THE UNIVERSITY *of* EDINBURGH
School of Physics
and Astronomy

# Accelerated Tempering Dynamics in HMC Simulations of Lattice Field Theory

**MPhys Project Report**

**Jack Frankland**

March 4, 2018

**Abstract**

Supervisors: Dr Roger Horsley, Dr Brian Pendleton

## Personal statement

## Acknowledgments

# Contents

**1 Introduction** **1**

**2 Background** **1**

2.1 Quantum Mechanics . . . . . 1

**3 Methods** **4**

3.1 Monte Carlo Methods . . . . . 4

3.2 Hybrid Monte Carlo . . . . . 5

3.2.1 Hamiltonian Dynamics . . . . . 5

3.2.2 Sampling and the Hamiltonian . . . . . 8

3.2.3 Steps of the HMC Algorithm . . . . . 9

**4 Results and Discussion** **11**

4.1 Quantum Oscillators . . . . . 11

4.1.1 Quantum Harmonic Oscillator . . . . . 12

4.1.2 Quantum Anharmonic Oscillator . . . . . 15

4.2 Accelerated Dynamics (Tempering) . . . . . 21

**5 Conclusion** **24**

**Appendices** **25**

**A Quantum Virial Theorem** **25**

**B Derivation of the discrete path integral for quantum harmonic oscillator** **26**

# 1 Introduction

# 2 Background

## 2.1 Quantum Mechanics

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

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

where we will use the notation:

$$\begin{aligned} Z_{ba} &= \langle x_b, t_b | x_a, t_a \rangle & (2) \\ &= \langle x_b | \, e^{-(t_b - t_a)\hat{H}/\hbar} \, | x_a \rangle \, . & (3) \end{aligned}$$

$Z_{ba}$ 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 $\int \mathcal{D}x\left(t\right)$ is an integral over all paths between $x_a$ and $x_b$ and $S_M\left[x\left(t\right)\right]$ is the Minkowski action of a particle of mass $m$ on the path $x\left(t\right)$ which is defined by:

$$S_M\left[x\left(t\right)\right] = \int_{t_a}^{t_b} dt \left[\frac{1}{2} m \left(\frac{dx}{dt}\right)^2 - V(x)\right], \tag{4}$$

where $x\left(t_a\right) = x_a$ and $x\left(t_b\right) = x_b$ are the boundary conditions.

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 [1] to get equation 1 into a form we can work with.

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 $x_i = x\left(t_i\right)$ where $i = 0, 1, \ldots, N$. We introduce the notation:

$$\boldsymbol{x} = \left(x_0, x_1, \ldots, x_N\right) \tag{5}$$

to denote a particular path on the lattice where each position coordinate has been specified and we refer to equation 5 as a *configuration* on the lattice. In our notation for the labelling of the position eigenstates in equation 1 we also have $x_b = x_N = x\left(t_N\right)$ and $x_a = x_0 = x\left(t_0\right)$ in

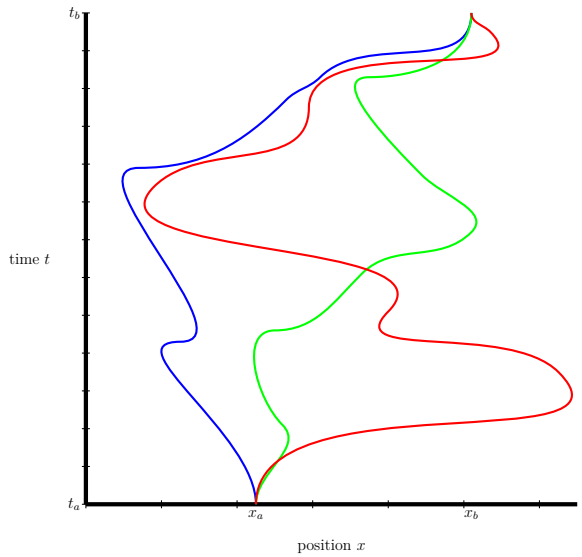



Figure 1. Three of the infinitely many paths from $(x_a, t_a)$ to $(x_b, t_b)$ that contribute to the path integral.

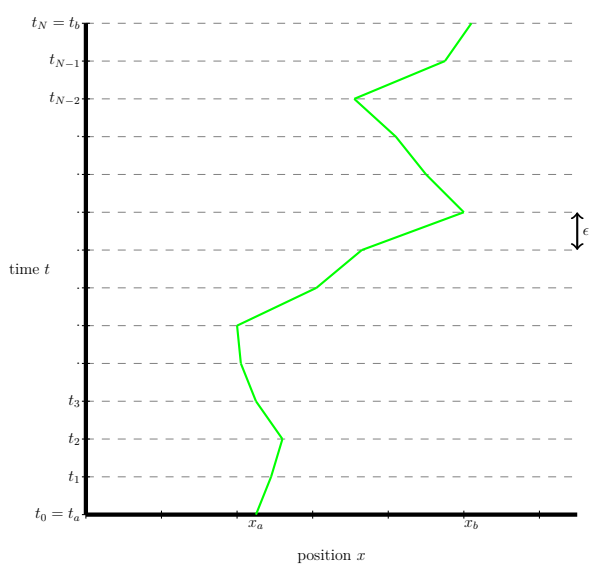



Figure 2. Discretising time and a path from $(x_a, t_a)$ to $(x_b, t_b)$ onto a lattice of time spacing $\epsilon$.

order to match with figure 2. $\epsilon$ 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 4 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:

$$S_M(\boldsymbol{x}) = \sum_{i=0}^{N-1} \epsilon \left[ \frac{1}{2} m \left( \frac{x_{i+1} - x_i}{\epsilon} \right)^2 - V(x_i) \right]. \tag{6}$$

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

$$\int_{x_a}^{x_b} \mathcal{D}x \sim \prod_{n=1}^{N-1} \int_{-\infty}^{\infty} dx_n, \tag{7}$$

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:

$$Z_{ba} \sim \int_{-\infty}^{+\infty} \prod_{i=1}^{N-1} dx_i \exp\left( \frac{i}{\hbar} S_M(\{x_i\}) \right). \tag{8}$$

In the limit that $N \to \infty$ (or equivalently $\epsilon \to 0$) we recover (up to normalisation) equation 1 from equation 8 exactly.

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:

$$Z = \int dx_a \int dx_b \delta(x_b - x_a) Z_{ba} \tag{9}$$

$$\sim \int_{-\infty}^{+\infty} \prod_{i=0}^{N-1} dx_i \exp\left( \frac{i}{\hbar} S_M\{x_i\} \right). \tag{10}$$

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:

$$\tau = it. \tag{11}$$

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 6:

$$S_M(\boldsymbol{x}) = i \sum_{i=0}^{N-1} a \left[ \frac{1}{2} m \left( \frac{x_{i+1} - x_i}{a} \right)^2 + V(x_i) \right] \tag{12}$$

$$= i S_E(\boldsymbol{x}), \tag{13}$$

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

$$S_E(\boldsymbol{x}) \equiv \sum_{i=0}^{N-1} a \left[ \frac{1}{2} m \left( \frac{x_{i+1} - x_i}{a} \right)^2 + V(x_i) \right] \tag{14}$$

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 $(x, y, z, \tau)$ and vice-versa:

$$ds_M^2 = -dt^2 + dx^2 + dy^2 + dz^2 \tag{15}$$
$$= d\tau^2 + dx^2 + dy^2 + dz^2 \tag{16}$$
$$= ds_E^2. \tag{17}$$

Equation 13 is a very useful result since upon substitution into the discrete path integral we find:

$$Z \sim \int_{-\infty}^{+\infty} \prod_{i=0}^{N-1} dx_i \exp\left(-\frac{1}{\hbar} S_E\left(\{x_i\}\right)\right) \tag{18}$$

which is known as the discrete euclidean path integral and will converge since the integrand is now exponentially suppressed. 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 $i = 0, 1, \ldots N-1$ then the canonical partition function is given by:

$$\mathcal{Z} \sim \int_{-\infty}^{+\infty} \prod_{i=0}^{N-1} dx_i \exp\left(-\beta H\left(\{x_i\}\right)\right), \tag{19}$$

with $\beta = \frac{1}{k_B T}$ where $T$ is the system temperature and $k_B$ is the Boltzmann constant. Comparing equation 18 to equation 19 we can see that the discretise Euclidean path integral is a classical canonical partition function on a system with $N$ degrees of freedom, provided that we take:

$$\frac{1}{\hbar} S_E\left(\{x_i\}\right) = \beta H\left(\{x_i\}\right). \tag{20}$$

We then have a Boltzmann factor given by $\exp\left(-\frac{1}{\hbar} S_E\left(\{x_i\}\right)\right)$. 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 Hamiltonian is given by $S_E\left(\{x_i\}\right)$.

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. The general idea is that we can approximate

the value of an $d$ dimensional integral:

$$I = \int_V d^d x f(\boldsymbol{x}) \tag{21}$$

by considering it as an average value of the function over the integral volume:

$$I \approx \frac{1}{N} \sum_i^N f(\boldsymbol{x}_i), \tag{22}$$

where $\boldsymbol{x}_i \in V$ are *samples* from the integration domain. The most naive method for generating the samples is to draw them from a uniform distribution over the domain. To see why uniform sampling is naive note that the error in the Monte Carlo estimate on the integral is:

$$\sigma_I \approx \frac{\sigma_f}{\sqrt{N}} = \sqrt{\frac{1}{N}\left[\frac{1}{N}\sum_{i=1}^N f(\boldsymbol{x}_i)^2 - \left(\frac{1}{N}\sum_{i=1}^N f(\boldsymbol{x}_i)\right)^2\right]}, \tag{23}$$

so that unless the function is very smooth and hence has a small variance, it will have a large error. It is equally likely that samples are taken from the subsets of the domain where $f(\boldsymbol{x}_i)$ is very small as when it is large, hence the integral converges very slowly. A solution to this problem is to use importance sampling. In importance sampling, rather than sample uniformly from the domain, samples are selected according to a probability distribution that gives preference to samples that give dominant contributions to the integral. To realise importance sampling we

## 3.2 Hybrid Monte Carlo

### 3.2.1 Hamiltonian Dynamics

Due to the fundamental importance of Hamiltonian dynamics in the Hybrid-Monte-Carlo 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**:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \tag{24}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \tag{25}$$

for $i = 1, 2, \ldots d$, provide the time evolution of the vectors $\boldsymbol{q}$ and $\boldsymbol{p}$. 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:

$$H(\boldsymbol{q}, \boldsymbol{p}) = U(\boldsymbol{q}) + K(\boldsymbol{p}), \tag{26}$$

Where $U(\boldsymbol{q})$ and $K(\boldsymbol{p})$ are known as the *potential* and *kinetic* terms respectively. Equations 24 and 25 then become:

$$\frac{dq_i}{dt} = \frac{\partial K}{\partial p_i} \tag{27}$$

$$\frac{dp_i}{dt} = -\frac{\partial U}{\partial q_i}. \tag{28}$$

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 Neal's review of Hamiltonian Monte-Carlo methods [2] it is argued that a successful method for this is *leapfrog integration*. In the leapfrog method to make the small time step $\epsilon$ from $t$ to $t + \epsilon$ on position and momentum we use the equations:

$$p_i(t + \epsilon/2) = p_i(t) + \epsilon/2 \frac{dp_i}{dt}(t) = p_i(t) - \epsilon/2 \frac{\partial U}{\partial q_i}(\boldsymbol{q}(t)), \tag{29}$$

$$q_i(t + \epsilon) = q_i(t) + \epsilon \frac{dq_i}{dt}(t + \epsilon/2) = q_i(t) + \epsilon \frac{\partial K}{\partial p_i}(\boldsymbol{p}(t + \epsilon/2)), \tag{30}$$

$$p_i(t + \epsilon) = p_i(t + \epsilon/2) + \epsilon/2 \frac{dp_i}{dt}(t + \epsilon/2) = p_i(t + \epsilon/2) - \epsilon/2 \frac{\partial U}{\partial q_i}(\boldsymbol{q}(t + \epsilon)), \tag{31}$$

where in the second equality in each step we have made the substitution for Hamilton's equations. In equation 29 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 30. We end in equation 31 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 $\epsilon$ 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:

$$p_i(\epsilon/2) = p_i(0) - \epsilon/2 \frac{\partial U}{\partial q_i}(\boldsymbol{q}(0)), \tag{32}$$

   and a full step in position:

$$q_i(\epsilon) = q_i(0) + \epsilon \frac{\partial K}{\partial p_i}(\boldsymbol{p}(\epsilon/2)). \tag{33}$$

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

$$p_i\left(n\epsilon+\epsilon/2\right)=p_i\left(n\epsilon-\epsilon/2\right)-\epsilon\frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(n\epsilon\right)\right), \tag{34}$$

$$q_i\left(n\epsilon+\epsilon\right)=q_i\left(n\epsilon\right)+\epsilon\frac{\partial K}{\partial p_i}\left(\boldsymbol{p}\left(n\epsilon+\epsilon/2\right)\right), \tag{35}$$

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

3. Then a final half step in momentum:

$$p_i\left(l\epsilon\right)=p_i\left(l\epsilon-\epsilon/2\right)-\epsilon/2\frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(l\epsilon\right)\right). \tag{36}$$

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 by Goldstein, Poole and Safko's in [3]. For the discrete dynamics integrated via the leap frog scheme we can easily verify that volume preservation is still the case. The leap frog equations 29, 30 and 31 which evolve the system a single step in time from $t$ to $t+\epsilon$ can be rewritten using matrices as:

$$\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & 0_{d\times d}\\ A & I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t\right)\end{bmatrix} \tag{37}$$

$$\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & \epsilon\\ 0_{d\times d} & I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix} \tag{38}$$

$$\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & 0_{d\times d}\\ B & I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix} \tag{39}$$

where the off diagonal $d\times d$ matrices $A$ and $B$ are determined by the partial derivatives of the potential evaluated at $t$ and $t+\epsilon$ respectively. The form of $A$ and $B$ is not important here since the other off diagonal term is always a zero matrix (the transformation is always a sheer) so these elements make no contribution to the determinant which we will see below leads to preservation of volume. Combing equations 37, 38 and 39 we have that:

$$\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon\right)\end{bmatrix}=T\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t\right)\end{bmatrix} \tag{40}$$

where the $2d\times 2d$ matrix $T$ is the Jacobian matrix of the transformation from $\left(\boldsymbol{q}\left(t\right),\boldsymbol{p}\left(t\right)\right)$ to $\left(\boldsymbol{q}\left(t+\epsilon\right),\boldsymbol{p}\left(t+\epsilon\right)\right)$ given by:

$$T=\begin{bmatrix}I_{d\times d} & 0_{d\times d}\\ A & I_{d\times d}\end{bmatrix}\begin{bmatrix}I_{d\times d} & \epsilon\\ 0_{d\times d} & I_{d\times d}\end{bmatrix}\begin{bmatrix}I_{d\times d} & 0_{d\times d}\\ B & I_{d\times d}\end{bmatrix} \tag{41}$$

from which it is clear that $\det(T) = 1$ and therefore we have preservation of volume in phase space for a single step of the leap frog method. Clearly if volume is preserved for a single step it must be preserved for an arbitrary number of steps as well. 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.

It is also obvious that the leapfrog integration scheme provides reversible dynamics; by negating the momenta at the end of a trajectory and applying the equations a second time we will return to the initial point in phase space.

### 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 of the system that depends on $d$ variables, the canonical Boltzmann distribution (probability density function for the states of the system) is given by:

$$P(\boldsymbol{x}) = \frac{1}{Z} \exp\left(-\frac{E(\boldsymbol{x})}{T}\right), \tag{42}$$

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

$$Z = \int \prod_{i=1}^{d} dx_i \exp\left(-\frac{E(\boldsymbol{x})}{T}\right). \tag{43}$$

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

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:

$$P(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{Z} \exp\left(-\frac{H(\boldsymbol{q}, \boldsymbol{p})}{T}\right), \tag{44}$$

with

$$Z = \int \prod_{i=1}^{d} dx_i dp_i \exp\left(\frac{-H(\boldsymbol{q}, \boldsymbol{p})}{T}\right) \tag{45}$$

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

$$P(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{Z_U} \exp\left(-\frac{U(\boldsymbol{q})}{T}\right) \frac{1}{Z_K} \exp\left(-\frac{K(\boldsymbol{p})}{T}\right) = P(\boldsymbol{q}) P(\boldsymbol{p}), \tag{46}$$

so the variables $\boldsymbol{q}$ and $\boldsymbol{p}$ and their respective probability distributions are independent. 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 $P(\boldsymbol{q}, \boldsymbol{p})$. Therefore, if we have some probability distribution $P(\boldsymbol{x})$ for states $\boldsymbol{x}$ we wish to sample, we may through equation 42 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 44, but by the factorization in equation 46 this gives samples of $\boldsymbol{x}$ according to out 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:

$$K(\boldsymbol{p}) = \sum_{i=1}^{d} \frac{p_i^2}{2} \tag{47}$$

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 then give a more detailed explanation of each stage and its implementation.

0. Provide an initial sample $\boldsymbol{q}$
1. Generate $\boldsymbol{p}$ from the multivariate Gaussian Distribution $\mathcal{N}(\boldsymbol{0}, I_{d\times d})$
2. Evolve the variables from $(\boldsymbol{q}, \boldsymbol{p})$ to $(\boldsymbol{q}^*, \boldsymbol{p}^*)$ using Hamilton's equations of motion.
3. Accept the proposed state $(\boldsymbol{q}^*, \boldsymbol{p}^*)$ with probability:

   $$\min\left[1, \exp\left(-H_{HMC}(\boldsymbol{q}^*, \boldsymbol{p}^*) + H_{HMC}(\boldsymbol{q}, \boldsymbol{p})\right)\right]$$

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

In step 0 we provide an initial state. This can be chosen or generated randomly; some starting configurations are better than others. However, since we normally disregard the samples generated in the early HMC iterations (see section ) this choice is not critical.

Good place for a reference

Reference section on thermalization

In step 1 we draw the fictitious momenta according to their probability distribution. The choice of kinetic energy in equation 47 and the form of equation 42 means that:

$$P(\boldsymbol{p}) = \frac{1}{Z_K} \exp\left(-\sum_{i=1}^{d} \frac{p_i^2}{2}\right) \tag{48}$$

$$= \frac{1}{Z_K} \prod_{i}^{d} \exp\left(-\frac{p_i^2}{2}\right) \tag{49}$$

and

$$Z_k = \int \prod_{i=1}^{d} dp_i \exp\left(-\frac{1}{2}\sum_{i=1}^{d} p_i^2\right) \tag{50}$$

$$= \int \prod_{i=1}^{d} dp_i \prod_{i=1}^{d} \exp\left(-\frac{p_i}{2}\right) \tag{51}$$

$$= \prod_{i=1}^{d} Z_{K_i} \tag{52}$$

such that

$$Z_{K_i} = \int dp_i \exp\left(-\frac{1}{2}p_i\right) = \sqrt{2\pi}. \tag{53}$$

We can then see that the canonical distribution function for the momentum factorises into the marginal distribution functions, that is:

$$P(\boldsymbol{p}) = \prod_{i=1}^{d} \frac{1}{Z_{K_i}} P(p_i) \tag{54}$$

where each marginal distribution $P(p_i)$ is given by:

$$P(p_i) = \frac{1}{Z_{k_i}} \exp\left(-\frac{p_i^2}{2}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{p_i^2}{2}\right) \tag{55}$$

which is a Gaussian distribution of mean 0 and variance 1. Each component $p_i$ of $\boldsymbol{p}$ is therefore drawn from a Gaussian $\mathcal{N}(0,1)$, so we draw $\boldsymbol{p}$ from a multivariate Gaussian distribution of mean $\mathbf{0}$ and covariance $I_{d\times d}$ .

In step 2 we evolve the variables $(\boldsymbol{q}, \boldsymbol{p})$ by integrating Hamilton's equations. In order to implement this in a computer simulation we use the leapfrog method described in section 3.2.1. We have a choice of the number of time steps $l$ and their size $\epsilon$; it is common practice in HMC to choose $\epsilon$ and $l$ such that $l\epsilon = 1$.

In step 3 we accept or reject the proposed state with a probability given by:

$$P(\boldsymbol{q}^*, \boldsymbol{p}^*, \boldsymbol{q}, \boldsymbol{p}) = \min\left[1, \exp\left(-H_{HMC}(\boldsymbol{q}^*, \boldsymbol{p}^*) + H_{HMC}(\boldsymbol{q}, \boldsymbol{p})\right)\right], \tag{56}$$

which is known as a *Metropolis update*. The interpretation of 56 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 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 the current state $\boldsymbol{q}$.

From the Metropolis update in equation 56 we can show that for smaller value of $\epsilon$ 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:

$$\frac{dH}{dt} = \sum_{i=1}^{d} \left[ \frac{dq_i}{dt}\frac{\partial H}{\partial q_i} + \frac{dp_i}{dt}\frac{\partial H}{\partial p_i} \right] = \sum_{i=1}^{d} \left[ \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial H}{\partial p_i} \right] = 0 \tag{57}$$

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 unity.

As $\epsilon$ 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 $\epsilon$ in the trajectory. Conversely if we have a fixed computation time, then for smaller $\epsilon$ 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 by Beskos et al. in [4] that an acceptance rate of $\sim 65.1\%$ provides an optimal balance between the cost of generating proposals and the total number of samples.

# 4 Results and Discussion

## 4.1 Quantum Oscillators

In section 2.1 we showed in equation 18 that by discretising the path integral and performing a Wick rotation into imaginary time the path integral of quantum mechanics is a partition function if $S_E(\{x_i\}) = H(x_i)$ (working now in units where $\hbar = T = k_B = 1$). The Hamiltonian is an energy function of the system and so we may use equation 42 with $E(\boldsymbol{x}) = S_E(\boldsymbol{x})$ to get the Boltzmann distribution:

$$P(\boldsymbol{x}) = \frac{1}{Z}\exp\left(-S_E(\boldsymbol{x})\right). \tag{58}$$

The probability distribution of equation 58 is the distribution we wish to draw are samples according to for the Monte Carlo simulations, therefore we follow the method of section 3.2.2 with this distribution, taking $\boldsymbol{q} \equiv \boldsymbol{x}$,

$$U(\boldsymbol{q}) \equiv S_E(\boldsymbol{x}) \tag{59}$$

$$= \sum_{i=0}^{N-1} a \left[ \frac{1}{2} m \left( \frac{x_{i+1} - x_i}{a} \right)^2 + V(x_i) \right] \tag{60}$$

and the kinetic energy $K(\boldsymbol{p})$ as in equation 47. We define our HMC Hamiltonian:

$$H_{hmc}(\boldsymbol{q},\boldsymbol{p}) \equiv K(\boldsymbol{p}) + U(\boldsymbol{q}) \tag{61}$$

$$= \sum_{i=0}^{N-1} \frac{p_i^2}{2} + S_E(\boldsymbol{q}) \tag{62}$$

$$= \sum_{i=0}^{N-1} \frac{p_i^2}{2} + \sum_{i=0}^{N-1} \epsilon \left[ \frac{1}{2} m \left( \frac{q_{i+1} - q_i}{\epsilon} \right)^2 + V(q_i) \right], \tag{63}$$

we can then use this Hamiltonian to generate samples with the HMC algorithm for any 1-dimensional quantum mechanical system, by specifying the potential $V(x)$. With the choice of kinetic energy given in equation 47 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 $\epsilon$ taking the form:

1. Half step in momentum:

$$p_i(\epsilon/2) = p_i(0) - \epsilon/2 \left[ \frac{m}{a} (2q_i(0) - q_{i-1}(0) - q_{i+1}(0)) + a \frac{\partial V(q_i(0))}{\partial q_i} \right], \tag{64}$$

   full step in position:

$$q_i(\epsilon) = q_i(0) + \epsilon p_i(\epsilon/2). \tag{65}$$

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

$$p_i(n\epsilon + \epsilon/2) = p_i(n\epsilon - \epsilon/2) - \epsilon \left[ \frac{m}{a} (2q_i(n\epsilon) - q_{i-1}(n\epsilon) - q_{i+1}(n\epsilon)) + a \frac{\partial V(q_i(n\epsilon))}{\partial q_i} \right], \tag{66}$$

$$q_i(n\epsilon + \epsilon) = q_i(n\epsilon) + \epsilon p_i(n\epsilon + \epsilon/2), \tag{67}$$

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

3. Final half step in momentum:

$$p_i(\epsilon l) = p_i(l\epsilon - \epsilon/2) - \epsilon/2 \left[ \frac{m}{a} (2q_i(\epsilon l) - q_{i-1}(\epsilon l) - q_{i+1}(\epsilon l)) + a \frac{\partial V(q_i(\epsilon l))}{\partial q_i} \right]. \tag{68}$$

### 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:

$$V(x) = \frac{1}{2}\mu^2 x^2 \tag{69}$$

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

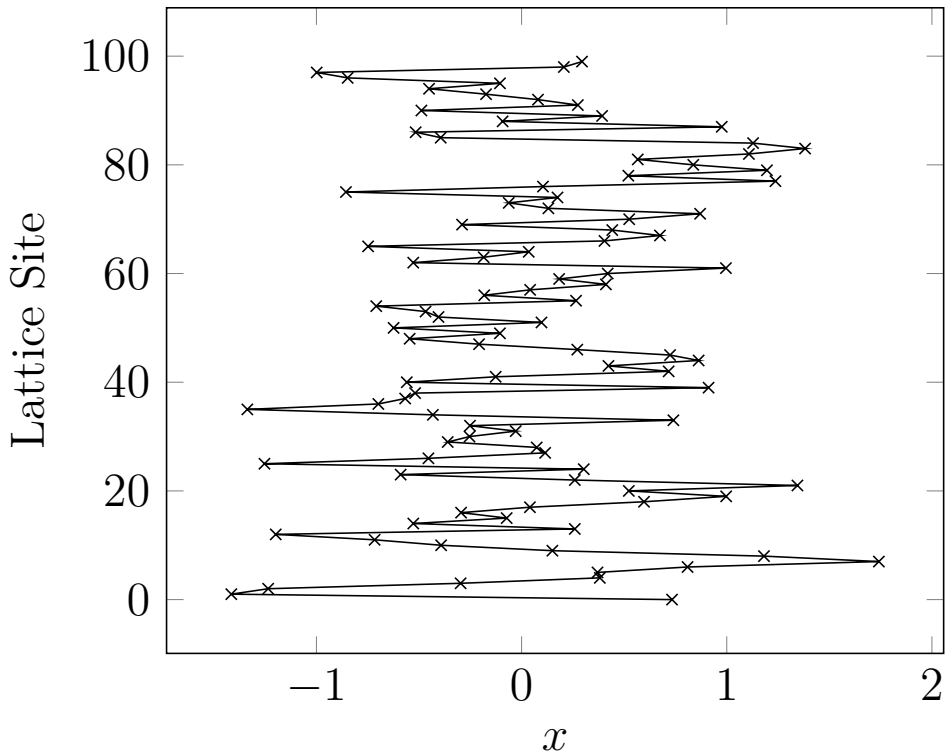



Figure 3. Typical configuration for the harmonic oscillator with $\mu^2 = 1, m = 1, a = 1, L = 1000, d = 0.1, N = 10,$ configurations $= 100000,$ burn period $= 1000$.

**Typical Trajectory** Figure 3 shows a typical path 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. 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:

$$\langle x^2 \rangle = \frac{1}{2\mu \left(m + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}}} \left(\frac{1 + R^N}{1 - R^N}\right), \tag{70}$$

where

$$R = 1 + \frac{a^2\mu^2}{2m} - \frac{a\mu}{\sqrt{m}} \left(1 + \frac{a^2\mu^2}{4m}\right)^{\frac{1}{2}}, \tag{71}$$

see [1] or appendix B for a full derivation of equation 70. Running the HMC simulation at a range of lattice spacings we obtain the results in figure 4.

**Energy Levels** In order to calculate the ground state energy of the system we use the virial theorem:

$$2 \langle \hat{T} \rangle = \langle x V'(x) \rangle \tag{72}$$

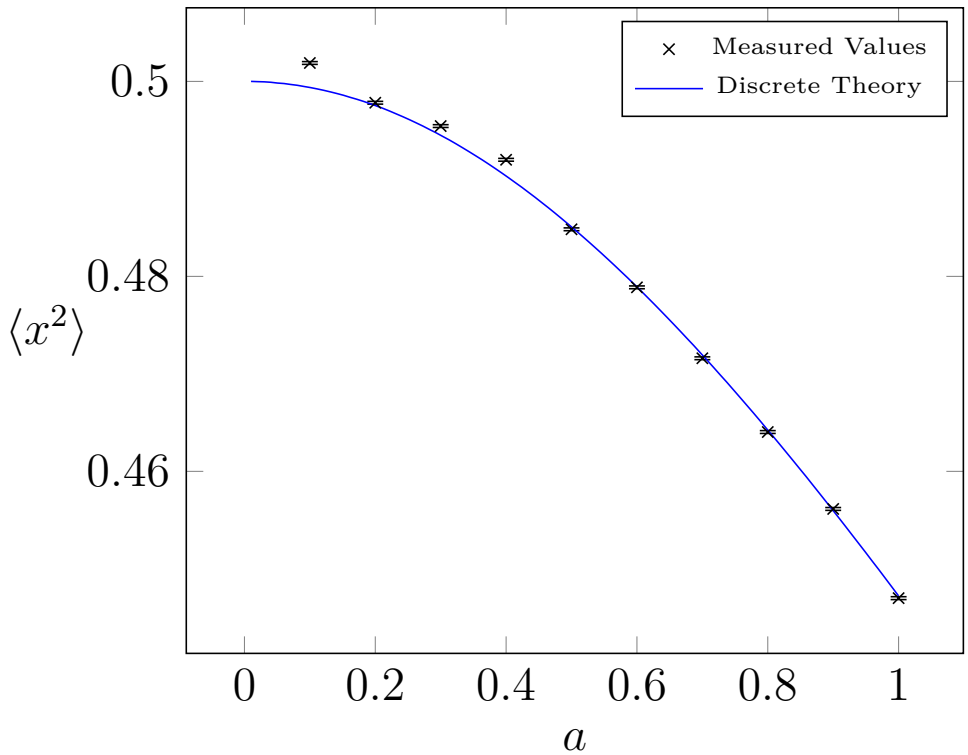



Figure 4. Relationship between lattice spacing and expectation of position squared for the harmonic oscillator with $\mu^2 = 1, m = 1$.

which relates the expectation value of kinetic energy to that of the potential and the derivation of which can be found in Binney and Skinner's The Physics of Quantum Mechanics [5] and was orginally proved by Fock in [6]. Applying this to the harmonic potential of equation 69 we get:

$$E = \langle \hat{T} \rangle + \langle V(x) \rangle \tag{73}$$
$$= m\mu^2 \langle x^2 \rangle , \tag{74}$$

so that for measurements made in the ground state:

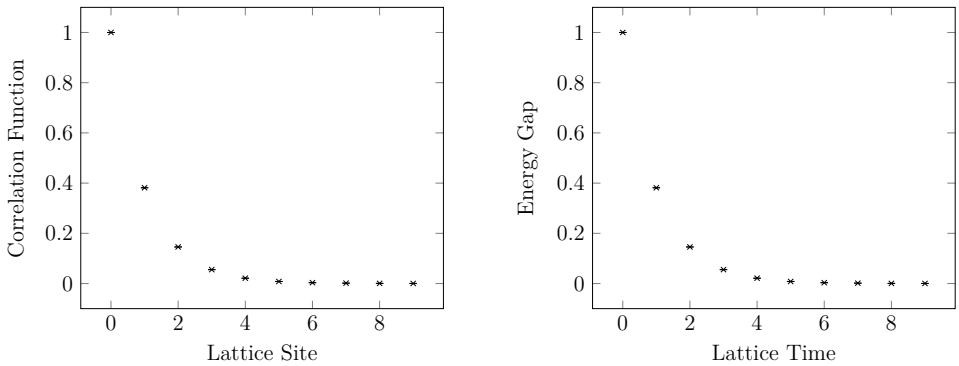
$$E_0 = m\mu^2 \langle x^2 \rangle \tag{75}$$

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. When we recover the continuum limit as $a \to 0$ and $N \to \infty$ we see that $E_0$ approaches the value of $\frac{1}{2}$ which is given by:

$$E_n = \mu \left( n + \frac{1}{2} \right) . \tag{76}$$

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

$$\Delta E(\delta\tau) = -\frac{1}{\delta\tau} \ln C(\delta\tau) , \tag{77}$$




(a) Correlation function. (b) Energy gap

Figure 5. Energy gap and correlation function for the harmonic oscillator with $a = 1$, $\mu = 1$ and $m = 1$.

where:

$$C(\delta\tau) = \frac{\sum_{i=0}^{N-1} x_i x_{i+\delta\tau}}{\sum_{i=0}^{N-1} x_i^2} \tag{78}$$

is the *correlation function*. In equation 78 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 77 see appendix Figures **??** and **??** show 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 76. 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 $\mu^2$. 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.

Is this okay?

Insert reference to appendix.

Why doesn't it stay constant?

Make data for the plot and include the plot

**Ground State Wave Function** Our method for measuring the ground state wave function follows that of [1]. We discretise position into $K$ bins of size $\Delta x$. 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 $x + \Delta x$ and as $\Delta x \to 0$, this becomes the probability density function that is the square amplitude of the wave function. On any configuration position coordinates are therefore binned according to:

### 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 the Hybrid Monte Carlo method for calculations we do not already know the answer to. Rather than use the more common form of the anharmonic quartic potential:

$$V(x) = \frac{1}{2} m\mu^2 x^2 + \lambda x^4 \tag{79}$$

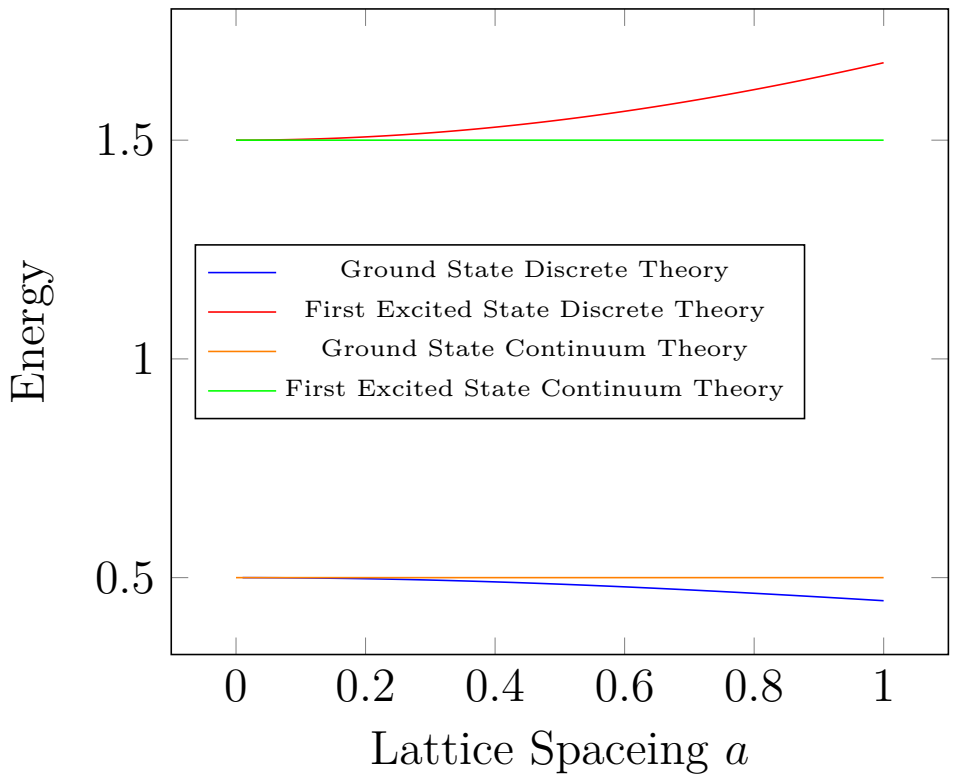



Figure 6. Ground and first excited state energies for the HMC simulation with the continuum and discrete theory results.

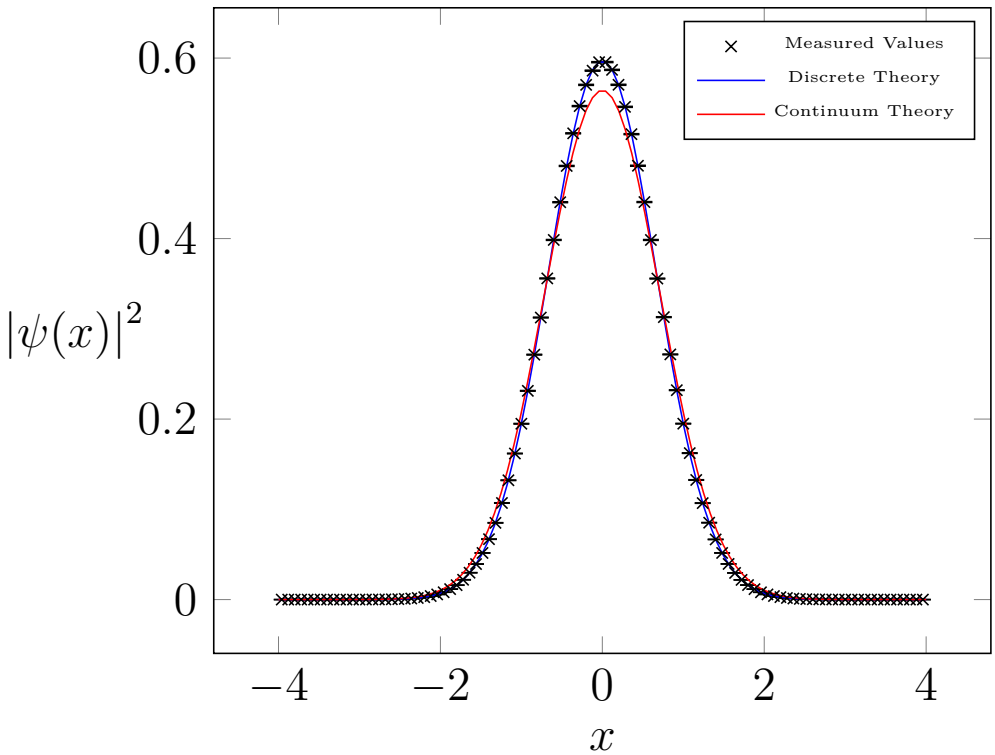



Figure 7. Continuum, discrete and measured wave functions for the harmonic oscillator with $\mu^2 = 1, m = 1, a = 1, L = 1000, d = 0.1, N = 10,$ configurations $= 100000,$ burn period $= 1000$.

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

$$V(x) = \lambda\left(x^2 - f^2\right) \tag{80}$$

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 potential 80 is equivalent to potential 79 up to the additive constant $\lambda f^4$, the action and consequently the HMC Hamiltonian for the system is the same up to the additive constant $\lambda f^4$ also. This makes no difference to the measured quantities 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.

**Typical Trajectory** Figure **??** shows a typical trajectory for the HMC simulation of the anharmonic oscillator. We can now observe the tunnelling behaviour we expect for a quantum 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

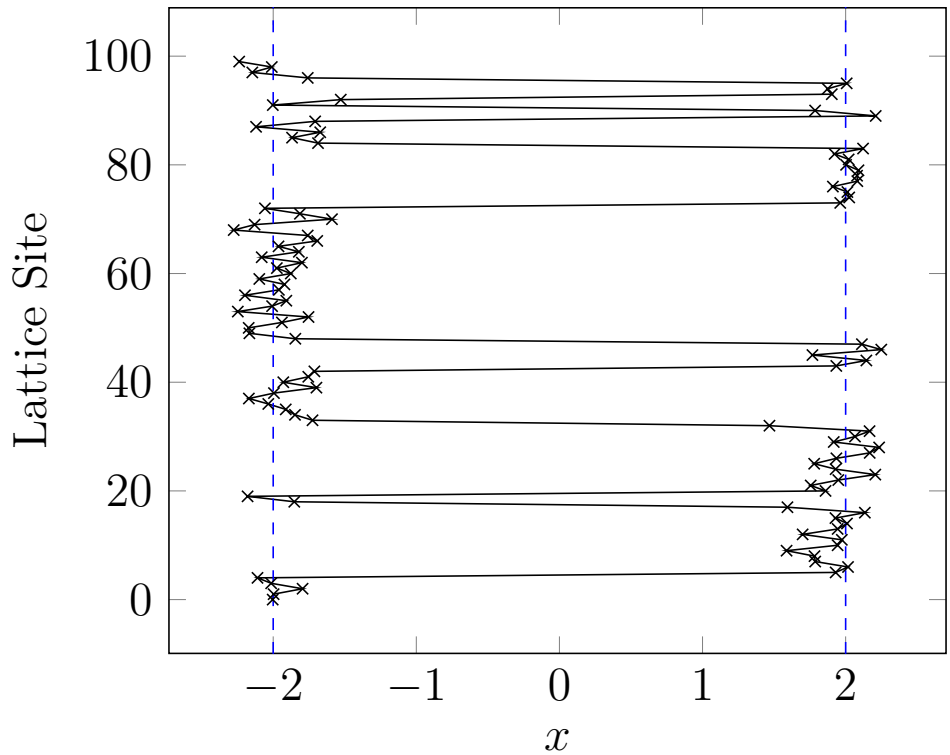



Figure 8. Typical configuration for the anharmonic oscillator with $\lambda = 1, f^2 = 4, m = 1, a = 1, L = 1000, d = 0.01, N = 100,$ configurations $= 100000,$ burn period $= 1000$.

theorem of equation 72 which gives:

$$E_0 = 3\lambda \langle x^4 \rangle - 4\lambda f^2 \langle x^2 \rangle + \lambda f^4. \tag{81}$$

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 **??** for a range of values of $f^2$ and lattice spacings.

How do we do this?

**Ground State Wave Function** To measure the ground state wave 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 10

**Isolated Modes** In figure 8 we saw the tunnelling behaviour displayed by the anharmonic oscillator as the particle tunnelled between the minima of the symmetric potential wells at $x = \pm f$. By the Boltzmann distribution of equation 42 the minima of the potential correspond to maxima of the Boltzmann distribution, which are separated by a region of low probability. We refer to the two regions around $x\pm = f$ where there is high probability of finding the particle as *isolated modes*. In figure **??** we investigate the effect of varying the value of $f^2$ which varies the distance between the isolated modes.

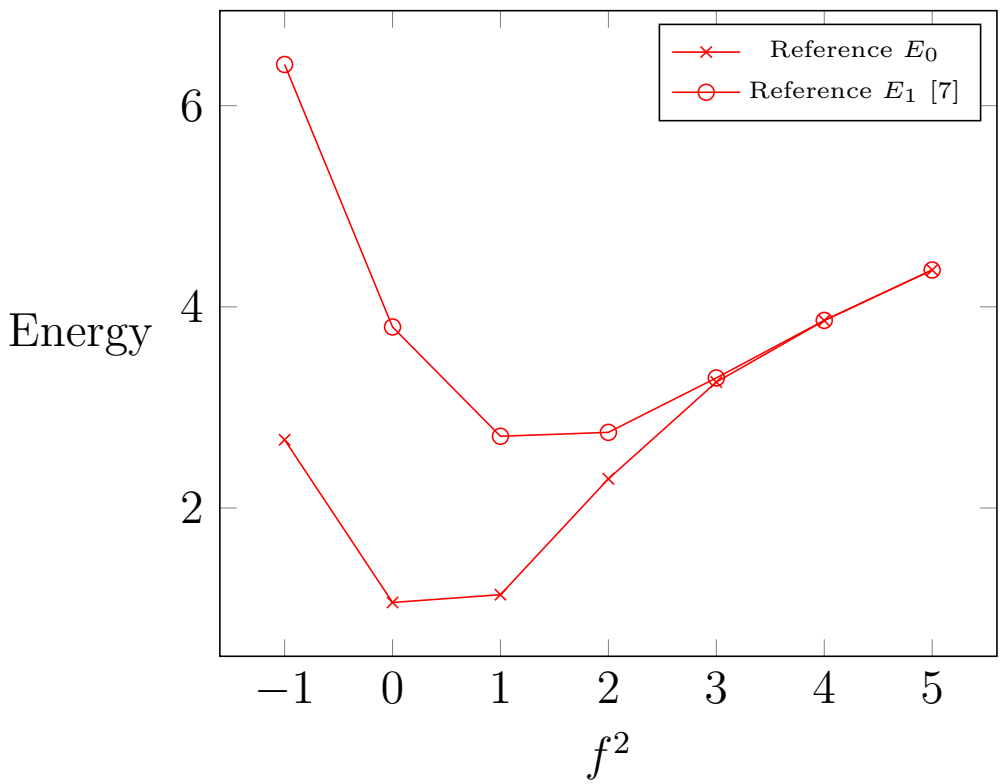



Figure 9. Ground and first excited state energies for the anharmonic oscillator with $\lambda = 1, m = 1, L = 1000 \text{configurations} = 100000, \text{burn period} = 1000$.

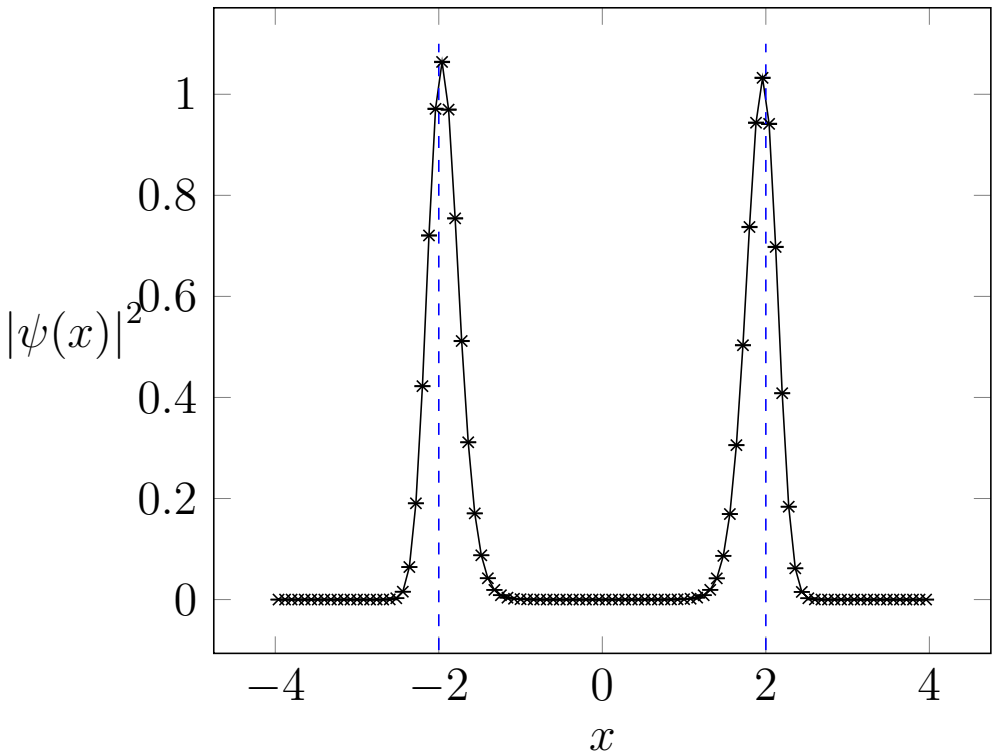
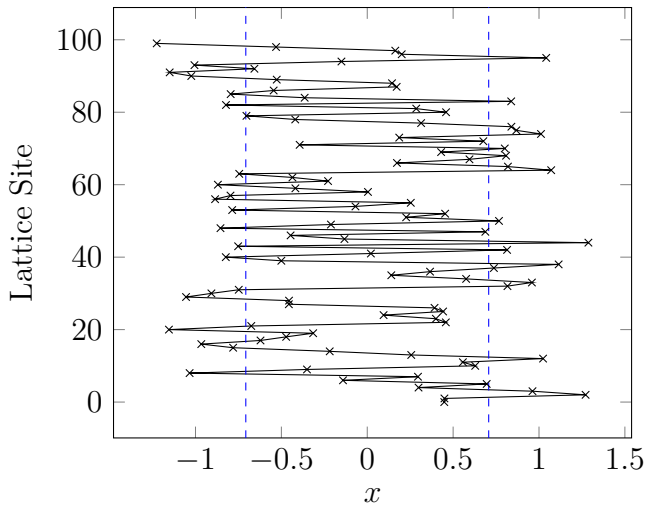



Figure 10. Measured wave function for the harmonic oscillator with $\lambda = 1, f^2 = 4m = 1, a = 1, L = 1000, d = 0.01, N = 100,$ configurations $= 100000,$ burn period $= 1000$.

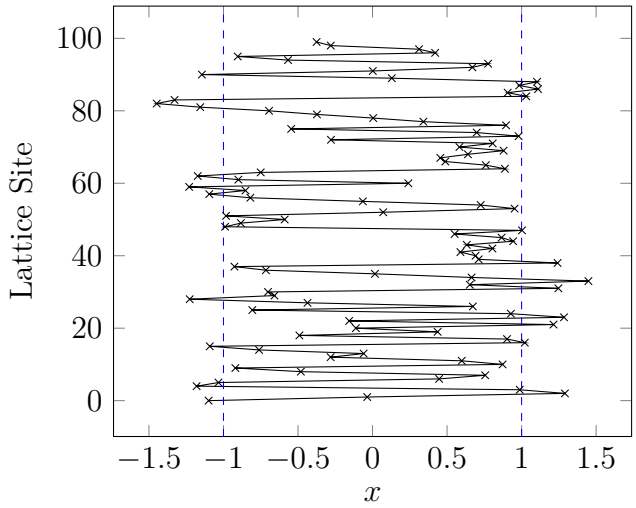
## 4.2 Accelerated Dynamics (Tempering)

We have seen above that Hybrid Monte Carlo much like other MCMC methods [2] 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. A solution to this problem proposed in [2] is to use *tempering* dynamics to try and sample more frequently from modes separated by a region of low probability. In HMC tempering is incorporated directly into the Hamiltonian dynamics and here we follow the method given in [2].

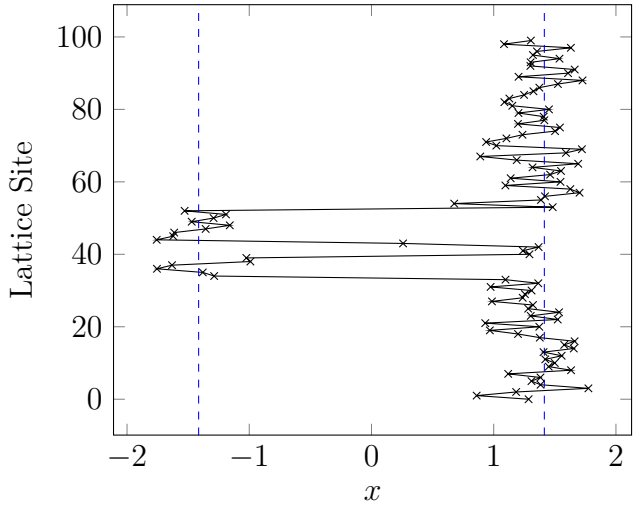
In order to implement tempering on a trajectory that is used to propose an update in the HMC algorithm we multiply each momentum variable in the first half of the trajectory by $1+\epsilon$ where $\epsilon$ is small, and divide each momentum variable in the second half of the trajectory by $1+\epsilon$. We will follow the symmetric method given in [2] and defining $\alpha = 1+\epsilon$ multiply each momentum variable before the first half step (equation 29) and after the second half step (equation 29) in the leapfrog update by $\sqrt{\alpha}$ in the first half of the trajectory and then in the second half of the trajectory to replace this multiplication with division $\sqrt{\alpha}$. If there is an odd number of steps in the leapfrog trajectory then on the middle step we 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 $\alpha$ is equal.




(a) $f^2 = 0.5$




(b) $f^2 = 1$




(c) $f^2 = 2.0$

Figure 11. Typical configuration for a harmonic oscillator with $a = 1$, $N = 100$ and $\lambda = 1$ when varying $f^2$.

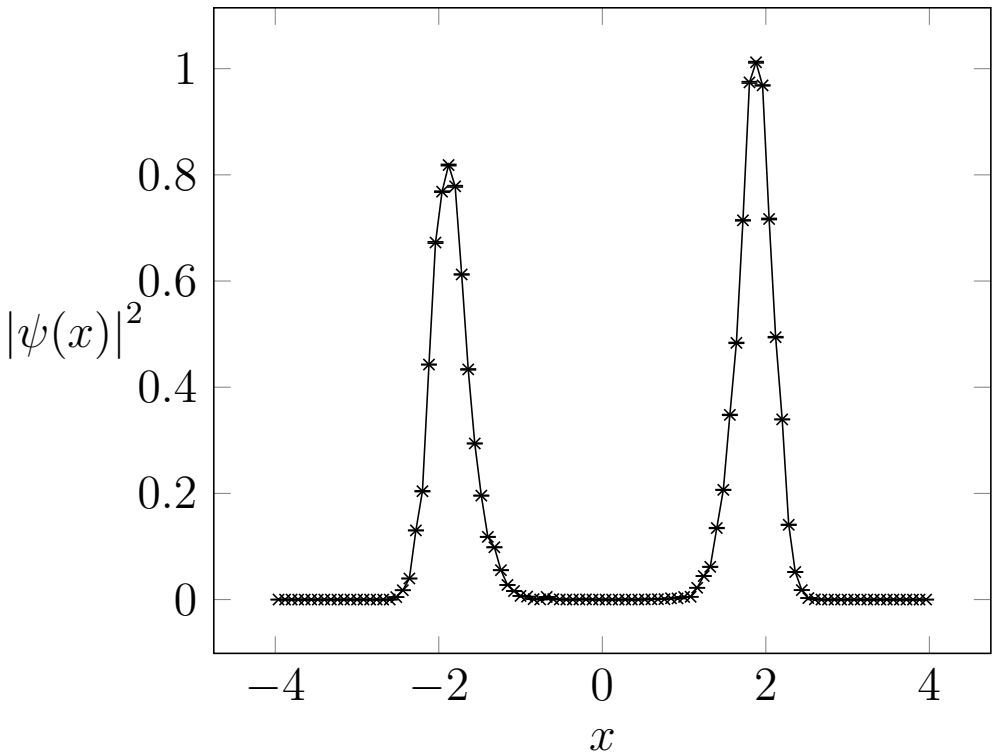



Figure 12. Measured wave function for the harmonic oscillator with $\lambda = 1, f^2 = 4m = 1, a = 1, L = 1000, d = 0.2, N = 5,$ configurations $= 20000,$ burn period $= 1000$.

With tempering introduced the leapfrog equations 29,30 and 31 become:

$$p_i\left(t+\epsilon/2\right) = \sqrt{\alpha}p_i\left(t\right) - \epsilon/2\frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t\right)\right), \tag{82}$$

$$q_i\left(t+\epsilon\right) = q_i\left(t\right) + \epsilon\frac{\partial K}{\partial p_i}\left(\boldsymbol{p}\left(t+\epsilon/2\right)\right) \tag{83}$$

$$p_i\left(t+\epsilon\right) = \sqrt{\alpha}\left(p_i\left(t+\epsilon/2\right) - \epsilon/2\frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t+\epsilon\right)\right)\right), \tag{84}$$

for steps in the first half of the trajectory and

$$p_i\left(t+\epsilon/2\right) = \frac{1}{\sqrt{\alpha}}p_i\left(t\right) - \epsilon/2\frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t\right)\right), \tag{85}$$

$$q_i\left(t+\epsilon\right) = q_i\left(t\right) + \epsilon\frac{\partial K}{\partial p_i}\left(\boldsymbol{p}\left(t+\epsilon/2\right)\right) \tag{86}$$

$$p_i\left(t+\epsilon\right) = \frac{1}{\sqrt{\alpha}}\left(p_i\left(t+\epsilon/2\right) - \epsilon/2\frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t+\epsilon\right)\right)\right), \tag{87}$$

in the second half of the trajectory, with the possible exception that for an odd number of

leap frog steps on the middle step:

$$p_i\left(t+\epsilon/2\right)=\sqrt{\alpha}p_i\left(t\right)-\epsilon/2\frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t\right)\right), \tag{88}$$

$$q_i\left(t+\epsilon\right)=q_i\left(t\right)+\epsilon\frac{\partial K}{\partial p_i}\left(\boldsymbol{p}\left(t+\epsilon/2\right)\right) \tag{89}$$

$$p_i\left(t+\epsilon\right)=\frac{1}{\sqrt{\alpha}}\left(p_i\left(t+\epsilon/2\right)-\epsilon/2\frac{\partial U}{\partial q_i}\left(\boldsymbol{q}\left(t+\epsilon\right)\right)\right). \tag{90}$$

Just as in section 3.2.1 we still require that volume is preserved by the leapfrog integration scheme, 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 write the new tempered leap frog equations in matrix form as as:

$$\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & 0_{d\times d}\\ A & \sqrt{\alpha}I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t\right)\end{bmatrix} \tag{91}$$

$$\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & \epsilon\\ 0_{d\times d} & I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix} \tag{92}$$

$$\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & 0_{d\times d}\\ B & \sqrt{\alpha}I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix} \tag{93}$$

$$\tag{94}$$

for single step from $t$ to $t+\epsilon$ in the first half of the trajectory and

$$\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & 0_{d\times d}\\ A & \frac{1}{\sqrt{\alpha}}I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t\right)\end{bmatrix} \tag{95}$$

$$\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & \epsilon\\ 0_{d\times d} & I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix} \tag{96}$$

$$\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & 0_{d\times d}\\ B & \frac{1}{\sqrt{\alpha}}I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix} \tag{97}$$

in the second half of the trajectory. Where as before $A$ will depend on the partial derivatives of $U\left(q\right)$ evaluated at $t$ and $B$ the partial derivatives at $t+\epsilon$ however since they are off diagonal terms with the opposite off diagonal terms being zero, there form is not relevant for the argument. It is also possible that for an odd number of steps the middle step will be of the form:

$$\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & 0_{d\times d}\\ A & \sqrt{\alpha}I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t\right)\end{bmatrix} \tag{98}$$

$$\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & \epsilon\\ 0_{d\times d} & I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix} \tag{99}$$

$$\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon\right)\end{bmatrix}=\begin{bmatrix}I_{d\times d} & 0_{d\times d}\\ B & \frac{1}{\sqrt{\alpha}}I_{d\times d}\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\left(t+\epsilon\right)\\ \boldsymbol{p}\left(t+\epsilon/2\right)\end{bmatrix}. \tag{100}$$

$$\tag{101}$$

From the above we can see that for a step of size $\epsilon$ in the first half of the trajectory the transformation that takes us from $(\boldsymbol{q}(t), \boldsymbol{p}(t))$ to $(\boldsymbol{q}(t), \boldsymbol{p}(t))$ given by:

$$\begin{bmatrix} \boldsymbol{q}(t+\epsilon) \\ \boldsymbol{p}(t+\epsilon) \end{bmatrix} = T \begin{bmatrix} \boldsymbol{q}(t) \\ \boldsymbol{p}(t) \end{bmatrix} \tag{102}$$

is

$$T = \begin{bmatrix} I_{d\times d} & 0_{d\times d} \\ B & \sqrt{\alpha} I_{d\times d} \end{bmatrix} \begin{bmatrix} I_{d\times d} & \epsilon \\ 0_{d\times d} & I_{d\times d} \end{bmatrix} \begin{bmatrix} I_{d\times d} & 0_{d\times d} \\ A & \sqrt{\alpha} I_{d\times d} \end{bmatrix} \tag{103}$$

the determinant of which is no longer unity $det(T) = \alpha$ so that a single step in the first half trajectory does not preserve volume. This is because with tempering the diagonal terms in the matrix are no longer all unity. For steps in the second half trajectory we may construct an equivalent Jacobian matrix:

$$S = \begin{bmatrix} I_{d\times d} & 0_{d\times d} \\ B & \frac{1}{\sqrt{\alpha}} I_{d\times d} \end{bmatrix} \begin{bmatrix} I_{d\times d} & \epsilon \\ 0_{d\times d} & I_{d\times d} \end{bmatrix} \begin{bmatrix} I_{d\times d} & 0_{d\times d} \\ A & \frac{1}{\sqrt{\alpha}} I_{d\times d} \end{bmatrix} \tag{104}$$

and we see $det(S) = 1/\alpha$. For an even number of steps of size $\epsilon$ it is clear that these factors will always cancel when we consider the Jacobian that takes us from $t$ to $t + 2N\epsilon$ formed by the product of an even number of matrices of the form $S$ and $T$ and take its determinant. For an odd number of steps there will be an extra term in the middle of the product of matrices of the form 98, however its determinant is unity so it does not contribute to any change in volume.

# 5 Conclusion

# References


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

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

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

[4] 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.

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

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

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

# 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:

$$\hat{H}\left(\hat{x},\hat{p}\right)=\hat{T}\left(\hat{p}\right)+\hat{V}\left(\hat{x}\right) \tag{105}$$

, where

$$\hat{T}\left(\hat{p}\right)=\frac{1}{2m}\hat{p}^{2}, \tag{106}$$

$$\hat{p}=-i\frac{d}{dx} \tag{107}$$

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

$$\hat{V}\left(\hat{x}\right)=V\left(x\right) \tag{108}$$

so the potential is time independent.

Using the identity:

$$\left[\hat{A},\hat{B}\hat{C}\right]=\left[\hat{A},\hat{B}\right]\hat{C}+\hat{B}\left[\hat{A},\hat{C}\right], \tag{109}$$

it can be shown that:

$$\begin{aligned}\left[\hat{H},\hat{x}\hat{p}\right]&=\left[\hat{H},\hat{x}\right]\hat{p}+\hat{x}\left[\hat{H},\hat{p}\right] &(110)\\ &=\frac{1}{2m}\left[\hat{p}^{2},\hat{x}\right]\hat{p}+\hat{x}\left[V\left(x\right),\hat{p}\right] &(111)\\ &=-i\frac{\hat{p}^{2}}{m}+ixV'\left(x\right), &(112)\end{aligned}$$

so that:

$$i\left[\hat{H},\hat{x}\hat{p}\right]=2\hat{T}\left(\hat{p}\right)-xV'\left(x\right). \tag{113}$$

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

$$\frac{d\hat{A}\left(t\right)}{dt}=-i\left[\hat{H},\hat{A}\left(t\right)\right], \tag{114}$$

so that equation 113 becomes:

$$\frac{d\left(\hat{x}\hat{p}\right)}{dt}=2\hat{T}\left(\hat{p}\right)-xV'\left(x\right) \tag{115}$$

Then taking the expectation value of both sides of equation 115 and noting that :

Why is this true?

$$\langle \frac{d\left(\hat{x}\hat{p}\right)}{dt} \rangle = 0, \tag{116}$$

we get the virial theorem:

$$2\langle \hat{T} \rangle = \langle xV'\left(x\right) \rangle \tag{117}$$

# B Derivation of the discrete path integral for quantum harmonic oscillator

Here we follow the derivation given in [1] for the exact result of the path integral for discrete theory and a particle of mass $m = 1$. Since in [1] 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 18 as:

FIX INDEXING

$$Z = \int_{-\infty}^{+\infty} \prod_{i=1}^{N-1} dx_i \exp\left( -\sum_{j=0}^{N-1} a \left[ \frac{1}{2} \left( \frac{x_{j+1} - x_j}{a} \right)^2 + \frac{1}{2}\mu^2 x_j^2 \right] \right). \tag{118}$$

We begin by defining an operator $T$ such that its matrix elements in the Schr'odinger picture are given by:

$$\langle x' | \hat{T} | x \rangle = \exp\left( -\frac{1}{2a}\left(x' - x\right)^2 - \frac{\mu^2 a}{4}\left(x^2 + x'^2\right) \right). \tag{119}$$

Then we can use the completeness relation that:

$$\hat{1} = \int_{-\infty}^{\infty} |x\rangle \langle x| . \tag{120}$$

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

$$Tr\left(\hat{T}^N\right), \tag{121}$$

and using the definition of the matrix elements in equation 119, we recover the path integral in equation 118, that is:

$$Z = \sim Tr\left(\hat{T}^N\right). \tag{122}$$

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

$$Tr\left(\hat{A}\right) = \int_{-\infty}^{\infty} dx \langle x | \hat{A} | x \rangle , \tag{123}$$

which is of course basis independent.

The next step is to make the ansatz that:

$$\hat{T} = \int_{-\infty}^{\infty} d\omega e^{\frac{-\mu^2 a}{4}\hat{x}^2} e^{-i\hat{p}\omega} e^{-1\frac{1}{2a}\omega^2} e^{\frac{-\mu^2 a}{4}\hat{x^2}}, \tag{124}$$

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

$$e^{-i\hat{p}\Delta} |x\rangle = |x + \Delta\rangle \tag{125}$$

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 119. We observe that the integral over $\omega$ in equation 124 is Gaussian; we may use the standard result:

$$\int_{-\infty}^{+\infty} dx e^{-\alpha x^2 + i\beta x} = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4a}}, \tag{126}$$

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 126 to equation 124 gives:

$$\hat{T} = \sqrt{2\pi a} e^{-\frac{-\mu^2 a}{4}\hat{x}^2} e^{-\frac{a}{2}\hat{p}^2} e^{-\frac{\mu^2 a}{4}\hat{x}^2}. \tag{127}$$

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:

$$[\hat{x}, \hat{p}] = i\hbar, \tag{128}$$

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

$$\left[\hat{A}, \hat{B}^n\right] = n\hat{B}^{n-1}\left[\hat{A}, \hat{B}\right] \tag{129}$$

and

$$\left[\hat{A}^n, \hat{B}\right] = n\hat{A}^{n-1}\left[\hat{A}, \hat{B}\right], \tag{130}$$

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:

$$\left[\hat{x}, e^{-\frac{1}{2}a\hat{p}^2}\right] = -ia\hat{p} \tag{131}$$

and

$$\left[\hat{p}, e^{-\frac{\mu^2 a}{4}\hat{x}^2}\right] = i\frac{\mu^2 a}{2}\hat{x} \tag{132}$$

Using identities 131 and 132 we can easily show that:

$$\hat{x}\hat{T} = \hat{T}\left[\left(1 + \frac{a^2\mu^2}{2}\right)\hat{x} - ia\hat{p}\right], \tag{133}$$

and

$$\hat{p}\hat{T} = \hat{T}\left[\left(1 + \frac{a^2\mu^2}{2}\right)\hat{p} + ia\mu^2\left(1 + \frac{a^2\mu^2}{4}\right)\hat{x}\right]. \tag{134}$$

Iterating equations 133 and 134 a second time gives:

$$\left[\hat{p}^2 + \mu^2\left(1 + \frac{a^2\mu^2}{4}\right)\hat{x^2}, \hat{T}\right] = 0. \tag{135}$$

Defining a new angular frequency parameter $\omega$ by:

$$\omega^2 = \mu^2\left(1 + \frac{a^2\mu^2}{4}\right), \tag{136}$$

then from equation 135 we have that $\hat{T}$ commutes with the simple harmonic oscillator Hamiltonian:

$$\hat{H} = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2\hat{x}^2. \tag{137}$$

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 $\omega$, therefore we may define the corresponding ladder operators:

$$\hat{a}^\dagger = \frac{1}{\sqrt{\omega}}\left(\hat{p} + i\omega\hat{x}\right) \tag{138}$$

and

$$\hat{a} = \frac{1}{\sqrt{\omega}}\left(\hat{p} - i\omega\hat{x}\right) \tag{139}$$

which allows us to write:

$$\hat{H} = \left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)\omega \tag{140}$$

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

$$\hat{a}\left|0\right\rangle = 0, \tag{141}$$

$$\left(\hat{a}^\dagger\right)^n\left|0\right\rangle = \left|n\right\rangle, \tag{142}$$

and

$$\left\langle n|n\right\rangle = n!. \tag{143}$$

Using the identities of equations 133 and 134 we can show that:

$$\hat{a}\hat{T} = \hat{T}\hat{a}\left(1 + \frac{a^2\mu^2}{2} - a\mu\left(1 + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}}\right). \tag{144}$$

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

$$\hat{T}\,|n\rangle = \lambda_n\,|n\rangle\,. \tag{145}$$

We can then use that $\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle$ and equation 144 to show the ratio:

$$\begin{aligned}
\frac{\lambda_n}{\lambda_{n-1}} &= 1 + \frac{a^2\mu^2}{2} - a\mu\left(1 + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}} && (146)\\
&= \left(\left(1 + \frac{a^2\mu^2}{4}\right)^{\frac{1}{2}} - \frac{1}{2}a\mu\right)^2 && (147)\\
&:= R, && (148)
\end{aligned}$$

so that

$$\lambda_n = R\lambda_{n-1} = R^2\lambda_{n-2} = \cdots = R^n\lambda_0, \tag{149}$$

which gives:

$$\begin{aligned}
\hat{T}\,|n\rangle &= \lambda_n\,|n\rangle && (150)\\
&= R^n\lambda_0\,|n\rangle && (151)\\
&= \lambda_0 e^{n\log R}\,|n\rangle\,. && (152)
\end{aligned}$$

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

$$\hat{H}\,|n\rangle = \left(n + \frac{1}{2}\right)\omega\,|n\rangle\,, \tag{153}$$

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

$$\hat{T} = Ce^{K\hat{H}}, \tag{154}$$

which can be shown to be consistent with equation 152 provided we take $K\omega = \log R$ and $Ce^{\frac{K\omega}{2}} = \lambda_0$ and can conclude that:

$$\begin{aligned}
\hat{T} &= Ce^{\frac{\hat{H}}{\omega}\log R} && (155)\\
&= CR^{\frac{1}{\omega}\hat{H}}. && (156)
\end{aligned}$$

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

$$\begin{aligned}
Tr\left(\hat{T}\right) &= \sum_n \langle n|\,\hat{T}\,|n\rangle && (157)\\
&= C\sum_n R^{n+\frac{1}{2}} && (158)\\
&= C\frac{R^{\frac{1}{2}}}{1-R} && (159)\\
&= C\frac{1}{a\mu} && (160)
\end{aligned}$$

where we have used that $\hat{H}\left|n\right\rangle = E_n\left|n\right\rangle = \left(n+\frac{1}{2}\right)\omega$ 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 123, and using equation 119:

$$Tr\left(\hat{T}\right) = \int dx \left\langle x\right|\hat{T}\left|n\right\rangle \tag{161}$$

$$= \int dx e^{-\frac{a\mu^2}{2}x^2} \tag{162}$$

$$= \frac{1}{\mu}\sqrt{\frac{2\pi}{a}}, \tag{163}$$

comparing equation 163 and 160 we conclude $C = \sqrt{2\pi a}$, and so

$$T = \sqrt{2\pi a}R^{\frac{\hat{H}}{\omega}}. \tag{164}$$

We may now explicitly compute the trace expression for the discrete path integral given in equation 121 using the result for $\hat{T}$ in equation 164. We again take the trace over the energy eigenstates and use the fact that $|R| < 1$ to compute the resulting sum, this gives:

$$Z = (2\pi a R)^{\frac{N}{2}}\frac{1}{1-R^N}. \tag{165}$$

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

The correlation functions are given by:

FIND OUT WHY THIS IS TRUE

$$\langle x_i x_{i+j}\rangle = \frac{1}{Z}Tr\left(\hat{x}\hat{T}^j\hat{x}\hat{T}^{N-j}\right) \tag{166}$$

$$= \frac{1}{2\mu\left(1+\frac{a^2\mu^2}{4}\right)^{\frac{1}{2}}\left(1-R^n\right)}\left(R^j + R^{N-j}\right). \tag{167}$$

Taking $j = 0$, equation 166 becomes:

$$\langle x^2\rangle = \frac{1}{2\omega}\left(\frac{1+R^N}{1-R^N}\right) \tag{168}$$

$$= \frac{1}{2\mu\left(1+\frac{a^2\mu^2}{4}\right)^{\frac{1}{2}}}\left(\frac{1+R^N}{1-R^N}\right) \tag{169}$$

We have shown in above that the discrete theory for the harmonic oscillator leads to a quantum system with the Hamiltonian given in equation 137. This is the Hamiltonian of a quantum harmonic oscillator of angular frequency $\omega$ rather than $\mu$ and hence we employ the standard result from quantum theory that for a a quantum harmonic oscillator of angular frequency $\omega$ the ground state wave function is given by:

CHECK THIS LOGIC IS CORRECT

$$\psi(x) = \left(\frac{\omega}{\pi}\right)^{\frac{1}{4}} \exp\left(-\frac{1}{2}\omega x^2\right) \tag{170}$$

$$= \left(\frac{\mu\left(1+\frac{a^2\mu^2}{4}\right)^{\frac{1}{2}}}{\pi}\right)^{\frac{1}{4}} \exp\left(-\frac{1}{2}\mu\left(1+\frac{a^2\mu^2}{4}\right)^{\frac{1}{2}} x^2\right) \tag{171}$$

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 $m$:

$$S_M(\boldsymbol{x}) = \sum_{i=0}^{N-1} a\left[\frac{1}{2}m\left(\frac{x_{i+1}-x_i}{a}\right)^2 + \frac{1}{2}\mu^2 x_i^2\right] \tag{172}$$

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

$$S_M(\boldsymbol{x}) = \sum_{i=0}^{N-1} a\left[\frac{1}{2}\left(\frac{x_{i+1}-x_i}{a}\right)^2 + \frac{1}{2}\eta^2 x_i^2\right]. \tag{173}$$

by making the replacements

$$x \to x\sqrt{m} \tag{174}$$

and

$$\mu \to \frac{\mu}{\sqrt{m}}. \tag{175}$$

Making these replacements in the above results gives:

$$\langle x_i x_{i+j}\rangle = \frac{1}{2\omega m}\left(\frac{R^j + R^{N-j}}{1-R^N}\right) \tag{176}$$

$$\langle x^2\rangle = \frac{1}{2m\omega}\left(\frac{1+R^N}{1-R^N}\right) \tag{177}$$

$$\psi(x) = \left(\frac{\omega}{\pi}\right)^{\frac{1}{4}} \exp\left(-\frac{1}{2}\omega x^2\right) \tag{178}$$

but where now:

$$R = 1 + \frac{a^2\mu^2}{2m} - \frac{a\mu}{\sqrt{m}}\left(1+\frac{a^2\mu^2}{4m}\right)^{\frac{1}{2}} \tag{179}$$

$$\omega = \frac{\mu}{\sqrt{m}}\sqrt{1+\frac{a^2\mu^2}{4m}}. \tag{180}$$