denote the marginal partition functions) so the variables $\boldsymbol{q}$ and $\boldsymbol{p}$ and their respective probability distributions are independent with energy functions $U(\boldsymbol{q})$ and $K(\boldsymbol{p})$. This means that in order to sample $\boldsymbol{q}$ from the marginal distribution $P(\boldsymbol{q})$ we can sample from the joint distribution $P(\boldsymbol{q}, \boldsymbol{p})$ and disregard the $\boldsymbol{p}$.

In the HMC algorithm we employ Hamiltonian dynamics to sample from the joint distribution $P(\boldsymbol{q}, \boldsymbol{p})$. If we have some probability distribution $P(\boldsymbol{x})$ for states $\boldsymbol{x}$ we wish to sample, we may through equation 64 define an energy function $E(\boldsymbol{x})$ for that distribution. We then define $\boldsymbol{q} \equiv \boldsymbol{x}$ and $U(\boldsymbol{q}) \equiv E(\boldsymbol{x})$, introduce "fictitious momenta" $\boldsymbol{p}$ and define a kinetic energy $K(\boldsymbol{p})$ and the Hamiltonian $H(\boldsymbol{q}, \boldsymbol{p}) = U(\boldsymbol{q}) + K(\boldsymbol{p})$. Running the HMC algorithm generates samples according to the distribution given by equation 66, but by the factorization in equation 68 this gives samples of $\boldsymbol{x}$ according to our original distribution $P(\boldsymbol{x})$. In the HMC algorithm, we think of the variables of interest $\boldsymbol{x}$ as position $\boldsymbol{q}$ with conjugate momenta $\boldsymbol{p}$ which of course matches up nicely in physical applications where we are actually sampling position, although the algorithm is equally applicable in non-physical cases.

Since the fictitious momenta $\boldsymbol{p}$ are introduced artificially, we need to choose their probability distribution $P(\boldsymbol{p})$ by specifying the kinetic energy function $K(\boldsymbol{p})$. For the purposes of our work we may take:

$$K(\boldsymbol{p}) = \sum_{i=1}^{d} \frac{p_i^2}{2} \tag{69}$$

and $T = 1$ from now on.

3.2.3 Steps of the HMC Algorithm

Here we will outline the main steps of the HMC algorithm as defined in [1], [22] and [3] then give a more detailed explanation of each stage and its implementation. The two key steps are 1 and 2. In step 1 momentum is changed, where as in step 2 may change position and momentum. The canonical joint distribution for $(\boldsymbol{q}, \boldsymbol{p})$ is left invariant in both steps, and therefore is also invariant under their combination. The steps of the algorithm are:

0. Provide an initial sample $\boldsymbol{q}$

1. Generate $\boldsymbol{p}$ from the multivariate Gaussian Distribution $\mathcal{N}\left(\boldsymbol{0}, I_{d \times d}\right)$

2. Evolve the variables from $(\boldsymbol{q}, \boldsymbol{p})$ to $(\boldsymbol{q}', \boldsymbol{p}')$ by simulating Hamiltonian dynamics for l steps of size ϵ and then negate the momentum variables at the end of the trajectory. Accept the proposed state $(\boldsymbol{q}', -\boldsymbol{p}')$ as the next state in the Markov chain with probability:
$$\min\left[1, \exp\left(-H_{HMC}\left(\boldsymbol{q}', -\boldsymbol{p}'\right) + H_{HMC}\left(\boldsymbol{q}, \boldsymbol{p}\right)\right)\right]$$
If proposed state rejected, record current state $\boldsymbol{q}$ as sample (i.e. state started with), if successful set current state to proposed state $\boldsymbol{q} = \boldsymbol{q}'$ and record it as a sample.

3. Return to step 1 with $\boldsymbol{q}$.

In step 0 we provide an initial state. This can be chosen or generated randomly, however, since we normally disregard the samples generated in the early HMC iterations as discussed above this choice is not critical.

In step 1 we draw the fictitious momenta according to their probability distribution. The choice of kinetic energy in equation 69 and the form of equation 64 means that:

$$P\left(\boldsymbol{p}\right) = \frac{1}{Z_K} \exp\left(-\sum_{i=1}^{d} \frac{p_i^2}{2}\right) \tag{70}$$

$$= \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{p_i^2}{2}\right) \tag{71}$$

$$= \prod_{i=1}^{d} P\left(p_i\right) \tag{72}$$

where each $P\left(p_i\right)$ is the density function of a Gaussian distribution of mean 0 and variance 1. So the d momentum variables are independent with each component p_i of $\boldsymbol{p}$ drawn from a Gaussian $\mathcal{N}(0, 1)$, so we draw $\boldsymbol{p}$ from a multivariate Gaussian distribution of mean $\boldsymbol{0}$ and covariance $I_{d \times d}$. In this step $\boldsymbol{q}$ is not changed, and $\boldsymbol{p}$ gets drawn from the it's correct conditional distribution given $\boldsymbol{q}$ which is it's marginal distribution by the above independence, so the canonical joint distribution is left invariant.

In step 2 we evolve the variables $(\boldsymbol{q}, \boldsymbol{p})$ by integrating Hamilton's equations, this is often referred to as the *molecular dynamics* and we will define the map that carries out this

evolution as $H : (\boldsymbol{q}, \boldsymbol{p}) \to (\boldsymbol{q}', \boldsymbol{p}')$. In order to implement this in a computer simulation we use the leapfrog method described in section 3.2.1, although in general any numerical integration scheme that is exactly area preserving and reversible is valid [22]. We have a choice of the number of time steps l and their size ϵ; it is common practice in HMC to choose ϵ and l such that $l\epsilon = 1$. At the end of the trajectory, negating the momentum variables makes the Metropolis proposal symmetric, which is needed for detailed balance to hold. However, in practice since $K(\boldsymbol{p}) = K(-\boldsymbol{p})$ and the momentum is replaced anyway in the next iteration, there is no need for the negation. We can define a mapping for the momentum flip as $F : (\boldsymbol{q}', \boldsymbol{p}') \to (\boldsymbol{q}', -\boldsymbol{p}')$ so that the mapping that gives the proposed update is $F \circ H : (\boldsymbol{q}, \boldsymbol{p}) \to (\boldsymbol{q}', -\boldsymbol{p}')$ since H preserves volume, and trivially negating momenta with F also preserves volume the mapping $F \circ H$ will preserve volume. We accept or reject the proposed state with a probability given by:

$$P(\boldsymbol{q}', \boldsymbol{p}', \boldsymbol{q}, \boldsymbol{p}) = \min\left[1, \exp\left(-H_{HMC}(\boldsymbol{q}', \boldsymbol{p}') + H_{HMC}(\boldsymbol{q}, \boldsymbol{p})\right)\right], \tag{73}$$

which is known as a *Metropolis update*. The interpretation of 73 is that if the Hamiltonian decreases (or stays constant) as a result of moving to the proposed state the update is accepted with certainty. However, if the Hamiltonian increases as a result of the moving to the proposed state, the proposal is accepted with a probability that decreases for larger increases in the Hamiltonian. If the state is accepted then it becomes the current state i.e. $\boldsymbol{q} = \boldsymbol{q}'$, otherwise the current state remains as it was in step 1. In either case we then record the current state as a sample and return to step 1 with that state.

From the Metropolis update in equation 73 we can show that for smaller value of ϵ and larger l (i.e. smaller time steps and more of them) the probability of accepting an update approaches unity. To see this note that as $\epsilon \to 0$ and $l \to \infty$ the leap frog method approaches an exact integration of Hamilton's equations. It is easily shown that for continuous time the Hamiltonian is a conserved quantity:

$$\frac{dH}{dt} = \sum_{i=1}^{d}\left[\frac{dq_i}{dt}\frac{\partial H}{\partial q_i} + \frac{dp_i}{dt}\frac{\partial H}{\partial p_i}\right] = \sum_{i=1}^{d}\left[\frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial H}{\partial p_i}\right] = 0 \tag{74}$$

where in the second equality we have used Hamilton's equations. Therefore as $\epsilon \to 0$ and $l \to \infty$, $H_{HMC}(\boldsymbol{q}, \boldsymbol{p}) \to H_{HMC}(\boldsymbol{q}', \boldsymbol{p}')$ and so the probability of accepting a state is always 1.

As ϵ decreases l increases provided $l\epsilon$ is kept fixed, therefore more steps and hence a larger computational cost is required for proposal states with smaller ϵ in the trajectory. Conversely if we have a fixed computation time, then for smaller ϵ we will generate less proposals with a higher acceptance rate and hence our the error in the Monte Carlo estimate on any observable will be larger. It is shown in [23] that an acceptance rate of $\sim 65.1\%$ provides an optimal balance between the cost of generating proposals and the total number of samples.

HMC will be typically ergodic [3] however proving ergodicity is highly non-trivial, we will therefore assume that ergodicity of the HMC algorithm holds for our simulation and refer the reader to the mathematical literature on the subject such as [24], [25] and [26] for justification.

We can however show that HMC obeys detailed balance using a proof similar to that given in [1]. In order to do this we will (for notational simplicity) adopt the field theory notation and set $\boldsymbol{q} \to \phi$ and $\boldsymbol{p} \to \pi$ and redefine the canonical distribution in equation 66 as $P(\boldsymbol{q}, \boldsymbol{p}) \to p(\phi, \pi)$ where we are using a lower case p to emphasise the fact we are dealing with a probability density. $p(\phi, \pi) \, d\phi d\pi$ is then the probability of being in the volume given by the intervals $[\phi, \phi + d\phi]$ and $[\pi, \pi + d\pi]$ and since the kinetic term in the Hamiltonian is quadratic and therefore symmetric in π, $p(\phi, \pi) = p(\phi, -\pi)$. If $P((\phi, \pi) \to (\phi', \pi'))$ is the probability of moving from (ϕ, π) to (ϕ', π') the detailed balance condition for HMC is given by equation 40:

$$d\phi d\pi p(\phi, \pi) P((\phi, \pi) \to (\phi', -\pi')) = d\phi' d(-\pi') p(\phi', -\pi') P((\phi', -\pi') \to (\phi, \pi)) \quad (75)$$
$$= d\phi' d\pi' p(\phi', \pi') P((\phi', -\pi') \to (\phi, \pi)), \quad (76)$$

where in the second equality we have used preservation of volume under negation of momenta and the symmetry of the density function. We have shown preservation of volume in phase space for the leapfrog method, so that $d\phi d\pi = d\phi' d\pi'$. We may decompose the second factor in the LHS of equation 75 as:

$$P((\phi, \pi) \to (\phi', -\pi')) = P_{F \circ H}((\phi, \pi) \to (\phi', -\pi')) P_A((\phi, \pi) \to (\phi', -\pi')) \quad (77)$$

where $P_{F \circ H}$ is the probability of proposing the update under the molecular dynamics and momentum flip and P_A is the probability of actually accepting that move. By the fact that leapfrog is reversible we have that

$$P_{F \circ H}((\phi, \pi) \to (\phi', -\pi')) = P_{F \circ H}((\phi', -\pi') \to (\phi, \pi)) \quad (78)$$

and since we are using the Metropolis update step and the kinetic term is symmetric:

$$P_A((\phi, \pi) \to (\phi', -\pi')) = \min\left[1, \exp\left(-\left(H(\phi', -\pi') - H(\phi, \pi)\right)\right)\right] \quad (79)$$
$$= \min\left[1, \exp\left(-\left(H(\phi', \pi') - H(\phi, \pi)\right)\right)\right] \quad (80)$$
$$= P_A((\phi, \pi) \to (\phi', \pi')). \quad (81)$$

Starting from the LHS of equation 75 we get that:

$$d\phi d\pi p(\phi, \pi) P((\phi, \pi) \to (\phi', -\pi')) \quad (82)$$
$$= d\phi d\pi \frac{1}{\mathcal{Z}} e^{-H(\phi, \pi)} P_{F \circ H}((\phi, \pi) \to (\phi', -\pi')) \min\left[1, e^{\left(-\left(H(\phi', \pi') - H(\phi, \pi)\right)\right)}\right] \quad (83)$$
$$= d\phi d\pi \frac{1}{\mathcal{Z}} e^{-H(\phi', \pi')} P_{F \circ H}((\phi', -\pi') \to (\phi, \pi)) \min\left[e^{\left(-\left(H(\phi, \pi) - H(\phi', \pi')\right)\right)}, 1\right] \quad (84)$$
$$= d\phi' d\pi' \frac{1}{\mathcal{Z}} e^{-H(\phi', \pi')} P_{F \circ H}((\phi', -\pi') \to (\phi, \pi)) \min\left[1, e^{\left(-\left(H(\phi, \pi) - H(\phi', \pi')\right)\right)}\right] \quad (85)$$
$$= d\phi' d\pi' p(\phi', \pi') P((\phi', -\pi') \to (\phi, \pi)) \quad (86)$$

where in the first equality we have just used the definitions, in the second we have used reversibility of the leapfrog integration scheme and factored out an exponent from the min function and in the third equality we have used the volume preservation, so detailed balance holds for HMC.

3.3 Data Analysis

The results of our simulation will ultimately be average values with accompanying statistical errors for several measured observables. Because Monte Carlo simulations generate samples using a Markov chain, subsequent samples in the chain can be correlated, in which case naive error estimates can be wrong. In what follows we will briefly explore this issue and explain our method for dealing with this problem in the simulation.

Uncorrelated Data Say we have via our Markov chain of Monte Carlo generated configurations at equilibrium calculated the values of $x_1, \ldots, x_N$ of some observable (e.g. the ground state energy, or position expectation in the ground state). Then since the Markov chain is a sequence of random variables, each sample x_i is the realisation of some random variable X_i. Since the lattice configurations and therefore the samples that are functions of those configurations are generated at equilibrium according to the canonical distribution, these random variables have the same expectation value and variance:

$$\langle X_i \rangle = \langle X \rangle, \tag{87}$$

$$\sigma_{X_i}^2 = \langle (X_i - \langle X_i \rangle)^2 \rangle = \sigma_X^2 \tag{88}$$

and unbiased estimators for these values are [19]

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{N} X_i \tag{89}$$

$$\overline{\sigma}_X^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(X_i - \overline{X} \right)^2. \tag{90}$$

For X_i uncorrelated and $i \neq j$, $\langle X_i X_j \rangle = \langle X_i \rangle \langle X_j \rangle = \langle X \rangle^2$ from which it can be shown [19], that:

$$\sigma_{\overline{X}}^2 = \frac{1}{N} \sigma_X^2, \tag{91}$$

so that the statistical error (standard deviation) in the observable $\overline{X}$ is $\sigma_{\overline{X}}$ and by approximating σ_X by the (now biased) estimator $\overline{\sigma}_X$, for N uncorrelated measurements of some observable X our final estimate will be:

$$\overline{X} \pm \frac{\overline{\sigma}_X}{\sqrt{N}} \tag{92}$$

from which we can see the error decreases as $1/\sqrt{N}$ with the number of measurements N.

Correlated Data Subsequent samples generated in the HMC simulation will be correlated due to the fact that they are generated via a Markov chain and hence the above error analysis will not be valid in general. In order to quantify the correlation between samples we can

use the *autocorrelation function*. If the random variables are correlated, they will have a non-vanishing autocorrelation function which is given by:

$$C_X\left(X_i, X_{i+t}\right) = \langle\left(X_i - \langle X_i\rangle\right)\left(X_{i+t} - \langle X_{i+t}\rangle\right)\rangle = \langle X_i X_{i+t}\rangle - \langle X_i\rangle\langle X_{i+t}\rangle. \tag{93}$$

For invariance in the shift of the index $\langle X_i\rangle = \langle X\rangle$ (which we can assume here since we are always sampling from the canonical distribution at equilibrium) the correlation function only depends on the separation time t between samples in the chain, and utilising the translational invariance in the index i we are able to get a good estimate on the correlation function from the sum:

$$C_X\left(t\right) \approx \frac{1}{N}\sum_{i=1}^{N} C_X\left(X_i, X_{i+t}\right). \tag{94}$$

It is useful to normalise the autocorrelation function by:

$$\Gamma_X\left(t\right) \equiv \frac{C_X\left(t\right)}{C_X\left(0\right)}, \tag{95}$$

since equation 95 typically takes the form of a sum of exponentials however for large t we may truncate to the asymptomatically leading term [19]:

$$\Gamma_X\left(t\right) \sim \exp\left(-\frac{t}{\tau_{X,\exp}}\right), \tag{96}$$

where $\tau_{X,\exp}$ is the *exponential autocorrelation time* for X, which gives us a measure of the number of configurations in the Markov chain, before measurements we are working with become uncorrelated. It is important to note that different observables will have in general, different autocorrelation times. For example, we may on each configuration of the lattice compute position expectation, and position squared expectation, which may have different autocorrelation times. Therefore in order to get an overall measure of when all measurements become uncorrelated we should compute:

$$\tau_{\exp} = \sup_X \tau_{X,\exp} \tag{97}$$

for all observables X. If the computer time between subsequent measurements is t then autocorrelations lead to systematic errors of $\mathcal{O}\left(\exp\left(-t/\tau_{\exp}\right)\right)$[19].

For correlated random variables, equation 91 is no longer valid, however it can be shown for the correlated case [19] that:

$$\sigma_{\bar{X}}^2 \approx \frac{\sigma_X^2}{N} 2\tau_{X,\text{int}} \tag{98}$$

where the *integrated autocorrelation time* is defined by:

$$\tau_{X,\text{int}} = \frac{1}{2} + \sum_{t=1}^{N} \Gamma_X\left(t\right), \tag{99}$$

and provides a second measure (like the exponential autocorrelation time) of the number of configurations before the samples become uncorrelated. As before, the integrated autocorrelation time will vary for different measured quantities, so to find the number of configurations

before measurements are completely uncorrelated one should consider $\sup_X \tau_{X,\mathrm{int}}$. For the case of correlated samples, our estimates will be now:

$$\overline{X} \pm \sqrt{\frac{1}{N} 2\tau_{X,\mathrm{int}} \overline{\sigma}_X^2}, \tag{100}$$

and we can see that the error will be larger than in equation 92 where we assumed the samples were uncorrelated.

In practice obtaining the values of either the exponential, or integrated autocorrelation times is not always easy. Due to the statistical noise that appears in the autocorrelation function at larger times one needs to first plot the autocorrelation function, then either fit an exponential for $\tau_{X,\mathrm{exp}}$, or compute the sum in $\tau_{X,\mathrm{int}}$ for an appropriate range. For the systems we simulated computational costs are (relatively) small, so rather than compute the autocorrelation times after the simulation to determine the errors in any quantities, we introduce into the simulation a parameter that determines at what frequency measurements are actually recorded from configurations in the Markov chain. By increasing this parameter on preliminary test runs until $\tau_{X,\mathrm{int}}$ and $\tau_{X,\mathrm{exp}}$ are both less than one, we know that our measurements are always uncorrelated and we are therefore free to use the naive error in equation 92 with confidence, since we know our samples are uncorrelated. This makes the simulation somewhat simpler, since it allows us to do statistical analysis during the simulation, rather than saving all generated configurations and performing data analysis after.

Finally in section 3.2.2 we claimed the optimum acceptance rate was $\sim 65.1\%$, in what follows we will always tune our input parameters so that the success rate is close to and/or greater than this value.

4 Results and Discussion

4.1 Quantum Oscillators

In section 2.1.1 we showed that by discretising the path integral and performing a Wick rotation into imaginary time the path integral of quantum mechanics is a partition function and working now in units where $\hbar = T = k_B = 1$ we have a statistical system with the Boltzmann distribution:

$$P(\boldsymbol{x}) = \frac{1}{Z} \exp\left(-S_E(\boldsymbol{x})\right). \tag{101}$$

The probability distribution of equation 101 is the distribution we wish to draw our samples according to for the Monte Carlo simulations, where the samples $\boldsymbol{x}_i = \left(x_{i_0}, \ldots, x_{i_{n-1}}\right)$ are lattice configurations, therefore we follow the method of section 3.2.2 with this distribution, taking $\boldsymbol{q} \equiv \boldsymbol{x}$,

$$U(\boldsymbol{q}) \equiv S_E(\boldsymbol{x}) \tag{102}$$

$$= \sum_{i=0}^{N-1} a \left[\frac{1}{2} m \left(\frac{x_{i+1} - x_i}{a}\right)^2 + V(x_i)\right] \tag{103}$$

and the kinetic energy $K(\boldsymbol{p})$ as in equation 69. We define our HMC Hamiltonian:

$$H_{hmc}(\boldsymbol{q}, \boldsymbol{p}) \equiv K(\boldsymbol{p}) + U(\boldsymbol{q}) \tag{104}$$

$$= \sum_{i=0}^{N-1} \frac{p_i^2}{2} + S_E(\boldsymbol{q}) \tag{105}$$

$$= \sum_{i=0}^{N-1} \frac{p_i^2}{2} + \sum_{i=0}^{N-1} \epsilon \left[\frac{1}{2} m \left(\frac{q_{i+1} - q_i}{\epsilon} \right)^2 + V(q_i) \right], \tag{106}$$

we can then use this Hamiltonian to generate samples with the HMC algorithm for any 1-dimensional quantum mechanical system, by specifying the potential $V(x)$. Once we reach equilibrium we can compute averages of any moments $\overline{\langle x^p \rangle}$ or correlation functions $\overline{\langle x_i x_{i+m} \rangle}$ which via the Monte Carlo estimate of equation 39 gives an approximation of the true expectation value under the Boltzmann distribution. We then know from section 2.1.1 that these statistical expectations correspond to quantum expectation values in the ground state, i.e. for a large enough lattice $\langle \hat{x}^p \rangle = \overline{\langle x^p \rangle}$ and $\langle \hat{x}(0)\hat{x}(\tau) \rangle = \overline{\langle x_i x_{i+m} \rangle}$ where $\tau = ma$. With the choice of kinetic energy given in equation 69 and $U(\boldsymbol{q}) \equiv S_E(\boldsymbol{x})$ this leads to the approximate solutions to Hamilton's equations in leapfrog implementation for l steps of size ϵ taking the form:

1. Half step in momentum:

$$p_i(\epsilon/2) = p_i(0) - \epsilon/2 \left[\frac{m}{a} (2q_i(0) - q_{i-1}(0) - q_{i+1}(0)) + a \frac{\partial V(q_i(0))}{\partial q_i} \right], \tag{107}$$

full step in position:
$$q_i(\epsilon) = q_i(0) + \epsilon p_i(\epsilon/2). \tag{108}$$

2. Make $l - 1$ alternating steps in momentum and position:

$$p_i(n\epsilon + \epsilon/2) = p_i(n\epsilon - \epsilon/2) - \epsilon \left[\frac{m}{a} (2q_i(n\epsilon) - q_{i-1}(n\epsilon) - q_{i+1}(n\epsilon)) + a \frac{\partial V(q_i(n\epsilon))}{\partial q_i} \right], \tag{109}$$

$$q_i(n\epsilon + \epsilon) = q_i(n\epsilon) + \epsilon p_i(n\epsilon + \epsilon/2), \tag{110}$$

for $n = 1, \ldots, l - 1$.

3. Final half step in momentum:

$$p_i(\epsilon l) = p_i(l\epsilon - \epsilon/2) - \epsilon/2 \left[\frac{m}{a} (2q_i(\epsilon l) - q_{i-1}(\epsilon l) - q_{i+1}(\epsilon l)) + a \frac{\partial V(q_i(\epsilon l))}{\partial q_i} \right], \tag{111}$$

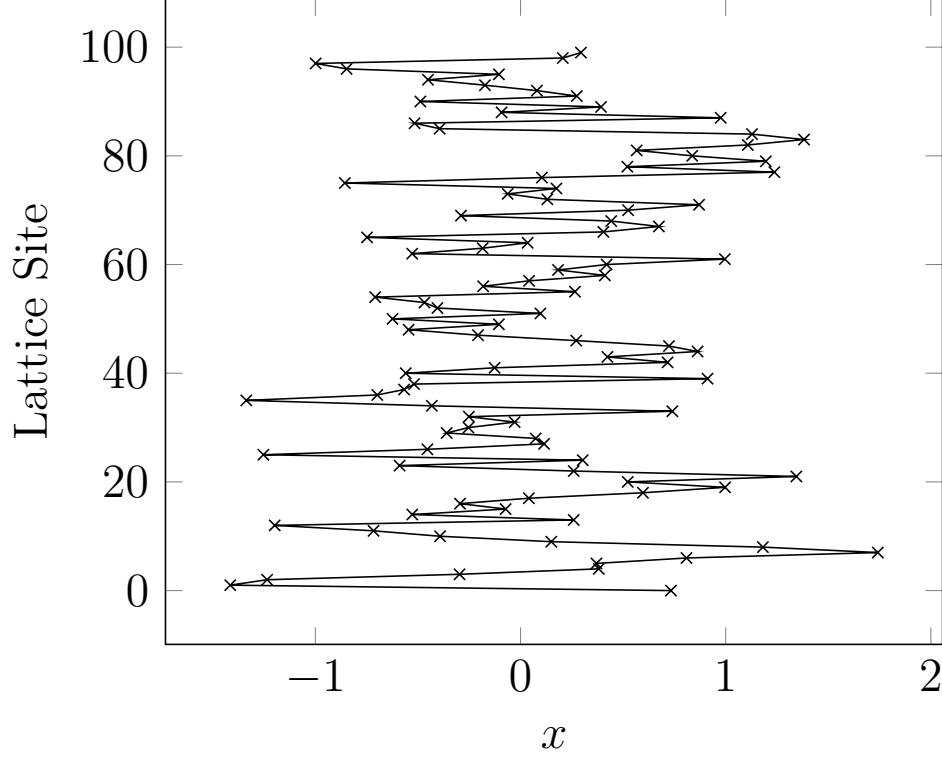where we are, as mentioned previously emposing the periodic boundary conditions that $q_N = q_0$.

Figure 3. Typical configuration for the harmonic oscillator with $\mu^2 = 1, m = 1, a = 1, L = 1000, d = 0.1, N = 10, \text{configurations} = 100000, \text{burn period} = 1000$.

4.1.1 Quantum Harmonic Oscillator

We begin by applying the HMC algorithm to the quantum harmonic oscillator. Since this system has an analytic solution it is a good testing ground for checking our simulation is correct. The potential for the quantum harmonic oscillator can be parametrised as:

$$V(x) = \frac{1}{2}\mu^2 x^2 \tag{112}$$

so that

$$H_{hmc}(\boldsymbol{q}, \boldsymbol{p}) = \sum_{i=0}^{N-1} \frac{p_i^2}{2} + \sum_{i=0}^{N-1} \epsilon \left[\frac{1}{2}m \left(\frac{q_{i+1} - q_i}{\epsilon} \right)^2 + \frac{1}{2}\mu^2 q_i^2 \right] \tag{113}$$

where $\mu \in \mathbb{R}$.

Typical Configuration Figure 3 shows a typical configuration for the HMC simulation of the harmonic oscillator. This is just a plot of the position variable at each lattice site.

Mean Square Position The first quantity we calculate is the mean square position which for the statistical mechanical system is $\langle x^2 \rangle$ and corresponds to $\langle 0| \hat{x}^2 |0\rangle$ for the quantum

22
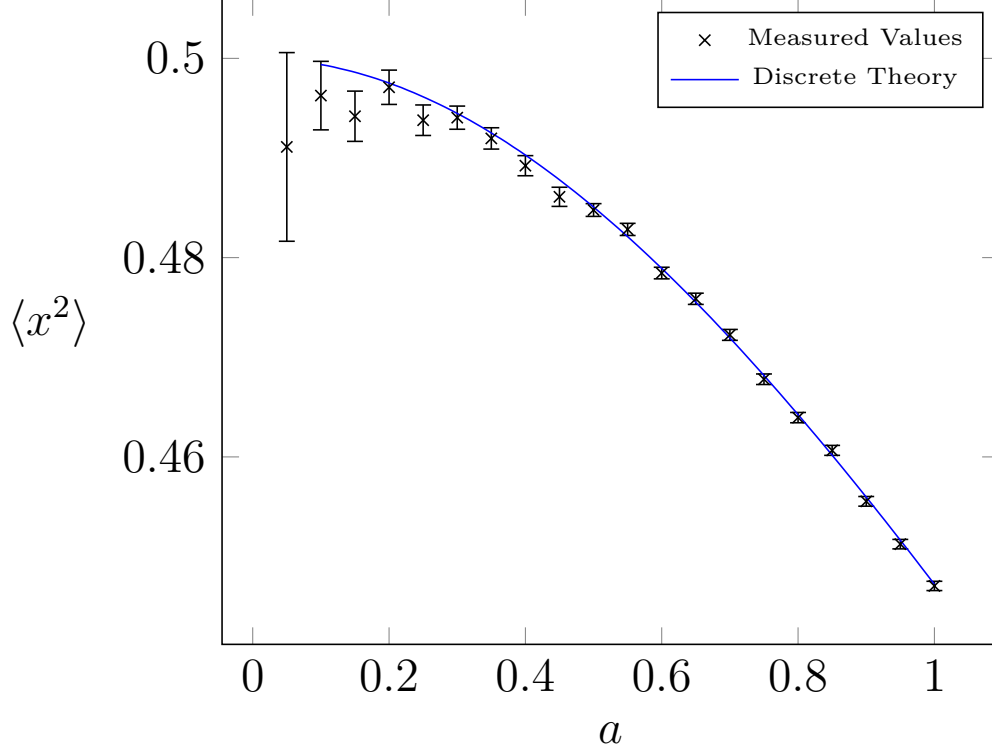
Figure 4. Relationship between lattice spacing and expectation of position squared for the harmonic oscillator with $\mu^2 = 1, m = 1$.

mechanical system. For the discrete theory, the value this quantity for an imaginary time lattice of spacing a and size N can be calculated exactly to be:

$$\langle x^2 \rangle = \frac{1}{2\mu \left(m + \frac{a^2 \mu^2}{4} \right)^{\frac{1}{2}}} \left(\frac{1 + R^N}{1 - R^N} \right), \tag{114}$$

where

$$R = 1 + \frac{a^2 \mu^2}{2m} - \frac{a\mu}{\sqrt{m}} \left(1 + \frac{a^2 \mu^2}{4m} \right)^{\frac{1}{2}}, \tag{115}$$

see [11], [14], [12] or appendix B for a full derivation of equation 114. Running the HMC simulation at a range of lattice spacings we obtain the results in figure 4. The errors grow as we approach the continuum limit which is due to the fact that for smaller lattice spacing the autocorrelation time increases, so we save samples at a lower frequency and hence for a fixed number of generated configurations, we have less samples.

Energy Levels In order to calculate the ground state energy of the system we use the virial theorem as discussed in section 2.1.1. Applied to the harmonic potential and remembering that measurements of moments $\langle x^p \rangle$ correspond to quantum expectations in the ground state the virial theorem gives:

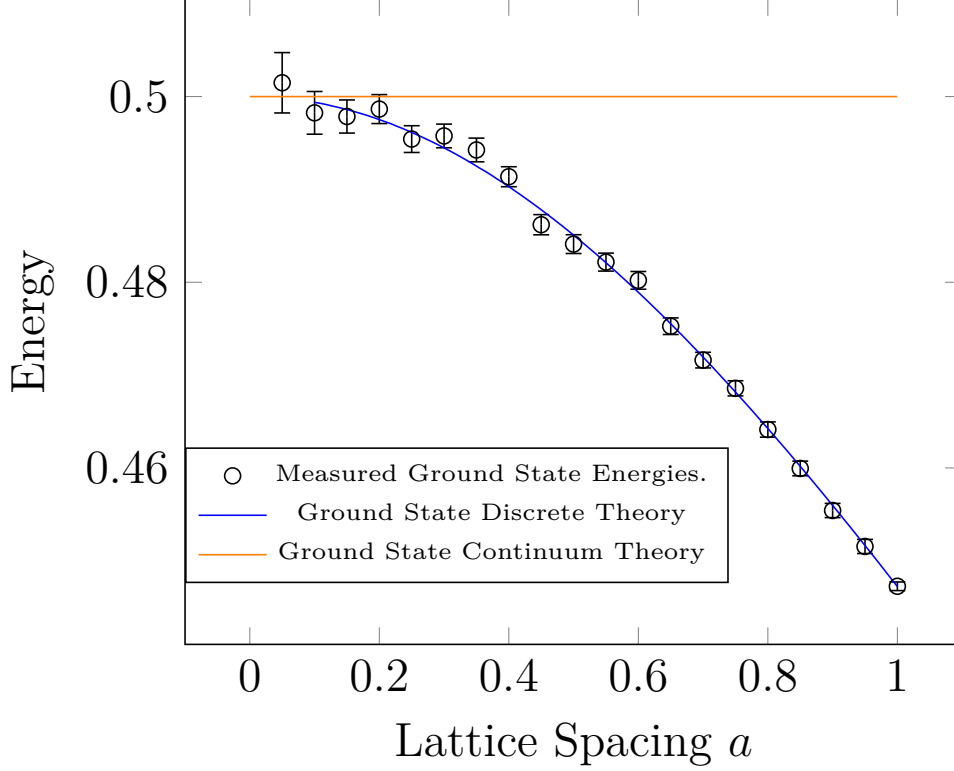$$E_0 = \mu^2 \langle x^2 \rangle \tag{116}$$

23

Figure 5. Ground state energies for the HMC simulation with the continuum and discrete theory results.

Of course when $\mu = m = 1$ the calculated values for the ground state energy are the same as those for $\langle x^2 \rangle$ in figure 4, however figure 5 shows the ground state energy on a larger lattice with the accompanying discrete and continuum theory results for comparison.

As $a \to 0$ and we approach the continuum we see that E_0 approaches the value of $\frac{1}{2}$ which is given by the formula for the energy levels of a harmonic oscillator:

$$E_n = \mu \left(n + \frac{1}{2} \right), \tag{117}$$

when $n = 0$.

To calculate the energy gap between the first and ground state we use:

$$\Delta E \left(\delta \tau \right) = -\frac{1}{\delta \tau} \ln C \left(\delta \tau \right), \tag{118}$$

where:

$$C \left(\delta \tau \right) = \frac{\sum_{i=0}^{N-1} x_i x_{i+\delta \tau}}{\sum_{i=0}^{N-1} x_i^2} \tag{119}$$

is the *correlation function*. In equation 119 we have utilized the periodic boundary conditions of the lattice and calculated the correlation at each lattice site then averaged over them to increase the statistics.For an derivation of equation 118 see appendix Figures **??** and **??** show

24

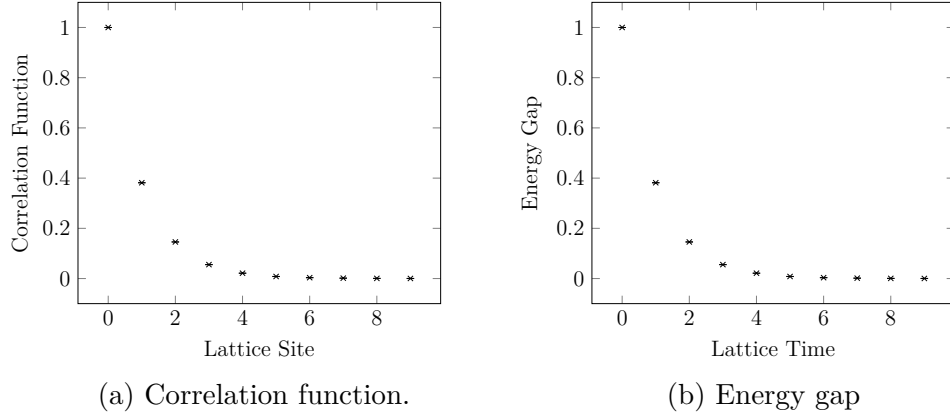(a) Correlation function.   (b) Energy gap

Figure 6. Energy gap and correlation function for the harmonic oscillator with $a = 1$, $\mu = 1$ and $m = 1$.

the energy gap and correlation function respectively on an HMC simulation of the harmonic oscillator. For small values of $\delta\tau$ we can see that $\Delta E \approx 1$ which is what we expect from equation 117. By averaging over the values for which $\Delta E \approx 1$ we can estimate the the first excited energies by subtracting the estimated ground state energy for various values of μ^2. These have been plotted in figure **??** alongside the ground state energies for a selection of lattice spacings along with the results from the discrete and continuous theory.

Why doesn't it stay constant?

Make data for the plot and include the plot

Ground State Probability Density Function Our method for measuring the density function of the ground state follows that of [11]. We discretise position into K bins of size Δx. On any given configuration we may then bin the position coordinate at each lattice site and create a histogram. Normalising the histogram gives the probability mass density for finding the particle in any interval $[x, x + \Delta x]$ and as $\Delta x \to 0$, this becomes the probability density function that is the amplitude squared of the wave function. Since we know that measurements made in our simulation correspond to ground state expectation values this wave function is that of the ground state. It can be shown (see appendix B or [11]) that for the discrete quantum harmonic oscillator the magnitude squared of the ground state wavefunction is given exactly by:

$$|\psi(x)|^2 = \left(\frac{\sqrt{5}}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{\sqrt{5}}{2}x^2\right) \tag{120}$$

when $m = \mu = 1$ and with a lattice spacing of $a = 1$ which should be contrast with the continuum result when $a \to 0$:

$$|\psi(x)|^2 = \frac{1}{\sqrt{\pi}} \exp -x^2. \tag{121}$$

Plots of the density functions for the discrete and continuum cases are shown in figure 8 along with the simulation results at the given parameters.
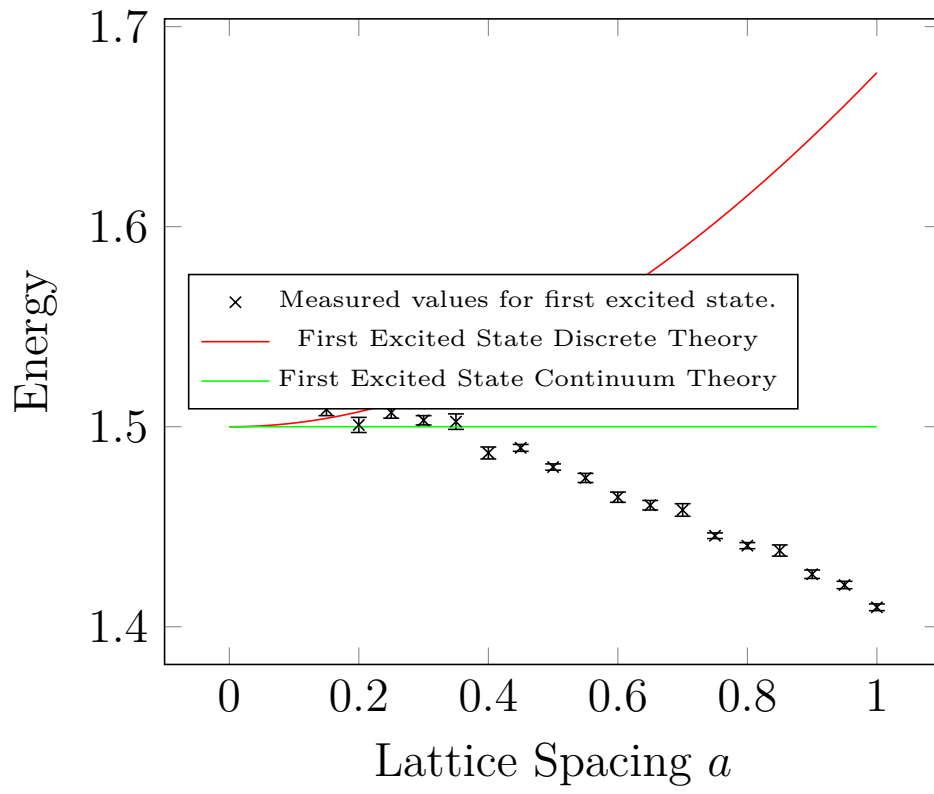
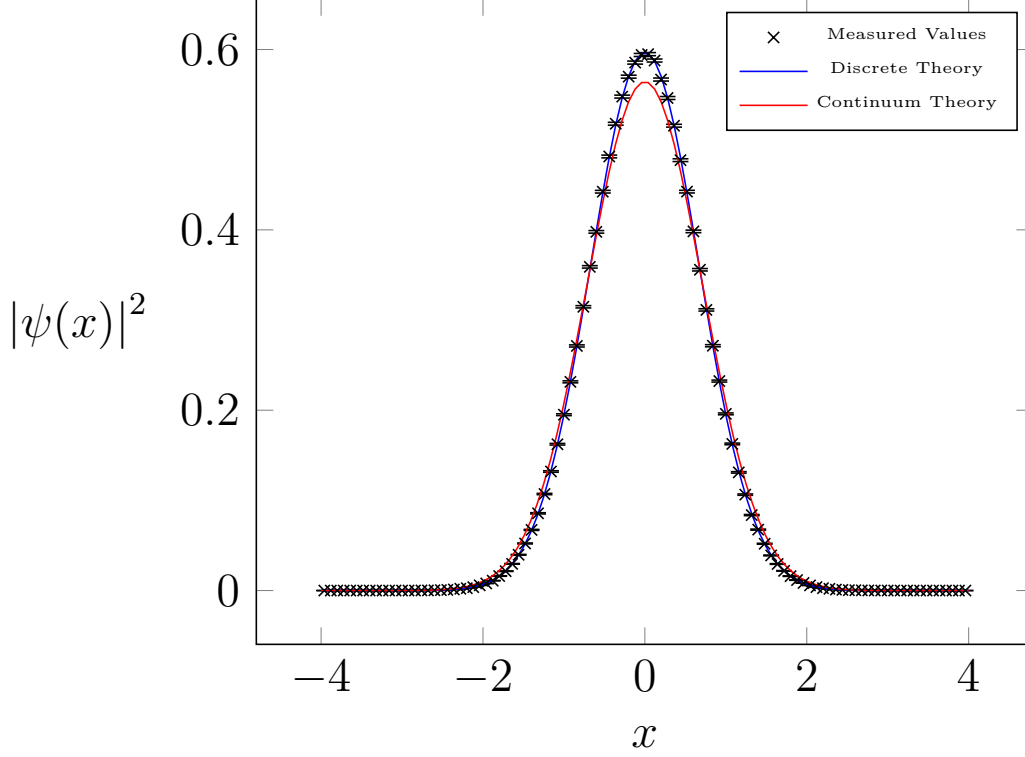Figure 7. First excited state of the harmonic oscillator on a 1000 site lattice of varying spacings.

Figure 8. Continuum, discrete and measured wave functions for the harmonic oscillator with $\mu^2 = 1, m = 1, a = 1, L = 1000, d = 0.1, N = 10,$ configurations $= 100000,$ burn period $= 1000$.

4.1.2 Quantum Anharmonic Oscillator

Since the anharmonic oscillator does not have an analytic solution for the properties we wish to calculate, it is a good first application of out HMC simulation for calculations we do not already know the answer to. Rather than use the more common form of the anharmonic quartic potential:

$$V\left(x\right) = \frac{1}{2}m\mu^2 x^2 + \lambda x^4 \tag{122}$$

where $\mu^2 \in \mathbb{R}$ and $\lambda \in \mathbb{R}_{>0}$, we use the potential:

$$V\left(x\right) = \lambda \left(x^2 - f^2\right) \tag{123}$$

with $\lambda \in \mathbb{R}_{>0}$ and $f^2 \in \mathbb{R}_{>0}$. This form of the potential has the advantage that its zeros coincide with its minima at $x = \pm f$. Since equation 123 is equivalent to equation 122 up to the additive constant λf^4, the action and consequently the HMC Hamiltonian for the system is the same up to the additive constant λf^4 also. Since the constant in the Hamiltonian vanishes when we take its derivatives in Hamilton's equations, and any constant term in the action amounts to rescaling the path integral in equation 1 by a complex exponential factor, which will therefore vanish when we take the magnitude to get the probability, this redefinition of the potential is valid. Of course it will have the effects of rescaling the energy eigen values by the additive constant λf^4.
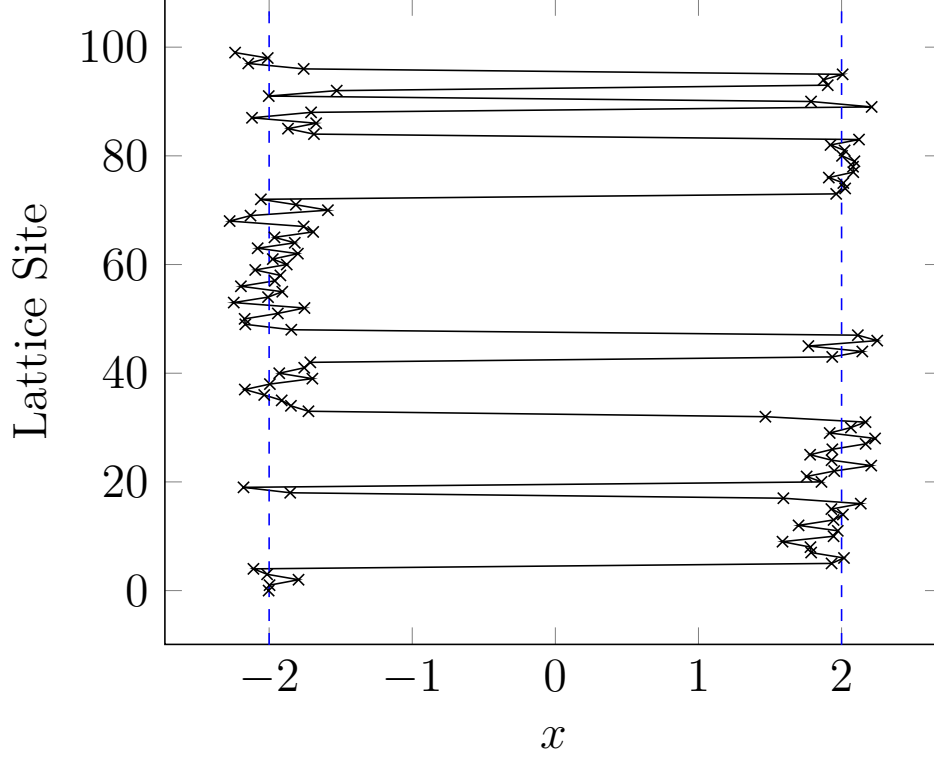
27

Figure 9. Typical configuration for the anharmonic oscillator with $\lambda = 1, f^2 = 4, m = 1, a = 1, L = 1000, d = 0.01, N = 100, \text{configurations} = 100000, \text{burn period} = 1000$.

Typical configuration   Figure 9 shows a typical configuration for the HMC simulation of the anharmonic oscillator. We now observe the tunnelling behaviour we expect for a quantum system double well system.

Energy Levels   In order to calculate the ground state energy we proceed with the same method we used in the case of the harmonic oscillator in section 4.1.1 and use the virial theorem of equation 33 which gives:

$$E_0 = 3\lambda \langle x^4 \rangle - 4\lambda f^2 \langle x^2 \rangle + \lambda f^4. \tag{124}$$

To calculate the first excited state we. The results of the HMC simulation of the anharmonic oscillator for the ground and first excited states are shown in 10 for a range of values of f^2.

Ground State Probability Density Function   To measure the ground state probability density function we follow the same method as we did for the harmonic oscillator in section 4.1.1. The results can be seen in figure 11

Isolated Modes   In figure 9 we saw the tunnelling behaviour displayed by the anharmonic oscillator as the particle tunnelled between the minima of the symmetric potential wells
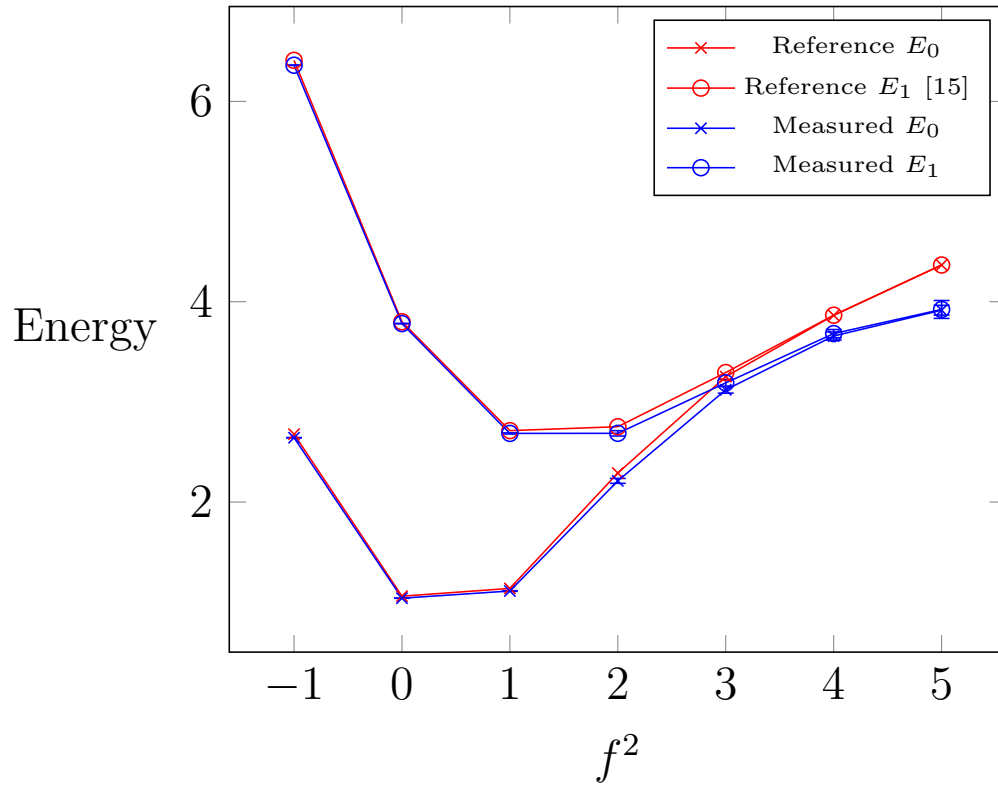
Figure 10. Ground and first excited state energies for the anharmonic oscillator with $\lambda = 1, m = 1, L = 1000, \text{configurations} = 100000, \text{burn period} = 1000$.
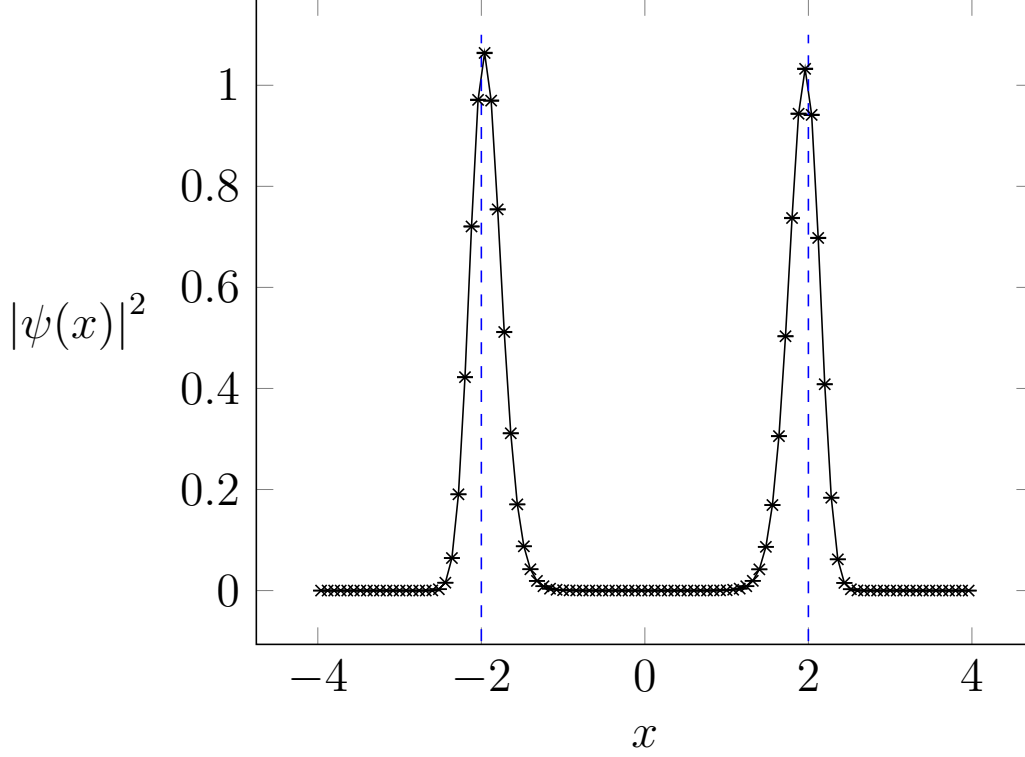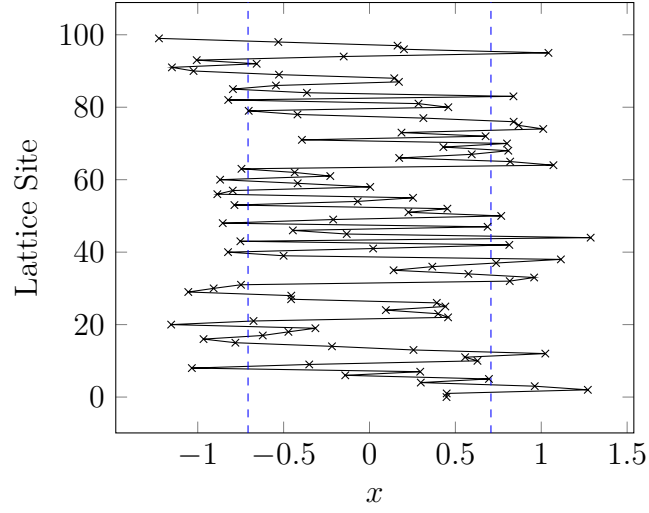
Figure 11. Measured wave function for the harmonic oscillator with $\lambda = 1, f^2 = 4m = 1, a = 1, L = 1000, d = 0.01, N = 100, \text{configurations} = 100000, \text{burn period} = 1000$.

at $x = \pm f$. By the Boltzmann distribution of equation 64 the minima of the potential correspond to maxima of the Boltzmann distribution, which are separated by a region of low probability. In figure 12 we investigate the effect of varying the value of f^2 which varies the distance between the isolated modes and we can clearly see that even a small increase in f^2 results in less tunnelling on a given configuration.

4.2   Accelerated Dynamics (Tempering)

We have seen above that Hybrid Monte Carlo much like other MCMC methods [3] has an issue in that it struggles to sample from isolated modes in a probability distribution, such as the wells in the double well potential. To solve this problem various methods, such as *parallel tempering* [27], [28], *simulated tempering* [29], *tempered transitions* [7] and *annealed importance sampling* [30] have been developed which involve sampling from a distribution that is more diffuse than the target distribution that has the isolated modes. These methods work by varying the temperature T in the canonical distribution of equation 64, so if $T = 1$ gives our target distributions, increasing T will give a more diffuse one. In HMC we can incorporate tempering directly into the Hamiltonian dynamics which proposes states and here we follow the method given in [7] and [3].
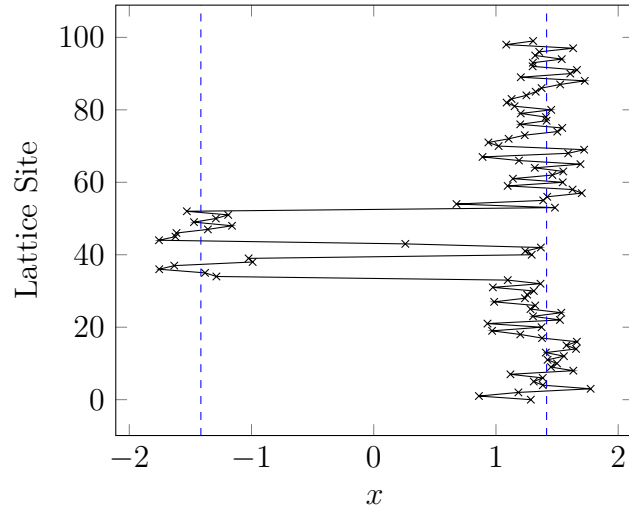
In order to implement tempering on a trajectory that is used to propose an update in the HMC algorithm we multiply each momentum variable in the first half of the trajectory by

(a) $f^2 = 0.5$
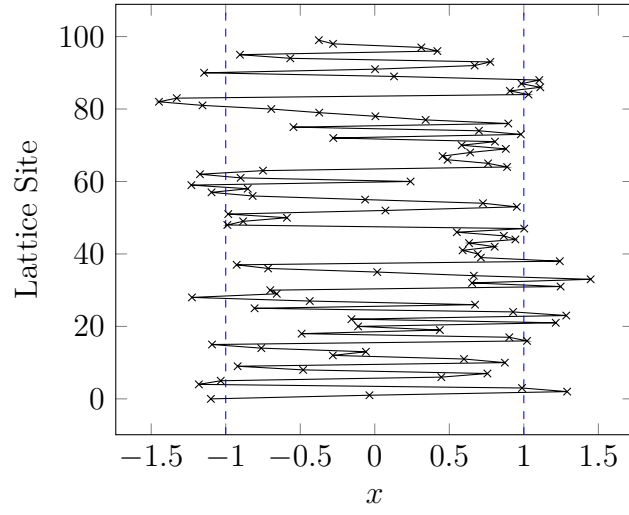
(b) $f^2 = 1$

(c) $f^2 = 2.0$

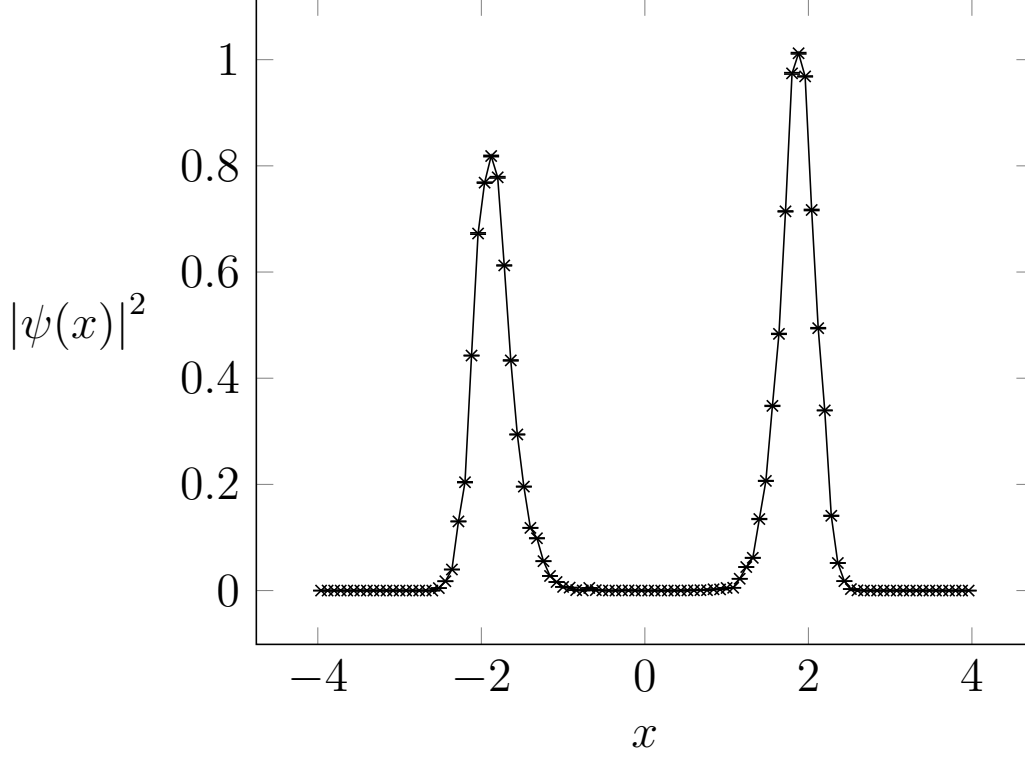Figure 12. Typical configurations for a harmonic oscillator with $a = 1$, $N = 100$ and $\lambda = 1$ when varying f^2.

31

Figure 13. Measured wave function for the harmonic oscillator with $\lambda = 1, f^2 = 4m = 1, a = 1, L = 1000, d = 0.2, N = 5, \text{configurations} = 20000, \text{burn period} = 1000$.

$1 + \delta$ where δ is small, and divide each momentum variable in the second half of the trajectory by $1 + \delta$. We will follow the symmetric method given in [3] and defining $\alpha = 1 + \delta$ multiply each momentum variable before the first half step (equation 46) and after the second half step (equation 46) in the leapfrog update by $\sqrt{\alpha}$ in the first half of the trajectory and then in the second half of the trajectory replace this multiplication with division by $\sqrt{\alpha}$. If there is an odd number of steps in the leapfrog trajectory then on the middle step before the first half step in momenta we multiply by $\sqrt{\alpha}$ and on the second half step we do a corresponding division so that the number of multiplications and divisions by α is equal.

With tempering introduced the leapfrog equations 46, 48 and 50 become:

$$p_i\left(t + \epsilon/2\right) = \sqrt{\alpha} p_i\left(t\right) - \epsilon/2 \frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t\right)\right), \tag{125}$$

$$q_i\left(t + \epsilon\right) = q_i\left(t\right) + \epsilon \frac{\partial K}{\partial p_i}\left(\boldsymbol{p}\left(t + \epsilon/2\right)\right) \tag{126}$$

$$p_i\left(t + \epsilon\right) = \sqrt{\alpha}\left(p_i\left(t + \epsilon/2\right) - \epsilon/2 \frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t + \epsilon\right)\right)\right), \tag{127}$$

for steps in the first half of the trajectory and

$$p_i\left(t + \epsilon/2\right) = \frac{1}{\sqrt{\alpha}} p_i\left(t\right) - \epsilon/2 \frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t\right)\right), \tag{128}$$

$$q_i\left(t + \epsilon\right) = q_i\left(t\right) + \epsilon \frac{\partial K}{\partial p_i}\left(\boldsymbol{p}\left(t + \epsilon/2\right)\right) \tag{129}$$

$$p_i\left(t + \epsilon\right) = \frac{1}{\sqrt{\alpha}} \left(p_i\left(t + \epsilon/2\right) - \epsilon/2 \frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t + \epsilon\right)\right)\right), \tag{130}$$

in the second half of the trajectory, with the possible exception that for an odd number of leap frog steps on the middle step:

$$p_i\left(t + \epsilon/2\right) = \sqrt{\alpha} p_i\left(t\right) - \epsilon/2 \frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t\right)\right), \tag{131}$$

$$q_i\left(t + \epsilon\right) = q_i\left(t\right) + \epsilon \frac{\partial K}{\partial p_i}\left(\boldsymbol{p}\left(t + \epsilon/2\right)\right) \tag{132}$$

$$p_i\left(t + \epsilon\right) = \frac{1}{\sqrt{\alpha}} \left(p_i\left(t + \epsilon/2\right) - \epsilon/2 \frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t + \epsilon\right)\right)\right). \tag{133}$$

Just as in section 3.2.1 we still require that volume is preserved by the leapfrog integration scheme since this is required for detailed balance, even with tempering present. The argument carries over from the non-tempered case, however on any given leapfrog step the determinant of the Jacobian will not necessarily be unity and therefore volume will not be preserved, it will only be the transformation for the whole trajectory that preserves volume. To see why this is the case note that as before we can define the mappings on phase space $T_{p_1}\left(\epsilon, \alpha\right)$: $(\boldsymbol{q}, \boldsymbol{p}) \to (\boldsymbol{q}', \boldsymbol{p}')$, $T_q\left(\epsilon\right) : (\boldsymbol{q}, \boldsymbol{p}) \to (\boldsymbol{q}', \boldsymbol{p}')$ (this is the same map as in the non-tempered case) and $T_{p_2}\left(\epsilon, \alpha\right) : (\boldsymbol{q}, \boldsymbol{p}) \to (\boldsymbol{q}', \boldsymbol{p}')$ such that:

$$T_{p_1}\left(\epsilon, \alpha\right) \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} = \begin{pmatrix} \boldsymbol{q} \\ \sqrt{\alpha}\boldsymbol{p} - \epsilon \underline{\nabla}_q U\left(\boldsymbol{q}\right) \end{pmatrix}, \tag{134}$$

$$T_q\left(\epsilon\right) \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} = \begin{pmatrix} \boldsymbol{q} + \epsilon \underline{\nabla}_p K\left(\boldsymbol{p}\right) \\ \boldsymbol{p} \end{pmatrix}, \tag{135}$$

$$T_{p_2}\left(\epsilon, \alpha\right) \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} = \begin{pmatrix} \boldsymbol{q} \\ \sqrt{\alpha}\left(\boldsymbol{p} - \epsilon \underline{\nabla}_q U\left(\boldsymbol{q}\right)\right) \end{pmatrix}, \tag{136}$$

where we have now had to define three maps due to the fact the tempering multiplication happens before and after the half updates in the momenta so the forms of the half steps in momentum in the leapfrog equations are no longer the same. Having defined these mappings we can calculate their Jacobians to be:

$$J_{p_1} = \det \begin{pmatrix} \delta_{ij} & 0 \\ -\epsilon \partial q_i \partial q_j U\left(\boldsymbol{q}\right) & \sqrt{\alpha}\delta_{ij} \end{pmatrix} = \alpha^{\frac{d}{2}} \tag{137}$$

$$J_q = \det \begin{pmatrix} \delta_{ij} & \epsilon \partial p_i \partial p_j K\left(\boldsymbol{p}\right) \\ 0 & \delta_{ij} \end{pmatrix} = 1 \tag{138}$$

$$J_{p_2} = \det \begin{pmatrix} \delta_{ij} & 0 \\ -\sqrt{\alpha}\epsilon \partial q_i \partial q_j U\left(\boldsymbol{q}\right) & \sqrt{\alpha}\delta_{ij} \end{pmatrix} = \alpha^{\frac{d}{2}}. \tag{139}$$

33

Let us assume that we have an even number l of leapfrog steps in our trajectory. A leapfrog step in the first half of the trajectory can be written in terms of the above mappings as $T(\alpha, \epsilon) = T_{p_2}(\alpha, \epsilon/2) T_q(\epsilon) T_{p_1}(\alpha, \epsilon/2)$, which from the compositions of the above maps will have Jacobian α^d and in the second half trajectory as $T(1/\alpha, \epsilon) = T_{p_2}(1/\alpha, \epsilon/2) T_q(\epsilon) T_{p_1}(1/\alpha, \epsilon/2)$ which will have Jacobian $1/\alpha^d$. The mapping which via the leapfrog equations takes us from the start to the end of the trajectory can be written as:

$$\mathrm{traj}(\epsilon, l) = [T(1/\alpha, \epsilon)]^{\frac{l}{2}} [T(\alpha, \epsilon)]^{\frac{l}{2}} \tag{140}$$

which using the Jacobians of the mappings in the composition will have Jacobian $J_{\mathrm{traj}} = \left[\alpha^d\right]^{\frac{l}{2}} \left[1/\alpha^d\right]^{\frac{l}{2}} = 1$ and so as in the case of non-tempered dynamics we have preservation of volume on phase space if we integrate Hamilton's equations via the leapfrog scheme and include a tempering factor. Note that for an odd number of steps we would have $\mathrm{traj}(\epsilon, l) = [T(1/\alpha, \epsilon)]^{\frac{l}{2}} T_{p_2}(1/\alpha, \epsilon/2) T_q(\epsilon) T_{p_1}(\alpha, \epsilon/2) [T(\alpha, \epsilon)]^{\frac{l}{2}}$ however the factors of alpha will still cancel in the Jacobian and so volume is still preserved.

The multiplication of the momenta variables by α will increase the value of the HMC Hamiltonian $H(\boldsymbol{q}, \boldsymbol{p})$, so that if at the beginning of the trajectory H has a typical value for the canonical distribution when $T = 1$ (which it will if taken at equilibrium), then after the multiplication H will have a typical value for a higher T; this is how the tempering is incorporated into the dynamics. As the momenta $\boldsymbol{p}$ are increased the initial increase in $H(\boldsymbol{q}, \boldsymbol{p}) = K(\boldsymbol{p}) + U(\boldsymbol{q})$ will be due to the increase in $K(\boldsymbol{p})$. However, successive leapfrog steps will distribute the increase in H between K and U [3], which then produces a more diffuse distribution for $\boldsymbol{q}$ than at $T = 1$. In fact, after lots of tempered steps and multiplications by α values for $\boldsymbol{q}$ can be visited that were very unlikely in the original distribution. In the second half of the trajectory divisions by $\sqrt{\alpha}$ return H to values for the original typical distribution at $T = 1$, but the $\boldsymbol{q}$ may now be in a different mode than before, so we may be able to tunnel between modes of the distribution separated by regions of low probability.

Applying this idea to our quantum mechanical double well potential we remember that our variables of interests are the lattice configurations, i.e. $\boldsymbol{q} \equiv \boldsymbol{x} = (x_0, \ldots x_{N-1})$ and the problem is that for a with minima separated by a large enough gap, the individual x_i variables have difficulty tunnelling between the wells, this leads to a large correlation between subsequent samples and hence a bad estimate on observables. We have the HMC Hamiltonian for the double well system given by:

$$H(\boldsymbol{q}, \boldsymbol{p}) = \sum_{i=0}^{N-1} \frac{p_i^2}{2} + \sum_{i=0}^{N-1} \epsilon \left[\frac{1}{2} m \left(\frac{q_{i+1} - q_i}{\epsilon} \right)^2 + \lambda \left(x^2 - f^2 \right) \right], \tag{141}$$

where the mass term in the action couples nearest neighbour variables on the lattice. Ignoring the mass term temporarily will give the Hamiltonian for a system of N an anharmonic oscillators each moving in its own double well potential each with momenta p_i and position q_i. As we evolve the system via the leapfrog integration we can imagine that by multiplying the p_is and increasing the kinetic energy of each oscillator we could in effect "kick" it into the other well. This picture is somewhat naive since it neglects the coupling term in the Hamiltonian which means the harmonic oscillator is actually moving in a more complex

potential where it is coupled to its nearest neighbours. Never the less, this intuitive picture provides a conceptual picture of how tempering would work in our quantum mechanical system. Alternatively we can think of the entire configuration $\boldsymbol{x}$ as being in a potential landscape given by the Euclidean action term in the Hamiltonian, where the isolated modes consist of the regions where different x_i move between the wells in the potential. By increasing the Hamiltonian through multiplication of the momentum variables as described above we may be able to move between these regions of the landscape, which correspond to the isolated modes in the distribution. Before the first tempered step we would expect each x_i variable to be localised around $\pm f$, as we begin to make tempered steps in the molecular dynamics trajectory we expect to see the variables x_i spread out and move far away from the potential minima by the time we reach half way through the trajectory. Then as we begin the divisions in the second half trajectory we should see the variables move back towards the minima at $\pm f$ however now they may be in a different minima to the one they started in.

In [3] it is shown that for large α the probability of accepting the proposed configuration at the end of the trajectory will be small, since $H\left(\boldsymbol{q}', \boldsymbol{p}'\right)$ will be much larger than $H\left(\boldsymbol{q}, \boldsymbol{p}\right)$. Here in lies a key difficulty in tempering, one needs to provide a large enough α such that the Hamiltonian increases enough on the trajectory such that the system is able to move between the isolated modes, however at the end of the trajectory the Hamiltonian needs to have been reduced again to a value similar to that of where it started, if it is to have any chance of being accepted in the Metropolis update. There is a further complication in that increasing the number of leapfrog steps may no longer increase the acceptance rate, since by doing more steps in the leapfrog algorithm we are in effect performing more multiplications by α and so doing more tempering, which may give an even larger Hamiltonian at the end of the trajectory.

To investigate the effects of tempering and confirm our intuition about its applicable to our quantum system we will work with an anharmonic potential where $\lambda = 1$ and $f^2 = 4$ since we know this exhibits the issues related to isolated modes. We first evolve our lattice of 100 sites at a spacing of 1 to equilibrium via non-tempered dynamics then from this configuration begin tempering for a range of α values. By seeding the pseudo random number generation with the same value on each run we can produce the same typical equilibrium trajectory for each tempering value we use. We evolve the trajectory for 200 leapfrog steps of size 0.005 so that $\epsilon l = 1$ but we are also able to observe the evolution of the Hamiltonian along the trajectory. Due to the complex nature of the potential the system is in, it is not clear how much the Hamiltonian would need to increase in order for it to begin tunnelling the variables x_i, however by observing the evolution of the configuration say at half way through the trajectory we will be able to see if the x_i have been able to move out of their minima.

Figures 14, 15 and 16 show the evolution of the Hamiltonian and the trajectory on the first equilibrated lattice configuration for tempering parameters 1.001, 1.01 and 1.05 respectively. In figure 14 we see the Hamiltonian returns to approximately the same place it started, and indeed in this example there is actually an overall decrease such that $\Delta H = -0.071$, so the proposed trajectory will be accepted with certainty. However, we can see from the plots of the trajectory evolution that no tunnelling has occurred, and on the middle leapfrog step, where the Hamiltonian is at its maximum value there is still no tunnelling in the
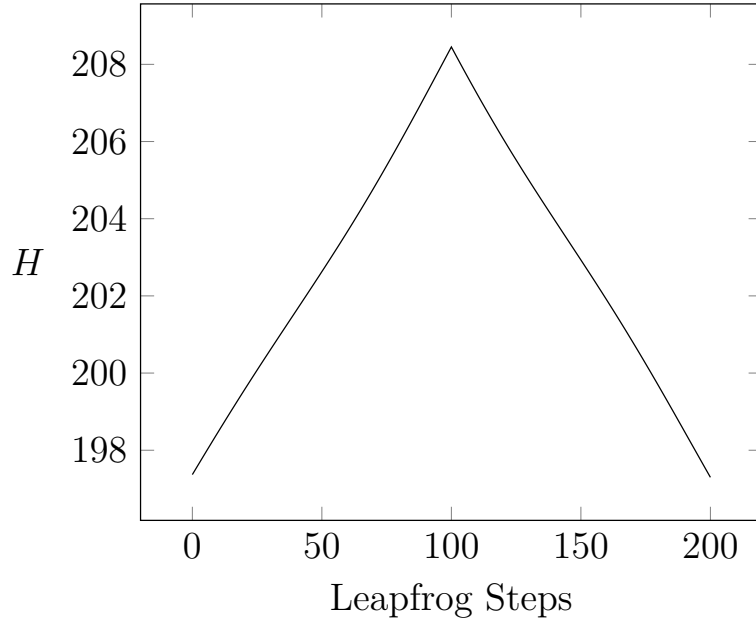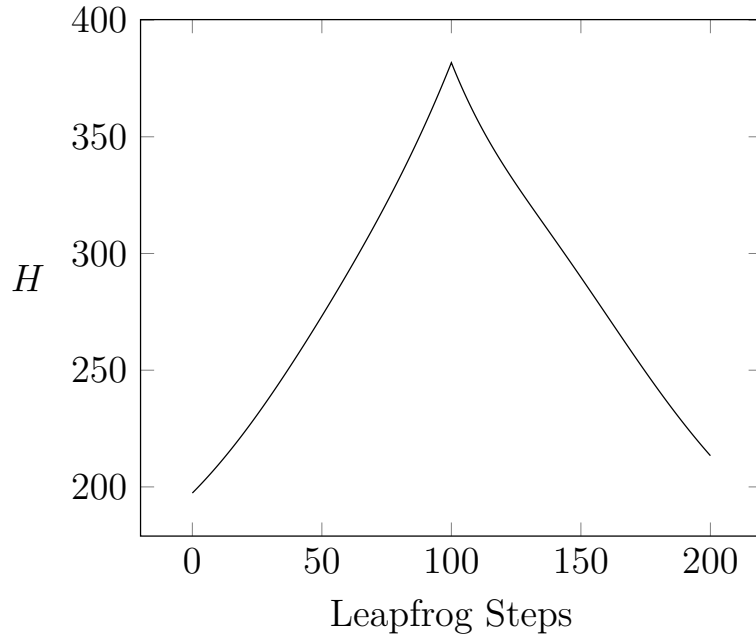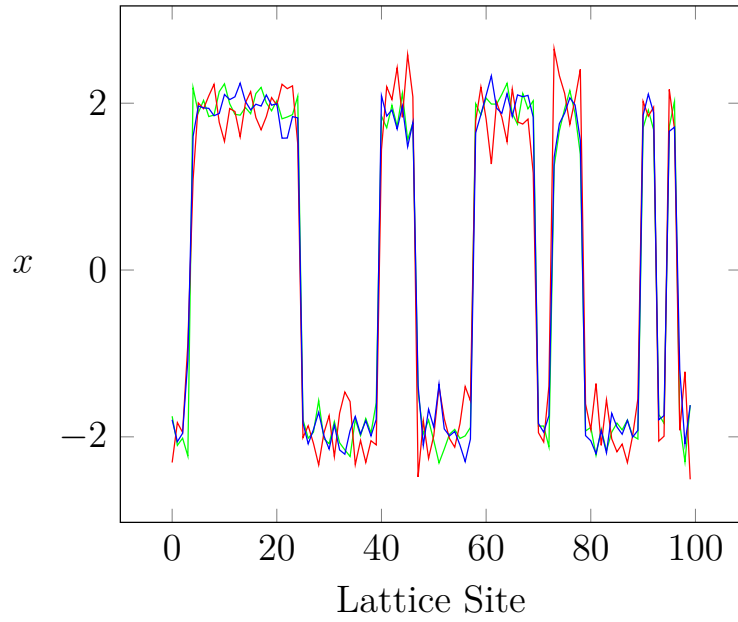
(a) Evolution of Hamiltonian.
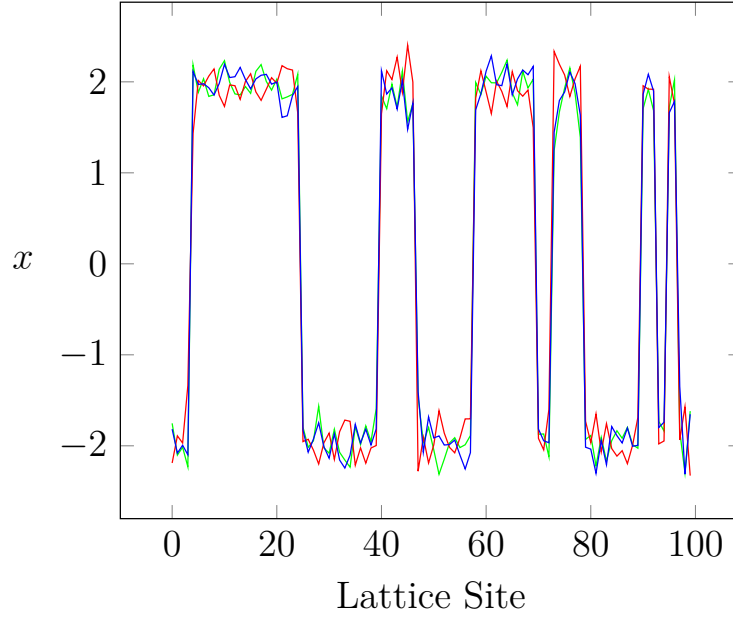
(b) Evolution of the configuration with initial configuration in green, after 100 leapfrog steps in red and at the end of the trajectory in blue.

Figure 14. Evolution of Hamiltonian and configuration proposal for tempering parameter $\alpha = 1.001$

(a) Evolution of Hamiltonian.

(b) Evolution of the configuration with initial configuration in green, after 100 leapfrog steps in red and at the end of the trajectory in blue.

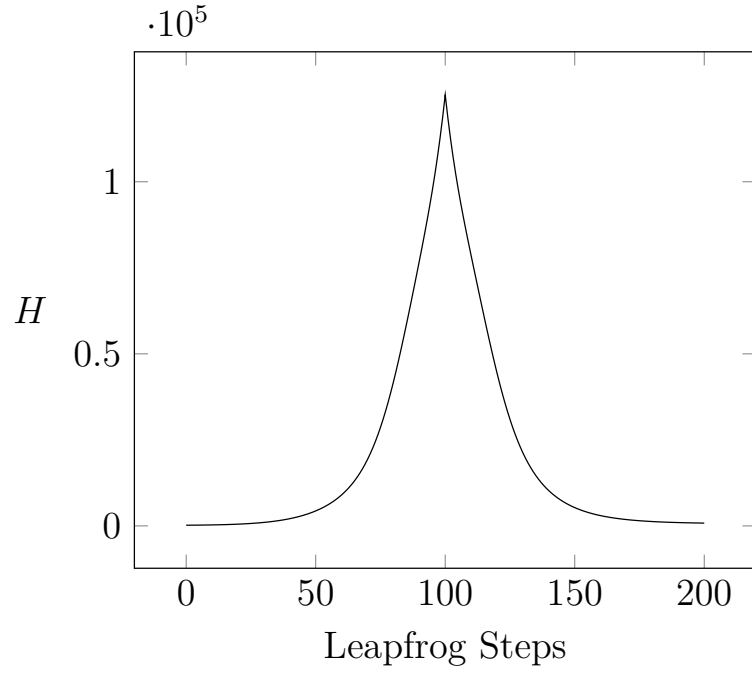Figure 15. Evolution of Hamiltonian and configuration proposal for tempering parameter $\alpha = 1.01$
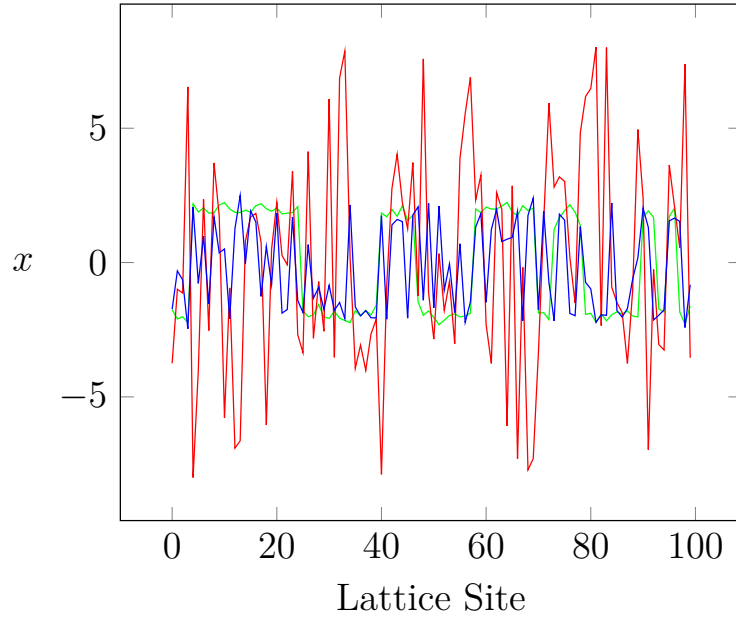
(a) Evolution of Hamiltonian.

(b) Evolution of the configuration with initial configuration in green, after 100 leapfrog steps in red and at the end of the trajectory in blue.

Figure 16. Evolution of Hamiltonian and configuration proposal for tempering parameter $\alpha = 1.05$

corresponding configuration. Figure 15 shows the same results with $\alpha = 1.01$ however now there is an increase in the Hamiltonian with $\Delta H = 16.005$ which corresponds to an acceptance probability of the order of 10^{-7} which is incredibly low, in fact the acceptance rate for this simulation was 1.588 ± 0.0790657. The issue that makes tempering as suggested in [3] and [7] inapplicable to our quantum mechanical system lies in this result. In the configuration evolution in figure 15 we can see as in the tempered case for $\alpha = 1.001$ there has still been no tunnelling despite the larger increase in H that occurs on the trajectory. The change in the Hamiltonian on the trajectory is however now too big and typical proposals will not be accepted, so even for tempering parameters that result in the Hamiltonian being too large at the end of the trajectory the configuration has been unable to move between the isolated modes. To see that the intuition around the effects of tempering on our simulation were valid, that is we can move the lattice variables x_i between isolated modes in the distribution observe figure 16. Here we can see that with $\alpha = 1.05$ the Hamiltonian increases massively and indeed on the middle leapfrog step the lattice variables have been able to move out of their respective wells. We can also see that at the end of the trajectory lattice variables that were in one well have been able to tunnel into the other and so the initial and proposal configurations are uncorrelated. However, with $\alpha = 1.05$ $\Delta H = 616.047$ which for all practical purposes gives an acceptance rate of zero.

5   Conclusion

References

[1] Simon Duane, A.d. Kennedy, Brian J. Pendleton, and Duncan Roweth. Hybrid monte carlo. *Physics Letters B*, 195(2):216–222, 1987.

[2] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6), Jun 1953.

[3] Radford Neal. Mcmc using hamiltonian dynamics. *Chapman & Hall/CRC Handbooks of Modern Statistical Methods Handbook of Markov Chain Monte Carlo*, Oct 2011.

[4] Radford M. Neal. Bayesian learning for neural networks. *Lecture Notes in Statistics*, 1996.

[5] Hemant Ishwaran. Applications of hybrid monte carlo to bayesian generalized linear models: Quasicomplete separation and neural networks. *Journal of Computational and Graphical Statistics*, 8(4):779, 1999.

[6] Mikkel N. Schmidt. Function factorization using warped gaussian processes. *Proceedings of the 26th Annual International Conference on Machine Learning - ICML 09*, 2009.

[7] Radford M. Neal. Sampling from multimodal distributions using tempered transitions. *Statistics and Computing*, 6(4):353–366, 1996.

[8] Wolfgang Bietenholz. Hadron physics from lattice qcd. *International Journal of Modern Physics E*, 25(07):1642008, 2016.

[9] Luigi Del Debbio, Haralambos Panagopoulos, and Ettore Vicari. Îÿ dependence of su(n) gauge theories. *Journal of High Energy Physics*, 2002(08):044–044, 2002.

[10] Luigi Del Debbio, Gian Mario Manca, and Ettore Vicari. Critical slowing down of topological modes. *Physics Letters B*, 594(3-4):315–323, 2004.

[11] M Creutz and B Freedman. A statistical approach to quantum mechanics. *Annals of Physics*, 132(2):427–462, 1981.

[12] Marise J. E. Westbroek, Peter R. King, Dimitri D. Vvedensky, and Stephan Durr. User's guide to monte carlo methods for evaluating path integrals, Dec 2017.

[13] Ronnnie Rodgers and Laura Raes. Monte carlo simulations of harmonic and anharmonic oscillators in discrete euclidean time, 2014.

[14] Aleksandra Slapik and William Serenone. Lattice monte carlo study of the harmonic oscillator in ..., 2012.

[15] R. Blankenbecler, T. Degrand, and R. L. Sugar. Moment method for eigenvalues and expectation values. *Physical Review D*, 21(4):1055–1061, 1980.

[16] Richard P. Feynman and A. R. Hibbs. *Quantum mechanics and path integrals R.P. Feynman A.R. Hibbs.* McGraw-Hill, 1965.

[17] James Binney and David Benjamin. Skinner. *The physics of quantum mechanics.* Oxford University Press, 2015.

[18] V. Fock. Bemerkung zum virialsatz. *Zeitschrift fï£ · r Physik*, 63(11-12):855–858, 1930.

[19] Christof Gattringer and Christian B. Lang. *Quantum Chromodynamics on the Lattice An Introductory Presentation.* Springer Berlin, 2013.

[20] A. Sokal. Monte carlo methods in statistical mechanics: Foundations and new algorithms. *Functional Integration NATO ASI Series*, pages 131–192, 1997.

[21] Herbert Goldstein, Charles P. Poole, and John Safko. *Classical mechanics.* Pearson, 2014.

[22] A.d. Kennedy and Brian Pendleton. Cost of the generalised hybrid monte carlo algorithm for free field theory. *Nuclear Physics B*, 607(3):456–510, 2001.

[23] Alexandros Beskos, Natesh Pillai, Gareth Roberts, Jesus-Maria Sanz-Serna, and Andrew Stuart. Optimal tuning of the hybrid monte carlo algorithm. *Bernoulli*, 19(5A):1501–1534, 2013.

[24] Nawaf Bou-Rabee and Jesús María Sanz-Serna. Randomized hamiltonian monte carlo. *The Annals of Applied Probability*, 27(4):2159–2194, 2017.

[25] Alain Durmus, Eric Moulines, and Eero Saksman. On the convergence of hamiltonian monte carlo, 2017.

[26] Samuel Livingstone, Michael Betancourt, Simon Byrne, and Mark Girolami. On the geometric ergodicity of hamiltonian monte carlo, 2016.

[27] Charles J. Geyer. Markov chain monte carlo maximum likelihood. *Interface Foundation of North America*, 1991.

[28] David J. Earl and Michael W. Deem. Parallel tempering: Theory, applications, and new perspectives. *Physical Chemistry Chemical Physics*, 7(23):3910, 2005.

[29] E Marinari and G Parisi. Simulated tempering: A new monte carlo scheme. *Europhysics Letters (EPL)*, 19(6):451–458, 1992.

[30] Radford M. Neal. Annealed importance sampling. *Statistics and Computing*, 11(2):125–139, 2001.

Appendices

Note — Appendices are provided for completeness only and any content included in them will be disregarded for the purposes of assessment.

A Quantum Virial Theorem

The quantum virial theorem relates the expectation value of the kinetic energy to that of the potential. For a single particle of mass m can be derived as follows. We begin with a Hamiltonian of the form:

$$\hat{H}\left(\hat{x}, \hat{p}\right) = \hat{T}\left(\hat{p}\right) + \hat{V}\left(\hat{x}\right) \tag{142}$$

, where

$$\hat{T}\left(\hat{p}\right) = \frac{1}{2m}\hat{p}^2, \tag{143}$$

$$\hat{p} = -i\frac{d}{dx} \tag{144}$$

(in units where $\hbar = 1$) and

$$\hat{V}\left(\hat{x}\right) = V\left(x\right) \tag{145}$$

so the potential is time independent.

Using the identity:

$$\left[\hat{A}, \hat{B}\hat{C}\right] = \left[\hat{A}, \hat{B}\right]\hat{C} + \hat{B}\left[\hat{A}, \hat{C}\right], \tag{146}$$

it can be shown that:

$$\left[\hat{H}, \hat{x}\hat{p}\right] = \left[\hat{H}, \hat{x}\right]\hat{p} + \hat{x}\left[\hat{H}, \hat{p}\right] \tag{147}$$

$$= \frac{1}{2m}\left[\hat{p}^2, \hat{x}\right]\hat{p} + \hat{x}\left[V\left(x\right), \hat{p}\right] \tag{148}$$

$$= -i\frac{\hat{p}^2}{m} + ixV'\left(x\right), \tag{149}$$

so that:

$$i\left[\hat{H}, \hat{x}\hat{p}\right] = 2\hat{T}\left(\hat{p}\right) - xV'\left(x\right). \tag{150}$$

The Heisenberg equation of motion gives for any operator $\hat{A}\left(t\right)$:

$$\frac{d\hat{A}\left(t\right)}{dt} = -i\left[\hat{H}, \hat{A}\left(t\right)\right], \tag{151}$$

so that equation 150 becomes:

$$\frac{d\left(\hat{x}\hat{p}\right)}{dt} = 2\hat{T}\left(\hat{p}\right) - xV'\left(x\right) \tag{152}$$

Then taking the expectation value of both sides of equation 152 and noting that : _____ [Why is this true?]

$$\langle\frac{d\left(\hat{x}\hat{p}\right)}{dt}\rangle = 0, \tag{153}$$

we get the virial theorem:

$$2\langle\hat{T}\rangle = \langle xV'\left(x\right)\rangle \tag{154}$$

B   Derivation of the discrete path integral for quantum harmonic oscillator

Here we follow the derivation given in [11] for the exact result of the path integral for discrete theory and a particle of mass $m = 1$. Since in [11] the derivation has several typographical mistakes and jumps in the algebra that make it difficult to follow for the reader, we have chosen to reproduce it with corrections and alterations.

For a quantum harmonic oscillator of mass $m = 1$, the discrete Euclidean path integral is given in equation 17 as: _____ [FIX INDEXING]

$$Z = \int_{-\infty}^{+\infty}\prod_{i=1}^{N-1}dx_i\exp\left(-\sum_{j=0}^{N-1}a\left[\frac{1}{2}\left(\frac{x_{j+1}-x_j}{a}\right)^2 + \frac{1}{2}\mu^2 x_j^2\right]\right). \tag{155}$$

We begin by defining an operator T such that its matrix elements in the Schŕodinger picture are given by:

$$\langle x'|\hat{T}|x\rangle = \exp\left(-\frac{1}{2a}\left(x'-x\right)^2 - \frac{\mu^2 a}{4}\left(x^2 + x'^2\right)\right). \tag{156}$$

Then we can use the completeness relation that:

$$\hat{1} = \int_{-\infty}^{\infty}|x\rangle\langle x|. \tag{157}$$

By inserting $N-1$ copies of this relation into the expression (one between each pair of $\hat{T}$s):

$$Tr\left(\hat{T}^N\right), \tag{158}$$

and using the definition of the matrix elements in equation 156, we recover the path integral in equation 155, that is:

$$Z =\sim Tr\left(\hat{T}^N\right). \tag{159}$$

Here we take the trace over the physical Hilbert space, so that for any operator $\hat{A}$ in the Schŕodinger eigenbasis, the trace is:

$$Tr\left(\hat{A}\right) = \int_{-\infty}^{\infty}dx\langle x|\hat{A}|x\rangle, \tag{160}$$

43

which is of course basis independent.

The next step is to make the ansatz that:

$$\hat{T} = \int_{-\infty}^{\infty} d\omega e^{\frac{-\mu^2 a}{4}\hat{x}^2} e^{-i\hat{p}\omega} e^{-1\frac{1}{2a}\omega^2} e^{\frac{-\mu^2 a}{4}\hat{x}^2}, \tag{161}$$

then, using that canonical momentum generates translations on position, that is:

$$e^{-i\hat{p}\Delta} |x\rangle = |x + \Delta\rangle \tag{162}$$

which can easily be shown by Fourier transforming the position eigenbasis into momentum space, acting with the momentum operator in the exponential then Fourier transforming back into position space; we can calculate the matrix element $\langle x'| T |x\rangle$ and recover equation 156. We observe that the integral over ω in equation 161 is Gaussian; we may use the standard result:

$$\int_{-\infty}^{+\infty} dx e^{-\alpha x^2 + i\beta x} = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4a}}, \tag{163}$$

this result follows from completing the square on x, analytically continuing the integral over x into the complex plane and creating a rectangular contour in the lower half plane, then applying residue theory, it is valid provided $\alpha \in \mathbb{R}_{>0}$ and $\beta \in \mathbb{R}$. Applying the identity in equation 163 to equation 161 gives:

$$\hat{T} = \sqrt{2\pi a} e^{-\frac{-\mu^2 a}{4}\hat{x}^2} e^{-\frac{a}{2}\hat{p}^2} e^{-\frac{\mu^2 a}{4}\hat{x}^2}. \tag{164}$$

Notice at this stage in our derivation we could apply the Baker-Campbell-Hausdorffformula to combine the exponentials; dropping $\mathcal{O}(a^2)$ would then give the harmonic oscillator Hamiltonian exponentiated. Since this system is exactly solvable in the discrete case, we will avoid doing this and keep all terms.

CHECK THAT THIS IS WHAT CREUTZ MEANS

The canonical commutation relation of quantum mechanics gives:

$$[\hat{x}, \hat{p}] = i\hbar, \tag{165}$$

which can be along with the identities that for operators $\hat{A}$ and $\hat{B}$:

$$\left[\hat{A}, \hat{B}^n\right] = n\hat{B}^{n-1}\left[\hat{A}, \hat{B}\right] \tag{166}$$

and

$$\left[\hat{A}^n, \hat{B}\right] = n\hat{A}^{n-1}\left[\hat{A}, \hat{B}\right], \tag{167}$$

if $\left[\hat{A}, \left[\hat{A}, \hat{B}\right]\right] = \left[\hat{B}, \left[\hat{A}, \hat{B}\right]\right] = 0$, to show that:

$$\left[\hat{x}, e^{-\frac{1}{2}a\hat{p}^2}\right] = -ia\hat{p} \tag{168}$$

and

$$\left[\hat{p}, e^{-\frac{\mu^2 a}{4}\hat{x}^2}\right] = i\frac{\mu^2 a}{2}\hat{x} \tag{169}$$

44

Using identities 168 and 169 we can easily show that:

$$\hat{x}\hat{T} = \hat{T}\left[\left(1 + \frac{a^2\mu^2}{2}\right)\hat{x} - ia\hat{p}\right],\tag{170}$$

and

$$\hat{p}\hat{T} = \hat{T}\left[\left(1 + \frac{a^2\mu^2}{2}\right)\hat{p} + ia\mu^2\left(1 + \frac{a^2\mu^2}{4}\right)\hat{x}\right].\tag{171}$$

Iterating equations 170 and 171 a second time gives:

$$\left[\hat{p}^2 + \mu^2\left(1 + \frac{a^2\mu^2}{4}\right)\hat{x}^2, \hat{T}\right] = 0.\tag{172}$$

Defining a new angular frequency parameter ω by:

$$\omega^2 = \mu^2\left(1 + \frac{a^2\mu^2}{4}\right),\tag{173}$$

then from equation 172 we have that $\hat{T}$ commutes with the simple harmonic oscillator Hamiltonian:

$$\hat{H} = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2\hat{x}^2.\tag{174}$$

Since $\hat{H}$ and $\hat{T}$ commute we know $\hat{T}$ is diagonalized by the eigenstates of $\hat{H}$.

The Hamiltonian is in the form of a harmonic oscillator with angular frequency ω, therefore we may define the corresponding ladder operators:

$$\hat{a}^\dagger = \frac{1}{\sqrt{\omega}}\left(\hat{p} + i\omega\hat{x}\right)\tag{175}$$

and

$$\hat{a} = \frac{1}{\sqrt{\omega}}\left(\hat{p} - i\omega\hat{x}\right)\tag{176}$$

which allows us to write:

$$\hat{H} = \left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)\omega\tag{177}$$

The eigenstates of $\hat{H}$ satisfy the standard relations:

$$\hat{a}\left|0\right\rangle = 0,\tag{178}$$

$$\left(\hat{a}^\dagger\right)^n\left|0\right\rangle = \left|n\right\rangle,\tag{179}$$

and

$$\langle n|n\rangle = n!.\tag{180}$$

Using the identities of equations 170 and 171 we can show that:

$$\hat{a}\hat{T} = \hat{T}\hat{a}\left(1 + \frac{a^2\mu^2}{2} - a\mu\left(1 + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}}\right).\tag{181}$$

Since the eigenstates of $\hat{H}$ given as $|n\rangle$ diagonalise $\hat{T}$, for λ_i the eigenvalues of $\hat{T}$:

$$\hat{T}|n\rangle = \lambda_n |n\rangle. \tag{182}$$

We can then use that $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$ and equation 181 to show the ratio:

$$\frac{\lambda_n}{\lambda_{n-1}} = 1 + \frac{a^2\mu^2}{2} - a\mu\left(1 + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}} \tag{183}$$

$$= \left(\left(1 + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}} - \frac{1}{2}a\mu\right)^2 \tag{184}$$

$$:= R, \tag{185}$$

so that

$$\lambda_n = R\lambda_{n-1} = R^2\lambda_{n-2} = \cdots = R^n\lambda_0, \tag{186}$$

which gives:

$$\hat{T}|n\rangle = \lambda_n |n\rangle \tag{187}$$

$$= R^n\lambda_0 |n\rangle \tag{188}$$

$$= \lambda_0 e^{n\log R}|n\rangle. \tag{189}$$

However, the energy eigenstates $|n\rangle$ are defined through:

$$\hat{H}|n\rangle = \left(n + \frac{1}{2}\right)\omega |n\rangle, \tag{190}$$

we therefore make the ansatz that for constants C and K:

$$\hat{T} = Ce^{K\hat{H}}, \tag{191}$$

which can be shown to be consistent with equation 189 provided we take $K\omega = \log R$ and $Ce^{\frac{K\omega}{2}} = \lambda_0$ and can conclude that:

$$\hat{T} = Ce^{\frac{\hat{H}}{\omega}\log R} \tag{192}$$

$$= CR^{\frac{1}{\omega}\hat{H}}. \tag{193}$$

We may then calculate C by first taking the trace of $\hat{T}$ over the energy eigenbasis $|n\rangle$:

$$Tr\left(\hat{T}\right) = \sum_n \langle n|\hat{T}|n\rangle \tag{194}$$

$$= C\sum_n R^{n+\frac{1}{2}} \tag{195}$$

$$= C\frac{R^{\frac{1}{2}}}{1 - R} \tag{196}$$

$$= C\frac{1}{a\mu} \tag{197}$$

where we have used that $\hat{H}|n\rangle = E_n|n\rangle = \left(n + \frac{1}{2}\right)\omega$ for the harmonic oscillator in units where $\hbar = 1$, and in order to compute the sum we have used that $|R| < 1$. We then take the trace of $\hat{T}$ over the position eigenbasis according to equation 160, and using equation 156:

$$Tr\left(\hat{T}\right) = \int dx \langle x|\,\hat{T}\,|n\rangle \tag{198}$$

$$= \int dx e^{-\frac{a\mu^2}{2}x^2} \tag{199}$$

$$= \frac{1}{\mu}\sqrt{\frac{2\pi}{a}}, \tag{200}$$

comparing equation 200 and 197 we conclude $C = \sqrt{2\pi a}$, and so

$$T = \sqrt{2\pi a}R^{\frac{\hat{H}}{\omega}}. \tag{201}$$

We may now explicitly compute the trace expression for the discrete path integral given in equation 158 using the result for $\hat{T}$ in equation 201. We again take the trace over the energy eigenstates and use the fact that $|R| < 1$ to compute the resulting sum, this gives:

$$Z = (2\pi a R)^{\frac{N}{2}}\frac{1}{1 - R^N}. \tag{202}$$

This is the exact expression for the discrete path integral of a quantum harmonic oscillator of mass $m = 1$, angular frequency μ on an imaginary time lattice of $N - 1$ sites with lattice spacing a.

The correlation functions are given by:

$$\langle x_i x_{i+j}\rangle = \frac{1}{Z}Tr\left(\hat{x}\hat{T}^j\hat{x}\hat{T}^{N-j}\right) \tag{203}$$

$$= \frac{1}{2\mu\left(1 + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}}(1 - R^n)}\left(R^j + R^{N-j}\right). \tag{204}$$

Taking $j = 0$, equation 203 becomes:

$$\langle x^2\rangle = \frac{1}{2\omega}\left(\frac{1 + R^N}{1 - R^N}\right) \tag{205}$$

$$= \frac{1}{2\mu\left(1 + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}}}\left(\frac{1 + R^N}{1 - R^N}\right) \tag{206}$$

We have shown in above that the discrete theory for the harmonic oscillator leads to a quantum system with the Hamiltonian given in equation 174. This is the Hamiltonian of a quantum harmonic oscillator of angular frequency ω rather than μ and hence we employ the standard result from quantum theory that for a a quantum harmonic oscillator of angular frequency ω the ground state wave function is given by:

47

$$\psi(x) = \left(\frac{\omega}{\pi}\right)^{\frac{1}{4}} \exp\left(-\frac{1}{2}\omega x^2\right) \tag{207}$$

$$= \left(\frac{\mu\left(1 + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}}}{\pi}\right)^{\frac{1}{4}} \exp\left(-\frac{1}{2}\mu\left(1 + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}} x^2\right) \tag{208}$$

In order order to get the discrete path integral results for a particle of any mass (not just unity), we note that the action for a particle of mass m:

$$S_M(\boldsymbol{x}) = \sum_{i=0}^{N-1} a \left[\frac{1}{2}m\left(\frac{x_{i+1} - x_i}{a}\right)^2 + \frac{1}{2}\mu^2 x_i^2\right] \tag{209}$$

can be obtained from of the discrete Euclidean action of a particle of mass unity:

$$S_M(\boldsymbol{x}) = \sum_{i=0}^{N-1} a \left[\frac{1}{2}\left(\frac{x_{i+1} - x_i}{a}\right)^2 + \frac{1}{2}\eta^2 x_i^2\right]. \tag{210}$$

by making the replacements

$$x \to x\sqrt{m} \tag{211}$$

and

$$\mu \to \frac{\mu}{\sqrt{m}}. \tag{212}$$

Making these replacements in the above results gives:

$$\langle x_i x_{i+j}\rangle = \frac{1}{2\omega m}\left(\frac{R^j + R^{N-j}}{1 - R^N}\right) \tag{213}$$

$$\langle x^2\rangle = \frac{1}{2m\omega}\left(\frac{1 + R^N}{1 - R^N}\right) \tag{214}$$

$$\psi(x) = \left(\frac{\omega}{\pi}\right)^{\frac{1}{4}} \exp\left(-\frac{1}{2}\omega x^2\right) \tag{215}$$

but where now:

$$R = 1 + \frac{a^2\mu^2}{2m} - \frac{a\mu}{\sqrt{m}}\left(1 + \frac{a^2\mu^2}{4m}\right)^{\frac{1}{2}} \tag{216}$$

$$\omega = \frac{\mu}{\sqrt{m}}\sqrt{1 + \frac{a^2\mu^2}{4m}}. \tag{217}$$

48