2010

EXTENSIONS FOR THE RIGID-BODY QUANTUM MONTE CARLO METHOD

Chris P. Smith
University of Rhode Island, chris.smith.p@gmail.com

EXTENSIONS FOR THE RIGID-BODY QUANTUM MONTE CARLO METHOD

BY

CHRIS P. SMITH

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

PHYSICS

UNIVERSITY OF RHODE ISLAND

2010

DOCTOR OF PHILOSOPHY DISSERTATION

OF

CHRIS P. SMITH

APPROVED:

Dissertation Committee:

Major Professor Dr. M.P. Nightingale

Dr. Dave Freeman

Dr. Leonard Kahn

Dr. Nasser Zawia

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2010

ABSTRACT

Solving the Schrödinger equation and finding excited states for quantum mechanical many-body systems is a fundamental problem. If this problem is formulated in terms of integrals, the dimensionality of the configuration space is often too high to perform the integrations required directly. This is because the volume of configuration space increases exponentially with dimension quickly, making its complete exploration impossible. Monte Carlo methods provide a way to estimate these integrals by statistically sampling a subset of configuration space, and these methods provide $\frac{1}{\sqrt{N}}$ convergence regardless of dimension.

For many problems involving the quantum mechanics of molecules there exists a large time scale separation between the high-frequency internal vibrations, and low-frequency intermolecular motions. This separation motivates the rigid-body approximation which freezes internal degrees of freedom in order to study intermolecular effects. The reduction in the dimensionality of configuration space further increases the range of accessible problems. This thesis is devoted to the construction and implementation of algorithms which incorporate the rigid-body approximation into existing Monte Carlo methods for solving the quantum mechanical many-body problem.

Monte Carlo estimators are constructed as averages over samples drawn from some probability distribution. If this distribution is known in closed form the samples can be generated by application of the Metropolis algorithm. In most cases the distribution is not known in closed form or even representable with a finite number of variables. If this distribution is the dominant eigenstate of some known operator then a stochastic implementation of the power method can be used to generate the required samples. For the quantum mechanical problem this operator can be taken to be the imaginary-time evolution operator and its application can

be represented in terms of random walks. For rigid bodies, this method involves the implementation of rotational Brownian motion. The use of quaternion algebra to implement this rotational motion enhances simplicity, performance, and numerical stability.

The method described above can be generalized to investigate excited state properties using correlation function Monte Carlo which is a Monte Carlo implementation of the Rayleigh-Ritz variational method. This generalization requires the construction of a trial subspace which is then subjected to the variational method resulting in a generalized eigenvalue problem.

It is important to represent optimized trial states accurately and efficiently. For this purpose, we write trial states as functions of invariant polynomials of the interparticle distances. By reducing the number of free interparticle distances, the rigid-body approximation greatly simplifies the construction of a basis used to represent the necessary trial states. The same applies to the so-called guiding function which is used to sample configurations that simultaneously represent all trial states.

A computer program is written to test these algorithms on a number of problems. Results for simple test problems are in agreement with exact solutions.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

CHAPTER 1

Introduction

In the earliest days of quantum mechanics Schrödinger pointed out the remarkable similarity between the Schrödinger equation in imaginary time, $\tau = it$,

$$\frac{\partial \psi}{\partial \tau} = \frac{\hbar}{2m}\nabla^2\psi - V\psi \tag{1}$$

and the diffusion equation [1].

$$\frac{\partial C}{\partial t} = D\nabla^2 C \tag{2}$$

This similarity suggests the possibility of finding solutions to quantum mechanical problems through the use of random walks.

Many years later at Los Alamos a group of physicists working on the Manhattan Project developed a statistical approach for the study of differential equations called the Monte Carlo method [2]. The birth of the Monte Carlo method can also be viewed as the rebirth of method called statistical sampling which was about to become much more useful with the arrival of the first computers. The first description of the diffusion Monte Carlo method appears as part of this work and is attributed to a suggestion by Fermi who stated:

> ... the time-independent Schrödinger equation
>
> $$\nabla^2\psi = (\mathcal{E} - V)\,\psi \tag{3}$$
>
> could be studied as follows. Re-introduce time dependence by considering
>
> $$u = \psi e^{-\mathscr{H}t} \tag{4}$$
>
> u will obey the equation
>
> $$\partial_t u = \nabla^2 u - Vu \tag{5}$$
>
> The last equation can be interpreted however as describing the behavior of a system of particles each of which performs a random walk, i.e., diffuses isotropically and at the same time is subject to multiplication, which is determined by the point value of V.

In this way, a population of walkers diffusing through configuration space can be used to generate samples from the unknown probability distribution corresponding to the ground state wavefunction. Specific methods for accomplishing this will be discussed in Chapter 2. These methods will be related to the quantum mechanical problem in Chapter 3 and will lead to a specific algorithm for estimating ground sate properties of rigid-bodies in Section 3.5.

The Metropolis method for generating a sequence of random samples from a known probability distribution was introduced in 1953 [3], and the method would be generalized by Hastings [4] in 1970. The Metropolis-Hastings method can be used to construct a Markov process with an arbitrary probability density as it's stationary state. In Section 2.2, I will show how to use this simple algorithm to construct estimators for expectation values as averages over random samples drawn form a known probability distribution.

The Rayleigh-Ritz variational method is an extension of the variational principle. The method is based on a linear expansion of the solution of the eigenvalue problem at hand where the expansion coefficients are determined by a variational procedure. Wave functions are written as linear combinations of basis states. Stationary points are obtained from equating derivatives to zero and can be found by solution of a generalized eigenproblem. It can be shown that the eigenvalues are upper bounds on the exact energy levels. In most cases the high dimensionality of the integrals required to evaluate the matrix elements leaves us no choice but to estimate their values using Monte Carlo methods. A review of this method is given in Section 3.6.

The variational calculation described above produces approximate eigenstates as linear combinations of the basis states with coefficients that are determined through solution of the generalized eigenproblem. These approximate eigenstates

are limited by the completeness of the basis used to represent them. A basis set that spans the same invariant subspace as the exact eigenstates will provide exact results with zero variance. In basis set that does not span this subspace will result in a variational bias which can be removed by application of diffusion Monte Carlo in the evaluation of matrix elements. This procedure, known as correlation function Monte Carlo [5], is discussed in Section 3.6.2.

One of the central problems in quantum Monte Carlo is the construction of a basis used for representing these approximate eigenstates. Evaluation of these states and their derivatives can be the most computationally expensive part of a quantum Monte Carlo program. In practice, the best way to improve upon results is to use a more complete basis in the sense that it has maximum overlap with the subspace spanned by the exact eigenstates. Therefore, what is needed is as complete a basis as possible while at the same time containing a minimal number of terms. Borrowing from the field of commutative algebra, such a basis can be constructed from the set of invariant polynomials associated with some symmetry group [6]. Depending on the symmetry group, this can result in a large reduction of the number of terms to be evaluated. A simple example of this is given in Section 4.3. The construction of these basis states is described in Section 4.1 and a simple example of how to construct the set of invariant polynomials is given in Section 4.3. In Section 4.3.3 it is demonstrated that the rigid-body approximation results in a further reduction of the number of terms required to construct these basis states.

The Schrödinger equation in imaginary time can be transformed into an equation for diffusion, branching, and drift with the introduction of a guiding function as shown in Section 3.5. This is known as the importance sampling transformation and it allows for the random sampling of configuration space to be guided into

areas that contribute more to the integral. In this case, Monte Carlo estimates take the form of a weighted average. Section 4.2 deals with the construction and optimization of this guiding function which can be carried out by maximizing the efficiency of this weighted average.

Trial states are written as functions of the basis of invariant polynomials with many linear and non-linear parameters. The linear parameters are found by solution of the generalized eigenproblem as described above. Section 4.1.1 describes the optimization of the non-linear parameters by minimization of the variance of the local energy, $E_L = \frac{\mathscr{H}\psi}{\psi}$ which is clearly a constant if ψ is an eigenstate. For each set of non-linear parameters the linear parameters must be recalculated. Full optimization of the trial states requires a linear optimization nested in a non-linear one.

Rigid-body diffusion Monte Carlo [7] is a variant of the diffusion Monte Carlo method that reduces the dimensionality of configuration space by decoupling high-frequency intramolecular vibrations from the low-frequency intermolecular motions, and treating molecules as quantum mechanical tops. This approximation greatly increases the scope of the problems that are accessible to Monte Carlo methods. Much of this thesis is devoted to the implementation of the methods described above within the rigid-body approximation.

In the case of quantum mechanical tops, the diffusion required for sampling takes the form of rotational Brownian motion. For performance, simplicity, and numerical stability, these rotations are implemented with quaternion algebra. Quaternions are an extension of complex numbers that can be used to implement rotation in three dimensions and are defined in Section 5.1. The representation of rotations with quaternion algebra is described in Section 5.2. Orthogonal matrices represent pure rotations, that is, they do not change the shape of the object they rotate.

Numerically, products of rotation matrices implemented with finite precision will eventually loose orthogonality, resulting in distortion of body being rotated. Unit quaternions can be used to represent pure rotations. Quaternions need only be normalized to avoid this problem which is less expensive than orthogonalization. The resulting numerical stability allows for the location to be calculated from a cumulative rotation product applied to the initial state, and removes the need to locate the principal axes by diagonalizing the moment of inertia tensor every step. The implementation of rigid-body diffusion Monte Carlo using quaternion algebra is described in Section 6.4.

In Chapter 7 specific examples for toy problems are presented. These problems are chosen to be solvable by other means in order to demonstrate the correctness of the methods.

This work is devoted to the development of the rigid-body diffusion Monte Carlo method. This development includes the adaptation of some existing quantum Monte Carlo methods to the rigid-body approximation, as well as the addition of some new algorithms. The primary contributions of this work to existing methods include an algorithm for the implementation correlation function Monte Carlo within the rigid-body approximation that incorporates trial states constructed from a basis of primary invariants and a new algorithm for implementing rotational Brownian motion that takes full advantage of quaternion algebra.

List of References

[1] E. Schrodinger, "Uber die umkehrung der naturgesetze," *Sitzber. Preuss. Akad. Wiss. Phys.-math. Kl*, pp. 144–153, 1931.

[2] N. Metropolis and S. Ulam, "The Monte Carlo method," *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, Sept. 1949.

[3] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.*, vol. 21, p. 1087, 1953.

[4] W. Hastings, "Monte carlo samping methods using markov chains and their applications," *Biometrika*, pp. 97–109, 1970.

[5] D. M. Ceperley and B. Bernu, "The calculation of excited state properties with quantum Monte Carlo," *J. Chem. Phys.*, vol. 89, p. 6316, 1988.

[6] A. Mushinski and M. P. Nightingale, "Many-body trial wave functions for atomic systems and ground states of small noble gas clusters," *J. Chem. Phys.*, vol. 101, p. 8831, 1994.

[7] V. Buch, "Treatment of rigid bodies by diffusion Monte Carlo: Application to the para-$H_2 \ldots H_2O$ and ortho-$H_2 \ldots$-H_2O clusters," *J. Chem. Phys.*, vol. 97, p. 726, 1992.

CHAPTER 2

Monte Carlo background

The purpose of this chapter is to review some of the relevant Monte Carlo methods that will be related to the quantum mechanics of rigid-bodies in chapter 3. Section 2.1 will show how to construct an estimator for some observable, $\mathcal{O}$, with a probability distribution, $\rho(x)$, when a set of configurations, $\{x_i\}$, sampled from that distribution is available and that observable is diagonal or near-diagonal in the chosen representation. Next, section 2.2, will describe the use of the Metropolis-Hastings method to generate sample configurations, $\{x_i\}$, from a known probability distribution, $\rho(x)$. Then, section 2.3, will show how to sample from an unknown distribution, ψ_0, which is an eigenstate of some operator, G, using projector methods. Finally, section 2.4 introduces a guiding function, ψ_g, which will allow us to sample from the mixed distribution, $\psi_g\psi_0$. In addition, the guiding function allows one to sample from two or more overlapping distributions simultaneously and even from distributions containing negative probability amplitudes. This transformation will enable us to sample selectively from more important areas of configuration space.

2.1 Estimators

Suppose a set of N configurations , $\{x_i\}$, sampled from a probability distribution, $\rho(x)$, is available. An estimate of an expectation value in that distribution can be constructed as an average over the configurations $\{x_i\}$.

$$\mathcal{O}_\rho = \int \mathcal{O}(x)\rho(x)dx = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \mathcal{O}(x_i) \tag{6}$$

2.1.1 Example

As a simple example consider the Monte Carlo integration for the area of a circle. Uniformly distributed points are selected at random inside a square of

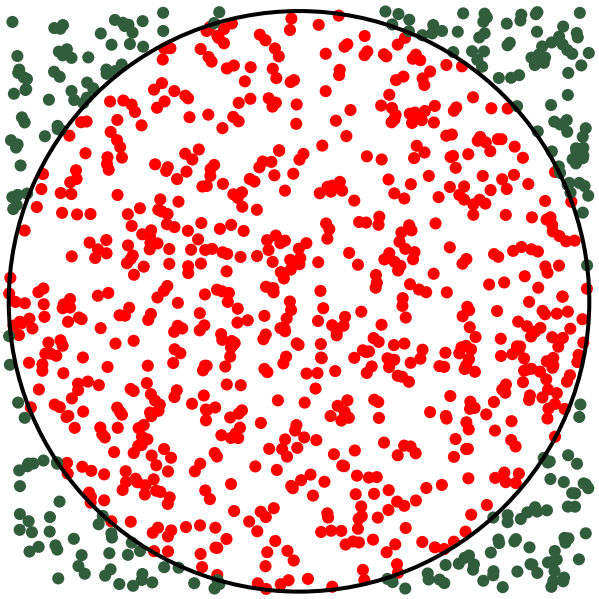


Figure 1: Monte Carlo integration for estimating π.

known area.

$$\rho(\boldsymbol{x}) = \frac{1}{\text{Area}} \tag{7}$$

The ratio of total number of points inside to points outside provides an estimate for the ratio of the area of the circle to the total sampling area.

$$\mathcal{O}(\boldsymbol{x}) = \begin{cases} 0, \text{outside} \\ \\ 1, \text{inside} \end{cases} \tag{8}$$

$$\pi \approx \frac{1}{N} \sum_{i=1}^{N} \mathcal{O}(\boldsymbol{x}_i) \tag{9}$$

For quantum mechanical calculations we are interested in estimating expectation values of the form

$$\mathcal{O}_{0,0} \equiv \frac{\langle \psi_0 | \hat{\mathcal{O}} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} = \frac{\int \psi_0^*(x) \hat{\mathcal{O}} \psi_0(x) dx}{\int |\psi_0(x)|^2 dx} \tag{10}$$

In general, the eigenstate, ψ_0, is not known exactly, and the dimensionality of the configuration space may make numerically exact computation impossible.

2.2 Metropolis-Hastings Method

The Metropolis-Hastings method [1] can be used to generate a set of configurations $\{x_i\}$, sampled from a known probability distribution, $\rho(x)$. Given some current configuration x, configuration y is proposed, with probability $T(y|x)$. The proposed configuration, is then accepted as $x' = y$, with probability $A(y|x)$. If a configuration is not accepted the previous configuration will be used, $x' = x$. Samples are generated by a transition matrix with elements

$$M(x'|x) = A(x'|x)T(x'|x) + \left\{1 - \int dy A(y|x)T(y|x)\right\}\delta(x' - x) \tag{11}$$

The acceptance matrix is chosen such that detailed balance is satisfied.

$$A(x'|x) = \min\left\{1, \frac{T(x|x')\rho(x')}{T(x'|x)\rho(x)}\right\} \tag{12}$$

$$M(x'|x)\rho(x) = M(x|x')\rho(x') \tag{13}$$

Repeated application of M will result in the unique stationary distribution $\rho(x)$. It is necessary that the transition matrix, M, is such that every point in configuration space can be reached from every other point in a finite number of steps. For discrete spaces what this means is clear, for continuous ones the situation is more complicated, but the technical details are beyond the scope of this thesis.

2.3 Projection

It is possible to generate samples from an unknown eigenstate , ψ_0, of some operator, G, by projection. The projected eigenstate, ψ_0, will be the state with the eigenvalue largest in magnitude the eigenstate of which has non-vanishing overlap with the initial state. To understand this, expand some initial state, ψ_T,

in eigenstates of that operator.

$$G|\psi_i\rangle = \lambda_i|\psi_i\rangle, \quad \lambda_0 > \lambda_1 > \cdots > \lambda_N \tag{14}$$

$$|\psi(\tau = 0)\rangle = |\psi_T\rangle = \sum_n a_n|\psi_n\rangle \tag{15}$$

For simplicity, we restrict the eigenvalues to be strictly positive. Repeated application of the operator, G, projects out the dominant eigenstate, provided there is some overlap with the initial state, $a_0 > 0$.

$$|\psi(\tau)\rangle = G^\tau(\sum_n a_n|\psi_n\rangle) = \sum_n a_n\lambda_n^\tau|\psi_n\rangle \approx a_0\lambda_0^\tau|\psi_0\rangle\left\{1 + \mathscr{O}\left(|\frac{\lambda_1}{\lambda_0}|^\tau\right)\right\} \tag{16}$$

If all G elements are non-negative this projection can be implemented in the following way. $\psi(x)$ can be represented by a population of walkers in configuration space. The probability of the walker having configuration x after τ steps is $p_\tau(x)$. If the dominant eigenvalue is one, a walk through configuration space, generated by application of $G(x'|x)$, produces an stationary probability distribution of walker configurations that approaches the dominant eigenstate in the limit of infinitely many steps.

$$\lambda_0 = 1, \qquad \lim_{\tau\to\infty} G^\tau(x|x')p_0(x) = \psi_0(x) \tag{17}$$

If the dominant eigenvalue is not one then, our walk does not conserve probability, and the distribution generated by repeated application of G is not stationary. In this case, it is possible to define a probability conserving walk whose stationary distribution can be reweighted to give $\psi_0(x)$. This process can be constructed by factoring G into a stochastic part, and non-stochastic part associated with a weight. The walkers in this more general method are assigned a weight, $\mathcal{W}$, as well as a configuration.

$$G(x'|x) = P(x'|x)w(x), \qquad w(x) = \sum_{x'} G(x'|x) > 0 \tag{18}$$

Walks are generated by application of the stochastic part of $G(x'|x)$.

$$x_{\tau+1} = x', \text{with probability} \quad P(x'|x_\tau) \tag{19}$$

While the weight of a walker is produced by the non-stochastic part.

$$\mathcal{W}_{\tau+1} = w(x_\tau)\mathcal{W}_\tau, \qquad \mathcal{W}_\tau \equiv \mathcal{W}_\tau(x_{\tau-1}, x_{\tau-2}, \cdots, x_0) = \prod_{i=0}^{\tau-1} w(x_i) \tag{20}$$

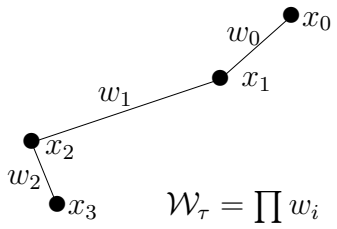


Figure 2: Weighted path

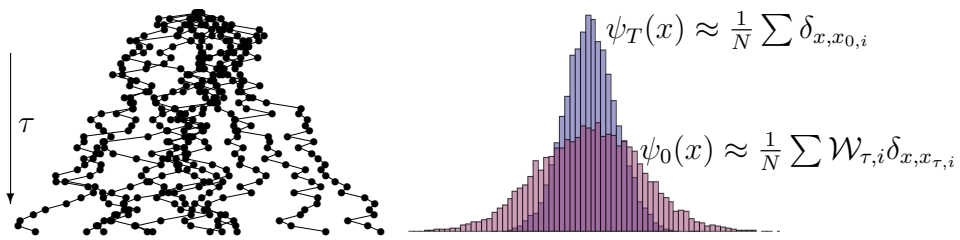
The probability distribution of walker configurations has the following relationship to the dominant eigenstate.

$$\lim_{\tau\to\infty} \mathcal{W}_\tau p_\tau(x) \propto \psi_0(x) \tag{21}$$

An estimate for the stationary distribution can be constructed as an average over samples.

$$\psi_0(x) \propto \lim_{\tau\to\infty} \sum_{x'} G^\tau(x|x')p_0(x') = \lim_{\tau\to\infty} \left\{ \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N} \mathcal{W}_\tau \delta_{x,x_\tau} \right\} \tag{22}$$

In general, the weight, $w(x)$, in Equation 18 cannot be computed directly. In that case, $w(x)$, is computed on the fly so that $G(x|x')/w(x)$ has eigenvalue unity. The fact that this value is known only after the fact introduces what is known as population control bias [2]. The presence of the fluctuating weights in these expressions increases statistical noise. There are a number of ways to mitigate this problem. One method is to advance a large collection of walkers simultaneously



(a) Walkers (b) Histogram of Walker Locations

Figure 3: A population of random walkers can be used to sample from an unknown probability distribution

that are capable of multiplying or vanishing depending on the weights, another is the importance sampling transformation of Section 2.4

In practice, the Metropolis method can be combined with projection. If the initial configuration for the walk is chosen by Metropolis method from $\psi_T(x)$, that is, initial configurations proposed with probability $P(x'|x)$ are accepted with probability

$$A(x'|x) = \min\left\{1, \frac{G(x|x')\psi_g(x')}{G(x'|x)\psi_g(x)}\right\} \tag{23}$$

The distribution sampled for zero projection time will be $\psi_g(x)$. Because it samples ψ_0 instead of ψ_0^2, this method is not yet able to estimate expectation values of the form in Eq. (10). In the next section I will introduce a similarity transformation to address this problem.

2.4 Importance Sampling Transformation

Expectation values of the form in Eq. (10) require us to sample from ψ_0^2 not ψ_0. A step in this direction is to sample from the "mixed distribution", $\psi_0\psi_g$, and construct the "mixed estimator", $\mathcal{O}_{0,g}$. Where ψ_g is the "guiding function", which needs to be positive, and have sufficient overlap with the desired eigenstate or subspace. In addition to enabling construction of the mixed estimator, the guiding function can be used to guide the walk to the more statistically important parts

of configuration space. This mixed estimator can be constructed by a change of basis using the similarity transform

$$\tilde{G} = \psi_g G \psi_g^{-1} = \tilde{P}\tilde{w} \tag{24}$$

The matrix giving this similarity transform is diagonal in the position representation, which is why it looks like a wavefunction instead of a matrix. If the eigenstate of G is ψ_0 then, the transformed operator $\tilde{G}$ has the eigenstate $\psi_g\psi_0$.

$$G\psi_0 = \lambda\psi_0, \qquad \tilde{G}\psi_g\psi_0 = \lambda\psi_g\psi_0 \tag{25}$$

In this case we are calculating expectation values of the form.

$$\begin{aligned}\mathcal{O}_{0,g} &\equiv \lim_{\tau\to\infty} \frac{\langle\psi_g|G^\tau\hat{\mathcal{O}}|\psi_g\rangle}{\langle\psi_g|G^\tau|\psi_g\rangle} = \frac{\langle\psi_0|\hat{\mathcal{O}}|\psi_g\rangle}{\langle\psi_0|\psi_g\rangle} \\ &= \frac{\int \psi_0(x)\mathcal{O}_g(x)\psi_g(x)dx}{\int \psi_0(x)\psi_g(x)dx} = \frac{\int \psi_0(x)\psi_g(x)\mathcal{O}_L(x)dx}{\int \psi_0(x)\psi_g(x)dx}\end{aligned} \tag{26}$$

Where the configurational eigenvalue, $\mathcal{O}_g(x)$, is defined by

$$\langle x|\hat{\mathcal{O}}|\psi_g\rangle \equiv \mathcal{O}_g(x)\psi_g(x) \tag{27}$$

and the local operator , $\mathcal{O}_L(x)$, by

$$\mathcal{O}_L(x) \equiv \psi_g^{-1}(x)\mathcal{O}_g(x)\psi_g(x) \tag{28}$$

The configurational eigenvalue can be calculated it the operator ids diagonal or near diagonal in the chosen representation. An estimator can be constructed as a weighted average of the local operator over N projected configurations.

$$\mathcal{O}_{0,g} = \lim_{\tau,N\to\infty} \left\{ \frac{\sum_{i=1}^N \tilde{\mathcal{W}}_{\tau,i}\mathcal{O}_L(x_{\tau,i})}{\frac{1}{N}\sum_{i=1}^N \tilde{\mathcal{W}}_{\tau,i}} \right\} \tag{29}$$

where $\tilde{W}_\tau = \prod \tilde{w}(x)$ as in Eq. (20). For Hermitian operators that commute with G, $\mathcal{O}_{0,g} = \mathcal{O}_{0,0}$ and the mixed estimator gives the unbiased expectation value of Equation 10 [2]. It is important to note that if the guiding function is an exact

eigenstate then the estimator will yield the exact result with zero variance. For guiding functions that are nearly eigenstates, this transformation results in much more stable weights and reduces the associated statistical noise. In section 3.5, this importance sampling transformation will be applied to the quantum mechanical many-body problem with the substitution: $G = \exp\left(-\tau\mathscr{H}\right)$.

List of References

[1] W. Hastings, "Monte carlo samping methods using markov chains and their applications," *Biometrika*, pp. 97–109, 1970.

[2] M. P. Nightingale and C. J. Umrigar, *Monte Carlo Eigenvalue Methods in Quantum Mechanics and Statistical Mechanics*, ser. Advances in Chemical Physics. New York: John Wiley & Sons, 1999, vol. 105, p. 65.

CHAPTER 3

Relation to Quantum Mechanics

In this section, I will relate the previous discussion to the quantum mechanics of molecules. The goal of this section is to construct estimators for ground state observables. In order to simplify the calculation, molecules will be approximated by rigid quantum mechanical tops with their internal degrees of freedom frozen out. Section 3.1 will review the variational method and show how to construct a Monte Carlo estimator in the case where the approximate eigenstate is known. Section 3.2 will relate the projector method from the previous section to the quantum mechanical problem. The operator, G, from the Chapter 2 is replaced by the quantum mechanical imaginary-time evolution operator, $G = \exp(-\tau\mathscr{H})$. The dominant eigenstate of this operator, corresponds to the ground state. Using the short-time approximation for G, described in Section 3.4, I will show that the application of the stochastic part of this operator on our walkers produces diffusion. In the context of the quantum mechanical rigid-body problem this diffusion takes the form of rotational Brownian motion. The part that does not conserve probability, is related to the potential and can be incorporated by assigning the walkers a weight. Next, Section 3.5 introduces the guiding function, and the change of basis that is the importance sampling transformation. This transformation serves to guide the sampling into areas of configuration space that contribute the most to the integral. Application of the importance sampled imaginary-time evolution operator causes our walkers not only to diffuse, but also to drift. An estimator for ground state observables is derived, and a specific algorithm given for the implementation of these methods on a system of quantum mechanical tops. Section 3.6 generalizes all of these ideas to apply to excited state properties and can be

viewed as a diffusion Monte Carlo implementation of the Rayleigh-Ritz variational method.

3.1 Variational Monte Carlo

The variational principle states that the expectation value of the Hamiltonian in some trial state, ψ_T, is always greater than or equal to the ground state energy, $\mathcal{E}_0$.

$$\mathcal{E}_0 \leq \frac{\langle \psi_T | \mathscr{H} | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} \tag{30}$$

where the reduced Hamiltinain in dimensionless units is

$$\mathscr{H} = -\frac{1}{2}\nabla^2 + V \tag{31}$$

If the problem is formulated in terms of integrals the equation to be minimized is

$$\mathcal{E}_0 \leq \frac{\int d\boldsymbol{x} \psi_T^*(\boldsymbol{x}) \mathscr{H} \psi_T(\boldsymbol{x})}{\int d\boldsymbol{x} \psi_T^*(\boldsymbol{x}) \psi_T(\boldsymbol{x})} \tag{32}$$

Variational Monte Carlo Estimator

In general, the dimensionality of configuration space often makes exact integration impossible. In these cases Monte Carlo estimators for these integrals can be constructed as averages over samples. Suppose we have an approximate ground state wavefunction, $\psi_T(x) \approx \psi_0(x)$, and a set of configurations sampled from $\rho(x) = \frac{|\psi_T(x)|^2}{\int |\psi_T(x)|^2 dx}$. As in section 2.4, an estimator for some observable in the trial state can be constructed in the following way. Define the configurational eigenvalue as

$$\mathcal{O}_T(x)\psi_T(x) \equiv \langle x | \hat{\mathcal{O}} | \psi_T \rangle \tag{33}$$

An estimator can be constructed by rewriting our observable in the form.

$$\begin{aligned} \mathcal{O}_{T,T} \equiv \frac{\langle \psi_T | \hat{\mathcal{O}} | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} &= \frac{\int \psi_T^*(x) \mathcal{O}_T(x) \psi_T(x) dx}{\int |\psi_T(x)|^2 dx} \\ &= \frac{\int |\psi_T(x)|^2 \left\{ \psi_T^{-1}(x) \mathcal{O}_T(x) \psi_T(x) \right\} dx}{\int |\psi_T(x)|^2 dx} \end{aligned} \tag{34}$$

Define the local operator $\mathcal{O}_L$, as

$$\mathcal{O}_L(x) = \psi_T(x)^{-1}\mathcal{O}_T(x)\psi_T(x) \tag{35}$$

The average over of the local operator over configurations is an estimator for the expectation value.

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathcal{O}_L(x_i) = \frac{\int|\psi_T(x)|^2\mathcal{O}_L(x)dx}{\int|\psi_T(x)|^2dx} = \mathcal{O}_{T,T} \approx \mathcal{O}_{0,0} \tag{36}$$

If the trial state is an eigenstate of $\hat{\mathcal{O}}$, then the local operator is a constant, and the estimate is exact. In most cases it is not possible to find or even represent the exact eigenstate. This limitation results in a variational bias. In the next section I will show how to use the projector method of Section 2.3 to systematically remove this variational bias by using the imaginary-time evolution operator to sample from an unknown eigenstate.

3.2 The Schrödinger Equation and Diffusion

Except for the term involving the potential energy, the Schrödinger equation in imaginary-time is the diffusion equation. This implies the possibility of finding solutions to the Schrödinger equation through the use of random walks.

Using the projector method described in Section 2.3, with the imaginary-time evolution operator acting as the projection operator in Eq. (16) it is possible to sample from the lowest non-orthogonal state. For short times, this imaginary-time evolution operator can be approximately split into a probability conserving part corresponding to diffusion, and a weight.

$$G = \exp\{-\tau\mathscr{H}\} \approx \underbrace{\exp\left\{\Delta\tau\nabla^2\right\}}_{\text{free particle diffusion}}\underbrace{\exp\left\{-\Delta\tau\hat{V}\right\}}_{\text{weight}} = Pw \tag{37}$$

Since the operators in the Hamiltonian do not commute, this approximation is only good for small $\Delta\tau$. On the other hand, $\Delta\tau$ has to be large enough to allow

the exploration of configuration space and the total projection time over a path must be adequate to remove the variational bias discussed in 3.1.

3.3 Rigid Bodies

Molecules can be represented by asymmetric tops interacting via some potential function. The Hamiltonian for N of these rigid bodies in dimensionless units is

$$\mathscr{H} = \sum_m \left\{ -\frac{1}{2M_m}\nabla_m^2 - \frac{1}{2I_{m,p}} \sum_p \partial_{\phi_{m,p}}^2 \right\} + \hat{V} \tag{38}$$

Rigid bodies are indexed by m, and their principal axes by p. $I_{m,p}$ is the moment of inertia of the m'th body, about its p'th principal axis. The angles of rotation about the principal axes are $\phi_{m,p}$. The Laplacian is with respect to the coordinate of the center of mass of body m.

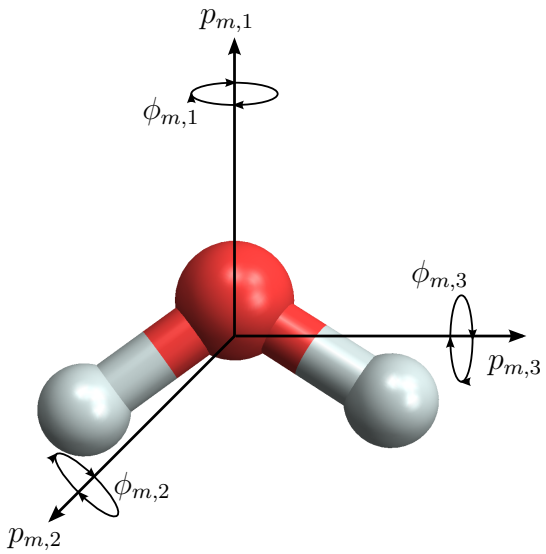


Figure 4: Water molecule as an example quantum mechanical top.

3.4 Short-Time Approximation

For short times, the propagator can be split into translational and rotational parts.

$$G_{\text{total}}(\vec{R}, \vec{R}', \vec{\Phi}, \vec{\Phi}'; \Delta\tau) \approx G_{\text{rot}}(\vec{\Phi}, \vec{\Phi}'; \Delta\tau) G_{\text{trans}}(\vec{R}, \vec{R}'; \Delta\tau) \tag{39}$$

where $\vec{R}$ is a vector that defines the Cartesian coordinates of the centers of mass, and $\vec{\Phi}$ are the angular coordinates defining the molecular orientations. Similarly, the rotational propagator can be separated by individual coordinates.

$$G_{\text{rot}}(\vec{\Phi}, \vec{\Phi}'; \Delta\tau) \approx \prod_i G(\phi_i, \phi_i'; \Delta\tau) \tag{40}$$

For small $\Delta\tau$ the propagator can be approximated by Trotter decomposition. This approximation reduces the time step error from $\mathcal{O}(\Delta\tau^2)$ to $\mathcal{O}(\Delta\tau^3)$.

$$e^{-\Delta\tau\mathscr{H}} = e^{-\Delta\tau(\hat{T}+\hat{V})} \approx e^{-\Delta\tau\frac{\hat{V}}{2}} e^{-\Delta\tau\hat{T}} e^{-\Delta\tau\frac{\hat{V}}{2}} + \mathcal{O}(\Delta\tau^3) \tag{41}$$

The term involving $\hat{V}$ can be associated with a non-conservation of probability, which we will implement as a weight. A short-time approximation for the propagator is that of free particle diffusion multiplied by a weight factor. For a single rotational coordinate in a potential that depends only on ϕ, this approximation is [1]

$$\begin{aligned} G(\phi', \phi; \Delta\tau) &= \langle\phi'| \exp\left\{\left(-d\partial_\phi^2 + \hat{V}\right)\Delta\tau\right\} |\phi\rangle \\ &\approx \underbrace{\frac{1}{\sqrt{4\pi d\Delta\tau}} \exp\left\{\frac{-(\phi'-\phi)^2}{4d\Delta\tau}\right\}}_{\text{free-particle diffusion}} \underbrace{\exp\left\{\frac{\Delta\tau}{2}\left(V(\phi') + V(\phi)\right)\right\}}_{\text{weight}} \end{aligned} \tag{42}$$

where $d = \frac{1}{2I_\phi}$. Since the propagator eventually propagates an initial state to the lowest non-orthogonal state as in Eq. (16), we can use this propagator to generate samples from the ground state, ψ_0.

3.5 Quantum Importance Sampling Transformation

In this section, I will incorporate the guiding function of Sec. (2.4) to define the mixed distribution, $\mathcal{F}$, and the importance sampling transformation. This

transformation can serve to improve the statistics of our estimators by guiding the sampling to more significant areas of configuration space.

$$\mathcal{F} \equiv \psi_0 \psi_g \tag{43}$$

In order to find the corresponding propagator

$$\tilde{G} = \psi_g G \psi_g^{-1} = e^{\psi_g \mathscr{H} \psi_g^{-1}} \tag{44}$$

The transformed Schrödinger equation is.

$$\left(\psi_g \mathscr{H} \psi_g^{-1}\right) \mathcal{F} = \psi_g \left\{ -\nabla^2 \frac{\mathcal{F}}{\psi_g} + V \frac{\mathcal{F}}{\psi_g} \right\} = -\partial_\tau \mathcal{F} \tag{45}$$

The transformed Schrödinger equation can be cast into an equation for diffusion, branching, and drift with the following definition of the "quantum force" $\vec{D}$, and local energy E_L.

$$\vec{D} = 2\frac{\nabla \psi_g}{\psi_g}, \qquad E_L = -\frac{\nabla^2 \psi_g}{\psi_g} + V \tag{46}$$

$$-\nabla^2 \mathcal{F} + \nabla \cdot \left(\mathcal{F} \vec{D}\right) + E_L \mathcal{F} = -\partial \tau \mathcal{F} \tag{47}$$

The term quantum force is unfortunate. The $\vec{D}$ term is better understood in this context as a drift velocity. The local energy, E_L, is associated with the violation of probability conservation in Eq. (18), and will simply replace the potential in the expression for the weight, w. It can be shown that an approximation for the importance sampled imaginary-time evolution operator for a single angular coordinate representing rotation about an axis with moment of inertia, I, is

$$\begin{aligned} &\tilde{G}_{\mathrm{rot}}(\phi', \phi; \Delta\tau) \approx \tilde{P}(\phi', \phi; \Delta\tau) \tilde{w}(\phi', \phi) \\ &= \underbrace{\frac{1}{\sqrt{4\pi d \Delta\tau}} \exp\left\{ \frac{-(\phi' - \phi - D(\phi)\Delta\tau)^2}{4 d \Delta\tau} \right\}}_{\text{drift/diffusion}} \underbrace{\exp\left\{ \frac{\Delta\tau}{2} \left[E_L(\phi') + E_L(\phi) \right] \right\}}_{\text{weight}} \end{aligned} \tag{48}$$

where $d = \frac{1}{2I}$. This propagator can be used to generate samples from the distribution $\psi_0 \psi_g$. This is the same as the propagator for translational coordinates with the

mass replacing the moment of inertia in the diffusion constant, d. An algorithm for estimating ground state observables that incorporates this transformation is given in Algorithm 1.

Algorithm 1 A simple algorithm for estimating ground state expectation values for m quantum mechanical rigid bodies

- Generate a set of N initial configurations

$$\{\vec{x}_{0,i}\} = \left\{\vec{R}_{0,i}, \vec{\Phi}_{0,i}\right\} \tag{49}$$

sampled from ψ_g^2, by Metropolis. Assign to each of these a weight $\mathcal{W}_{0,i} = 1$.

- Propagate each configuration with the importance sampled propagator.

$$\tilde{G}(x', x; \Delta\tau) = \tilde{P}(x', x; \Delta\tau)\tilde{w}(x, x; \Delta\tau) \tag{50}$$

The application of $\tilde{P}(x', x; \Delta\tau)$ on an configuration consists of a Gaussian distributed random step ξ, for all coordinates with $\sigma = \sqrt{2d\Delta\tau} = \left\{\sqrt{\frac{\Delta\tau}{I_{m,p}}}, \sqrt{\frac{\Delta\tau}{M_m}}\right\}$, and a drift step, $D\Delta\tau$.

  - For a Cartesian center of mass coordinate $\vec{r}_m$, of rigid-body m, with mass M_m.

$$\vec{r}_m(\tau + \Delta\tau) = \vec{r}_m(\tau) + \vec{\xi}_m + 2\frac{\vec{\nabla}_m\psi_g(x_\tau)}{\psi_g(x_\tau)}\Delta\tau \tag{51}$$

  - For an angular coordinate of rigid-body m, about principal axis p, with moment of inertia $I_{m,p}$.

$$\phi_{m,p}(\tau + \Delta\tau) = \phi_{m,p}(\tau) + \xi_{m,p} + 2\frac{\partial_{\phi_{m,p}}\psi_g(x_\tau)}{\psi_g(x_\tau)}\Delta\tau \tag{52}$$

  - The weight of the configuration $\mathcal{W}$, is multiplied by w.

$$\mathcal{W}_\tau(x_\tau, x_{\tau-\Delta\tau}, \cdots, x_0) = \mathcal{W}_{\tau-\Delta\tau}e^{\frac{\Delta\tau}{2}\left[E_{L,\psi_g}(x_\tau)+E_{L,\psi_g}(x_{\tau-\Delta\tau})\right]} \tag{53}$$

where the local energy is

$$E_{L,\psi_g}(x) = -\sum_m\left(\frac{1}{2M_m}\frac{\nabla_m^2\psi_g(x)}{\psi_g(x)} + \sum_p\frac{1}{2I_{m,p}}\frac{\partial_{m,p}^2\psi_g(x)}{\psi_g(x)}\right) + V(x) \tag{54}$$

- Estimate observables as weighted averages over projected configurations. For example, an estimate of the ground state energy is

$$\mathcal{E}_0 = \lim_{\tau,N\to\infty;\Delta\tau\to 0}\left\{\frac{\sum_{i=1}^N \mathcal{W}_{\tau,i}E_L(x_{\tau,i})}{\sum_{i=1}^N \mathcal{W}_{\tau,i}}\right\} \tag{55}$$

3.5.1 Radial Coordinate

Consider the two-body Hamiltonian where the kinetic energy due to the motion of the center of mass has been separated out, and what remains is the same as the Hamiltonian of a single particle of reduced mass μ.

$$\mathcal{H} = -\frac{1}{2\mu}\frac{1}{r^2}\partial_r r^2 \partial_r + \hat{L}^2 + V \tag{56}$$

There is no reason why the radial part of our rotors Hamiltonian has to be discarded. There may be circumstances where the differing time scales of vibrational and rotational motions justify separate treatment, but there may also be cases where the advantage of spherical center of mass coordinates goes beyond that. In Cartesian coordinates, if the two masses of the rotor differ greatly, then the Brownian motion required for diffusion Monte Carlo to generate its samples has a time-step that is limited by the smaller mass. As a practical matter this means that it becomes impossible for the system to effectively explore it's configuration space in a reasonable number of Monte Carlo steps. If there was a way to propose configurations based on the diffusion in spherical coordinates, then these two masses could be treated democratically. All that is needed is the short time propagator for the radial motion. Following the method of Section (2.4), start with the mixed distribution.

$$\mathcal{F}(r) \equiv \psi_0(r)\psi_g(r) \tag{57}$$

And find the corresponding propagator

$$\tilde{G} = \psi_g(r) G \psi_g^{-1}(r) = e^{\psi_g(r)\mathcal{H}\psi_g^{-1}(r)} \tag{58}$$

The transformed Schrödinger equation is.

$$\left\{\psi_g(r)\mathcal{H}\psi_g(r)^{-1}\right\}\mathcal{F}(r) = \psi_g(r)\left\{-\frac{1}{2\mu r^2}\partial_r r^2 \partial_r\left[\frac{\mathcal{F}(r)}{\psi_g(r)}\right] + V\frac{\mathcal{F}(r)}{\psi_g(r)}\right\} = -\partial_\tau \mathcal{F} \tag{59}$$

With the use of the identity,

$$\frac{1}{r^2}\partial_r r^2 \partial_r \left[\frac{\mathcal{F}(r)}{\psi_g(r)}\right] = \frac{1}{r}\partial_r^2 \left[r\frac{\mathcal{F}(r)}{\psi_g(r)}\right], \tag{60}$$

the transformed Schrödinger equation can be cast into an equation for diffusion, branching, and drift. The drift D, and local energy E_L are

$$D(r) = \frac{1}{2\mu}\left\{2\frac{\partial_r\left(\frac{\psi_g(r)}{r}\right)}{\left(\frac{\psi_g(r)}{r}\right)}\right\} = \frac{1}{2\mu}\left(2\frac{\partial_r\psi_g(r)}{\psi_g(r)} - \frac{2}{r}\right) \tag{61}$$

$$E_L(r) = -\frac{1}{2\mu}\left\{\frac{\partial_r^2\left(\frac{\psi_g(r)}{r}\right)}{\left(\frac{\psi_g(r)}{r}\right)}\right\} + V = -\frac{1}{2\mu}\left\{\frac{\partial_r^2\psi_g(r)}{\psi_g(r)} - \frac{2\partial_r\psi_g(r)}{r\psi_g(r)} + \frac{2}{r^2}\right\} + V \tag{62}$$

$$-\frac{1}{2\mu}\partial_r^2\mathcal{F} + \partial_r\left(\mathcal{F}D\right) + E_L\mathcal{F} = -\partial\tau\mathcal{F} \tag{63}$$

If the solution to the one-dimensional problem is known to be $\varphi(r)$, then the choice for the guiding function $\psi_g(r) = r\varphi(r)$ will result in a constant local energy equal to the exact energy of the one-dimensional problem. The short time propagator for the radial coordinate is

$$\tilde{G}_{\rm rot}(r', r; \Delta\tau) \approx \frac{1}{\sqrt{4\pi d\Delta\tau}} \exp\left\{\frac{-(r' - r - D(r)\Delta\tau)^2}{4d\Delta\tau}\right\} \exp\left\{\frac{\Delta\tau}{2}\left[E_L(r') + E_L(r)\right]\right\}. \tag{64}$$

where $d = \frac{1}{2\mu}$ is our new democratic diffusion constant. This "local energy" only exists to help generate samples from the desired distribution, and should not be confused with the configurational eigenvalue which will be used to estimate the energy.

3.6 Excited States

In this section I will review the variational method as it applies to excited states, and combine it with the projector methods above. The idea is to solve $\mathscr{H}|\Psi\rangle = \mathcal{E}|\Psi\rangle$ by expansion in basis functions. This transforms the Schrödinger

equation, a differential equation, into a linear algebra problem, the generalized eigenvalue problem. Once this is done the matrix elements can evaluated in the mixed distribution in order to eliminate the variational bias resulting from the incomplete trial basis set.

3.6.1 Reyleigh-Ritz Variational Method and the Generalized Eigenproblem

Begin with an arbitrary state built from some trial basis set, $\{|b_i\rangle\}$ where, for simplicity, $a_i \in \mathbb{R}$. The results can easily be generalized to complex a_i.

$$|\Psi\rangle = \sum_i a_i |b_i\rangle \tag{65}$$

Define the matrix elements

$$N_{i,j} = \langle b_i | b_j \rangle, \qquad H_{i,j} = \langle b_i | \hat{\mathscr{H}} | b_j \rangle \tag{66}$$

Upper bounds on the true eigenvalues are found through extremization the energy with respect to the expansion coefficients [2].

$$\mathcal{E} = \frac{\langle \Psi | \hat{\mathscr{H}} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{i,j=1}^{N} a_i a_j H_{i,j}}{\sum_{i,j=1}^{N} a_i a_j N_{i,j}} \tag{67}$$

$$d\mathcal{E} = \sum_{i=1}^{N} \left\{ \partial_{a_i} \left(\frac{\sum_{i,j=1}^{N} a_i a_j H_{i,j}}{\sum_{i,j=1}^{N} a_i a_j N_{i,j}} \right) da_i \right\} = 0 \tag{68}$$

Yields the set of N secular equations

$$\sum_j \left(H_{i,j} - \mathcal{E} N_{ij} \right) a_j = 0, \qquad \forall i = 1, N \tag{69}$$

These can be rewritten as an generalized eigenproblem.

$$H\vec{a} = \mathcal{E} N \vec{a} \tag{70}$$

$\vec{a}$ is the column-vector representation of $|\Psi\rangle$ in the basis, $\{|b_i\rangle\}$. It is an important quality of this method that if the subspace spanned by the trial basis includes that

spanned by the eigenstates then the method will yield an exact solution. In the next section I will apply the projector methods described in Sections 2.3 and 2.4 to remove the variational bais of the estimators of the individual matrix elements.

3.6.2 Subspace Projection and Correlation Function Monte Carlo

The variational bias resulting from an incomplete basis set can be removed by projection of the matrix elements onto the desired subspace. The time dependent matrix elements are defined as

$$N_{ij}(\tau) = \langle b_i | G^\tau | b_j \rangle = \int dx_1 dx_2 b_j(x_2) G(x_2, x_1; \tau) b_i(x_1) \tag{71}$$

$$H_{ij}(\tau) = \langle b_i | \mathscr{H} G^\tau | b_j \rangle = \int dx_1 dx_2 b_j(x_2) \mathscr{H}_i(x_2) G(x_2, x_1; \tau) b_i(x_1) \tag{72}$$

where $\mathscr{H}_i(x) b_i(x) \equiv \langle x | \hat{\mathscr{H}} | b_i \rangle$ is the configurational eigenvalue as in Eq. (33). Once these are evaluated they can be inserted into the time dependent generalized eigenproblem.

$$\sum_{j=1}^{N} \left[H_{i,j}(\tau) - \mathcal{E}_k(\tau) N_{i,j}(\tau) \right] a_{k,j}(\tau) = 0 \tag{73}$$

as long as there is some overlap between the basis $\{b_i\}$ and the eigenstates $\{\varphi_i\}$. $\mathcal{E}_k(\tau)$ converges exponentially and monotonically to the exact energy levels [3].

$$\lim_{\tau \to \infty} \mathcal{E}_k(\tau) = E_k; \qquad 1 \leq k \leq N \tag{74}$$

Transformation of the basis set ($b' = \hat{T} b, \quad \det(T) \neq 0$) has no effect on the eigenvalues. If H, N are symmetric it is possible to transform to a basis where $N = \mathbb{1}$ and H is diagonal with elements $\mathcal{E}_k(\tau)$. In this basis the projected basis sets converge exponentially to the exact eigenfunctions.

$$\lim_{\tau \to \infty} e^{-\tau \mathscr{H}} b_i = \tilde{b_i} = \varphi_i \tag{75}$$

All that remains is to construct estimators for these matrix elements. The $N_{i,j}$ and $H_{i,j}$ matrices can be rewritten in terms of the importance sampled imaginary-time

evolution operator. In order for all matrix elements to be collected over the same random walk, the matrix elements are given for the case where the distribution sampled from for 0 projection time is $\psi_g^2(x_1)$. As a practical matter, this can be guaranteed by a Metropolis accept reject step.

$$N_{ij}(\tau) = \int dx_1 dx_2 B_j(x_2) \tilde{G}(x_1, x_2; \tau) B_i(x_1) \psi_g^2(x_1) \tag{76}$$

$$H_{ij}(\tau) = \int dx_1 dx_2 B_j(x_2) \mathscr{H}_j(x_2) \tilde{G}(x_1, x_2; \tau) B_i(x_1) \psi_g^2(x_1) \tag{77}$$

$$\tilde{G}(x_1, x_2; \tau) = \psi_g(x_2) \langle x_2 | e^{-\tau \mathscr{H}} | x_1 \rangle \psi_g^{-1}(x_1), \quad B_i(x) = \frac{b_i(x)}{\psi_g(x)} \tag{78}$$

Estimators for the $N_{i,j}(\tau)$ and $H_{i,j}(\tau)$ can be written as averages over samples.

$$N_{ij}(k\Delta\tau) = \langle B_i(x_n) \mathcal{W}_{n,n+k} B_j(x_{n+k}) \rangle \tag{79}$$

$$H_{ij}(k\Delta\tau) = \langle B_i(x_n) \mathcal{W}_{n,n+k} B_j(x_{n+k}) \mathscr{H}_j(x_{n+k}) \rangle \tag{80}$$

The required weights depend on the local energy of the guiding function.

$$\begin{aligned} \mathcal{W}_{n,n+k} &= \prod_{j=n}^{k+n-1} \tilde{w}(x_{j+1}, x_j; \Delta\tau) \\ &= \exp\left\{ -\frac{\Delta\tau}{2} \sum_{j=n}^{n+k-1} \left[E_{L,\psi_g}(x_j) + E_{L,\psi_g}(x_{j+1}) \right] \right\} \end{aligned} \tag{81}$$

where the local energy of the guiding function is defined as

$$E_{L,\psi_g}(x) = \frac{\mathscr{H}_g(x) \psi_g(x)}{\psi_g(x)} \tag{82}$$

N and H are just averages because the probability distribution of the path is.

$$\rho(x_n) = \psi_g^2(x_n) \prod_{j=n}^{n+k-1} G_d(x_{j+1}, x_j, \tau) \tag{83}$$

In the long projection limit $\rho(x)$ becomes the mixed distribution $\psi_g(x)\psi_0(x)$.

$$\lim_{\tau \to \infty} \left\{ \psi_g^2(x_n) \prod_{j=n}^{n+k-1} G_d(x_{j+1}, x_j, \tau) \right\} = \psi_g(x_n) \psi_0(x_n) \tag{84}$$

Although $H_{i,j}$ is symmetric in i, j, one should resist the temptation to symmetrize the corresponding estimator as it destroys the zero-variance property and adds noise to the calculation [4].

As the matrix elements are projected in imaginary-time the information about the excited states decays exponentially. Eventually the excited states will one by one disappear into the statistical noise and the matrix pencil, $(H - \mathscr{E}N)$, will become singular. Methods for dealing with this instability include singular value decomposition and generalized upper triangular decomposition [5] and are beyond the scope of this document. This problem is made worse by the fluctuating weights which have the effect of amplifying the statistical noise. The best results are obtained by taking projection times that are long enough to remove the variational bias but not so long as to encounter numerical instability. In the next chapter, I will discuss the problem of constructing basis states and choosing a guiding function in such a way as to minimize these problems.

List of References

[1] A. Viel, M. V. Patela, P. Niyaza, and K. B. Whaley, "Importance sampling in rigid body diffusion monte carlo," *Computer Physics Communications*, vol. 145, pp. 24–47, 2000.

[2] J. K. L. MacDonald, "Successive approximations by the rayleigh-ritz variation method," *Phys. Rev.*, vol. 43, p. 830, 1933.

[3] D. M. Ceperley and B. Bernu, "The calculation of excited state properties with quantum Monte Carlo," *J. Chem. Phys.*, vol. 89, p. 6316, 1988.

[4] M. P. Nightingale and V. Melik-Alaverdian, "Optimization of ground and excited state wavefunctions and van der Waals clusters," *Phys. Rev. Lett.*, vol. 87, p. 043401, 2001.

[5] J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, *Templates for the solution of algebraic eigenvalue problems: a practical guide*, Z. Bai, Ed. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2000.

CHAPTER 4

Construction of Basis States, Trial States and the Guiding Function

In this chapter I will describe, in the context of the rigid-body approximation, the construction and optimization of the basis states, trial states and guiding function. Section 4.1 introduces the basic form of the trial states and defines an algorithm for their optimization. Section 4.2 describes a method for constructing an optimized guiding function based on the efficiency of the weighted average used to estimate observables. The invariant polynomials from which the basis states are built are described in section 4.3. A simple example of how they are constructed is given and there is a discussion of the reduction in the number of terms associated with the rigid-body approximation.

4.1 Basis States and Trial Functions

The reduction of variational bias by projection described in chapter 3 comes at the cost of statistical error. The projector method causes information about the excited states to decay exponentially and eventually become overwhelmed by the ground state. Keeping the necessary projection time as small as possible by optimizing the basis states allows us to retrieve the information before it disappears. This is best accomplished by optimization of the trial states. Trial states are constructed from the basis states with linear and non-linear variational parameters.

$$\psi_T(x) \propto \left(\sum_i a_i b_i(x) \right) \exp \left(\sum_i \alpha_i b_i(x) \right) \tag{85}$$

where the a_i and α_i are constants and the b_i are configuration space functions.

4.1.1 Optimization of the Trial States

The optimal values for the non-linear variational parameters are found by minimizing the variance of the local energy over a small Monte Carlo sample,

which is kept fixed during the optimization. The linear parameters are optimized by solution of the generalized eigenproblem and must be recalculated for each new set of non-linear parameters as is described in the algorithm given. This generalized eigenproblem is ill-conditioned in the sense that as the projection time increases information about excited states becomes overwhelmed by the ground state as described in 3.6.1. In addition, the basis states may be close to numerically dependent even if they are mathematically independent. Consider, for example, the functions x^n with $n = 0, 1, \cdots$ on the interval $x = (0, 1)$. As a function of n these rapidly produce $x^n \approx 0$.

An algorithm for optimizing trial wavefunctions is given in Algorithm 2. The

---

Algorithm 2 Algorithm for optimizing trial wavefunctions,$\{\psi_T\}$

- Generate a sample of configurations from a relative probability density function $\psi_g(R)^2$, the guiding function.
- Solve ill-conditioned generalized eigenproblem for linear parameters.
- non-linear parameters are optimized by minimization of the variance of the local energy of individual linearly optimized excited states.

$$\chi^2 = \frac{\sum_\sigma \left\{ \frac{\mathscr{H}\psi_k(x_\sigma)}{\psi_g(x_\sigma)} - \mathscr{E}_k \frac{\psi_k(x_\sigma)}{\psi_g(x_\sigma)} \right\}^2}{\sum_\sigma \frac{\psi_k(x_\sigma)^2}{\psi_g(x_\sigma)^2}} \tag{86}$$

- For each choice of the non-linear parameters, new optimized linear parameters have to be computed. Full optimization of all parameters consists of a linear optimization nested in a non-linear one.

---

factors of ψ_g in the denominators are due to the fact that we are calculating expectation values in the distribution ψ_k^2 as averages over samples taken from ψ_g^2.

4.2 Guiding Function

the need for a guiding function that allow one to sample simultaneously from all excited states in the computation has been mentioned before. I now discuss in

more detail how this function can be designed for high sampling efficiency. The final ingredient in the Monte Carlo calculation is the guiding function, ψ_g. In an excited state calculation, this function is chosen to have sufficient overlap with all of the states we wish to sample, and must be positive everywhere. Since estimators are constructed as weighted averages over samples taken from ψ_g^2, optimization of this function involves maximizing the efficiency of a weighted average (or minimizing it's inefficiency) as described below.

4.2.1 Efficiency of a weighted average

Estimators take the form of a weighted average. What follows assumes that the weights are constants which they are not. The approximation of constant weights is justified by fact that any breakdown resulting from the approximation affects only the efficiency of the process and not the ultimate result.

$$\hat{\mu} = \frac{\sum_i w_i x_i}{\sum_i w_i} \tag{87}$$

The variance of a weighted average can be used to obtain an expression for the effective number of measurements.

$$\hat{\sigma}^2 = \underbrace{\left\{ \frac{\left(\sum_i w_i\right)^2}{\left(\sum_i w_i\right)^2 - \sum_i w_i^2} \right\}}_{\frac{N_{\mathrm{Eff}}}{N_{\mathrm{Eff}}-1}} \frac{\sum_i w_i \left(x_i - \hat{\mu}\right)^2}{\sum_i w_i} \tag{88}$$

$$N_{\mathrm{Eff}} = \frac{\left(\sum_i w_i\right)^2}{\sum_i w_i^2} \tag{89}$$

In turn this can be used to define the efficiency of a weighted average.

$$\eta = \frac{N_{\mathrm{Eff}}}{N} \tag{90}$$

4.2.2 Object function

The parameters of a guiding function can be found by minimizing an object function consisting of inefficiencies, η^{-1}. Each trial state, $\tilde{\psi}_k$, will contribute a

term, $\eta_k(w_\sigma^k)$. The weights for this term are defined as

$$w_\sigma^k = \left(\frac{\tilde{\psi}^k(R_\sigma)}{\psi_g(R_\sigma)} \right)^2 \tag{91}$$

and the efficiency

$$\eta_k = \frac{\left(\sum_\sigma w_\sigma^k\right)^2}{\sum_\sigma (w_\sigma^k)^2} \tag{92}$$

In the process of removing the variational bias of the trial states, there is an additional source of fluctuating weights that reduce the efficiency. As a simple model we consider the special case of a weighted sum with only two different weights. One of these, without loss of generality is set to unity. The projection weights are given by

$$w = \exp\left\{ -\frac{\tau}{2} \left(\mathscr{E}_g(R) + \mathscr{E}_g(R') - 2E_G \right) \right\} \tag{93}$$

Consider a special case where the weights can only take on the values one and w with $\frac{N(\text{weight}=1)}{N(\text{weight}=w)} = x$. The smallest efficiency, i.e. the most pessimistic estimate, corresponds to

$$x = \frac{w^2}{1+w^2} \tag{94}$$

The efficiency at this ratio is

$$\eta = \frac{4w^2}{(1+w^2)^2} \tag{95}$$

A typical projection time will be on the order of $\frac{1}{E_g}$, where E_g is the energy associated with the guiding function. An approximate ratio of largest to smallest projected weights

$$w_p \approx \frac{e^{E_g+\sigma_g}}{e^{E_g-\sigma_g}} = \exp\left\{ \frac{2\sigma_g}{E_g} \right\} \tag{96}$$

These approximations give the following expression for diffusion Monte Carlo efficiency.

$$\eta_p = \frac{4w_p^2}{1+w_p^2} \tag{97}$$

For stability, we insist that the new guiding function is not very different from the old one by defining the efficiency of the guiding function, η_g, in terms of the weights.

$$w_\sigma^g = \left(\frac{\psi_g'(R_\sigma)}{\psi_g(R_\sigma)} \right)^2 \tag{98}$$

$$\eta_g = \frac{\left(\sum_\sigma w_\sigma^g\right)^2}{\sum_\sigma (w_\sigma^g)^2} \tag{99}$$

An algorithm for optimizing the guiding function is given in Algorithm 3.

Algorithm 3 Optimization of the Guiding Function

- Generate initial sample from simple diffusion Monte Carlo.
- Construct optimized ground state.
- Construct corresponding guiding function.
- Add states and update the guiding function.
- The optimized guiding function will be one that minimizes the object function

$$\mathscr{O} = \sum_{i=1}^{n} \eta_i^{-1} + \eta_p^{-1} + \eta_g^{-1} \tag{100}$$

4.3 Invariants

The evaluation of the trial states and their derivatives is typically the most computationally expensive part of the calculation. The problem is how to write these states with as few terms as possible such that the basis is as complete as it can be. One solution for this problem is to construct a basis of fundamental invariants from the set of interparticle distances $\{r_{i,j}\}$. A set of invariants $\{I_i\}$, can be constructed by application of the symmetry group operations $\{P_i\}$, on the interparticle distances. Algorithm 4 provides a method for generating these primary invariants.

4.3.1 Algorithm for Generating a Basis From Primary Invariants

Algorithm 4 Construction of a Basis of Primary Invariants

for D=0 **to** Maximum Degree **do**

Generate monomials $\{m_i\}$, from the interparticle distances, $\{r_i\}$.

$$\{m_i\} = \left\{ \prod_i r_i^{a_i} | \sum a_i = D \right\} \tag{101}$$

for all m_i **do**

Apply group operations, $\{P_i\}$, to generate invariant polynomial, x_i:

$$x_i = \sum_j P_j m_i \tag{102}$$

if x_i cannot be written in terms of the existing invariants, $\{I_i\}$ **then**

add x_i to the list of invariants:

$I_n = x_i$

$n = n + 1$

end if

end for

end for

Ensure: The number of invariants is in agreement with the Molien series.

This set of primary invariants is not unique, and the algorithm given in this section is not the only method for obtaining them. The number of independent invariants for a given degree is unique and can be obtained via the Molien series [1].

4.3.2 Simple Example

As a simple example, consider a single rigid rotor interacting with a particle as shown in Figure 5. For simplicity, we assume all particles have bosonic symmetry. There are only two non-fixed interparticle distances, $r_{a,c}$ and $r_{b,c}$, as in fig. (5). For the bosonic case, our wavefunctions must be symmetric under interchange of

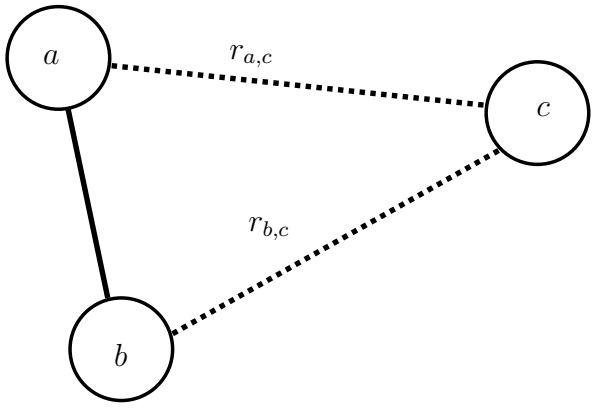


Figure 5: Simple Example: Basis of Fundamental Invariants.

a and b. There is only one symmetry operation, P.

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{103}$$

Apply the symmetry operations to the variables to generate invariant polynomials.

$$x_1 = r_{a,c} + Pr_{a,c} = r_{a,c} + r_{b,c} = I_1 \tag{104}$$

$$x_2 = r_{b,c} + Pr_{b,c} = r_{b,c} + r_{a,c} = I_1 \tag{105}$$

In this example there is only one invariant of degree one.

$$I_1 = r_{a,c} + r_{b,c} \tag{106}$$

This procedure is then carried out for monomials of higher degrees. New invariants that can be written as sums of products of existing invariants are thrown out. The list of monomials of degree two is

$$\{m_i\} = \left\{ r_{a,c}r_{b,c}, \quad r_{a,c}^2, \quad r_{b,c}^2 \right\} \tag{107}$$

Application of P yields

$$x_1 = r_{a,c}r_{b,c} + Pr_{a,c}r_{b,c} = 2r_{a,c}r_{b,c} = I_2 \tag{108}$$

$$x_2 = r_{a,c}^2 + Pr_{a,c}^2 = r_{a,c}^2 + r_{b,c}^2 = I_1^2 - \frac{I_2}{2} \tag{109}$$

$$x_3 = r_{b,c}^2 + Pr_{b,c}^2 = x_2 \tag{110}$$

Only one of these terms can be added to the list as I_2. The other will be writable in terms of I_1 and I_2. We choose the polynomial with the fewest terms, $r_{a,c}r_{b,c}$. For this simple case, it can be shown that all other invariants can be written in terms of I_1 and I_2. A basis consisting of products of powers of these two invariants is complete.

$$\{I_i\} = \{r_{a,c} + r_{b,c}, r_{a,c}r_{b,c}\} \tag{111}$$

$$\{b_i\} = \left\{I_1, I_2, I_1I_2, I_1^2I_2, I_1I_2^2, I_1^3I_2, I_1^2I_2^2, I_1I_2^3, \cdots\right\} \tag{112}$$

4.3.3 Effect of the Rigid-Body Approximation on the Construction of Basis States

The procedure for calculating invariants is limited in part by our ability to enumerate the symmetry operations. For a bosonic cluster the number of permutations is $n!$ which very quickly becomes prohibitive. For a cluster of rigid molecules consisting of identical atoms we ignore permutations that move atoms from one molecule to another reducing the number of group elements to $N = m!(a!)^m$, where m is the number of molecules and a is the number of atoms. A plot of $\log N$ vs. ma is shown for $a = 1, 2$ in Figure 6.

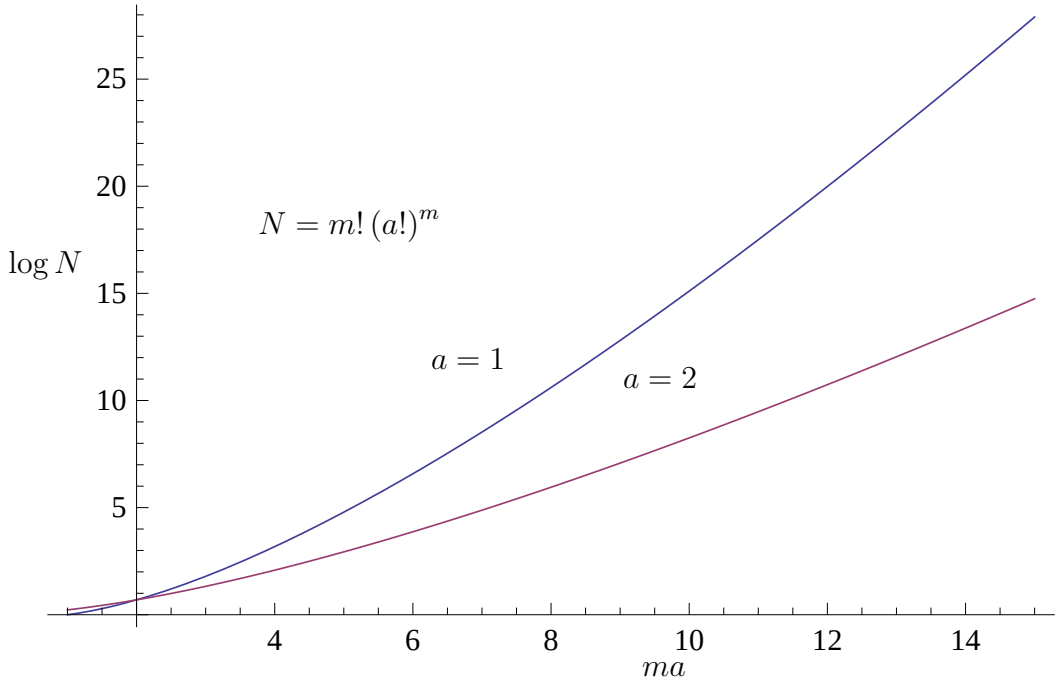


Figure 6: Logarithm of the number of symmetry permutations for bosonic clusters of one and two atom molecules.

List of References

[1] B. Sturmfels, *Algorithms in invariant theory; 2nd ed.*, ser. Texts and Monographs in Symbolic Computation. Dordrecht: Springer, 2007.

CHAPTER 5

Quaternions

In this chapter I will introduce quaternion numbers and the associated notation as it relates to representing rotation in three dimensions. This framework is a prerequisite for the following chapter where I will describe the use of quaternion algebra to implement the Brownian rotational motion generated by the imaginary-time evolution operator of Eq. (48).

5.1 Basics and Notation

Quaternions are a non-commutative extension of complex numbers. Instead of the usual $i = \sqrt{-1}$, we consider solutions of

$$x = \sqrt{-\mathbb{1}} = \sqrt{-\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} \tag{113}$$

and extend the complex plane to a four dimensional space with one real, and three imaginary basis vectors. The set of equations, and the multiplication table are

$$i^2 = j^2 = k^2 = ijk = -\mathbb{1}, \qquad \begin{array}{rclrcl} ij & = & k, & ji & = & -k, \\ jk & = & i, & kj & = & -i, \\ ki & = & j, & ik & = & -j. \end{array} \tag{114}$$

A basis can be written in terms of the Pauli matrices, $-i\sigma_1, -i\sigma_2, -i\sigma_3$, since $\sigma_i\sigma_j = \delta_{i,j} + i\epsilon_{i,j,k}\sigma_k$.

$$\hat{i} = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}, \quad \hat{j} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \hat{k} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \tag{115}$$

As a matter of notation a quaternion is written as a scalar real part and vector imaginary part.

$$q \equiv q_0 + q_1 i + q_2 j + q_3 k = q_0 + \vec{q} = (q_0, \vec{q}) \tag{116}$$

where addition is defined as

$$(q_0+q_1i+q_1j+q_3k)+(r_0+r_1i+r_2j+r_3k) = (q_0+r_0)+(q_1+r_1)i+(q_2+r_2)j+(q_3+r_3)k \tag{117}$$

Multiplication is defined as

$$qr = (q_0 + \vec{q})\,(r_0 + \vec{r}) = q_0r_0 + q_0\vec{r} + r_0\vec{q} - \vec{q}\cdot\vec{r} + \vec{q}\times\vec{r} \tag{118}$$

A quaternion with no real part is called pure, not to be confused with a pure rotation which refers to a rotation by a perfectly orthonormal matrix or unit quaternion.

$$q_{\text{pure}} \equiv (0, \vec{q}) \tag{119}$$

5.2 Quaternion Rotation

The goal is to find a formula in analogy to complex numbers that expresses rotation in three dimensions using quaternion multiplication. Represent a regular vector $\vec{r}$ as a pure quaternion $q = (0, \vec{r})$. Rotating a vector should yield a vector, but multiplying pure quaternion with an arbitrary quaternion r may result in a non-pure quaternion. We can cancel the real part of q' through conjugation of q by a unit quaternion R.

$$q' = RqR^{-1} \tag{120}$$

Unit quaternions, R, of the following form are used for rotation.

$$R(\theta, \vec{\omega}) = (R_0, \vec{R}) = (\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\hat{\omega}) = e^{\frac{\theta}{2}\hat{\omega}} \tag{121}$$

$$q' = (0, \vec{r'}) = e^{\frac{\theta}{2}\hat{\omega}}(0, \vec{r})e^{-\frac{\theta}{2}\hat{\omega}} \tag{122}$$

For unit quaternions the complex conjugate is the inverse, $R^{-1} = R^*$. A counterclockwise rotation of a vector $\vec{r}$, through an angle θ, about an axis $\vec{\omega}$, can be

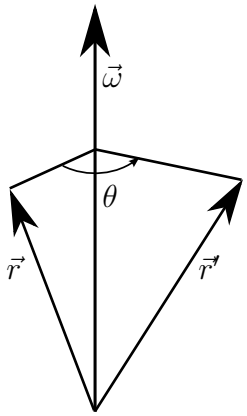


Figure 7: Quaternion Rotation

represented by conjugation of a pure quaternion, $q = (0, \vec{r})$. The general formula for quaternion rotation is

$$\begin{aligned} RqR^* &= (R_0^2 - \vec{R} \cdot \vec{R})\vec{q} + 2R_0(\vec{R} \times \vec{q}) + 2\vec{R}(\vec{R} \cdot \vec{q}) \\ &= \cos(\theta)\vec{q} + \sin(\theta)(\hat{\omega} \times \vec{q}) + \{1 - \cos(\theta)\}\,(\hat{\omega} \cdot \vec{q})\hat{\omega} \end{aligned} \tag{123}$$

This equation is easily recognized as the Rodrigues formula for rotation about an arbitrary axis.

CHAPTER 6

Matrix vs. Quaternion Implementation of Rotational Brownian Motion

To sample from the probability distributions required to construct estimators for the quantum mechanical rigid-body problem it is necessary to implement rotational diffusion as was described in Chapter 3. In the past this has done by producing a new configuration from the previous one. Ignoring translational motion, this new configuration is generated by random walk of the angles about the principal axes of the rigid body . For this reason, it is required that the transformation to the principal axes frame be available. This transformation is typically calculated for every step by diagonalizing the moment of inertia tensor. We would like a way to represent the transformation to the principal axes frame as a cumulative rotation product applied to some initial state and avoid explicitly calculating the transformation for every step. In this chapter, I will show that the numerical stability associated with the quaternion representation or rotation allows for a simple, efficient solution to this problem which is not available by matrix methods. Section 6.1 is intended to introduce and clarify some necessary concepts related to rotation. Section 6.3 provides an overview of more conventional matrix based approach and is intended to illustrate the associated difficulties. Finally, in Section 6.4, I present a simple quaternion based implementation and discuss it's advantages.

6.1 Numerical Stability and Pure Rotation

A proper rotation is one that preserves dot products and therefore does not change the shape of any object (collection of vectors) being rotated. Numerically, this property is lost upon repeated rotations due to the presence of roundoff error. The term "pure rotation" is used to describe proper, orthogonal rotations

that correspond to orthonormal matrices within the accuracy of the real number arithmetic.

The numerical stability provided by the quaternion representation of rotation suggests an algorithmic implementation of rotational diffusion based on a cumulative product of rotations. The equivalent cumulative matrix product is numerically unstable, that is, roundoff error causes it to eventually become non-pure and the orthonormalization required to recover purity is computationally expensive. Quaternion products on the other hand require only simple normalization to guarantee purity. For this reason, a quaternion based algorithm provides a simple mechanism to avoid the need to diagonalize the moment of inertia tensor and transform to the principal axis frame with every step. All rotations are pure cumulative rotation products applied to an initial state in which the principal axes are known as opposed to stepwise rotations applied to an intermediate state where the principal axes need to be calculated.

6.2 Definitions

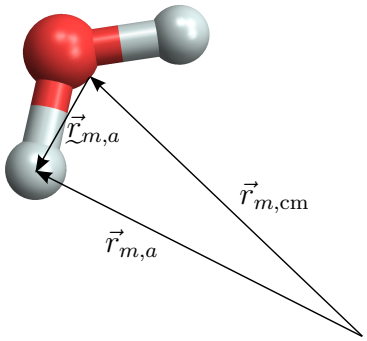


Figure 8: Definitions for Rotation of Rigid-Body

The vector $\vec{r}_{m,a}$, denotes the position of atom a, on molecule m, in the lab frame. $\vec{r}_{m,cm}$, denotes the position of the center of mass of molecule m, in the lab

frame. $\vec{\underset{\sim}{r}}_{m,a}$, is the position vector of atom a, on molecule m, in the reference frame that has it's origin at the center of mass of molecule m with it's axes parallel to those of the lab frame.

6.3 Matrix Based Implementation of Rotational Brownian Motion

The drift and diffusion resulting from propagation by $\tilde{G}(\phi', \phi; \Delta\tau)$, can be implemented using matrices. One way to do this [1] begins with the transformation of $\vec{r}_{m,a}$ to the principal axes frame, $\vec{x}_{m,a}$, with the moment of inertia tensor, I_m, and center of mass coordinates, $\vec{r}_{m,\mathrm{cm}}$.

$$\vec{\underset{\sim}{r}}_{m,a} \equiv \vec{r}_{m,a} - \vec{r}_{m,\mathrm{cm}}, \qquad \vec{x}_{m,a} = I_m \vec{\underset{\sim}{r}}_{m,a} I_m^T \tag{124}$$

Given the angles for rotations about the principal axes, ϕ_{m,p_i}, for molecule m, a rotation matrix, R_m, can be constructed in the following way.

$$R_m \equiv R_{m,p_1}(\phi_{m,p_1}) R_{m,p_2}(\phi_{m,p_2}) R_{m,p_3}(\phi_{m,p_3}) \tag{125}$$

The new position of atom a, on molecule m, in the principal axes frame is given by

$$\vec{x'}_{m,a} = R_m \vec{x}_{m,a} \tag{126}$$

The new position of atom a, on molecule m, in the lab frame is given by

$$\vec{r}_{m,a}(\tau + \Delta\tau) = I_m^T R_m(\Delta\tau) I_m \vec{\underset{\sim}{r}}_{m,a}(\tau) + \vec{r}_{m,\mathrm{cm}}(\tau) + \Delta\vec{r}_{m,\mathrm{cm}} \tag{127}$$

For small rotations, the commutation error resulting from the order of rotations can be ignored.

As described, calculation of the moment of inertia tensor every step to locate the principal axes is expensive, and avoidable. If a cumulative rotation product were available, one could find the current configuration by applying it to an initial configuration with no need to continually change reference frames.

6.4 Quaternion Based Implementation of Rotational Brownian Motion

In this section I will describe the use of quaternions for the implementation of rotational Brownian motion. One way to do this is to rewrite the matrix based approach from the previous section in the language of quaternions as has been done in [2],[3]. Although this simple translation is computationally less expensive than a matrix based approach, it does not take full advantage of the quaternion representation and needlessly propagates complications associated with a matrix implementation. In this section, I will instead describe a modification of this procedure that takes advantage of quaternion normalization to maintain numerically stable rotation products which can be applied to an initial state to generate new configurations.

Let the quaternion $R_{m,p_i}(\phi_{m,p_i}, \hat{p}_i)$, represent rotation about an axis $\hat{p}_i$, on rigid-body m, by angle ϕ_{m,p_i}.

$$R_{m,p_i}(\phi_{m,p_i}, \hat{p}_i) = (\cos\frac{\phi_{m,p_i}}{2}, \sin\frac{\phi_{m,p_i}}{2}\hat{p}_i) = e^{\frac{\phi_{m,p_i}}{2}\hat{p}_i} \tag{128}$$

An approximate quaternion for combined rotations about the three principal axes p_i, for rigid-body m is

$$R_m(\Delta\tau) \approx \prod_i R_{m,p_i} \approx \exp\left\{\sum_i \left(\frac{\phi_{m,p_i}}{2}\hat{p}_i\right)\right\} = \exp\left\{\frac{\theta_m}{2}\hat{N}_m\right\} \tag{129}$$

where

$$\hat{N}_m \equiv \frac{\phi_{m,p_1}\hat{p}_1 + \phi_{m,p_2}\hat{p}_2 + \phi_{m,p_3}\hat{p}_3}{\sqrt{\phi_{m,p_1}^2 + \phi_{m,p_2}^2 + \phi_{m,p_3}^2}}, \qquad \theta_m \equiv \sqrt{\phi_{m,p_1}^2 + \phi_{m,p_2}^2 + \phi_{m,p_3}^2} \tag{130}$$

Alternatively, a Trotter approximation can be used to reduce the errors associated with the non-commutation of the angular momentum operators.

$$R_m(\Delta\tau) \approx e^{\frac{\phi_{m,p_1}}{4}\hat{p}_1} e^{\frac{\phi_{m,p_2}}{4}\hat{p}_2} e^{\frac{\phi_{m,p_3}}{2}\hat{p}_3} e^{\frac{\phi_{m,p_2}}{4}\hat{p}_2} e^{\frac{\phi_{m,p_1}}{4}\hat{p}_1} \tag{131}$$

Imposing normalization

$$
\begin{aligned}
R_m(\tau) &= (R_{m,0}(\tau), \vec{R_m}(\tau)) \\
|R_m(\tau)| &= \sqrt{R_{m,0}(\tau)^2 + \vec{R_m}(\tau) \cdot \vec{R_m}(\tau)} \\
R_m(\tau) &= \frac{R_m(\tau)}{|R_m(\tau)|}
\end{aligned} \tag{132}
$$

allows for a stable product of rotations, that is a product of rotations that remains pure in the presence of roundoff error.

$$
R_m(\tau) = \prod_{t=0}^{\frac{\tau}{\Delta\tau}} R_m(\Delta\tau), \qquad R(0) = 1 \tag{133}
$$

This product can be used to find the position of a point on the rigid-body from its initial position, and the center of mass.

$$
\underbrace{\left\{0, \vec{\underset{\sim}{r}}_{m,a}(\tau)\right\}}_{\text{pure quaternion}}, = R_m(\tau) \left\{0, \vec{\underset{\sim}{r}}_{m,a}(0)\right\} R_m^*(\tau) \tag{134}
$$

$$
\vec{r}_{m,a}(\tau) = \vec{\underset{\sim}{r}}_{m,a}(\tau) + \vec{r}_{\text{cm}}(\tau), \tag{135}
$$

Because all configurations are generated by application of a cumulative rotation to an initial configuration, there is no need to translate to and from the principal axis frame with the moment of inertia tensor.

List of References

[1] V. Buch, "Treatment of rigid bodies by diffusion Monte Carlo: Application to the para-$H_2 \ldots H_2O$ and ortho-$H_2 \ldots$-H_2O clusters," *J. Chem. Phys.*, vol. 97, p. 726, 1992.

[2] D. M. Benoit and D. C. Clary, "Quaternion formulation of diffusion quantum monte carlo for the rotation of rigid molecules in clusters," *J. Chem. Phys.*, vol. 113(13), pp. 5193–5202, 2000.

[3] A. Viel, M. V. Patela, P. Niyaza, and K. B. Whaley, "Importance sampling in rigid body diffusion monte carlo," *Computer Physics Communications*, vol. 145, pp. 24–47, 2000.

CHAPTER 7

Results for Example Problems

In this chapter I will test these methods on some example problems. These examples are designed to demonstrate accuracy not precision and are chosen to be exactly solvable by other means. In addition, these problems are intended to serve as validation tests for the computer code. The problems are organized in the following way. First, Section 7.1 provides an exact solution for a few problems involving rigid rotors interacting via some potential function. Next, Section 7.2 provides the rigid-body diffusion Monte Carlo solution and compares results to the exact values. Section 7.3, provides the rigid-body diffusion Monte Carlo solution for the more complicated problem of two rigid rotors interacting at multiple atomic sited via a Lennard-Jones potential. Although an exact solution is not available for this problem, the results should approach that of the harmonic approximation as the mass increases. This problem also provides a more realistic framework to validate all aspects of the method and code..

7.1 Exact Solution

We consider the energy levels of a single rigid rotor in an external field and of a pair of mutually interacting rigid rotors. The rigid rotors come in two forms: AA (homo- nuclear or molecular) and AB (hetero- nuclear or molecular). In case of the former, we are interested in the case of bosons. Undoubtedly, one can prove that the bosonic ground state is nodeless, and therefore bosonic; Feynman's argument [1] can be applied to the current case.

It turns out that the simple form of the pair interaction that we consider is separable, in the sense that it can be written as sum over products of single-particle matrix elements. As a consequence, the problem reduces to summing the results

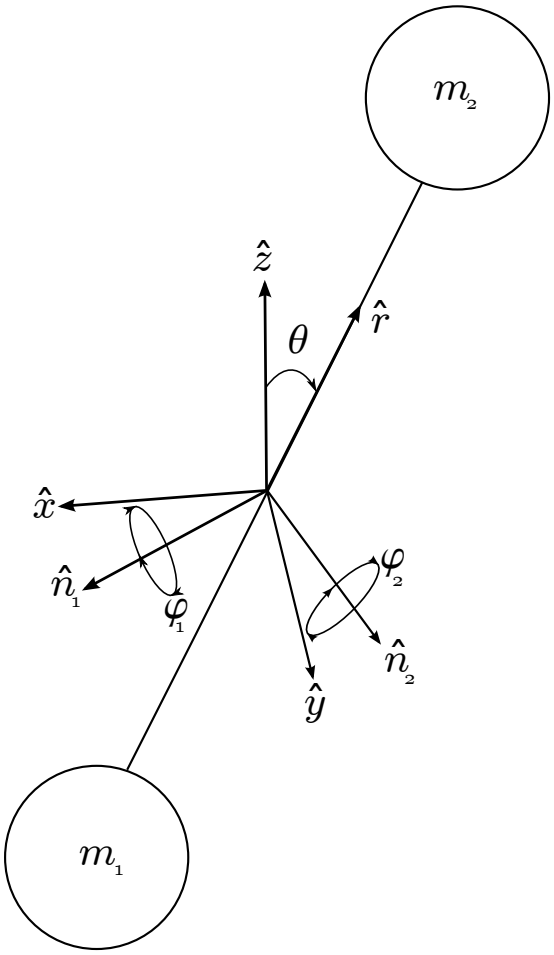


Figure 9: Rotor

of a small number of independent particle problems. We calculate these matrix elements, and along the way we introduce the various different forms of potential we use.

7.1.1 Matrix elements

Let the dimensionless one-particle Hamiltonian of the rigid rotator be

$$\mathscr{H} = \frac{1}{2}\mathbf{L}^2. \tag{136}$$

The eigenvalues are

$$E_{l,m} = \frac{1}{2}l(l+1) \tag{137}$$

with $l = 0, 1, \ldots$ and $m = -l, -l+1, \ldots, l$; the corresponding eigenkets are $|lm\rangle$, which in the polar-coordinate representation take the form of spherical harmonics $Y_{lm}(\theta, \phi)$.

We consider interactions that depend on the angle between the axis of a rotor and an external field or the axis of another rotor. As will be explained below, it turns out that the potential energy of the system can be written in terms of components of one-particle spherical tensors of rank one or two, depending on the specific form of the interaction. Using the spherical harmonics as a basis, we represent the potential and calculate the required matrix elements of these tensors.

Recall that given the unit Cartesian vector $\vec{r} = (x, y, z)$ one can construct a spherical tensor of unit rank:

$$T^{(1)}_{-1}(\vec{r}) = \frac{x - iy}{\sqrt{2}} = e^{-i\phi} \sin\theta \tag{138}$$

$$T^{(1)}_{0}(\vec{r}) = z = \cos\theta \tag{139}$$

$$T^{(1)}_{1}(\vec{r}) = -\frac{x + iy}{\sqrt{2}} = -e^{i\phi} \sin\theta. \tag{140}$$

Note that

$$T^{(1)}_{m} = 2\sqrt{\frac{\pi}{3}} Y_{1m}(\theta, \phi). \tag{141}$$

Clearly, apart from an interaction coefficient, the tensor component $T^{(1)}_0$ describes the interaction of a single AB rotor with an external field.

A direct product of tensor components can be contracted to form a rotationally invariant quantity. The contraction process requires Clebsch-Gordan coefficients, as does the calculation of matrix elements of tensor components by means of the Wigner-Eckart theorem. For the Clebsch-Gordan coefficients we use the Mathematica notation. In other words, $C[\{j_1, m_1\}, \{j_2, m_2\}, \{j, m\}]$ is the coefficient of the total angular momentum state $|jm\rangle$ for the direct product of $|j_1 m_1\rangle$ and $|j_2 m_2\rangle$, which coefficient vanishes unless $m = m_1 + m_2$ and total angular momentum quantum number j satisfies the triangle inequality $|j_2 - j_1| \leq j \leq j_1 + j_2$.

The matrix elements $\langle l_1, m_1 | T^{(k)}_m | l_2, m_2 \rangle$ can be computed using the Wigner-Eckart theorem, which states that the reduced matrix element

$$\langle l_1 || T^{(k)}_1 || l_2 \rangle \equiv \sqrt{2l_2+1} \frac{\langle l_1, m_1 | T^{(k)}_m | l_2, m_2 \rangle}{C[\{k,m\},\{l_2,m_2\},\{l_1,m_1\}]} \tag{142}$$

depends only on l_1 and l_2, as indicated.

Recall that properties of the Clebsch-Gordan coefficients can be expressed more easily in terms of Wigner $3j$-symbols [2],[3].

$$C[\{l_1,m_1\},\{l_2,m_2\},\{l_3,m_3\}] = (-1)^{l_1-l_2+m_3}\sqrt{1+2l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} \tag{143}$$

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} j_{1'} & j_{2'} & j_{3'} \\ m_{1'} & m_{2'} & m_{3'} \end{pmatrix} \tag{144}$$

for any even permutation $1'2'3'$ of 123; and

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_{1''} & j_{2''} & j_{3''} \\ m_{1''} & m_{2''} & m_{3''} \end{pmatrix} \tag{145}$$

for any odd permutation $1''2''3''$ of 123.

Wigner $3j$-symbols vanish unless the following two conditions are satisfied: (1) the triangle inequality $|j_1 - j_2| \leq j_3 \leq j_1 + j_2$ –in all possible permutations of the sub-indices– and (2) $m_1 + m_2 + m_3 = 0$.

The extra factor $\sqrt{2l_2+1}$ in the definition of the reduced matrix element is a matter of convention; its inclusion results in restoring some of the symmetry of the reduced matrix elements. For instance one has

$$\frac{C[\{l_1,0\},\{1,0\},\{l_2,0\}]}{\sqrt{2l_2+1}} = -\frac{C[\{l_2,0\},\{1,0\},\{l_1,0\}]}{\sqrt{2l_1+1}}. \tag{146}$$

Also

$$C[\{l,m\},\{l,-m\},\{0,0\}] = (-1)^{l-m}\sqrt{\frac{1}{2l+1}}. \tag{147}$$

The contraction theorem for spherical tensors implies that if S and S' are spherical tensors of rank l, then $\sum_{q=-l}^{l}(-1)^q S_q S'_{-q}$ is a scalar.

In particular, for the tensor given in Eq. (140) one has

$$\sum_{q=-1}^{1}(-1)^q T_q^{(1)}(\vec{r}_1)T_{-q}^{(1)}(\vec{r}_2) = \cos\theta_1\cos\theta_2+\cos(\phi_1-\phi_2)\sin\theta_1\sin\theta_2 = \cos\theta_{12}, \quad (148)$$

where θ_{12} is the angle between $\vec{r}_1$ and $\vec{r}_2$. This result, Eq. (148), expresses the potential energy of two mutually interacting AB particles in terms of terms of tensor components, and it shows that this interaction is separable, i.e., it can be written as a sum of products of single particle operators.

$$T_{-2}^{(2)} = \frac{1}{2}e^{-2i\phi}\sin^2\theta, \quad (149)$$

$$T_{-1}^{(2)} = e^{-i\phi}\cos\theta\sin\theta, \quad (150)$$

$$T_0^{(2)} = \frac{3\cos^2\theta-1}{\sqrt{6}}, \quad (151)$$

$$T_1^{(2)} = -e^{i\phi}\cos\theta\sin\theta, \quad (152)$$

$$T_2^{(2)} = \frac{1}{2}e^{2i\phi}\sin^2\theta. \quad (153)$$

Note that

$$T_m^{(2)} = 2\sqrt{2\frac{\pi}{15}}Y_{2m}(\theta,\phi), \quad (154)$$

while

$$\sqrt{\frac{2}{3}}T_0^{(2)} = \cos^2\theta - \frac{1}{3}. \quad (155)$$

The latter equation shows that the potential energy of an AA rotor interacting with an external field, which is proportional to $\cos^2\theta - 1/3$, can be expressed in a form to which, once again, the Wigner-Eckart theorem can be applied conveniently.

As we did before, we can apply the contraction theorem to construct the scalar

$$\sum_{q=-2}^{2}(-1)^q T_q^{(2)}(\vec{r}_1)T_{-q}^{(2)}(\vec{r}_2) = \cos^2\theta_{12} - \frac{1}{3} \quad (156)$$

Once again, this shows that the interaction potential of two rigid AA rotors is separable, i.e., it can be expressed as a sum of products of one-particle matrix elements.

By applying the Wigner-Eckart theorem, we find that the reduced matrix elements for $T^{(1)}$ are given by

$$\langle l_1||T^{(1)}||l_2\rangle = \begin{cases} \operatorname{sgn}(l_1 - l_2)\sqrt{\max(l_1, l_2)\frac{2l_2+1}{2l_1+1}} & \text{if } |l_1 - l_2| = 1 \\ 0 & \text{otherwise} \end{cases} \tag{157}$$

The reduced matrix elements for $T^{(2)}$ are given by

$$\langle l_1||T^{(2)}||l_2\rangle = \begin{cases} \sqrt{\frac{2}{3}}\sqrt{\frac{(l-1)l(2l+1)}{(2l-3)(2l-1)}} & \text{if } l_1 = l_2 - 2 \\ -\frac{2}{3}\sqrt{\frac{l(l+1)(2l+1)}{(2l-1)(2l+3)}} & \text{if } l_1 = l_2 \\ \sqrt{\frac{2}{3}}\sqrt{\frac{(l-1)l(2l-3)}{4l^2-1}} & \text{if } l_1 = l_2 + 2 \\ 0 & \text{otherwise} \end{cases} \tag{158}$$

where $l = \min(l_1, l_2)$.

We have no rigorous proof that these expressions for the reduced matrix elements are correct. Instead, we used Mathematica to calculate them explicitly and this shows that the expressions are correct for any value of l we tried. We found the expressions by fitting the squares of the reduced matrix elements to Padé approximants in l. In this way, the values of the coefficients can be found for small values of l, and, as it turns out, this yields correct results for much larger values of l.

7.2 Quaternion Based Rigid-Body Diffusion MC Solution

In this section I will describe in detail the rigid-body diffusion Monte Carlo solution to the problems solved exactly in the previous section and present numerical results.

7.2.1 AB Rotor

Consider a rigid rotor in an external potential defined as

$$V(\theta) = P\cos(\theta) = P\hat{r}\cdot\hat{z}, \tag{159}$$

where P is a parameter governing the strength of the interaction. This potential corresponds to a single AA rotor in an external field. Because rotation about the

axis of the rotor has no meaning, the rotor has only two principal axes, $\{\hat{n}_1, \hat{n}_2\}$, which are arbitrarily chosen in the plane perpendicular to the axis defining the orientation of the rotor. Ignoring the translational part, the Hamiltonian can be written in terms of the angles about the principal axes, $\{\varphi_1, \varphi_2\}$ and the interaction angle θ.

$$\mathscr{H} = -\frac{1}{2I}\left(\partial_{\varphi_1}^2 + \partial_{\varphi_2}^2\right) + V(\theta) \tag{160}$$

A suitable basis can be constructed from normalized powers of $\cos(\theta)$.

$$b_i(\theta) = a_i \cos^i(\theta) = a_i(\hat{r}\cdot\hat{z})^i, \qquad a_i = \sqrt{\frac{1+2i}{4\pi}} \tag{161}$$

A simple guiding function is

$$\psi_g(\theta) = \sqrt{\sum_{i=0}^{N} b_i^2(\theta)}. \tag{162}$$

Derivatives with respect to the angles about the principal axes are needed for the local energy and drift. One simple way to find these is to apply an infinitesimal quaternion rotation to the unit vector that defines the orientation of the rotor.

$$\partial_{\varphi_i}(\cos(\theta)) = \partial_{\varphi_i}(\hat{r}\cdot\hat{z}) = \frac{\hat{r}(\varphi_i + d\varphi_i)\cdot\hat{z} - \hat{r}(\varphi_i)\cdot\hat{z}}{d\varphi_i} \tag{163}$$

$$\{0, \hat{r}(\varphi_i + d\varphi_i)\} = \exp\left(\frac{d\varphi_i}{2}\hat{n}_i\right)\{0, \hat{r}(\varphi_i)\}\exp\left(-\frac{d\varphi_i}{2}\hat{n}_i\right) \tag{164}$$

Following the formula for quaternion rotation Eq. (164).

$$\hat{r}(\varphi_i + d\varphi_i) = \hat{r}(\varphi_i) + (\hat{n}_i \times \hat{r}(\varphi_i))\, d\varphi_i + \mathcal{O}\left(d\varphi_i^2\right) \tag{165}$$

$$\partial_{\varphi_i}\cos(\theta) = \partial_{\varphi_i}(\hat{r}\cdot\hat{z}) = (-1)^i \hat{n}_j \cdot \hat{z} \tag{166}$$

And a similar calculation for the second derivative gives

$$\partial_{\varphi_i}^2\cos(\theta) = \partial_{\varphi_i}^2(\hat{r}\cdot\hat{z}) = -(\hat{r}\cdot\hat{z}) = -\cos(\theta) \tag{167}$$

$$\vec{D} = 2\frac{\partial_{\varphi_1}\psi_g(\theta)\hat{\varphi}_1 + \partial_{\varphi_2}\psi_g(\theta)\hat{\varphi}_2}{\psi_g(\theta)} = 2\frac{\frac{1}{\sin(\theta)}\partial_\theta\psi_g(\theta)}{\psi_g(\theta)}\begin{pmatrix} -\hat{n_2}\cdot\hat{z} \\ \hat{n_1}\cdot\hat{z} \end{pmatrix} \tag{168}$$

The local energies are

$$E_{L,b_i}(\theta) = \frac{\mathscr{H}(x)b_i(x)}{b_i(x)} = \frac{-\frac{1}{\sin(\theta)}\partial_\theta\left(\sin(\theta)\partial_\theta b(\theta)\right)}{b(\theta)} + P\cos(\theta) \tag{169}$$

$$E_{L,\psi_g}(\theta) = \frac{\mathscr{H}(x)\psi_g(x)}{\psi_g(x)} = \frac{-\frac{1}{\sin(\theta)}\partial_\theta\left(\sin(\theta)\partial_\theta \psi_g(\theta)\right)}{\psi_g(\theta)} + P\cos(\theta) \tag{170}$$

Calculations were carried out for three basis states, and three values for the interaction parameter. No attempt was made to optimize the guiding function or the basis states. On the contrary, the basis states and guiding function are intentionally poor in order to produce results typical of a more sophisticated problem and to demonstrate the effectiveness of projection. Because the basis states are not optimized, the true ground state energy is well below the confidence interval defined by the variational estimate as shown in fig. (11). This feature allows for the demonstration that the projection in imaginary-time removes this variational bias. As the projection time increases, the statistical error also increases as the result of the fluctuating weights. This effect could have been minimized by proper choice of the guiding function, but is only eliminated in the limit of perfect importance sampling.

Figure (10) shows the energy of the second excited state of the AB rotor as a function of projection time, $\mathcal{E}_2(\tau)$. The Monte Carlo estimate for zero projection time is the variational estimate. This variational estimate can be verified by simple application of the Reyleigh-Ritz variational procedure described in Section 3.6.1. We are looking for the point with a long enough projection time so that the variational bias has been projected out but not so long as to be numerically unstable. The best point is selected based on maximum overlap with the points of greater projection time. Typically, this is the point where the uncertainty begins to increase faster than the projected observable decreases and is chosen based on overlap with points to the right. The exact value is obtained from the results of

the previous section and is shown for comparison.

In order to remove the time step bias, each of the points on the plot of projected energy vs. imaginary-time Fig. (10) is the result of extrapolation to zero $\Delta\tau$ for a number of runs performed with different time-steps. The plot of this extrapolation for the selected interval is shown in fig. (11). The shaded region corresponds to the selected interval. The largest eigenvalue and interaction parameter was chosen for these figures to further exaggerate the variational bias and statistical noise.

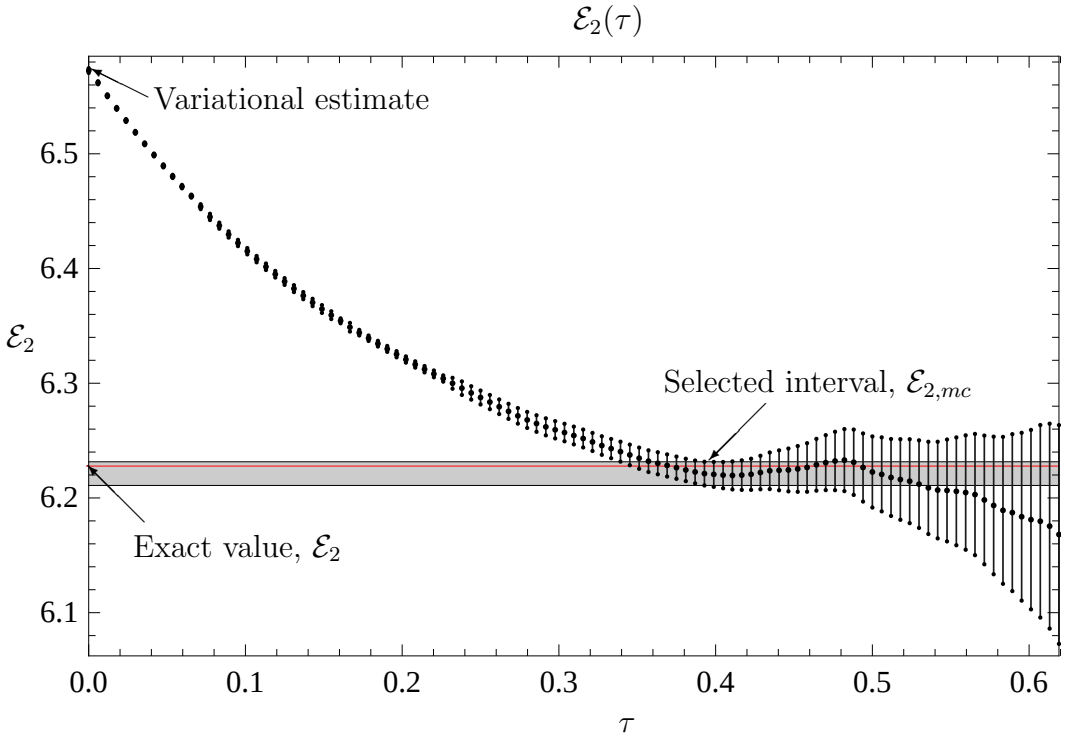


Figure 10: Plot of projected energy for the largest calculated eigenvalue of the AB rotor with $V = 3\cos(\theta)$. Each point is the extrapolation to zero time step of the same point on similar plots of finite time step. $\mathcal{E}_2$ is the exact energy.

Table (1) provides the numerical results of the rigid-body diffusion Monte Carlo calculation of the first three energy levels of an AA rotor.

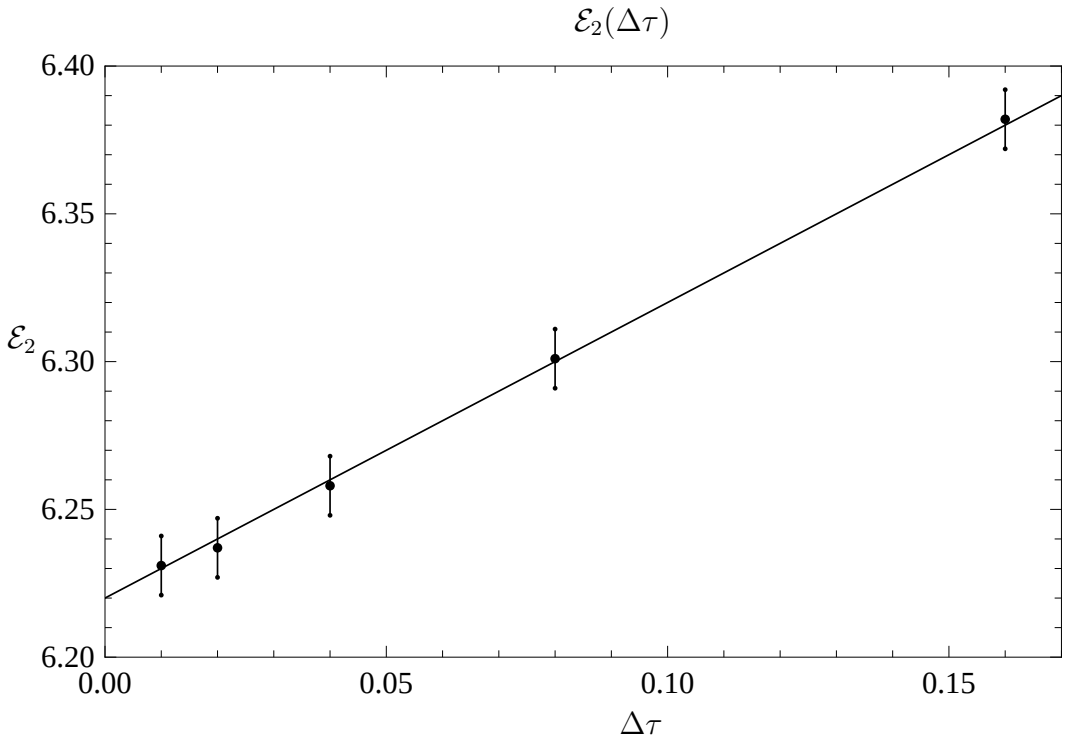


Figure 11: Extrapolation to zero projection time for selected interval in fig. (10).

Table 1: Comparison with exact results for the first three energies in dimensionless units for the AB rigid-rotor in a $P\cos(\theta)$ potential.

P	$\mathcal{E}_{0,\mathrm{mc}}$	$\mathcal{E}_0$	$\mathcal{E}_{1,\mathrm{mc}}$	$\mathcal{E}_1$	$\mathcal{E}_{2,\mathrm{mc}}$	$\mathcal{E}_2$
1	-0.15766(1)	-0.1576634	2.09063(8)	2.09076	6.021(6)	6.02415
2	-0.5573(2)	-0.55728	2.287(1)	2.2871	5.97(5)	6.098
3	-1.0926(2)	1.09267	2.477(1)	2.4779	6.22(1)	6.227

7.2.2 AA Rotor

Calculations were also carried out for rigid rotor in an external potential that cannot tell one particle from the other. In this case the potential is given by

$$V(\theta) = P\cos(2\theta) \tag{171}$$

Table (2) provides the numerical results of the rigid-body diffusion Monte Carlo calculation of the first three energy levels of an AA rotor.

Table 2: Comparison with exact results for the first three energies in dimensionless units for the AA rigid rotor.

P	$\mathcal{E}_{0,\mathrm{mc}}$	$\mathcal{E}_0$	$\mathcal{E}_{1,\mathrm{mc}}$	$\mathcal{E}_1$	$\mathcal{E}_{2,\mathrm{mc}}$	$\mathcal{E}_2$
1	-0.3887(1)	-0.38868	2.174(2)	2.1721	6.086(2)	6.0842
2	-0.8725(4)	-0.87226	2.287(6)	2.2871	6.225(8)	6.2257
3	-1.430(1)	-1.42884	2.35(1)	2.3439	6.41(1)	6.4019

7.2.3 Two rotors

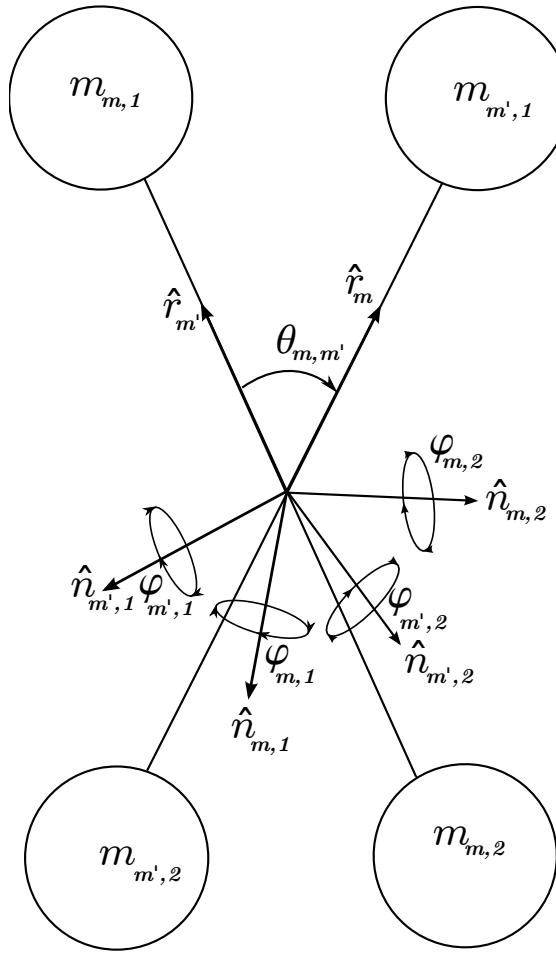


Figure 12: Two mutually interacting rotors

In the next example the same calculations are applied to a system of two rigid rotors experiencing a potential that depends on the angle between them. Each rotor has it's own orientation vector, $\hat{r}_m$, and principal axes, $\hat{n}_{m,i}$ where m

indicates which rotor, and i indicates which principal axis.

$$V(\theta_{m,m'}) = P\cos(\theta_{m,m'}) = \hat{r}_m \cdot \hat{r}_{m'} \tag{172}$$

As in the case of the single rotor, a basis can be constructed from normalized powers of $\cos(\theta_{m,m'})$.

$$b_i(\theta_{m,m'}) = a_i \cos^i(\theta_{m,m'}) = a_i(\hat{r}\cdot\hat{z})^i, \qquad a_i = \frac{\sqrt{1+2i}}{4\pi} \tag{173}$$

The derivation for local energy and drift is almost identical to that of the single rigid rotor even though the physics is very different. As before, derivatives required for the local energies and drift can be found by application of infinitesimal rotation quaternion.

$$\partial_{\varphi_{m,i}} \cos(\theta_{m,m'}) = (-1)^i \hat{n}_{m,j} \cdot \hat{r}_{m'} \tag{174}$$

$$\partial^2_{\varphi_i} \cos(\theta_{m,m'}) = -\cos(\theta_{m,m'}) = -\hat{r}_m \cdot \hat{r}_{m'} \tag{175}$$

The drift is then given by

$$\begin{aligned} \vec{D}_m &= 2\frac{\partial_{\varphi_{1,m}}\psi_g(\theta_{m,m'})\hat{\varphi_{1,m}} + \partial_{\varphi_{2,m}}\psi_g(\theta_{m,m'})\hat{\varphi_{2,m}}}{\psi_g(\theta_{m,m'})} \\ &= 2\frac{\frac{1}{\sin(\theta_{m,m'})}\partial_{\theta_{m,m'}}\psi_g(\theta_{m,m'})}{\psi_g(\theta_{m,m'})} \begin{pmatrix} -\hat{n}_{m,2}\cdot\hat{r}_{m'} \\ \hat{n}_{m,1}\cdot\hat{r}_{m'} \end{pmatrix} \end{aligned} \tag{176}$$

and the local energies by

$$E_{L,\psi}(\theta_{m,m'}) = \frac{-\frac{2}{\sin(\theta_{m,m'})}\partial_{\theta_{m,m'}}\left(\sin(\theta_{m,m'})\partial_{\theta_{m,m'}}\psi(\theta_{m,m'})\right)}{\psi(\theta_{m,m'})} + P\cos(\theta_{m,m'}), \tag{177}$$

7.2.4 Two AB Rotors

Table (3) provides the numerical results of the rigid-body diffusion Monte Carlo calculation of the first three energy levels of two AB rotors interacting via $V(\theta_{m,m'}) = P\cos(\theta_{m,m'})$.

Table 3: First three energies in dimensionless units for the two AB rigid rotors interacting via $V(\theta_{m,m'}) = P\cos(\theta_{m,m'})$.

P	$\mathcal{E}_{0,\mathrm{mc}}$	$\mathcal{E}_{0,\mathrm{var}}$	$\mathcal{E}_{1,\mathrm{mc}}$	$\mathcal{E}_{1,\mathrm{var}}$	$\mathcal{E}_{2,\mathrm{mc}}$	$\mathcal{E}_{2,\mathrm{var}}$
1	-0.08211(1)	-0.0821016	4.0488(1)	4.04874	12.011(1)	12.0119
2	-0.3154(1)	-0.315327	4.1813(2)	4.18152	12.048(1)	12.0484
3	-0.6699(1)	-0.66980	4.6754(3)	4.6752	12.109(3)	12.1108

7.2.5 Two AA Rotors

As before calculations were also carried out for the case of distinguishable and identical particles. Table (4) provides the numerical results of the rigid-body diffusion Monte Carlo calculation of the first three energy levels of two AA rotors interacting via $V(\theta_{m,m'}) = P\cos(2\theta_{m,m'})$.

Table 4: First three energies in dimensionless units for the two AA rigid rotors interacting via $V(\theta_{m,m'}) = P\cos(2\theta_{m,m'})$.

P	$\mathcal{E}_{0,\mathrm{mc}}$	$\mathcal{E}_{0,\mathrm{var}}$	$\mathcal{E}_{1,\mathrm{mc}}$	$\mathcal{E}_{1,\mathrm{var}}$	$\mathcal{E}_{2,\mathrm{mc}}$	$\mathcal{E}_{2,\mathrm{var}}$
1	-0.36201(3)	-0.36198	4.184(2)	4.1861	12.062(2)	12.076
2	-0.7777(2)	-0.77715	4.339(2)	4.34436	12.153(6)	12.206
3	-1.2404(3)	-1.2391	4.67(5)	4.4741	12.271(5)	12.382

7.2.6 Two AB Rotors and a Harmonic Potential

As a check of the energy associated with the center of mass motion the above calculations were carried out with the rotor centers interacting by a harmonic oscillator potential. Although this is a trivial case for the exact calculations because the angular and center of mass coordinates are separable, it is a non-trivial verification of the computer code for which separability is not a simplifying feature.

$$V(\theta_{m,m'}, R_{m,m'}) = P\cos(\theta_{m,m'}) + \frac{R^2_{m,m'}}{2} \tag{178}$$

The Hamiltonian in this case is

$$\mathscr{H} = -\frac{1}{2I}\left(\partial^2_{\varphi_1} + \partial^2_{\varphi_2}\right) - \frac{1}{2\mu}\partial^2_{R_{m,m'}} + V(\theta) \tag{179}$$

where $R_{m,m'}$ is the distance between the centers of mass and μ is the reduced mass for the pair of rotors. Since the motion of the centers of mass is independent of the rotational motion it is easy to see that this interaction simply adds $\mathscr{E}_{HO} = (n + \frac{3}{2})$ to the rotational energy. It is required that the guiding function prevents the random walk from separating the rotors. For this reason the guiding function used is that of the previous section, $\psi_g(\theta_{m,m'})$ eq. (162), multiplied by a simple form factor.

$$\psi_g(\theta_{m,m'}, R_{m,m'}) = \psi_g(\theta_{m,m'}) \exp\left\{-\frac{R_{m,m'}}{2}\right\} \tag{180}$$

Since the $R_{m,m'}$ dependence is not built into the basis states, all calculations are effectively performed in the ground oscillator state $n = 0$. Table (5) provides the numerical results of the rigid-body diffusion Monte Carlo calculation of the first three energy levels of two AA rotors interacting via $V(\theta_{m,m'}) = P\cos(2\theta_{m,m'}) + \frac{R^2_{m,m'}}{2}$.

Table 5: First three energies in dimensionless units for two AB rigid rotors interacting via $V(\theta_{m,m'}) = P\cos(2\theta_{m,m'}) + \frac{R^2_{m,m'}}{2}$.

P	$\mathcal{E}_{0,\mathrm{mc}}$	$\mathcal{E}_{0,\mathrm{var}}$	$\mathcal{E}_{1,\mathrm{mc}}$	$\mathcal{E}_{1,\mathrm{var}}$	$\mathcal{E}_{2,\mathrm{mc}}$	$\mathcal{E}_{2,\mathrm{var}}$
1	1.41788(1)	1.41789	5.5487(1)	5.54874	13.513(1)	13.5119
2	1.1847(1)	1.18467	5.6814(2)	5.68152	13.548(1)	13.5484
3	0.8303(1)	0.83020	6.1755(3)	6.1752	13.609(3)	13.6108

7.2.7 Two AA Rotors and a Harmonic Potential

As before, the calculation was also carried out for the AA rotor case. Table (6) provides the numerical results of the rigid-body diffusion Monte Carlo calculation of the first three energy levels of two AA rotors interacting via $V(\theta_{m,m'}) = P\cos(2\theta_{m,m'}) + \frac{R^2_{m,m'}}{2}$.

Table 6: First three energies in dimensionless units for the two interacting AA rigid rotors interacting via eq. (178).$V(\theta_{m,m'}) = P\cos(2\theta_{m,m'}) + \frac{R^2_{m,m'}}{2}$.

P	$\mathcal{E}_{0,\mathrm{mc}}$	$\mathcal{E}_{0,\mathrm{var}}$	$\mathcal{E}_{1,\mathrm{mc}}$	$\mathcal{E}_{1,\mathrm{var}}$	$\mathcal{E}_{2,\mathrm{mc}}$	$\mathcal{E}_{2,\mathrm{var}}$
1	1.13800(3)	1.13802	5.685(2)	5.6861	13.577(2)	13.576
2	0.7227(2)	0.72285	5.845(2)	5.8443	13.70(6)	13.706
3	0.2611(3)	0.2609	5.99(5)	5.9741	13.88(5)	13.882

7.3 Two Rotors Interacting at Multiple Atomic Sites via Lennard-Jones Potential

In this section I will present the rigid-body diffusion Monte Carlo solution for the case of two rigid rotors interacting at atomic sites via Lennard-Jones potential. An exact solution for this problem is unavailable, however, as the masses increase, the results should approach those given by harmonic approximation. This calculation is presented for the ground state only.

For N rotors with n atoms each interacting via Lennard-Jones potential the Hamiltonian can be written as.

$$\mathscr{H} = \sum_{m=1}^{N}\left\{-\frac{1}{8\mu}\nabla_m^2 - \frac{1}{2\mu l^2}\sum_p \partial^2_{\phi_{m,p}} + \sum_{m'>m;a,a'=1}^{n}\left(\frac{1}{r^{12}_{m_a,m'_{a'}}} - \frac{2}{r^{6}_{m_a,m'_{a'}}}\right)\right\} \quad (181)$$

where a, a' index atoms on molecules m, m',and ∇^2_m is with respect to the center of mass of rotor m and $\mu = 2^{\frac{1}{3}}M\epsilon\sigma/\hbar^2$ is the reduced mass of a single rotor. l is the length of a rotor and is taken to be one in this example. $\phi_{m,p}$ represents the angle about principal axis p on rotor m. $r_{m_a,m'_{a'}}$ is the distance between atom a on molecule m ans atom a on molecule m'

A simple guiding function is required to prevent the rotors from separating and is chosen such that there is no drift at the equilibrium configuration.

$$\psi_g = \exp\left\{-\sqrt{2\mu}\sum_{m,m'>m}\left(\sum_{a,a'}\frac{r^{-5}_{m_a,m'_a}}{5} + 2\sqrt{2}R_{m,m'}\right)\right\} \quad (182)$$

$R_{m,m'}$ is the distance between the center of mass locations of the rotors.

7.3.1 Interparticle Distance Derivatives

Derivatives of the interparticle distances with respect to angles about the principal axes are required to evaluate the expressions for the drift and local energy. In this subsection I present general expressions for these derivatives derived with the help of quaternion algebra.

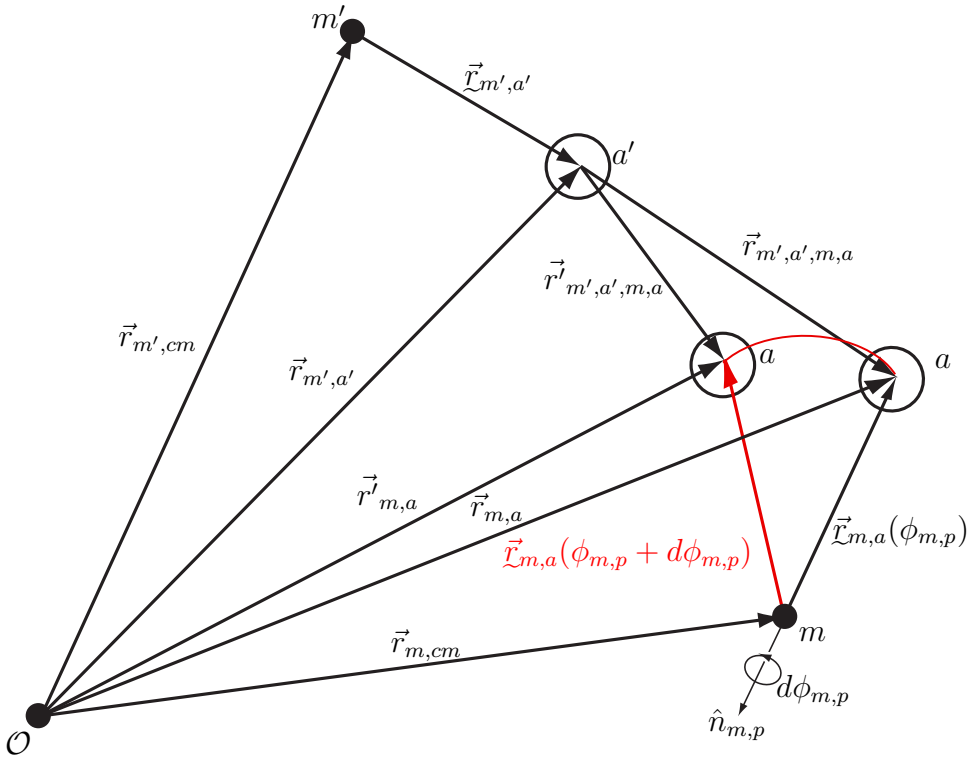


Figure 13: Sketch showing definitions of vectors involved in calculating derivatives of interparticle distances with respect to angles about the principal axes

Figure 7.3.1 shows the geometry associated with the problem of calculating derivatives of the interparticle distances with respect to angles about the principle axes. The interparticle distance vector can be written as

$$\vec{r}_{m,a,m',a'} = \vec{r}_{m,cm} - \vec{r}_{m',cm} + \vec{r}_{m,a} - \vec{r}_{m',a'} \tag{183}$$

The derivative of the square of the interparticle distance with respect to the angle about the principal axis, $\hat{n}_{m,p}$, can be defined as

$$\partial_{\phi_{m,p}}\left|\vec{r}_{m,a,m',a'}\right|^2 = \frac{\left|\vec{r}_{m,a,m',a'}(\phi_{m,p}+d\phi_{m,p})\right|^2 - \left|\vec{r}_{m,a,m',a'}(\phi_{m,p})\right|^2}{d\phi_{m,p}} \tag{184}$$

$\underset{\sim}{\vec{r}}_{m,a}(\phi_{m,p}+d\phi_{m,p})$ can be found by application of an infinitesimal rotation quaternion.

$$e^{\frac{d\phi_{m,p}}{2}\hat{n}_{m,p}} = \left(\cos(\frac{d\phi}{2}), \sin(\frac{d\phi}{2})\right) \approx \left(1, \frac{d\phi}{2}\hat{n}\right) \tag{185}$$

$$\begin{aligned}\underset{\sim}{\vec{r}}_{m,a}(\phi_{m,p}+d\phi_{m,p}) &= \mathrm{Im}\left\{ e^{\frac{d\phi_{m,p}}{2}\hat{n}_{m,p}} \underbrace{\left(0, \underset{\sim}{\vec{r}}_{m,a}\right)}_{\text{pure quaternion}} e^{-\frac{d\phi_{m,p}}{2}\hat{n}_{m,p}} \right\} \\ &= \underset{\sim}{\vec{r}}_{m,a} + \left(\hat{n}_{m,p} \times \underset{\sim}{\vec{r}}_{m,a}\right) d\phi_{m,p} + \mathcal{O}\left(d\phi_{m,p}^2\right)\end{aligned} \tag{186}$$

Ignoring higher order terms, the new interparticle distance vector is given by

$$\vec{r}_{m,a,m',a'}(\phi_{m,p}+d\phi_{m,p}) = \vec{r}_{m,cm} - \vec{r}_{m',cm} + \underset{\sim}{\vec{r}}_{m,a} + \left(\hat{n}_{m,p} \times \underset{\sim}{\vec{r}}_{m,a}\right) d\phi_{m,p} - \underset{\sim}{\vec{r}}_{m',a'} \tag{187}$$

Plugging in the expressions above gives

$$\partial_{\phi_{m,p}}\left|\vec{r}_{m,a,m',a'}\right|^2 = 2\vec{r}_{m,a,m',a'} \cdot \left(\hat{n}_{m,p} \times \underset{\sim}{\vec{r}}_{m,a}\right) \tag{188}$$

$$\partial_{\phi}\left|\vec{r}_{m,a,m',a'}\right| = \hat{r}_{m,a,m',a'} \cdot \left(\hat{n}_{m,p} \times \underset{\sim}{\vec{r}}_{m,a}\right) \tag{189}$$

which we probably should have been able to guess from the start. A similar procedure for the second derivative gives

$$\partial^2_{\phi_{m,p}}\left|\vec{r}_{m,a,m',a'}\right|^2 = 2\left\{\vec{r}_{m,a,m',a'} \cdot \left[\hat{n}_{m,p} \times \left(\hat{n}_{m,p} \times \underset{\sim}{\vec{r}}_{m,a}\right)\right] + \left|\underset{\sim}{\vec{r}}_{m,a}\right|^2\right\} \tag{190}$$

7.3.2 Drift and Local Energy

Now that the derivatives of the interparticle distances are available we can calculate the expressions for the drift and local energy. For the rotational coordinates

the drift is

$$\vec{D}_{\text{rot}} = \frac{2}{\psi_g}\sum_{m',p}\frac{1}{2\mu}\left(\partial_{\alpha_{m',p}}\psi_g\right)\hat{\alpha}_{m',p}$$
$$= \sqrt{\frac{2}{\mu}}\sum_{m',p}\left\{\sum_{m\neq m',a,a'} r^{-6}_{m_a,m',a'}\left[\hat{r}_{m,a,m',a'}\cdot\left(\hat{n}_{m',p}\times\underset{\sim}{\vec{r}}_{m',a'}\right)\right]\right\}\hat{\alpha}_{m',p} \quad (191)$$

and for the translational coordinates

$$\vec{D}_{\text{trans},m'} = \frac{2}{\psi_g}\sum_{m'}\frac{1}{8\mu}\vec{\nabla}_{m'}\psi_g$$
$$= \frac{1}{2\sqrt{2\mu}}\sum_{m\neq m'}\left[\sum_{a,a'} r^{-6}_{m,a,m',a'}\hat{r}_{m,a,m',a'} - 2\sqrt{2}\hat{R}_{m,m'}\right] \quad (192)$$

The rotational part of local energy is.

$$E_{L,\text{rot}} = -\frac{\sum_{m',p}\left(\frac{1}{2\mu}\right)\left(\partial^2_{\alpha_{m',p}}\psi_g\right)}{\psi_g}$$
$$= \frac{1}{\sqrt{2\mu}}\sum_{m',p}\sum_{m\neq m',a,a'} r^{-7}_{m,a,m',a'}\left[7\left(\partial_{m',p}r_{m_a,m',a'}\right)^2 - \frac{\left(\partial^2_{m',p}r^2_{m,a,m',a'}\right)}{2}\right] - \frac{2\mu D^2_{rot}}{2}$$

and the translational part

$$E_{L,\text{trans}} = -\frac{\sum_{m'}\left(\frac{1}{8\mu}\right)\left(\nabla^2_{m'}\psi_g\right)}{\psi_g}$$
$$= \frac{1}{\sqrt{2\mu}}\sum_{m',m\neq m'}\left\{\sum_{a,a'} r^{-7}_{m,a,m',a'} + \frac{P}{2R_{m,m'}}\right\} - 2\mu D^2_{trans} \quad (193)$$

$$E_L = E_{L,trans} + E_{L,rot} + V \quad (194)$$

Calculations were performed for multiple masses and time steps. For each mass value results were extrapolated to zero time step.

Figure (14) shows the plot of ground sate energy vs. projection time for the smallest mass and time step. As stated in Section 7.2.1, the best interval is selected as the interval that has maximum overlap with points to the right and is typically near the point where the uncertainty starts increasing faster than the exponential decay.

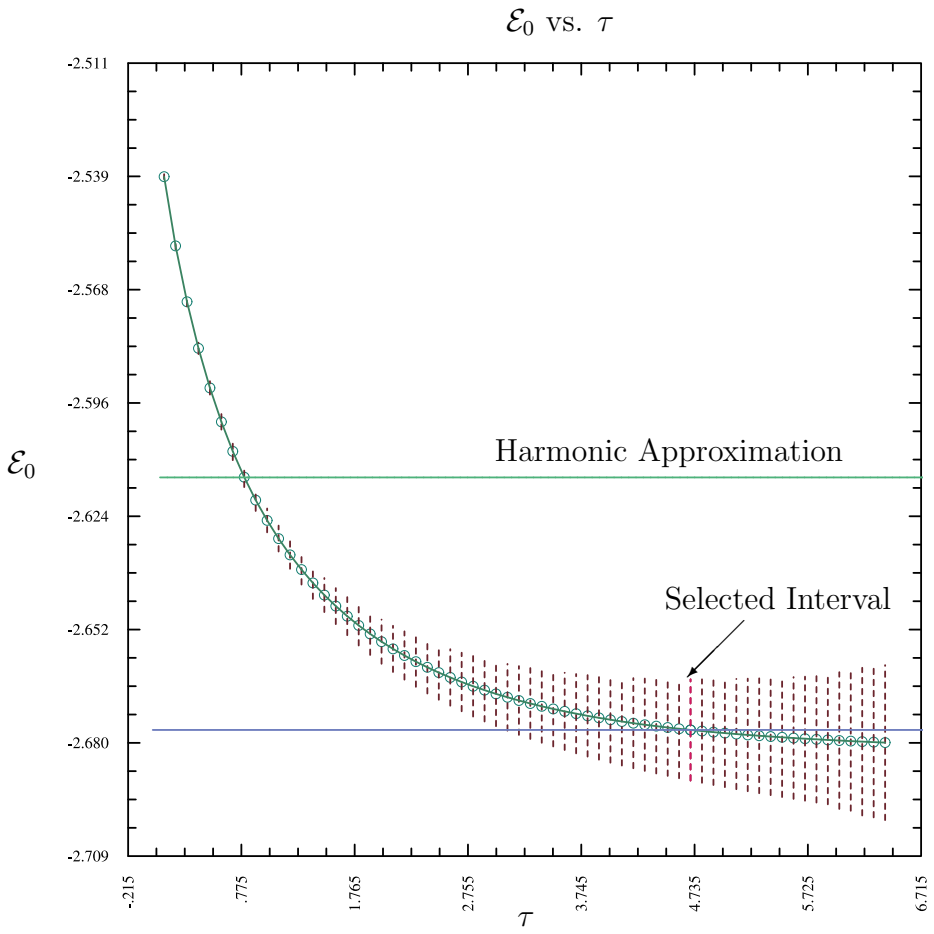


Figure 14: $\mathcal{E}_0$ vs. τ for a pair of rotors interacting via Lennard-Jones potential. The parameters for this calculation are: $\Delta\tau = 0.1$, $\mu = 75$.

Figure (15) shows a plot of $-\left|\mathcal{E}_0\right|^{\frac{1}{2}}$ vs. $\mu^{-\frac{1}{2}}$. Each point on this plot is an extrapolation to zero time step of a number of intervals derived from plots like fig. (14). This plot is becomes linear and approaches the harmonic approximation in the limit of large mass.

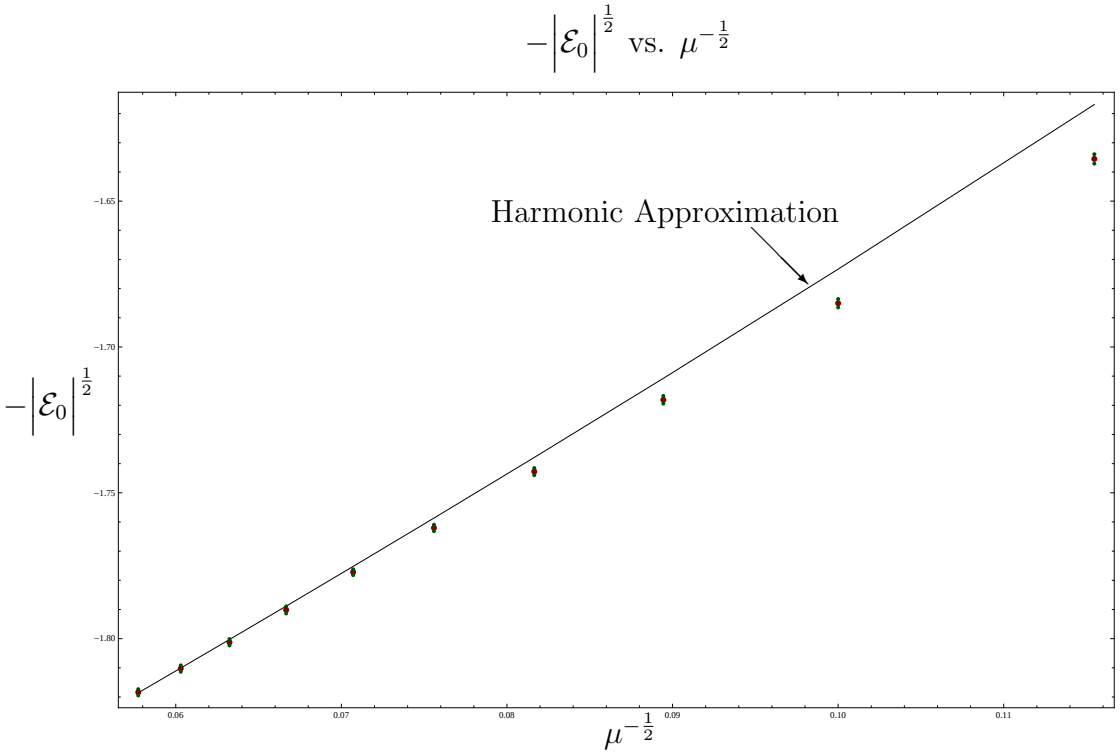


Figure 15: $-\left|\mathcal{E}_0\right|^{\frac{1}{2}}$ vs. $\mu^{-\frac{1}{2}}$ for a pair of rotors interacting via Lennard-Jones potential. The line represents the harmonic approximation

List of References

[1] R. Feynman, *Statistical Mechanics: a set of lectures*, ser. Advanced book classics. Perseus Books, 1998.

[2] Wikipedia, "Clebsch-gordon coefficients — Wikipedia, the free encyclopedia," 2010. [Online]. Available: http://en.wikipedia.org/wiki/Clebsch-Gordan_coefficients#Symmetry_properties

[3] Wikipedia, "3-jm symbol — Wikipedia, the free encyclopedia," 2010. [Online]. Available: http://en.wikipedia.org/wiki/3-jm_symbol

CHAPTER 8

Conclusion

This work is intended to contribute to the usefulness of the rigid-body diffusion Monte Carlo method and the scope end of problems to which it can be applied. I have provided an algorithm for the implementation correlation function Monte Carlo within the rigid-body approximation that incorporates trial states constructed from a basis of primary invariants and a new algorithm for implementing rotational Brownian motion that takes full advantage of quaternion algebra. This contribution allows for the Monte Carlo calculation of excited state properties within the rigid-body approximation in a way that was not previously available. In addition, example problems are given that compare results with exact and approximate solutions. Results for example problems clearly demonstrate the correctness of the method.

A brief introduction to the relevant Monte Carlo methods is presented. In order to sample an unknown ground state, I defined a stochastic process which uses the imaginary-time evolution operator to generate samples from an unknown ground state. This stochastic process involves the implementation of rotational Brownian motion. The use of quaternion algebra for implementing this rotational motion in the context of rigid-body diffusion Monte Carlo is described, and compared to matrix methods. The main advantage of quaternions in this application is the ability to guarantee pure rotation by simple normalization. This provides a mechanism where new configurations can be generated from an initial configuration by application of a numerically stable cumulative rotation product and eliminates the need to compute the principal axes from the moment of inertia tensor every step.

Correlation function Monte Carlo is incorporated for the calculation of excited state properties. This method is an extension of the Rayleigh-Ritz variational method where the matrix elements are evaluated by Monte Carlo methods. The construction of trial states and guiding functions from a basis of primary invariants is discussed. It is shown that the rigid-body approximation greatly reduces the number of terms required for the construction this basis.

It is my hope that these tools are simple and powerful enough to be accessible and useful for many applications.

BIBLIOGRAPHY

Benoit, D. M. and Clary, D. C., "Quaternion formulation of diffusion quantum monte carlo for the rotation of rigid molecules in clusters," *J. Chem. Phys.*, vol. 113(13), pp. 5193–5202, 2000.

Buch, V., "Treatment of rigid bodies by diffusion Monte Carlo: Application to the para-$H_2 \ldots H_2O$ and ortho-$H_2 \ldots$-H_2O clusters," *J. Chem. Phys.*, vol. 97, p. 726, 1992.

Ceperley, D. M. and Bernu, B., "The calculation of excited state properties with quantum Monte Carlo," *J. Chem. Phys.*, vol. 89, p. 6316, 1988.

Demmel, J., Dongarra, J., Ruhe, A., and van der Vorst, H., *Templates for the solution of algebraic eigenvalue problems: a practical guide*, Bai, Z., Ed. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2000.

Feynman, R., *Statistical Mechanics: a set of lectures*, ser. Advanced book classics. Perseus Books, 1998.

Hastings, W., "Monte carlo samping methods using markov chains and their applications," *Biometrika*, pp. 97–109, 1970.

MacDonald, J. K. L., "Successive approximations by the rayleigh-ritz variation method," *Phys. Rev.*, vol. 43, p. 830, 1933.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E., "Equation of state calculations by fast computing machines," *J. Chem. Phys.*, vol. 21, p. 1087, 1953.

Metropolis, N. and Ulam, S., "The Monte Carlo method," *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, Sept. 1949.

Mushinski, A. and Nightingale, M. P., "Many-body trial wave functions for atomic systems and ground states of small noble gas clusters," *J. Chem. Phys.*, vol. 101, p. 8831, 1994.

Nightingale, M. P. and Melik-Alaverdian, V., "Optimization of ground and excited state wavefunctions and van der Waals clusters," *Phys. Rev. Lett.*, vol. 87, p. 043401, 2001.

Nightingale, M. P. and Umrigar, C. J., *Monte Carlo Eigenvalue Methods in Quantum Mechanics and Statistical Mechanics*, ser. Advances in Chemical Physics. New York: John Wiley & Sons, 1999, vol. 105, p. 65.

Schrodinger, E., "Uber die umkehrung der naturgesetze," *Sitzber. Preuss. Akad. Wiss. Phys.-math. Kl*, pp. 144–153, 1931.

Sturmfels, B., *Algorithms in invariant theory; 2nd ed.*, ser. Texts and Monographs in Symbolic Computation. Dordrecht: Springer, 2007.

Viel, A., Patela, M. V., Niyaza, P., and Whaley, K. B., "Importance sampling in rigid body diffusion monte carlo," *Computer Physics Communications*, vol. 145, pp. 24–47, 2000.

Wikipedia, "3-jm symbol — Wikipedia, the free encyclopedia," 2010. [Online]. Available: http://en.wikipedia.org/wiki/3-jm_symbol

Wikipedia, "Clebsch-gordon coefficients — Wikipedia, the free encyclopedia," 2010. [Online]. Available: http://en.wikipedia.org/wiki/Clebsch-Gordan_coefficients#Symmetry_properties