THEORY, INTERPRETATION AND IMPLEMENTATION OF THE PATH INTEGRAL METHODOLOGY

University of Waterloo

AYAN CHATTOPADHYAY

Dept. of Chemistry

Indian Institute of Technology, Kharagpur

Under the supervision of:

Prof. MARCEL NOOIJEN

Dept. of Chemistry

University of Waterloo
# TABLE OF CONTENTS

1. Introduction ........................................ 3

2. History ................................................ 4

3. Theory .................................................. 5
   i. Path Integrals .................................... 5
      a. Dirac’s work .................................. 5
      b. Feynman’s path integral formulation ... 5
      c. Time slicing .................................. 6
      d. The approximate path integration formula ... 6
      e. Path Integral and Statistical Mechanics .. 7
         1. The path integral representation of the propagator ... 7
         2. The time evolution function and the quantum density operator ... 9
         3. The Trotter Factorization .................... 9
   ii. Path Integral Formulation for non-adiabatic systems ... 10
      a. The Mapping Hamiltonian .................. 10
      b. The Stock Thoss Hamiltonian ............ 10
      c. The path integral discretization ........ 10

4. Implementation of the PIMC .......................... 13
   A. Initialization of the theorem ............... 13
   B. Functions to be computed .................. 13
   C. Importance Sampling ....................... 15
   D. Calculating the nuclear distribution function .... 15
   E. Calculating other properties ............. 15

5. Results ................................................ 16
   A. Results for the Born Oppenheimer case ...... 16
   B. Results for the non-adiabatic case ......... 18
I. INTRODUCTION

The path integral formulation in quantum mechanics is a generalization of the action principle of classical mechanics in quantum terms, where an integral over infinite trajectories gives us the quantum amplitude rather than one single classical trajectory.

It is a crucial idea in quantum mechanics as it is symmetric between time and space, allowing physicists to switch the co-ordinates quite freely between very different canonical descriptions of the same quantum state.
II. HISTORY

The path integral methodology evolved over time through the works of the following people:

Wiener introduced the use of Weiner integrals for studying diffusion processes and Brownian motion. Later Paul Dirac extended the idea with the introduction of the use of Lagrangian in quantum mechanics. However, path integrals as a methodology was completely developed by Richard Feynman as a part of his doctoral work with Archibald Wheeler.
III. THEORY

i. Path Integrals

A. Dirac’s work:

Dirac’s work gives us the quantum time-evolution operator in the S representation as \( e^{iS} \). In the limit of large powers of this operator one can reconstruct the full quantum evolution between two given states. Dirac however did not provide any precise way of calculating the sum over paths or did not show how to solve the Schrodinger equation with this theory.

B. Feynman’s path integral formulation:

Feynman attempted to explain quantum mechanics with his path integral methodology with the help of these 3 postulates:
1. Probability of an event is the square of a complex quantity called the probability amplitude
2. This probability amplitude is given by adding the contributions of all possible histories in the configuration space.
3. The contribution of a history to the probability amplitude is given by \( e^{iS/h} \), where S is the action of the given history computed by the time integral of the Lagrangian along the corresponding trajectory.

The path integral assigns all of the possible histories equal magnitudes but different phases. The contributions from histories that are far away from the classical one are cancelled away by interference of similar histories with opposite phase. Thus in the classical limit the sum over an infinite possible histories give us the value corresponding to the classical action.
C. Time slicing

The time interval from time $t_a$ to $t_b$ can be divided into $n+1$ equal time segments, given by: $\Delta t = \frac{t_b - t_a}{n+1}$. The process is called time slicing. The time slicing is important when we express the path integral analogously to the Riemann integral.

D. The approximate path integral formula:

$$\int_{x_a}^{x_b} dx_1 \cdots \int_{x_a}^{x_b} dx_n \exp \left( \frac{i}{\hbar} \int_{t_a}^{t_b} L(x(t), v(t), t) \, dt \right)$$

Figure 1 Figure showing three of the possible paths contributing to the quantum amplitude of a particle moving from point A to B
E. Path integrals and Statistical Mechanics

The path integral representation of the propagator $e^{iHt/\hbar}$ is of profound importance in deriving semi classical approximations to quantum dynamical phenomena.

1. Path integral representation of the propagator

The time independent N-dimensional Hamiltonian operator is given by:

$$H(x, p) = \frac{p^2}{2m} + V(x) = \sum \frac{p_i^2}{2m} + V(x_1, x_2, \ldots, x_n) = T + V$$

The kernel or green function corresponding to the time evolution operator is:

$$K(x_f, x_0; t) = \langle x_f | e^{-iHt/\hbar} | x_0 \rangle$$

This can be shown to satisfy the time-dependent Schrödinger equation:

$$HK = i\hbar \frac{\partial K}{\partial t}$$

Therefore $K$ is the quantum mechanical amplitude for a particle to get to a position $x_f$ in time $t$ if initially it was located at $x_0$. The time evolution of a system initially prepared in state $|\psi_0\rangle$ can be expressed as a simple integral of the propagator:

$$\psi(x, t) = \langle x | \psi(t) \rangle = \int dx' \langle x | e^{-iHt/\hbar} | x' \rangle \langle x' | \psi_0 \rangle = \int dx' K(x, x', t) \psi_0(x')$$

where we make use of the completeness property of the position eigenstates.

The propagator for an arbitrary Hamiltonian is not analytically known, except for a few cases like the free particle, the linear potential, the harmonic oscillator and a few 1-d potentials. However the short time propagator for any system can be well approximated by simple analytical expressions. Path integration provides the link between the propagator for any arbitrary long time $t$ and the known short time propagator.
For representing the long time propagator in terms of the short time propagator we define an integer $N$ which is large enough such that $t / N \equiv \Delta t$ is a short time for our system of interest. We can refer $N$ as the number of time slices. Next we factor the long time propagator into products of the short time propagators as follows:

$$e^{-iHt/h} = e^{-iH\Delta t/h} e^{-iH\Delta t/h} \ldots e^{-iH\Delta t/h} = \prod_{k=1}^{N} e^{-iH\Delta t/h}$$

where $\Delta t = t_{n} - t_{n-1}$. Now using the completeness property for the energy eigenstates we arrive at the result:

$$\langle x_{f} | e^{-iHt/h} | x_{0} \rangle = \int dx_{1} \ldots \int dx_{N-1} \prod_{k=1}^{N} \langle x_{k} | e^{-iHt/h} | x_{k-1} \rangle$$

The expression derived till this point is an exact expression i.e. it does not have any approximations. The propagator is thus derived by integrating over all possible paths connecting the initial and final points in time $t$, i.e. by performing a path integral.

Figure 2  Figure showing a discretized path that connects points $x_{0}$ and $x_{t}$ in time $t$. 

\[8\]
2. **The time evolution function and the quantum density operator**

The path integral formalism of the time evolution operator generalizes in a straightforward way to the quantum density operator (or the Boltzmann operator) given by $e^{-\beta H}$. The analogy is obvious since the density operator can be thought of as the quantum propagator for the imaginary time $t = -i\beta \hbar$. This allows us to write the quantum density matrix elements in the path integral form.

3. **The Trotter factorization**

It can be easily shown that the kinetic and potential energy operators do not commute with each other. i.e. $[\hat{K}, \hat{V}] \neq 0$. Hence we cannot break the Hamiltonian i.e. $e^{-\beta H} \neq e^{-\beta K} e^{-\beta V}$. A very useful and important approximation in this case is the Trotter factorization which is given by:

$$e^{-\beta H} \approx e^{-\beta K} e^{-\beta V} + O(\beta^2)$$

The error in the approximation varies quadratically with $\beta$, which is related to the time interval $t$. Thus we see that the error in the calculation of the propagator can be greatly reduced if we introduce time slicing (i.e. more beads during the path integration) thus reducing the error as:

$$O(\beta^2) \rightarrow O(\beta_p^2)$$

where $P$ is the number of time slices and $\beta_p$ is given by $\beta / P$. The approximation converges to the exact value with increasing number of beads.
ii. Path Integral Formulation for non-adiabatic system

A. The Mapping Hamiltonian

We start with the general N-level Hamiltonian, which is given by the following equation:

\[ \hat{H} = h_0(\vec{R}, \vec{P}) + \sum_{n,m=1}^{N} V_{n,m}(\vec{R}) \langle \psi_n | \psi_m \rangle \]

where \((\vec{R}, \vec{P})\) denotes the nuclear positions and momenta, and \(h_0(\vec{R}, \vec{P}) = T(\vec{P}) + V_0(\vec{R})\) where \(T(\vec{P})\) is the nuclear kinetic energy operator and \(V_0(\vec{R})\) is a state-independent part of the potential energy. \(\{\psi_n\}\) is the basis for the electronic states and \(\{V_{n,m}(\vec{R})\}\) is the matrix elements set for the potential energy.

B. The Stock-Thoss Mapping

The Stock-Thoss mapping exactly maps the discrete quantum variables onto continuous degrees of freedom. The mapping relations used are the following:

\[ |\psi_n\rangle \langle \psi_m| \rightarrow a_n^+ a_m \]

\[ |\psi_n\rangle \rightarrow |0,0,\ldots,1_n,\ldots,0_N\rangle \]

Also using the Cartesian representation of the Hamiltonian operator we get the mapping Hamiltonian as:

\[ \hat{H} = h_0(\vec{R}, \vec{P}) + \frac{1}{2} \sum_{n,m=1}^{N} (x_n x_m + p_n p_m - \delta_{nm}) V_{nm}(\vec{R}) \]
C. PATH INTEGRAL DISCRETIZATION

The canonical partition function is obtained from the trace of the Boltzmann operator,

$$ Z = Tr[ e^{-\beta H} ] $$

where the states span the electronic and nuclear DoFs. The resolution of identity operator for this space can be expressed as:

$$ I = \int dR \sum_{n=1}^{N} |R,n\rangle \langle R,n| $$

where R indicates the nuclear positions and n indicates the SEO states. Repeated insertion of the completeness relation yields a path integral discretization of the partition function,

$$ Z = \int d\{R_\alpha\} \sum_{\{n_\alpha=1\}}^{N} \prod_{\alpha=1}^{P} |R_\alpha n_\alpha e^{-\beta_\rho H}| R_{\alpha+1} n_{\alpha+1} \rangle $$

where the notation $\int d\{R_\alpha\} \equiv \left( \prod_{\alpha=1}^{P} \int dR_\alpha \right)$. The standard Trotter approximation is then used to factorize the matrix elements in this equation yielding,

$$ Z = \lim_{P \to \infty} \int d\{R_\alpha\} \prod_{\alpha=1}^{P} \left( \frac{MP}{2\beta} \right)^{P/2} e^{-\beta_\rho V_0(R_\alpha)} \times \exp \left[ -\frac{MP}{2\beta} (R_\alpha - R_{\alpha+1})^T (R_\alpha - R_{\alpha+1}) \right] \times \sum_{\{n_\alpha=1\}}^{N} \prod_{\alpha=1}^{P} |n_\alpha e^{-\beta_\nu V(R_\alpha)}| n_{\alpha+1} \rangle $$

By using a projection operator the SEO basis is transformed into a Cartesian coordinate basis. The last term is then transformed to the form:

$$ \prod_{\alpha=1}^{P} \sum_{n,m=1}^{N} \langle x_\alpha | n \rangle \langle n | e^{-\beta_\nu V(R)} | n \rangle \langle n | x_\alpha \rangle $$
It is found that although a short time approximation to the electronic states could be done directly in the Cartesian coordinate basis, a more numerically stable result is obtained by using the SEO basis, such that

\[
\langle x | e^{-\beta_p v(R)} | x' \rangle = \sum_{n,m=1}^{N} \langle x | n \rangle M_{nm}(R) \langle m | x' \rangle
\]

where \( M_{nm}(R) = \langle n | e^{-\beta_p v(R)} | m \rangle \).

Finally the expression for the partition function is obtained as:

\[
Z = \lim_{p \to \infty} \left( \frac{2MP}{\beta \pi^{N+1}} \right)^{p/2} \int d\{R_{\alpha}\} \int d\{x_{\alpha}\} \prod_{\alpha=1}^{P} A_{\alpha} F_{\alpha} G_{\alpha}
\]

where

\[
A_{\alpha} = e^{-\frac{MP}{2\beta} (R_{\alpha} - R_{\alpha+1})^T (R_{\alpha} - R_{\alpha+1})} e^{-\beta_p V_0(R_{\alpha})}
\]

\[
F_{\alpha} = x_{\alpha}^T M(R_{\alpha}) x_{\alpha+1}^T
\]

\[
G_{\alpha} = e^{-x_{\alpha}^T x_{\alpha}}
\]

The partition function is used to calculate a number of equilibrium properties for nonadiabatic systems.
IV. IMPLEMENTATION OF PIMC

A. Initialization of the system

The nuclear co-ordinates and the electronic states are all dimensionless co-ordinates. The initial values of the nuclear and electronic co-ordinates are all random numbers between 0.5 and -0.5 as they can be either positive or negative.

B. Functions to be computed

The following functions are required for computing the partition function for the system:

\[ A_\alpha, F_\alpha, G_\alpha \] (Calculated for each bead)

1. Calculating \( A_\alpha \)

The A function is given by the following formula:

\[
A_\alpha = e^{-\frac{MP}{2\beta}(R_\alpha - R_{\alpha+1})^T(R_\alpha - R_{\alpha+1})} \exp(-\beta P V_0(R_\alpha))
\]

The A function has two parts:

a. The kinetic part (variable Ax1 and Ay1 in the code)
b. The potential part (variables Ay1 and Ay2 in the code). However this part has not been calculated since the potential function is included in the \( M(R_\alpha) \) matrix later on.

The kinetic functions for each bead are summed and the exponential is taken after all the kinetic terms have been summed up as computing the exponential is computationally more expensive than computing a sum.

2. Calculating \( F_\alpha \)

\[
F_\alpha = x_\alpha^T M(R_\alpha) x_{\alpha+1}^T
\]
First we are to calculate $M(R_\alpha)$ from the potential matrix. The potential matrix is given by:

$$V(q) = \begin{bmatrix} E + \lambda q + \frac{1}{2}wq^2 & \mu q \\ \mu q & E + \lambda q + \frac{1}{2}wq^2 \end{bmatrix}$$

This matrix is diagonalized and two eigenvalues are obtained. The exponents are taken and this gives us:

$$e^{-\beta_p \nu(R)} = \begin{bmatrix} e^{-\beta_p e_1} & 0 \\ 0 & e^{-\beta_p e_2} \end{bmatrix}$$

Where $e_1$ and $e_2$ are the eigenvalues. This gives us the diagonalized $M$ matrix, which is converted to the SEO basis by multiplication with the eigenvector matrix by the help of the operation:

$$M_{nm}(R_\alpha) = UM_d U^T$$

Where $U$ is the eigenvector matrix.

In the code $M$ matrix subroutine computes the matrix $M$. The a3eig call is used for diagonalization of the matrices and dgemm call is used for matrix multiplication. The eigenvectors are obtained in the diagonalization step automatically (as the 3$^{rd}$ argument of the a3eig call).

**3. Calculating $G_\alpha$**

This is pretty straightforward and involves just a matrix multiplication step.
C. Importance sampling

Importance sampling technique is used so that our measurements are constrained in the important region of the sampling, i.e. we get our properties calculated in the zone with the highest probability of occurrence.

D. Calculation of the nuclear probability distribution

The nuclear probability distribution is given by the following function:

\[
P(R) = \frac{\left\langle \delta(R - R_p) \text{sgn}(F) \right\rangle_w}{\left\langle \text{sgn}(F) \right\rangle_w}
\]

The nuclear distribution is in essence the distribution for the accepted nuclear coordinates.

E. Calculation of other properties

The other properties that have been calculated are the average potential energies and total energies. The formulae for calculating these are given by:

\[
\langle E \rangle = \frac{\left\langle \left( \frac{P}{2\beta} + \tilde{F} - \frac{\partial A}{\partial \beta} \right) \text{sgn}(F) \right\rangle_w}{\left\langle \text{sgn}(F) \right\rangle_w}
\]

where,

\[
\tilde{F} = \sum_{\alpha=1}^{P} x_\alpha^T - \frac{\partial M(R_\alpha)}{\partial \beta} x_\alpha
\]

The first and third terms in brackets in the expression for average total energy are the kinetic terms and the second term is the potential term. The sign of F is the product over all the signs for each bead. The sign of F is essential because the function F is the only function that can have a negative value out of three functions in the integrand, and we require our weight function to be positive.
V. Results

The results of my calculation are as follows:

A. The Born Oppenheimer calculations

The model parameters used in the Born Oppenheimer calculations are:

\[ E = 0.1 \text{eV} \]

\[ \lambda_1 = 0.2 \]

\[ \lambda_2 = -0.1 \]

\[ w = 0.1 \]

\[ \mu = 0.1 \]

1) The results for the coupling parameter '\textbf{\mu}=0.1' are:

<table>
<thead>
<tr>
<th>Temp (in K)</th>
<th>V (in eV)</th>
<th>U (in eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>-0.139</td>
<td>-0.113</td>
</tr>
<tr>
<td>600</td>
<td>-0.126</td>
<td>-0.093</td>
</tr>
<tr>
<td>1000</td>
<td>-0.101</td>
<td>-0.054</td>
</tr>
<tr>
<td>1500</td>
<td>-0.073</td>
<td>-0.006</td>
</tr>
<tr>
<td>2000</td>
<td>-0.050</td>
<td>0.037</td>
</tr>
</tbody>
</table>

The adiabatic potential curve is given by:
2) The results for the coupling parameter \( \mu = 0.2 \) are:

<table>
<thead>
<tr>
<th>Temp (in K)</th>
<th>V (in eV)</th>
<th>U (in eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>-0.324</td>
<td>-0.298</td>
</tr>
<tr>
<td>600</td>
<td>-0.314</td>
<td>-0.281</td>
</tr>
<tr>
<td>1000</td>
<td>-0.289</td>
<td>-0.242</td>
</tr>
<tr>
<td>1500</td>
<td>-0.254</td>
<td>-0.191</td>
</tr>
<tr>
<td>2000</td>
<td>-0.228</td>
<td>0.140</td>
</tr>
</tbody>
</table>

The adiabatic potential curve is given by:

3) The results for the coupling parameter \( \mu = 0.3 \) are:

<table>
<thead>
<tr>
<th>Temp (in K)</th>
<th>V (in eV)</th>
<th>U (in eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>-0.617</td>
<td>-0.590</td>
</tr>
<tr>
<td>600</td>
<td>-0.610</td>
<td>-0.575</td>
</tr>
<tr>
<td>1000</td>
<td>-0.588</td>
<td>-0.539</td>
</tr>
<tr>
<td>1500</td>
<td>-0.551</td>
<td>-0.481</td>
</tr>
<tr>
<td>2000</td>
<td>-0.512</td>
<td>0.422</td>
</tr>
</tbody>
</table>
The adiabatic potential curve is given by:

These results match perfectly with the DVR calculations for the same model.

**B. The non-adiabatic calculations**

The **nuclear probability distribution**

The distribution obtained in the PIMC simulation is the following:
The DVR approach gives the distribution as:

\[
\rho^{DVR}(q_i) = \rho(q_i)
\]

The two graphs cannot be exactly compared because we have the question whether \(\rho^{DVR}(q_i) = \rho(q_i)\)

However we note that we have the peaks roughly in the same point in both curves. Thus the nuclear distribution is assumed to be quite all right although there are still a few errors. Finally the calculation of the properties yield the following results:

<table>
<thead>
<tr>
<th>Temp (in K)</th>
<th>V (in eV)</th>
<th>U (in eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>600</td>
<td>-0.71</td>
<td>-0.092</td>
</tr>
<tr>
<td>1000</td>
<td>-0.68</td>
<td>-0.049</td>
</tr>
<tr>
<td>1500</td>
<td>-0.64</td>
<td>0.006</td>
</tr>
<tr>
<td>2000</td>
<td>-0.61</td>
<td>0.059</td>
</tr>
</tbody>
</table>

The results do not match with the DVR calculations possibly because of some bug in the codes.
REFERENCES