

Chapter 11

Quantum Monte Carlo methods

If, in some cataclysm, all scientific knowledge were to be destroyed, and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words? I believe it is the atomic hypothesis (or atomic fact, or whatever you wish to call it) that all things are made of atoms, little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another. In that one sentence you will see an enormous amount of information about the world, if just a little imagination and thinking are applied.

Richard Feynman, The Laws of Thermodynamics.

11.1 Introduction

The aim of this chapter is to present examples of applications of Monte Carlo methods in studies of simple quantum mechanical systems. We study systems such as the harmonic oscillator, the hydrogen atom, the hydrogen molecule, the helium atom and more complicated atoms. Systems with many interacting fermions and bosons such as liquid ^4He and Bose Einstein condensation of atoms are discussed in chapter 18. Most quantum mechanical problems of interest in for example atomic, molecular, nuclear and solid state physics consist of a large number of interacting electrons and ions or nucleons. The total number of particles N is usually sufficiently large that an exact solution cannot be found. In quantum mechanics we can express the expectation value for a given \hat{O} operator for a system of N particles as

$$\langle \hat{O} \rangle = \frac{\int d\mathbf{R}_1 d\mathbf{R}_2 \dots d\mathbf{R}_N \Psi^*(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \hat{O}(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \Psi(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N)}{\int d\mathbf{R}_1 d\mathbf{R}_2 \dots d\mathbf{R}_N \Psi^*(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \Psi(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N)}, \quad (11.1)$$

where $\Psi(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N)$ is the wave function describing a many-body system. Although we have omitted the time dependence in this equation, it is in general intractable. As an example from the nuclear many-body problem, we can write Schrödinger's equation as a differential equation with the energy operator \hat{H} (the so-called energy Hamiltonian) acting on the wave function as

$$\hat{H}\Psi(\mathbf{r}_1, \dots, \mathbf{r}_A, \alpha_1, \dots, \alpha_A) = E\Psi(\mathbf{r}_1, \dots, \mathbf{r}_A, \alpha_1, \dots, \alpha_A)$$

where

$$\mathbf{r}_1, \dots, \mathbf{r}_A,$$

are the coordinates and

$$\alpha_1, \dots, \alpha_A,$$

are sets of relevant quantum numbers such as spin and isospin for a system of A nucleons ($A = N + Z$, N being the number of neutrons and Z the number of protons). There are

$$2^A \times \binom{A}{Z}$$

coupled second-order differential equations in $3A$ dimensions. For a nucleus like ^{10}Be this number is 215040. This is a truly challenging many-body problem.

Eq. (11.1) is a multidimensional integral. As such, Monte Carlo methods are ideal for obtaining expectation values of quantum mechanical operators. Our problem is that we do not know the exact wavefunction $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_A, \alpha_1, \dots, \alpha_N)$. We can circumvent this problem by introducing a function which depends on selected variational parameters. This function should capture essential features of the system under consideration. With such a trial wave function we can then attempt to perform a variational calculation of various observables, using Monte Carlo methods for solving Eq. (11.1).

The present chapter aims therefore at giving you an overview of the variational Monte Carlo approach to quantum mechanics. We limit the attention to the simple Metropolis algorithm, without the inclusion of importance sampling. Importance sampling and diffusion Monte Carlo methods are discussed in chapters 18 and 16.

However, before we proceed we need to recapitulate some of the postulates of quantum mechanics. This is done in the next section. The remaining sections deal with mathematical and computational aspects of the variational Monte Carlo methods, with applications from atomic and molecular physics.

11.2 Postulates of Quantum Mechanics

11.2.1 Mathematical Properties of the Wave Functions

Schrödinger's equation for a one-dimensional onebody problem reads

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(x,t) + V(x,t)\Psi(x,t) = i\hbar\frac{\partial\Psi(x,t)}{\partial t}, \quad (11.2)$$

where $V(x,t)$ is a potential acting on the particle. The first term is the kinetic energy. The solution to this partial differential equation is the wave function $\Psi(x,t)$. The wave function itself is not an observable (or physical quantity) but it serves to define the quantum mechanical probability, which in turn can be used to compute expectation values of selected operators, such as the kinetic energy or the total energy itself. The quantum mechanical probability $P(x,t)dx$ is defined as¹

$$P(x,t)dx = \Psi(x,t)^*\Psi(x,t)dx, \quad (11.3)$$

representing the probability of finding the system in a region between x and $x + dx$. It is, as opposed to the wave function, always real, which can be seen from the following definition of the wave function, which has real and imaginary parts,

$$\Psi(x,t) = R(x,t) + iI(x,t), \quad (11.4)$$

yielding

$$\Psi(x,t)^*\Psi(x,t) = (R - iI)(R + iI) = R^2 + I^2. \quad (11.5)$$

¹This is Max Born's postulate on how to interpret the wave function resulting from the solution of Schrödinger's equation. It is also the commonly accepted and operational interpretation.

The variational Monte Carlo approach uses actually this definition of the probability, allowing us thereby to deal with real quantities only. As a small digression, if we perform a rotation of time into the complex plane, using $\tau = it/\hbar$, the time-dependent Schrödinger equation becomes

$$\frac{\partial \Psi(x, \tau)}{\partial \tau} = \frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, \tau)}{\partial x^2} - V(x, \tau) \Psi(x, \tau). \quad (11.6)$$

With $V = 0$ we have a diffusion equation in complex time with diffusion constant

$$D = \frac{\hbar^2}{2m}.$$

This is the starting point for the Diffusion Monte Carlo method discussed in chapter 18. In that case it is the wave function itself, given by the distribution of random walkers, that defines the probability. The latter leads to conceptual problems when we have anti-symmetric wave functions, as is the case for particles with the spin being a multiple of $1/2$. Examples of such particles are various leptons such as electrons, muons and various neutrinos, baryons like protons and neutrons and quarks such as the up and down quarks.

The Born interpretation constrains the wave function to belong to the class of functions in L^2 . Some of the selected conditions which Ψ has to satisfy are

1. Normalization

$$\int_{-\infty}^{\infty} P(x, t) dx = \int_{-\infty}^{\infty} \Psi(x, t)^* \Psi(x, t) dx = 1 \quad (11.7)$$

meaning that

$$\int_{-\infty}^{\infty} \Psi(x, t)^* \Psi(x, t) dx < \infty \quad (11.8)$$

2. $\Psi(x, t)$ and $\partial \Psi(x, t) / \partial x$ must be finite
3. $\Psi(x, t)$ and $\partial \Psi(x, t) / \partial x$ must be continuous.
4. $\Psi(x, t)$ and $\partial \Psi(x, t) / \partial x$ must be single valued

11.2.2 Important Postulates

We list here some of the postulates that we will use in our discussion.

Postulate I

Any physical quantity $A(\mathbf{r}, \mathbf{p})$ which depends on position \mathbf{r} and momentum \mathbf{p} has a corresponding quantum mechanical operator by replacing $\mathbf{p} \rightarrow i\hbar \nabla$, yielding the quantum mechanical operator

$$\hat{\mathbf{A}} = A(\mathbf{r}, -i\hbar \nabla).$$

Quantity	Classical definition	QM operator
Position	\mathbf{r}	$\hat{\mathbf{r}} = \mathbf{r}$
Momentum	\mathbf{p}	$\hat{\mathbf{p}} = -i\hbar \nabla$
Orbital momentum	$\mathbf{L} = \mathbf{r} \times \mathbf{p}$	$\hat{\mathbf{L}} = \mathbf{r} \times (-i\hbar \nabla)$
Kinetic energy	$T = (\mathbf{p})^2 / 2m$	$\hat{\mathbf{T}} = -(\hbar^2 / 2m)(\nabla)^2$
Total energy	$H = (\mathbf{p}^2 / 2m) + V(\mathbf{r})$	$\hat{\mathbf{H}} = -(\hbar^2 / 2m)(\nabla)^2 + V(\mathbf{r})$

Postulate II

The only possible outcome of an ideal measurement of the physical quantity A are the eigenvalues of the corresponding quantum mechanical operator $\hat{\mathbf{A}}$.

$$\hat{\mathbf{A}}\psi_\nu = a_\nu\psi_\nu,$$

resulting in the eigenvalues a_1, a_2, a_3, \dots as the only outcomes of a measurement. The corresponding eigenstates $\psi_1, \psi_2, \psi_3 \dots$ contain all relevant information about the system.

Postulate III

Assume Φ is a linear combination of the eigenfunctions ψ_ν for $\hat{\mathbf{A}}$,

$$\Phi = c_1\psi_1 + c_2\psi_2 + \dots = \sum_\nu c_\nu\psi_\nu.$$

The eigenfunctions are orthogonal and we get

$$c_\nu = \int (\Phi)^*\psi_\nu d\tau.$$

From this we can formulate the third postulate:

When the eigenfunction is Φ , the probability of obtaining the value a_ν as the outcome of a measurement of the physical quantity A is given by $|c_\nu|^2$ and ψ_ν is an eigenfunction of $\hat{\mathbf{A}}$ with eigenvalue a_ν .

As a consequence one can show that:
when a quantal system is in the state Φ , the mean value or expectation value of a physical quantity $A(\mathbf{r}, \mathbf{p})$ is given by

$$\langle A \rangle = \int (\Phi)^*\hat{\mathbf{A}}(\mathbf{r}, -i\hbar\nabla)\Phi d\tau.$$

We have assumed that Φ has been normalized, viz., $\int (\Phi)^*\Phi d\tau = 1$. Else

$$\langle A \rangle = \frac{\int (\Phi)^*\hat{\mathbf{A}}\Phi d\tau}{\int (\Phi)^*\Phi d\tau}.$$

Postulate IV

The time development of a quantal system is given by

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{\mathbf{H}}\Psi,$$

with $\hat{\mathbf{H}}$ the quantal Hamiltonian operator for the system.

11.3 First Encounter with the Variational Monte Carlo Method

The required Monte Carlo techniques for variational Monte Carlo are conceptually simple, but the practical application may turn out to be rather tedious and complex, relying on a good starting point for the variational wave functions. These wave functions should include as much as possible of the inherent physics to the problem, since they form the starting point for a variational calculation of the expectation value of the hamiltonian H . Given a hamiltonian H and a trial wave function Ψ_T , the variational principle states that the expectation value of $\langle H \rangle$, defined through Postulate III

$$\langle H \rangle = \frac{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) H(\mathbf{R}) \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R})}, \quad (11.9)$$

is an upper bound to the ground state energy E_0 of the hamiltonian H , that is

$$E_0 \leq \langle H \rangle. \quad (11.10)$$

To show this, we note first that the trial wave function can be expanded in the eigenstates of the hamiltonian since they form a complete set, see again Postulate III,

$$\Psi_T(\mathbf{R}) = \sum_i a_i \Psi_i(\mathbf{R}), \quad (11.11)$$

and assuming the set of eigenfunctions to be normalized, insertion of the latter equation in Eq. (11.9) results in

$$\langle H \rangle = \frac{\sum_{mn} a_m^* a_n \int d\mathbf{R} \Psi_m^*(\mathbf{R}) H(\mathbf{R}) \Psi_n(\mathbf{R})}{\sum_{mn} a_m^* a_n \int d\mathbf{R} \Psi_m^*(\mathbf{R}) \Psi_n(\mathbf{R})} = \frac{\sum_{mn} a_m^* a_n \int d\mathbf{R} \Psi_m^*(\mathbf{R}) E_n(\mathbf{R}) \Psi_n(\mathbf{R})}{\sum_n a_n^2}, \quad (11.12)$$

which can be rewritten as

$$\frac{\sum_n a_n^2 E_n}{\sum_n a_n^2} \geq E_0. \quad (11.13)$$

In general, the integrals involved in the calculation of various expectation values are multi-dimensional ones. Traditional integration methods such as the Gauss-Legendre will not be adequate for say the computation of the energy of a many-body system. The fact that we need to sample over a multi-dimensional density and that the probability density is to be normalized by the division of the norm of the wave function, suggests that e.g., the Metropolis algorithm may be appropriate.

We could briefly summarize the above variational procedure in the following three steps.

1. Construct first a trial wave function $\psi_T(\mathbf{R}; \alpha)$, for say a many-body system consisting of N particles located at positions $\mathbf{R} = (\mathbf{R}_1, \dots, \mathbf{R}_N)$. The trial wave function depends on α variational parameters $\alpha = (\alpha_1, \dots, \alpha_N)$.
2. Then we evaluate the expectation value of the hamiltonian H

$$\langle H \rangle = \frac{\int d\mathbf{R} \Psi_T^*(\mathbf{R}; \alpha) H(\mathbf{R}) \Psi_T(\mathbf{R}; \alpha)}{\int d\mathbf{R} \Psi_T^*(\mathbf{R}; \alpha) \Psi_T(\mathbf{R}; \alpha)}.$$

3. Thereafter we vary α according to some minimization algorithm and return to the first step.

The above loop stops when we reach the minimum of the energy according to some specified criterion. In most cases, a wave function has only small values in large parts of configuration space, and a straightforward procedure which uses homogeneously distributed random points in configuration space will most likely lead to poor results. This may suggest that some kind of importance sampling combined with e.g., the Metropolis algorithm may be a more efficient way of obtaining the ground state energy. The hope is then that those regions of configurations space where the wave function assumes appreciable values are sampled more efficiently.

The tedious part in a variational Monte Carlo calculation is the search for the variational minimum. A good knowledge of the system is required in order to carry out reasonable variational Monte Carlo calculations. This is not always the case, and often variational Monte Carlo calculations serve rather as the starting point for so-called diffusion Monte Carlo calculations. Diffusion Monte Carlo is a way of solving exactly the many-body Schrödinger equation by means of a stochastic procedure. A good guess on the binding energy and its wave function is however necessary. A carefully performed variational Monte Carlo calculation can aid in this context. Diffusion Monte Carlo is discussed in depth in chapter 18.

11.4 Variational Monte Carlo for quantum mechanical systems

The variational quantum Monte Carlo has been widely applied to studies of quantal systems. Here we expose its philosophy and present applications and critical discussions.

The recipe, as discussed in chapter 8 as well, consists in choosing a trial wave function $\psi_T(\mathbf{R})$ which we assume to be as realistic as possible. The variable \mathbf{R} stands for the spatial coordinates, in total $3N$ if we have N particles present. The trial wave function serves then, following closely the discussion on importance sampling in section 8.4, as a mean to define the quantal probability distribution

$$P(\mathbf{R}; \alpha) = \frac{|\psi_T(\mathbf{R}; \alpha)|^2}{\int |\psi_T(\mathbf{R}; \alpha)|^2 d\mathbf{R}; \alpha}. \quad (11.14)$$

This is our new probability distribution function (PDF).

The expectation value of the energy Hamiltonian is given by

$$\langle \hat{\mathbf{H}} \rangle = \frac{\int d\mathbf{R} \Psi^*(\mathbf{R}) H(\mathbf{R}) \Psi(\mathbf{R})}{\int d\mathbf{R} \Psi^*(\mathbf{R}) \Psi(\mathbf{R})}, \quad (11.15)$$

where Ψ is the exact eigenfunction. Using our trial wave function we define a new operator, the so-called local energy,

$$\hat{\mathbf{E}}_L(\mathbf{R}; \alpha) = \frac{1}{\psi_T(\mathbf{R}; \alpha)} \hat{\mathbf{H}} \psi_T(\mathbf{R}; \alpha), \quad (11.16)$$

which, together with our trial PDF allows us to compute the expectation value of the local energy

$$\langle E_L(\alpha) \rangle = \int P(\mathbf{R}; \alpha) \hat{\mathbf{E}}_L(\mathbf{R}; \alpha) d\mathbf{R}. \quad (11.17)$$

This equation expresses the variational Monte Carlo approach. We compute this integral for a set of values of α and possible trial wave functions and search for the minimum of the function $E_L(\alpha)$. If the trial wave function is close to the exact wave function, then $\langle E_L(\alpha) \rangle$ should approach $\langle \hat{\mathbf{H}} \rangle$. Eq. (11.17) is solved using techniques from Monte Carlo integration, see the discussion below. For most hamiltonians,

H is a sum of kinetic energy, involving a second derivative, and a momentum independent and spatial dependent potential. The contribution from the potential term is hence just the numerical value of the potential. A typical Hamiltonian reads thus

$$\hat{\mathbf{H}} = -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2 + \sum_{i=1}^N V_{\text{onebody}}(\mathbf{r}_i) + \sum_{i<j}^N V_{\text{int}}(|\mathbf{r}_i - \mathbf{r}_j|). \quad (11.18)$$

where the sum runs over all particles N . We have included both a onebody potential $V_{\text{onebody}}(\mathbf{r}_i)$ which acts on one particle at the time and a twobody interaction $V_{\text{int}}(|\mathbf{r}_i - \mathbf{r}_j|)$ which acts between two particles at the time. We can obviously extend this to more complicated three-body and/or many-body forces as well. The main contributions to the energy of physical systems is largely dominated by one- and two-body forces. We will therefore limit our attention to such interactions only.

Our local energy operator becomes then

$$\hat{\mathbf{E}}_L(\mathbf{R}; \alpha) = \frac{1}{\psi_T(\mathbf{R}; \alpha)} \left(-\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2 + \sum_{i=1}^N V_{\text{onebody}}(\mathbf{r}_i) + \sum_{i<j}^N V_{\text{int}}(|\mathbf{r}_i - \mathbf{r}_j|) \right) \psi_T(\mathbf{R}; \alpha), \quad (11.19)$$

resulting in

$$\hat{\mathbf{E}}_L(\mathbf{R}; \alpha) = \frac{1}{\psi_T(\mathbf{R}; \alpha)} \left(-\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2 \right) \psi_T(\mathbf{R}; \alpha) + \sum_{i=1}^N V_{\text{onebody}}(\mathbf{r}_i) + \sum_{i<j}^N V_{\text{int}}(|\mathbf{r}_i - \mathbf{r}_j|), \quad (11.20)$$

The numerically time-consuming part in the variational Monte Carlo calculation is the evaluation of the kinetic energy term. The potential energy, as long as it has a simple r -dependence adds only a simple term to the local energy operator.

In our discussion below, we base our numerical Monte Carlo solution on the Metropolis algorithm. The implementation is rather similar to the one discussed in connection with the Ising model, the main difference residing in the form of the PDF. The main test to be performed is a ratio of probabilities. Suppose we are attempting to move from position \mathbf{R} to \mathbf{R}' . Then we perform the following two tests.

1. If

$$\frac{P(\mathbf{R}'; \alpha)}{P(\mathbf{R}; \alpha)} > 1,$$

where \mathbf{R}' is the new position, the new step is accepted, or

2.

$$r \leq \frac{P(\mathbf{R}'; \alpha)}{P(\mathbf{R}; \alpha)},$$

where r is random number generated with uniform PDF such that $r \in [0, 1]$, the step is also accepted.

In the Ising model we were flipping one spin at the time. Here we change the position of say a given particle to a trial position \mathbf{R}' , and then evaluate the ratio between two probabilities. We note again that we do not need to evaluate the norm² $\int |\psi_T(\mathbf{R}; \alpha)|^2 d\mathbf{R}$ (an in general impossible task), since we are only computing ratios.

²This corresponds to the partition function Z in statistical physics.

When writing a variational Monte Carlo program, one should always prepare in advance the required formulae for the local energy E_L in Eq. (11.17) and the wave function needed in order to compute the ratios of probabilities in the Metropolis algorithm. These two functions are almost called as often as a random number generator, and care should therefore be exercised in order to prepare an efficient code.

If we now focus on the Metropolis algorithm and the Monte Carlo evaluation of Eq. (11.17), a more detailed algorithm is as follows

- Initialisation: Fix the number of Monte Carlo steps and thermalization steps. Choose an initial \mathbf{R} and variational parameters α and calculate $|\psi_T(\mathbf{R}; \alpha)|^2$. Define also the value of the stepsize to be used when moving from one value of \mathbf{R} to a new one.
- Initialise the energy and the variance.
- Start the Monte Carlo calculation with a loop over a given number of Monte Carlo cycles

1. Calculate a trial position $\mathbf{R}_p = \mathbf{R} + r * step$ where r is a random variable $r \in [0, 1]$.
2. Use then the Metropolis algorithm to accept or reject this move by calculating the ratio

$$w = P(\mathbf{R}_p)/P(\mathbf{R}).$$

If $w \geq s$, where s is a random number $s \in [0, 1]$, the new position is accepted, else we stay at the same place.

3. If the step is accepted, then we set $\mathbf{R} = \mathbf{R}_p$.
 4. Update the local energy and the variance.
- When the Monte Carlo sampling is finished, we calculate the mean energy and the standard deviation. Finally, we may print our results to a specified file.

Note well that the way we choose the next step $\mathbf{R}_p = \mathbf{R} + r * step$ is not determined by the wave function. The wave function enters only the determination of the ratio of probabilities, similar to the way we simulated systems in statistical physics. This means in turn that our sampling of points may not be very efficient. We will return to an efficient sampling of integration points in our discussion of diffusion Monte Carlo in chapter 18. This leads to the concept of importance sampling. As such, we limit ourselves in this chapter to the simplest possible form of the Metropolis algorithm, and relegate both importance sampling and advanced optimization techniques to chapter 18.

The best way however to understand the above algorithm and a specific method is to study selected examples.

11.4.1 First illustration of variational Monte Carlo methods

The harmonic oscillator in one dimension lends itself nicely for illustrative purposes. The hamiltonian is

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} kx^2, \quad (11.21)$$

where m is the mass of the particle and k is the force constant, e.g., the spring tension for a classical oscillator. In this example we will make life simple and choose $m = \hbar = k = 1$. We can rewrite the above equation as

$$H = -\frac{d^2}{dx^2} + x^2, \quad (11.22)$$

The energy of the ground state is then $E_0 = 1$. The exact wave function for the ground state is

$$\Psi_0(x) = \frac{1}{\pi^{1/4}} e^{-x^2/2}, \quad (11.23)$$

but since we wish to illustrate the use of Monte Carlo methods, we choose the trial function

$$\Psi_T(x) = \frac{\sqrt{\alpha}}{\pi^{1/4}} e^{-x^2\alpha^2/2}. \quad (11.24)$$

Inserting this function in the expression for the local energy in Eq. (11.16), we obtain the following expression for the local energy

$$E_L(x) = \alpha^2 + x^2(1 - \alpha^4), \quad (11.25)$$

with the expectation value for the hamiltonian of Eq. (11.17) given by

$$\langle E_L \rangle = \int_{-\infty}^{\infty} |\psi_T(x)|^2 E_L(x) dx, \quad (11.26)$$

which reads with the above trial wave function

$$\langle E_L \rangle = \frac{\int_{-\infty}^{\infty} dx e^{-x^2\alpha^2} \alpha^2 + x^2(1 - \alpha^4)}{\int_{-\infty}^{\infty} dx e^{-x^2\alpha^2}}. \quad (11.27)$$

Using the fact that

$$\int_{-\infty}^{\infty} dx e^{-x^2\alpha^2} = \sqrt{\frac{\pi}{\alpha^2}},$$

we obtain

$$\langle E_L \rangle = \frac{\alpha^2}{2} + \frac{1}{2\alpha^2}. \quad (11.28)$$

and the variance

$$\sigma^2 = \frac{(\alpha^4 - 1)^2}{2\alpha^4}. \quad (11.29)$$

In solving this problem we can choose whether we wish to use the Metropolis algorithm and sample over relevant configurations, or just use random numbers generated from a normal distribution, since the harmonic oscillator wave functions follow closely such a distribution. The latter approach is easily implemented in few lines, namely

```

... initialisations , declarations of variables
... mcs = number of Monte Carlo samplings
// loop over Monte Carlo samples
for ( i=0; i < mcs; i++) {
// generate random variables from gaussian distribution
  x = normal_random(&idum)/sqrt2/alpha;
  local_energy = alpha*alpha + x*x*(1-pow(alpha,4));
  energy += local_energy;
  energy2 += local_energy*local_energy;
// end of sampling
}
// write out the mean energy and the standard deviation
cout << energy/mcs << sqrt((energy2/mcs-(energy/mcs)**2)/mcs);

```

This variational Monte Carlo calculation is rather simple, we just generate a large number N of random numbers corresponding to the gaussian PDF $\sim |\Psi_T|^2$ and for each random number we compute the local energy according to the approximation

$$\langle \hat{\mathbf{E}}_L \rangle = \int P(\mathbf{R}) \hat{\mathbf{E}}_L(\mathbf{R}) d\mathbf{R} \approx \frac{1}{N} \sum_{i=1}^N E_L(x_i), \quad (11.30)$$

and the energy squared through

$$\langle \hat{\mathbf{E}}_L^2 \rangle = \int P(\mathbf{R}) \hat{\mathbf{E}}_L^2(\mathbf{R}) d\mathbf{R} \approx \frac{1}{N} \sum_{i=1}^N E_L^2(x_i). \quad (11.31)$$

In a certain sense, this is nothing but the importance Monte Carlo sampling discussed in chapter 8. Before we proceed however, there is an important aside which is worth keeping in mind when computing the local energy. We could think of splitting the computation of the expectation value of the local energy into a kinetic energy part and a potential energy part. If we are dealing with a three-dimensional system, the expectation value of the kinetic energy is

$$- \frac{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \nabla^2 \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R})}, \quad (11.32)$$

and we could be tempted to compute, if the wave function obeys spherical symmetry, just the second derivative with respect to one coordinate axis and then multiply by three. This will most likely increase the variance, and should be avoided, even if the final expectation values are similar. Recall that one of the subgoals of a Monte Carlo computation is to decrease the variance.

Another shortcut we could think of is to transform the numerator in the latter equation to

$$\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \nabla^2 \Psi_T(\mathbf{R}) = - \int d\mathbf{R} (\nabla \Psi_T^*(\mathbf{R})) (\nabla \Psi_T(\mathbf{R})), \quad (11.33)$$

using integration by parts and the relation

$$\int d\mathbf{R} \nabla (\Psi_T^*(\mathbf{R}) \nabla \Psi_T(\mathbf{R})) = 0, \quad (11.34)$$

where we have used the fact that the wave function is zero at $\mathbf{R} = \pm\infty$. This relation can in turn be rewritten through integration by parts to

$$\int d\mathbf{R} (\nabla \Psi_T^*(\mathbf{R})) (\nabla \Psi_T(\mathbf{R})) + \int d\mathbf{R} \Psi_T^*(\mathbf{R}) \nabla^2 \Psi_T(\mathbf{R}) = 0. \quad (11.35)$$

The rhs of Eq. (11.33) is easier and quicker to compute. However, in case the wave function is the exact one, or rather close to the exact one, the lhs yields just a constant times the wave function squared, implying zero variance. The rhs does not and may therefore increase the variance.

If we use integration by part for the harmonic oscillator case, the new local energy is

$$E_L(x) = x^2(1 + \alpha^4), \quad (11.36)$$

and the variance

$$\sigma^2 = \frac{(\alpha^4 + 1)^2}{2\alpha^4}, \quad (11.37)$$

which is larger than the variance of Eq. (11.29).

11.5 Variational Monte Carlo for atoms

The Hamiltonian for an N -electron atomic system consists of two terms

$$\hat{H}(\mathbf{x}) = \hat{T}(\mathbf{x}) + \hat{V}(\mathbf{x}); \quad (11.38)$$

the kinetic and the potential energy operator. Here $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ is the spatial and spin degrees of freedom associated with the different particles. The classical kinetic energy

$$T = \frac{\mathbf{P}^2}{2m} + \sum_{j=1}^N \frac{\mathbf{p}_j^2}{2m}$$

is transformed to the quantum mechanical kinetic energy operator by operator substitution of the momentum ($p_k \rightarrow -i\hbar\partial/\partial x_k$)

$$\hat{T}(\mathbf{x}) = -\frac{\hbar^2}{2M}\nabla_0^2 - \sum_{i=1}^N \frac{\hbar^2}{2m}\nabla_i^2. \quad (11.39)$$

Here the first term is the kinetic energy operator of the nucleus, the second term is the kinetic energy operator of the electrons, M is the mass of the nucleus and m is the electron mass. The potential energy operator is given by

$$\hat{V}(\mathbf{x}) = -\sum_{i=1}^N \frac{Ze^2}{(4\pi\epsilon_0)r_i} + \sum_{i=1, i < j}^N \frac{e^2}{(4\pi\epsilon_0)r_{ij}}, \quad (11.40)$$

where the r_i 's are the electron-nucleus distances and the r_{ij} 's are the inter-electronic distances.

We seek to find controlled and well understood approximations in order to reduce the complexity of the above equations. The *Born-Oppenheimer approximation* is a commonly used approximation, in which the motion of the nucleus is disregarded.

11.5.1 The Born-Oppenheimer Approximation

In a system of interacting electrons and a nucleus there will usually be little momentum transfer between the two types of particles due to their differing masses. The forces between the particles are of similar magnitude due to their similar charge. If one assumes that the momenta of the particles are also similar, the nucleus must have a much smaller velocity than the electrons due to its far greater mass. On the time-scale of nuclear motion, one can therefore consider the electrons to relax to a ground-state given by the Hamiltonian of Eqs. (11.38), (11.39) and (11.40) with the nucleus at a fixed location. This separation of the electronic and nuclear degrees of freedom is known as the Born-Oppenheimer approximation.

In the center of mass system the kinetic energy operator reads

$$\hat{T}(\mathbf{x}) = -\frac{\hbar^2}{2(M + Nm)}\nabla_{CM}^2 - \frac{\hbar^2}{2\mu}\sum_{i=1}^N \nabla_i^2 - \frac{\hbar^2}{M}\sum_{i>j}^N \nabla_i \cdot \nabla_j, \quad (11.41)$$

while the potential energy operator remains unchanged. Note that the Laplace operators ∇_i^2 now are in the center of mass reference system.

The first term of Eq. (11.41) represents the kinetic energy operator of the center of mass. The second term represents the sum of the kinetic energy operators of the N electrons, each of them having their

mass m replaced by the reduced mass $\mu = mM/(m + M)$ because of the motion of the nucleus. The nuclear motion is also responsible for the third term, or the *mass polarization* term.

The nucleus consists of protons and neutrons. The proton-electron mass ratio is about 1/1836 and the neutron-electron mass ratio is about 1/1839, so regarding the nucleus as stationary is a natural approximation. Taking the limit $M \rightarrow \infty$ in Eq. (11.41), the kinetic energy operator reduces to

$$\hat{T} = - \sum_{i=1}^N \frac{\hbar^2}{2m} \nabla_i^2 \quad (11.42)$$

The Born-Oppenheimer approximation thus disregards both the kinetic energy of the center of mass as well as the mass polarization term. The effects of the Born-Oppenheimer approximation are quite small and they are also well accounted for. However, this simplified electronic Hamiltonian remains very difficult to solve, and analytical solutions do not exist for general systems with more than one electron. We use the Born-Oppenheimer approximation in our discussion of atomic and molecular systems.

The first term of Eq. (11.40) is the nucleus-electron potential and the second term is the electron-electron potential. The inter-electronic potentials are the main problem in atomic physics. Because of these terms, the Hamiltonian cannot be separated into one-particle parts, and the problem must be solved as a whole. A common approximation is to regard the effects of the electron-electron interactions either as averaged over the domain or by means of introducing a density functional, such as by Hartree-Fock (HF) or Density Functional Theory (DFT). These approaches are actually very efficient, and about 99% or more of the electronic energies are obtained for most HF calculations. Other observables are usually obtained to an accuracy of about 90 – 95% (ref. [73]).

11.5.2 The hydrogen Atom

The spatial Schrödinger equation for the three-dimensional hydrogen atom can be solved analytically, see for example Ref. [74] for details. To achieve this, we rewrite the equation in terms of spherical coordinates using

$$x = r \sin\theta \cos\phi, \quad (11.43)$$

$$y = r \sin\theta \sin\phi, \quad (11.44)$$

and

$$z = r \cos\theta. \quad (11.45)$$

The reason we introduce spherical coordinates is the spherical symmetry of the Coulomb potential

$$\frac{e^2}{4\pi\epsilon_0 r} = \frac{e^2}{4\pi\epsilon_0 \sqrt{x^2 + y^2 + z^2}}, \quad (11.46)$$

where we have used $r = \sqrt{x^2 + y^2 + z^2}$. It is not possible to find a separable solution of the type

$$\psi(x, y, z) = \psi(x)\psi(y)\psi(z). \quad (11.47)$$

as we can with the harmonic oscillator in three dimensions. However, with spherical coordinates we can find a solution of the form

$$\psi(r, \theta, \phi) = R(r)P(\theta)F(\phi) = RPF. \quad (11.48)$$

These three coordinates yield in turn three quantum numbers which determine the energy of the systems. We obtain three sets of ordinary second-order differential equations which can be solved analytically, resulting in

$$\frac{1}{F} \frac{\partial^2 F}{\partial \phi^2} = -C_\phi^2, \quad (11.49)$$

$$C_r \sin^2(\theta)P + \sin(\theta) \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial P}{\partial \theta} \right) = C_\phi^2 P, \quad (11.50)$$

and

$$\frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + \frac{2mrke^2}{\hbar^2} + \frac{2mr^2}{\hbar^2} E = C_r, \quad (11.51)$$

where C_r and C_ϕ are constants. The angle-dependent differential equations result in the spherical harmonic functions as solutions, with quantum numbers l and m_l . These functions are given by

$$Y_{lm_l}(\theta, \phi) = P(\theta)F(\phi) = \sqrt{\frac{(2l+1)(l-m_l)!}{4\pi(l+m_l)!}} P_l^{m_l}(\cos(\theta)) \exp(im_l\phi), \quad (11.52)$$

with $P_l^{m_l}$ being the associated Legendre polynomials. They can be rewritten as

$$Y_{lm_l}(\theta, \phi) = \sin^{|m_l|}(\theta) \times (\text{polynom}(\cos\theta)) \exp(im_l\phi), \quad (11.53)$$

with the following selected examples

$$Y_{00} = \sqrt{\frac{1}{4\pi}}, \quad (11.54)$$

for $l = m_l = 0$,

$$Y_{10} = \sqrt{\frac{3}{4\pi}} \cos(\theta), \quad (11.55)$$

for $l = 1$ og $m_l = 0$,

$$Y_{1\pm 1} = \sqrt{\frac{3}{8\pi}} \sin(\theta) \exp(\pm i\phi), \quad (11.56)$$

for $l = 1$ og $m_l = \pm 1$, and

$$Y_{20} = \sqrt{\frac{5}{16\pi}} (3\cos^2(\theta) - 1) \quad (11.57)$$

for $l = 2$ og $m_l = 0$. The quantum numbers l and m_l represent the orbital momentum and projection of the orbital momentum, respectively and take the values

1.

$$l \geq 0$$

2.

$$l = 0, 1, 2, \dots$$

3.

$$m_l = -l, -l+1, \dots, l-1, l$$

The spherical harmonics for $l \leq 3$ are listed in Table 11.1.

We concentrate on the radial equation, which can be rewritten as

$$-\frac{\hbar^2 r^2}{2m} \left(\frac{\partial}{\partial r} \left(r^2 \frac{\partial R(r)}{\partial r} \right) \right) - \frac{ke^2}{r} R(r) + \frac{\hbar^2 l(l+1)}{2mr^2} R(r) = ER(r). \quad (11.58)$$

Introducing the function $u(r) = rR(r)$, we can rewrite the last equation as

Spherical Harmonics

$m_l \backslash l$	0	1	2	3
+3				$-\frac{1}{8}(\frac{35}{\pi})^{1/2} \sin^3\theta e^{+3i\phi}$
+2			$\frac{1}{4}(\frac{15}{2\pi})^{1/2} \sin^2\theta e^{+2i\phi}$	$\frac{1}{4}(\frac{105}{2\pi})^{1/2} \cos\theta \sin^2\theta e^{+2i\phi}$
+1		$-\frac{1}{2}(\frac{3}{2\pi})^{1/2} \sin\theta e^{+i\phi}$	$-\frac{1}{2}(\frac{15}{2\pi})^{1/2} \cos\theta \sin\theta e^{+i\phi}$	$-\frac{1}{8}(\frac{21}{2\pi})^{1/2} (5\cos^2\theta - 1) \sin\theta e^{+i\phi}$
0	$\frac{1}{2\pi^{1/2}}$	$\frac{1}{2}(\frac{3}{\pi})^{1/2} \cos\theta$	$\frac{1}{4}(\frac{5}{\pi})^{1/2} (3\cos^2\theta - 1)$	$\frac{1}{4}(\frac{7}{\pi})^{1/2} (2 - 5\sin^2\theta) \cos\theta$
-1		$+\frac{1}{2}(\frac{3}{2\pi})^{1/2} \sin\theta e^{-i\phi}$	$+\frac{1}{2}(\frac{15}{2\pi})^{1/2} \cos\theta \sin\theta e^{-i\phi}$	$+\frac{1}{8}(\frac{21}{2\pi})^{1/2} (5\cos^2\theta - 1) \sin\theta e^{-i\phi}$
-2			$\frac{1}{4}(\frac{15}{2\pi})^{1/2} \sin^2\theta e^{-2i\phi}$	$\frac{1}{4}(\frac{105}{2\pi})^{1/2} \cos\theta \sin^2\theta e^{-2i\phi}$
-3				$+\frac{1}{8}(\frac{35}{\pi})^{1/2} \sin^3\theta e^{-3i\phi}$

Table 11.1: Spherical harmonics Y_{lm_l} for the lowest l and m_l values.

The radial Schrödinger equation for the hydrogen atom can be written as

$$-\frac{\hbar^2}{2m} \frac{\partial^2 u(r)}{\partial r^2} - \left(\frac{ke^2}{r} - \frac{\hbar^2 l(l+1)}{2mr^2} \right) u(r) = Eu(r), \quad (11.59)$$

where m is the mass of the electron, l its orbital momentum taking values $l = 0, 1, 2, \dots$, and the term ke^2/r is the Coulomb potential. The first terms is the kinetic energy. The full wave function will also depend on the other variables θ and ϕ as well. The energy, with no external magnetic field is however determined by the above equation. We can then think of the radial Schrödinger equation to be equivalent to a one-dimensional movement conditioned by an effective potential

$$V_{\text{eff}}(r) = -\frac{ke^2}{r} + \frac{\hbar^2 l(l+1)}{2mr^2}. \quad (11.60)$$

The radial equation can also be solved analytically resulting in the quantum numbers n in addition to lm_l . The solution R_{nl} to the radial equation is given by the Laguerre polynomials. The analytic solutions are given by

$$\psi_{nlm_l}(r, \theta, \phi) = \psi_{nlm_l} = R_{nl}(r)Y_{lm_l}(\theta, \phi) = R_{nl}Y_{lm_l} \quad (11.61)$$

The ground state is defined by $n = 1$ og $l = m_l = 0$ and reads

$$\psi_{100} = \frac{1}{a_0^{3/2} \sqrt{\pi}} e^{-r/a_0}, \quad (11.62)$$

where we have defined the Bohr radius $a_0 = 0.05$ nm

$$a_0 = \frac{\hbar^2}{mke^2}. \quad (11.63)$$

The first excited state with $l = 0$ is

$$\psi_{200} = \frac{1}{4a_0^{3/2} \sqrt{2\pi}} \left(2 - \frac{r}{a_0} \right) e^{-r/2a_0}. \quad (11.64)$$

For states with with $l = 1$ and $n = 2$, we can have the following combinations with $m_l = 0$

$$\psi_{210} = \frac{1}{4a_0^{3/2}\sqrt{2\pi}} \left(\frac{r}{a_0}\right) e^{-r/2a_0} \cos(\theta), \quad (11.65)$$

and $m_l = \pm 1$

$$\psi_{21\pm 1} = \frac{1}{8a_0^{3/2}\sqrt{\pi}} \left(\frac{r}{a_0}\right) e^{-r/2a_0} \sin(\theta) e^{\pm i\phi}. \quad (11.66)$$

The exact energy is independent of l and m_l , since the potential is spherically symmetric.

The first few non-normalized radial solutions of equation are listed in Table 11.2. A problem with the

Hydrogen-Like Atomic Radial Functions

$l \setminus n$	1	2	3
0	e^{-Zr}	$(2-r)e^{-Zr/2}$	$(27-18r+2r^2)e^{-Zr/3}$
1		$re^{-Zr/2}$	$r(6-r)e^{-Zr/3}$
2			$r^2e^{-Zr/3}$

Table 11.2: The first few radial functions of the hydrogen-like atoms.

spherical harmonics of table 11.1 is that they are complex. The introduction of *solid harmonics* allows the use of real orbital wave-functions for a wide range of applications. The complex solid harmonics $\mathcal{Y}_{lm_l}(\mathbf{r})$ are related to the spherical harmonics $Y_{lm_L}(\mathbf{r})$ through

$$\mathcal{Y}_{lm_l}(\mathbf{r}) = r^l Y_{lm_l}(\mathbf{r}).$$

By factoring out the leading r -dependency of the radial-function

$$\mathcal{R}_{nl}(\mathbf{r}) = r^{-l} R_{nl}(\mathbf{r}),$$

we obtain

$$\Psi_{nlm_l}(r, \theta, \phi) = \mathcal{R}_{nl}(\mathbf{r}) \cdot \mathcal{Y}_{lm_l}(\mathbf{r}).$$

For the theoretical development of the *real solid harmonics* see Ref. [75]. Here Helgaker *et al* first express the complex solid harmonics, C_{lm_l} , by (complex) Cartesian coordinates, and arrive at the real solid harmonics, S_{lm_l} , through the unitary transformation

$$\begin{pmatrix} S_{lm_l} \\ S_{l,-m_l} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} (-1)_l^m & 1 \\ -(-1)_l^m i & i \end{pmatrix} \begin{pmatrix} C_{lm_l} \\ C_{l,-m_l} \end{pmatrix}.$$

This transformation will not alter any physical quantities that are degenerate in the subspace consisting of opposite magnetic quantum numbers (the angular momentum l is equal for both these cases). This means for example that the above transformation does not alter the energies, unless an external magnetic field is applied to the system. Henceforth, we will use the solid harmonics, and note that changing the spherical

Real Solid Harmonics

$m_l \backslash l$	0	1	2	3
+3				$\frac{1}{2}\sqrt{\frac{5}{2}}(x^2 - 3y^2)x$
+2			$\frac{1}{2}\sqrt{3}(x^2 - y^2)$	$\frac{1}{2}\sqrt{15}(x^2 - y^2)z$
+1		x	$\sqrt{3}xz$	$\frac{1}{2}\sqrt{\frac{3}{2}}(5z^2 - r^2)x$
0	1	y	$\frac{1}{2}(3z^2 - r^2)$	$\frac{1}{2}(5z^2 - 3r^2)x$
-1		z	$\sqrt{3}yz$	$\frac{1}{2}\sqrt{\frac{3}{2}}(5z^2 - r^2)y$
-2			$\sqrt{3}xy$	$\sqrt{15}xyz$
-3				$\frac{1}{2}\sqrt{\frac{5}{2}}(3x^2 - y^2)y$

Table 11.3: The first-order real solid harmonics \mathcal{Y}_{lm_l} .

potential beyond the Coulomb potential will not alter the solid harmonics. The lowest-order real solid harmonics are listed in table 11.3.

When solving equations numerically, it is often convenient to rewrite the equation in terms of dimensionless variables. One reason is the fact that several of the constants may differ largely in value, and hence result in potential losses of numerical precision. The other main reason for doing this is that the equation in dimensionless form is easier to code, sparing one for eventual typographic errors. In order to do so, we introduce first the dimensionless variable $\rho = r/\beta$, where β is a constant we can choose. Schrödinger's equation is then rewritten as

$$-\frac{1}{2} \frac{\partial^2 u(\rho)}{\partial \rho^2} - \frac{mke^2\beta}{\hbar^2 \rho} u(\rho) + \frac{l(l+1)}{2\rho^2} u(\rho) = \frac{m\beta^2}{\hbar^2} E u(\rho). \quad (11.67)$$

We can determine β by simply requiring³

$$\frac{mke^2\beta}{\hbar^2} = 1 \quad (11.68)$$

With this choice, the constant β becomes the famous Bohr radius $a_0 = 0.05 \text{ nm}$ $a_0 = \beta = \hbar^2/mke^2$.

As a petit digression, we list here the standard units used in atomic physics and molecular physics calculations. It is common to scale atomic units by setting $m = e = \hbar = 4\pi\epsilon_0 = 1$, see table 11.4.

We introduce thereafter the variable λ

$$\lambda = \frac{m\beta^2}{\hbar^2} E, \quad (11.69)$$

and inserting β and the exact energy $E = E_0/n^2$, with $E_0 = 13.6 \text{ eV}$, we have that

$$\lambda = -\frac{1}{2n^2}, \quad (11.70)$$

³Remember that we are free to choose β .

Atomic Units

Quantity	SI	Atomic unit
Electron mass, m	$9.109 \cdot 10^{-31}$ kg	1
Charge, e	$1.602 \cdot 10^{-19}$ C	1
Planck's reduced constant, \hbar	$1.055 \cdot 10^{-34}$ Js	1
Permittivity, $4\pi\epsilon_0$	$1.113 \cdot 10^{-10}$ C ² J ⁻¹ m ⁻¹	1
Energy, $\frac{e^2}{4\pi\epsilon_0 a_0}$	27.211 eV	1
Length, $a_0 = \frac{4\pi\epsilon_0 \hbar^2}{me^2}$	$0.529 \cdot 10^{-10}$ m	1

Table 11.4: Scaling from SI to atomic units

n being the principal quantum number. The equation we are then going to solve numerically is now

$$-\frac{1}{2} \frac{\partial^2 u(\rho)}{\partial \rho^2} - \frac{u(\rho)}{\rho} + \frac{l(l+1)}{2\rho^2} u(\rho) - \lambda u(\rho) = 0, \quad (11.71)$$

with the hamiltonian

$$H = -\frac{1}{2} \frac{\partial^2}{\partial \rho^2} - \frac{1}{\rho} + \frac{l(l+1)}{2\rho^2}. \quad (11.72)$$

The ground state of the hydrogen atom has the energy $\lambda = -1/2$, or $E = -13.6$ eV. The exact wave function obtained from Eq. (11.71) is

$$u(\rho) = \rho e^{-\rho}, \quad (11.73)$$

which yields the energy $\lambda = -1/2$. Sticking to our variational philosophy, we could now introduce a variational parameter α resulting in a trial wave function

$$u_T^\alpha(\rho) = \alpha \rho e^{-\alpha \rho}. \quad (11.74)$$

Inserting this wave function into the expression for the local energy E_L of Eq. (11.16) yields (check it!)

$$E_L(\rho) = -\frac{1}{\rho} - \frac{\alpha}{2} \left(\alpha - \frac{2}{\rho} \right). \quad (11.75)$$

For the hydrogen atom, we could perform the variational calculation along the same lines as we did for the harmonic oscillator. The only difference is that Eq. (11.17) now reads

$$\langle H \rangle = \int P(\mathbf{R}) E_L(\mathbf{R}) d\mathbf{R} = \int_0^\infty \alpha^2 \rho^2 e^{-2\alpha \rho} E_L(\rho) \rho^2 d\rho, \quad (11.76)$$

since $\rho \in [0, \infty]$. In this case we would use the exponential distribution instead of the normal distribution, and our code would contain the following elements

```

...  initialisations , declarations of variables
...  mcs = number of Monte Carlo samplings

//  loop over Monte Carlo samples
for ( i=0; i < mcs; i++) {

```

```

// generate random variables from the exponential
// distribution using ran1 and transforming to
// to an exponential mapping  $y = -\ln(1-x)$ 
    x=ran1(&idum);
    y=-log(1.-x);
// in our case  $y = \rho*\alpha*2$ 
    rho = y/alpha/2;
    local_energy = -1/rho -0.5*alpha*(alpha-2/rho);
    energy += (local_energy);
    energy2 += local_energy*local_energy;
// end of sampling
}
// write out the mean energy and the standard deviation
cout << energy/mcs << sqrt((energy2/mcs-(energy/mcs)**2)/mcs);

```

As for the harmonic oscillator case we just need to generate a large number N of random numbers corresponding to the exponential PDF $\alpha^2 \rho^2 e^{-2\alpha\rho}$ and for each random number we compute the local energy and variance.

11.5.3 Metropolis sampling for the hydrogen atom and the harmonic oscillator

We present in this subsection results for the ground states of the hydrogen atom and harmonic oscillator using a variational Monte Carlo procedure. For the hydrogen atom, the trial wave function

$$u_T^\alpha(\rho) = \alpha\rho e^{-\alpha\rho},$$

depends only on the dimensionless radius ρ . It is the solution of a one-dimensional differential equation, as is the case for the harmonic oscillator as well. The latter has the trial wave function

$$\Psi_T(x) = \frac{\sqrt{\alpha}}{\pi^{1/4}} e^{-x^2\alpha^2/2}.$$

However, for the hydrogen atom we have $\rho \in [0, \infty]$, while for the harmonic oscillator we have $x \in [-\infty, \infty]$.

This has important consequences for the way we generate random positions. For the hydrogen atom we have a random position given by e.g.,

```
r_old = step_length*(ran1(&idum))/alpha;
```

which ensures that $\rho \geq 0$, while for the harmonic oscillator we have

```
r_old = step_length*(ran1(&idum)-0.5)/alpha;
```

in order to have $x \in [-\infty, \infty]$. This is however not implemented in the program below. There, importance sampling is not included. We simulate points in the x , y and z directions using random numbers generated by the uniform distribution and multiplied by the step length. Note that we have to define a step length in our calculations. Here one has to play around with different values for the step and as a rule of thumb (one of the golden Monte Carlo rules), the step length should be chosen so that roughly 50% of all new moves are accepted. In the program at the end of this section we have also scaled the random position with the variational parameter α . The reason for this particular choice is that we have an external loop over the variational parameter. Different variational parameters will obviously yield different acceptance rates if we use the same step length. An alternative to the code below is to perform the

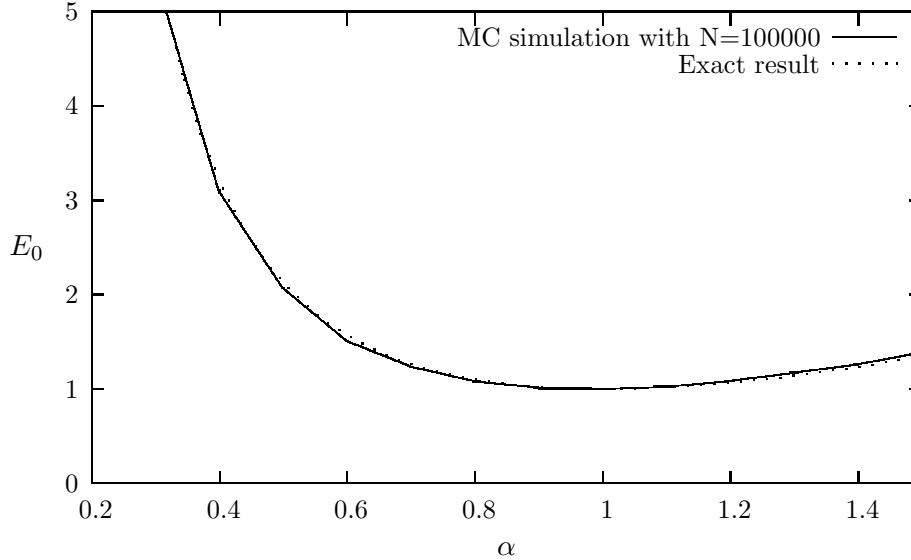


Figure 11.1: Result for ground state energy of the harmonic oscillator as function of the variational parameter α . The exact result is for $\alpha = 1$ with an energy $E = 1$. See text for further details

Monte Carlo sampling with just one variational parameter, and play around with different step lengths in order to achieve a reasonable acceptance ratio. Another possibility is to include a more advanced test which restarts the Monte Carlo sampling with a new step length if the specific variational parameter and chosen step length lead to a too low acceptance ratio.

In Figs. 11.1 and 11.2 we plot the ground state energies for the one-dimensional harmonic oscillator and the hydrogen atom, respectively, as functions of the variational parameter α . These results are also displayed in Tables 11.5 and 11.6. In these tables we list the variance and the standard deviation as well. We note that at α we obtain the exact result, and the variance is zero, as it should. The reason is that we then have the exact wave function, and the action of the hamiltonian on the wave function

$$H\psi = \text{constant} \times \psi,$$

yields just a constant. The integral which defines various expectation values involving moments of the hamiltonian becomes then

$$\langle H^n \rangle = \frac{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) H^n(\mathbf{R}) \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R})} = \text{constant} \times \frac{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R})} = \text{constant}. \quad (11.77)$$

This explains why the variance is zero for $\alpha = 1$. However, the hydrogen atom and the harmonic oscillator are some of the few cases where we can use a trial wave function proportional to the exact one. These two systems are also some of the few examples of cases where we can find an exact solution to the problem. In most cases of interest, we do not know a priori the exact wave function, or how to make a good trial wave function. In essentially all real problems a large amount of CPU time and numerical experimenting is needed in order to ascertain the validity of a Monte Carlo estimate. The next examples deal with such problems.

Table 11.5: Result for ground state energy of the harmonic oscillator as function of the variational parameter α . The exact result is for $\alpha = 1$ with an energy $E = 1$. The energy variance σ^2 and the standard deviation σ/\sqrt{N} are also listed. The variable N is the number of Monte Carlo samples. In this calculation we set $N = 100000$ and a step length of 2 was used in order to obtain an acceptance of $\approx 50\%$.

α	$\langle H \rangle$	σ^2	σ/\sqrt{N}
5.00000E-01	2.06479E+00	5.78739E+00	7.60749E-03
6.00000E-01	1.50495E+00	2.32782E+00	4.82475E-03
7.00000E-01	1.23264E+00	9.82479E-01	3.13445E-03
8.00000E-01	1.08007E+00	3.44857E-01	1.85703E-03
9.00000E-01	1.01111E+00	7.24827E-02	8.51368E-04
1.00000E+00	1.00000E+00	0.00000E+00	0.00000E+00
1.10000E+00	1.02621E+00	5.95716E-02	7.71826E-04
1.20000E+00	1.08667E+00	2.23389E-01	1.49462E-03
1.30000E+00	1.17168E+00	4.78446E-01	2.18734E-03
1.40000E+00	1.26374E+00	8.55524E-01	2.92493E-03
1.50000E+00	1.38897E+00	1.30720E+00	3.61553E-03

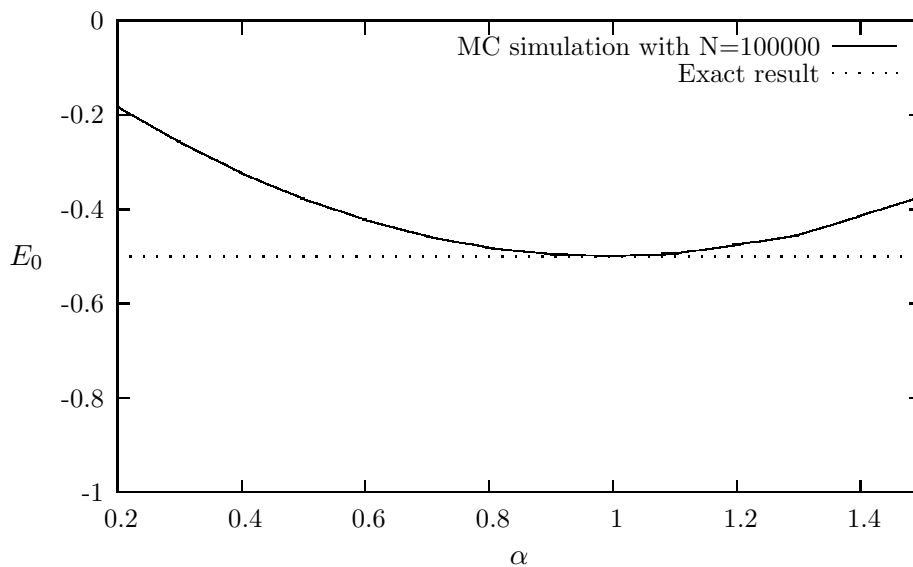


Figure 11.2: Result for ground state energy of the hydrogen atom as function of the variational parameter α . The exact result is for $\alpha = 1$ with an energy $E = -1/2$. See text for further details

Table 11.6: Result for ground state energy of the hydrogen atom as function of the variational parameter α . The exact result is for $\alpha = 1$ with an energy $E = -1/2$. The energy variance σ^2 and the standard deviation σ/\sqrt{N} are also listed. The variable N is the number of Monte Carlo samples. In this calculation we fixed $N = 100000$ and a step length of 4 Bohr radii was used in order to obtain an acceptance of $\approx 50\%$.

α	$\langle H \rangle$	σ^2	σ/\sqrt{N}
5.00000E-01	-3.76740E-01	6.10503E-02	7.81347E-04
6.00000E-01	-4.21744E-01	5.22322E-02	7.22718E-04
7.00000E-01	-4.57759E-01	4.51201E-02	6.71715E-04
8.00000E-01	-4.81461E-01	3.05736E-02	5.52934E-04
9.00000E-01	-4.95899E-01	8.20497E-03	2.86443E-04
1.00000E+00	-5.00000E-01	0.00000E+00	0.00000E+00
1.10000E+00	-4.93738E-01	1.16989E-02	3.42036E-04
1.20000E+00	-4.75563E-01	8.85899E-02	9.41222E-04
1.30000E+00	-4.54341E-01	1.45171E-01	1.20487E-03
1.40000E+00	-4.13220E-01	3.14113E-01	1.77232E-03
1.50000E+00	-3.72241E-01	5.45568E-01	2.33574E-03

11.5.4 The helium atom

Most physical problems of interest in atomic, molecular and solid state physics consist of a number of interacting electrons and ions. The total number of particles N is usually sufficiently large that an exact solution cannot be found. Typically, the expectation value for a chosen hamiltonian for a system of N particles is

$$\langle H \rangle = \frac{\int d\mathbf{R}_1 d\mathbf{R}_2 \dots d\mathbf{R}_N \Psi^*(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) H(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \Psi(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N)}{\int d\mathbf{R}_1 d\mathbf{R}_2 \dots d\mathbf{R}_N \Psi^*(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \Psi(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N)}, \quad (11.78)$$

an in general intractable problem. Controlled and well understood approximations are sought to reduce the complexity to a tractable level. Once the equations are solved, a large number of properties may be calculated from the wave function. Errors or approximations made in obtaining the wave function will be manifest in any property derived from the wave function. Where high accuracy is required, considerable attention must be paid to the derivation of the wave function and any approximations made.

The helium atom consists of two electrons and a nucleus with charge $Z = 2$. In setting up the hamiltonian of this system, we need to account for the repulsion between the two electrons as well.

A common and very reasonable approximation used in the solution of equation of the Schrödinger equation for systems of interacting electrons and ions is the Born-Oppenheimer approximation discussed above. In a system of interacting electrons and nuclei there will usually be little momentum transfer between the two types of particles due to their greatly differing masses. The forces between the particles are of similar magnitude due to their similar charge. If one then assumes that the momenta of the particles are also similar, then the nuclei must have much smaller velocities than the electrons due to their far greater mass. On the time-scale of nuclear motion, one can therefore consider the electrons to relax to a ground-state with the nuclei at fixed locations. This separation of the electronic and nuclear degrees of freedom is the the Born-Oppenheimer approximation we discussed previously in this chapter. But even this simplified electronic Hamiltonian remains very difficult to solve. No analytic solutions exist for general systems with more than one electron.

If we label the distance between electron 1 and the nucleus as r_1 . Similarly we have r_2 for electron 2. The contribution to the potential energy due to the attraction from the nucleus is

$$-\frac{2ke^2}{r_1} - \frac{2ke^2}{r_2}, \quad (11.79)$$

and if we add the repulsion arising from the two interacting electrons, we obtain the potential energy

$$V(r_1, r_2) = -\frac{2ke^2}{r_1} - \frac{2ke^2}{r_2} + \frac{ke^2}{r_{12}}, \quad (11.80)$$

with the electrons separated at a distance $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$. The hamiltonian becomes then

$$\hat{\mathbf{H}} = -\frac{\hbar^2 \nabla_1^2}{2m} - \frac{\hbar^2 \nabla_2^2}{2m} - \frac{2ke^2}{r_1} - \frac{2ke^2}{r_2} + \frac{ke^2}{r_{12}}, \quad (11.81)$$

and Schrödinger's equation reads

$$\hat{\mathbf{H}}\psi = E\psi. \quad (11.82)$$

Note that this equation has been written in atomic units *a.u.* which are more convenient for quantum mechanical problems. This means that the final energy has to be multiplied by a $2 \times E_0$, where $E_0 = 13.6$ eV, the binding energy of the hydrogen atom.

A very simple first approximation to this system is to omit the repulsion between the two electrons. The potential energy becomes then

$$V(r_1, r_2) \approx -\frac{Zke^2}{r_1} - \frac{Zke^2}{r_2}. \quad (11.83)$$

The advantage of this approximation is that each electron can be treated as being independent of each other, implying that each electron sees just a centrally symmetric potential, or central field.

To see whether this gives a meaningful result, we set $Z = 2$ and neglect totally the repulsion between the two electrons. Electron 1 has the following hamiltonian

$$\hat{\mathbf{h}}_1 = -\frac{\hbar^2 \nabla_1^2}{2m} - \frac{2ke^2}{r_1}, \quad (11.84)$$

with pertinent wave function and eigenvalue

$$\hat{\mathbf{h}}_1 \psi_a = E_a \psi_a, \quad (11.85)$$

where $a = \{n_a l_a m_{l_a}\}$, are its quantum numbers. The energy E_a is

$$E_a = -\frac{Z^2 E_0}{n_a^2}, \quad (11.86)$$

med $E_0 = 13.6$ eV, being the ground state energy of the hydrogen atom. In a similar way, we obtain for electron 2

$$\hat{\mathbf{h}}_2 = -\frac{\hbar^2 \nabla_2^2}{2m} - \frac{2ke^2}{r_2}, \quad (11.87)$$

with wave function

$$\hat{\mathbf{h}}_2 \psi_b = E_b \psi_b, \quad (11.88)$$

and $b = \{n_b l_b m_{l_b}\}$, and energy

$$E_b = \frac{Z^2 E_0}{n_b^2}. \quad (11.89)$$

Since the electrons do not interact, we can assume that the ground state wave function of the helium atom is given by

$$\psi = \psi_a \psi_b, \quad (11.90)$$

resulting in the following approximation to Schrödinger's equation

$$\left(\widehat{\mathbf{h}}_1 + \widehat{\mathbf{h}}_2\right) \psi = \left(\widehat{\mathbf{h}}_1 + \widehat{\mathbf{h}}_2\right) \psi_a(\mathbf{r}_1) \psi_b(\mathbf{r}_2) = E_{ab} \psi_a(\mathbf{r}_1) \psi_b(\mathbf{r}_2). \quad (11.91)$$

The energy becomes then

$$\left(\widehat{\mathbf{h}}_1 \psi_a(\mathbf{r}_1)\right) \psi_b(\mathbf{r}_2) + \left(\widehat{\mathbf{h}}_2 \psi_b(\mathbf{r}_2)\right) \psi_a(\mathbf{r}_1) = (E_a + E_b) \psi_a(\mathbf{r}_1) \psi_b(\mathbf{r}_2), \quad (11.92)$$

yielding

$$E_{ab} = Z^2 E_0 \left(\frac{1}{n_a^2} + \frac{1}{n_b^2} \right). \quad (11.93)$$

If we insert $Z = 2$ and assume that the ground state is determined by two electrons in the lowest-lying hydrogen orbit with $n_a = n_b = 1$, the energy becomes

$$E_{ab} = 8E_0 = -108.8 \text{ eV}, \quad (11.94)$$

while the experimental value is -78.8 eV . Clearly, this discrepancy is essentially due to our omission of the repulsion arising from the interaction of two electrons.

Choice of trial wave function

The choice of trial wave function is critical in variational Monte Carlo calculations. How to choose it is however a highly non-trivial task. All observables are evaluated with respect to the probability distribution

$$P(\mathbf{R}) = \frac{|\psi_T(\mathbf{R})|^2}{\int |\psi_T(\mathbf{R})|^2 d\mathbf{R}}. \quad (11.95)$$

generated by the trial wave function. The trial wave function must approximate an exact eigenstate in order that accurate results are to be obtained. Improved trial wave functions also improve the importance sampling, reducing the cost of obtaining a certain statistical accuracy.

Quantum Monte Carlo methods are able to exploit trial wave functions of arbitrary forms. Any wave function that is physical and for which the value, gradient and laplacian of the wave function may be efficiently computed can be used. The power of Quantum Monte Carlo methods lies in the flexibility of the form of the trial wave function.

It is important that the trial wave function satisfies as many known properties of the exact wave function as possible. A good trial wave function should exhibit much of the same features as does the exact wave function. Especially, it should be well-defined at the origin, that is $\Psi(|\mathbf{R}| = 0) \neq 0$, and its derivative at the origin should also be well-defined. One possible guideline in choosing the trial wave function is the use of constraints about the behavior of the wave function when the distance between one electron and the nucleus or two electrons approaches zero. These constraints are the so-called “cusp conditions” and are related to the derivatives of the wave function.

To see this, let us single out one of the electrons in the helium atom and assume that this electron is close to the nucleus, i.e., $r_1 \rightarrow 0$. We assume also that the two electrons are far from each other and that $r_2 \neq 0$. The local energy can then be written as

$$E_L(\mathbf{R}) = \frac{1}{\psi_T(\mathbf{R})} H \psi_T(\mathbf{R}) = \frac{1}{\psi_T(\mathbf{R})} \left(-\frac{1}{2} \nabla_1^2 - \frac{Z}{r_1} \right) \psi_T(\mathbf{R}) + \text{finite terms.} \quad (11.96)$$

Writing out the kinetic energy term in the spherical coordinates of electron 1, we arrive at the following expression for the local energy

$$E_L(R) = \frac{1}{\mathcal{R}_T(r_1)} \left(-\frac{1}{2} \frac{d^2}{dr_1^2} - \frac{1}{r_1} \frac{d}{dr_1} - \frac{Z}{r_1} \right) \mathcal{R}_T(r_1) + \text{finite terms,} \quad (11.97)$$

where $\mathcal{R}_T(r_1)$ is the radial part of the wave function for electron 1. We have also used that the orbital momentum of electron 1 is $l = 0$. For small values of r_1 , the terms which dominate are

$$\lim_{r_1 \rightarrow 0} E_L(R) = \frac{1}{\mathcal{R}_T(r_1)} \left(-\frac{1}{r_1} \frac{d}{dr_1} - \frac{Z}{r_1} \right) \mathcal{R}_T(r_1), \quad (11.98)$$

since the second derivative does not diverge due to the finiteness of Ψ at the origin. The latter implies that in order for the kinetic energy term to balance the divergence in the potential term, we must have

$$\frac{1}{\mathcal{R}_T(r_1)} \frac{d\mathcal{R}_T(r_1)}{dr_1} = -Z, \quad (11.99)$$

implying that

$$\mathcal{R}_T(r_1) \propto e^{-Zr_1}. \quad (11.100)$$

A similar condition applies to electron 2 as well. For orbital momenta $l > 0$ we have (show this!)

$$\frac{1}{\mathcal{R}_T(r)} \frac{d\mathcal{R}_T(r)}{dr} = -\frac{Z}{l+1}. \quad (11.101)$$

Another constraint on the wave function is found for two electrons approaching each other. In this case it is the dependence on the separation r_{12} between the two electrons which has to reflect the correct behavior in the limit $r_{12} \rightarrow 0$. The resulting radial equation for the r_{12} dependence is the same for the electron-nucleus case, except that the attractive Coulomb interaction between the nucleus and the electron is replaced by a repulsive interaction and the kinetic energy term is twice as large. We obtain then

$$\lim_{r_{12} \rightarrow 0} E_L(R) = \frac{1}{\mathcal{R}_T(r_{12})} \left(-\frac{4}{r_{12}} \frac{d}{dr_{12}} + \frac{2}{r_{12}} \right) \mathcal{R}_T(r_{12}), \quad (11.102)$$

with still $l = 0$. This yields the so-called 'cusp'-condition

$$\frac{1}{\mathcal{R}_T(r_{12})} \frac{d\mathcal{R}_T(r_{12})}{dr_{12}} = \frac{1}{2}, \quad (11.103)$$

while for $l > 0$ we have

$$\frac{1}{\mathcal{R}_T(r_{12})} \frac{d\mathcal{R}_T(r_{12})}{dr_{12}} = \frac{1}{2(l+1)}. \quad (11.104)$$

For general systems containing more than two electrons, we have this condition for each electron pair ij .

Based on these consideration, a possible trial wave function which ignores the 'cusp'-condition between the two electrons is

$$\psi_T(\mathbf{R}) = e^{-\alpha(r_1+r_2)}, \quad (11.105)$$

where $r_{1,2}$ are dimensionless radii and α is a variational parameter which is to be interpreted as an effective charge.

A possible trial wave function which also reflects the 'cusp'-condition between the two electrons is

$$\psi_T(\mathbf{R}) = e^{-\alpha(r_1+r_2)} e^{r_{12}/2}. \quad (11.106)$$

The last equation can be generalized to

$$\psi_T(\mathbf{R}) = \phi(\mathbf{r}_1)\phi(\mathbf{r}_2) \dots \phi(\mathbf{r}_N) \prod_{i<j} f(r_{ij}), \quad (11.107)$$

for a system with N electrons or particles. The wave function $\phi(\mathbf{r}_i)$ is the single-particle wave function for particle i , while $f(r_{ij})$ account for more complicated two-body correlations. For the helium atom, we placed both electrons in the hydrogenic orbit $1s$. We know that the ground state for the helium atom has a symmetric spatial part, while the spin wave function is anti-symmetric in order to obey the Pauli principle. In the present case we need not to deal with spin degrees of freedom, since we are mainly trying to reproduce the ground state of the system. However, adopting such a single-particle representation for the individual electrons means that for atoms beyond helium, we cannot continue to place electrons in the lowest hydrogenic orbit. This is a consequence of the Pauli principle, which states that the total wave function for a system of identical particles such as fermions, has to be anti-symmetric. The program we include below can use either Eq. (11.105) or Eq. (11.106) for the trial wave function. One or two electrons can be placed in the lowest hydrogen orbit, implying that the program can only be used for studies of the ground state of hydrogen or helium.

A similar approach can be applied to the cusp resulting from bringing two electrons close together. The mathematical trick is to expand the wave functions in spherical coordinates centered on one of the two electrons. Let us denote the distance between the two electrons as r_{ij} . Remembering that the Coulomb potential is now repulsive and the kinetic term is twice as large (because both electrons give kinetic contributions) we get:

$$\left(2 \frac{d^2}{dr_{ij}^2} + \frac{4}{r_{ij}} \frac{d}{dr_{ij}} + \frac{2}{r_{ij}} - \frac{l(l+1)}{r_{ij}^2} + 2E \right) R_{ij} = 0$$

where l is now equal 0 if the spins of the two electrons are anti-parallel and 1 if they are parallel. Repeating the argument for the electron-nucleus cusp with the factorization of the leading r -dependency, we get the similar cusp condition:

$$\frac{d\mathcal{R}(r_{ij})}{dr_{ij}} = -\frac{1}{2(l+1)} \mathcal{R}(r_{ij}) \quad r_{ij} \rightarrow 0$$

resulting in:

$$R(l=0) \propto \begin{cases} e^{r_{ij}/2} & \text{for anti-parallel spins, } l=0 \\ e^{r_{ij}/4} & \text{for parallel spins, } l=1 \end{cases}$$

We do not have to pay any attention to the exact values of the coefficients in such exponentials to realize that in order to incorporate the proper inter-particle correlations, we need our wave functions to

depend explicitly on the inter-particle distances. Thus we are motivated to extend our product-state based wave functions by a correlation function. Higher order cusp conditions can also be incorporated to deal with the cases where more than two particles bundle up. These would consist of factors dependent on the coordinates of multiplets of particles, both absolute distances and inter-particle distances. For the numerical experiments exemplified in this thesis, we will restrict ourselves to the simplest correlations dependent only on particle pairs.

Several forms of the correlation function exist in the literature and we will mention only a selected few to give the general idea of how they are constructed. A form given by Hylleraas that had great success for the helium atom was the series expansion:

$$H = e^{\epsilon s} \sum_k c_k r^{l_k} s^{m_k} t^{n_k}$$

where the inter-particle separation r_{ij} for simplicity is written as r . In addition $s = r_i + r_j$ and $t = r_i - r_j$ with r_i and r_j being the two electron-nucleus distances. All the other quantities are free parameters. Notice that the cusp condition is satisfied by the exponential. Unfortunately the convergence of this function turned out to be quite slow. For example, to pinpoint the He-energy to the fourth decimal digit a nine term function would suffice. To double the number of digits, one needed almost 1100 terms.

The so called Padé-Jastrow form, however, is more suited for larger systems. It is based on an exponential function with a rational exponent:

$$J = e^U$$

In its general form, U is a potential series expansion on both the absolute particle coordinates r_i and the inter-particle coordinates r_{ij} :

$$U = \sum_{i < j}^N \left(\frac{\sum_k \alpha_k r_i^k}{1 + \sum_k \alpha'_k r_i^k} \right) + \sum_i^N \left(\frac{\sum_k \beta_k r_{ij}^k}{1 + \sum_k \beta'_k r_{ij}^k} \right)$$

A very typical Padé-Jastrow function used for QMC calculations of molecular and atomic systems is:

$$\exp \left(\frac{\alpha r_{ij}}{2(1 + \beta r_{ij})} \right)$$

where α and β are adjustable parameters. Later on, when testing our program, we will employ this Jastrow function together with a Slater determinant of Hartree-Fock optimized Slater-orbitals for a VMC simulation of the four electron beryllium (Be) atom.

11.5.5 Program example for atomic systems

The variational Monte Carlo algorithm consists of two distinct phases. In the first a walker, a single electron in our case, consisting of an initially random set of electron positions is propagated according to the Metropolis algorithm, in order to equilibrate it and begin sampling. In the second phase, the walker continues to be moved, but energies and other observables are also accumulated for later averaging and statistical analysis. In the program below, the electrons are moved individually and not as a whole configuration. This improves the efficiency of the algorithm in larger systems, where configuration moves require increasingly small steps to maintain the acceptance ratio.

The main part of the code contains calls to various functions, setup and declarations of arrays etc. Note that we have defined a fixed step length h for the numerical computation of the second derivative of the kinetic energy. Furthermore, we perform the Metropolis test when we have moved all electrons. This should be compared to the case where we move one electron at the time and perform the Metropolis test. The latter is similar to the algorithm for the Ising model discussed in the previous chapter. A more detailed discussion and better statistical treatments and analyses are discussed in chapters 18 and 16.

<http://www.fys.uio.no/compphys/cp/programs/FYS3150/chapter11/cpp/program1.cpp>

```
// Variational Monte Carlo for atoms with up to two electrons
#include <iostream>
#include <fstream>
#include <iomanip>
#include "lib.h"
using namespace std;
// output file as global variable
ofstream ofile;
// the step length and its squared inverse for the second derivative
#define h 0.001
#define h2 1000000

// declaraton of functions

// Function to read in data from screen, note call by reference
void initialise(int&, int&, int&, int&, int&, int&, double&);

// The Mc sampling for the variational Monte Carlo
void mc_sampling(int, int, int, int, int, int, double, double *, double *);

// The variational wave function
double wave_function(double **, double, int, int);

// The local energy
double local_energy(double **, double, double, int, int, int);

// prints to screen the results of the calculations
void output(int, int, int, double *, double *);

// Begin of main program

//int main()
int main(int argc, char* argv[])
{
    char *outfilename;
    int number_cycles, max_variations, thermalization, charge;
    int dimension, number_particles;
    double step_length;
    double *cumulative_e, *cumulative_e2;

    // Read in output file, abort if there are too few command-line arguments
    if( argc <= 1 ){
        cout << "Bad Usage: " << argv[0] <<
```

```
        " read also output file on same line" << endl;
    exit(1);
}
else{
    outfile=argv[1];
}
ofile.open(outfile);
// Read in data
initialise(dimension, number_particles, charge,
           max_variations, number_cycles,
           thermalization, step_length);
cumulative_e = new double[max_variations+1];
cumulative_e2 = new double[max_variations+1];

// Do the mc sampling
mc_sampling(dimension, number_particles, charge,
            max_variations, thermalization,
            number_cycles, step_length, cumulative_e, cumulative_e2);
// Print out results
output(max_variations, number_cycles, charge, cumulative_e, cumulative_e2)
;
delete [] cumulative_e; delete [] cumulative_e2;
ofile.close(); // close output file
return 0;
}
```

The implementation of the brute force Metropolis algorithm is shown in the next function. Here we have a loop over the variational variables α . It calls two functions, one to compute the wave function and one to update the local energy.

```
// Monte Carlo sampling with the Metropolis algorithm

void mc_sampling(int dimension, int number_particles, int charge,
                int max_variations,
                int thermalization, int number_cycles, double step_length,
                double *cumulative_e, double *cumulative_e2)
{
    int cycles, variate, accept, dim, i, j;
    long idum;
    double wfnew, wfold, alpha, energy, energy2, delta_e;
    double **r_old, **r_new;
    alpha = 0.5*charge;
    idum=-1;
    // allocate matrices which contain the position of the particles
    r_old = (double **) matrix( number_particles, dimension, sizeof(double));
    r_new = (double **) matrix( number_particles, dimension, sizeof(double));
    for (i = 0; i < number_particles; i++) {
        for (j=0; j < dimension; j++) {
            r_old[i][j] = r_new[i][j] = 0;
        }
    }
    // loop over variational parameters
    for (variate=1; variate <= max_variations; variate++){
```

```

// initialisations of variational parameters and energies
alpha += 0.1;
energy = energy2 = 0; accept =0; delta_e=0;
// initial trial position, note calling with alpha
// and in three dimensions
for (i = 0; i < number_particles; i++) {
    for (j=0; j < dimension; j++) {
        r_old[i][j] = step_length*(ran1(&idum)-0.5);
    }
}
wfold = wave_function(r_old, alpha, dimension, number_particles);
// loop over monte carlo cycles
for (cycles = 1; cycles <= number_cycles+thermalization; cycles++){
    // new position
    for (i = 0; i < number_particles; i++) {
        for (j=0; j < dimension; j++) {
            r_new[i][j] = r_old[i][j]+step_length*(ran1(&idum)-0.5);
        }
    }
    wfnew = wave_function(r_new, alpha, dimension, number_particles);
    // Metropolis test
    if(ran1(&idum) <= wfnew*wfnew/wfold/wfold ) {
        for (i = 0; i < number_particles; i++) {
            for (j=0; j < dimension; j++) {
                r_old[i][j]=r_new[i][j];
            }
        }
        wfold = wfnew;
        accept = accept+1;
    }
    // compute local energy
    if ( cycles > thermalization ) {
        delta_e = local_energy(r_old, alpha, wfold, dimension,
                               number_particles, charge);

        // update energies
        energy += delta_e;
        energy2 += delta_e*delta_e;
    }
} // end of loop over MC trials
cout << "variational parameter= " << alpha
     << " accepted steps= " << accept << endl;
// update the energy average and its squared
cumulative_e[variate] = energy/number_cycles;
cumulative_e2[variate] = energy2/number_cycles;

} // end of loop over variational steps
free_matrix((void **) r_old); // free memory
free_matrix((void **) r_new); // free memory
} // end mc_sampling function

```

The wave function is in turn defined in the next function. Here we limit ourselves to a function which consists only of the product of single-particle wave functions.

```

// Function to compute the squared wave function, simplest form

double wave_function(double **r, double alpha, int dimension, int
    number_particles)
{
    int i, j, k;
    double wf, argument, r_single_particle, r_12;

    argument = wf = 0;
    for (i = 0; i < number_particles; i++) {
        r_single_particle = 0;
        for (j = 0; j < dimension; j++) {
            r_single_particle += r[i][j]*r[i][j];
        }
        argument += sqrt(r_single_particle);
    }
    wf = exp(-argument*alpha) ;
    return wf;
}

```

Finally, the local energy is computed using a numerical derivation for the kinetic energy. We use the familiar expression derived in Eq. (3.4), that is

$$f_0'' = \frac{f_h - 2f_0 + f_{-h}}{h^2},$$

in order to compute

$$-\frac{1}{2\psi_T(\mathbf{R})}\nabla^2\psi_T(\mathbf{R}). \quad (11.108)$$

The variable h is a chosen step length. For helium, since it is rather easy to evaluate the local energy, the above is an unnecessary complication. However, for many-electron or other many-particle systems, the derivation of an analytic expression for the kinetic energy can be quite involved, and the numerical evaluation of the kinetic energy using Eq. (3.4) may result in a simpler code and/or even a faster one.

```

// Function to calculate the local energy with num derivative

double local_energy(double **r, double alpha, double wfold, int dimension,
    int number_particles, int charge)
{
    int i, j, k;
    double e_local, wfminus, wfplus, e_kinetic, e_potential, r_12,
        r_single_particle;
    double **r_plus, **r_minus;

    // allocate matrices which contain the position of the particles
    // the function matrix is defined in the progam library
    r_plus = (double **) matrix( number_particles, dimension, sizeof(double));
    r_minus = (double **) matrix( number_particles, dimension, sizeof(double));
    ;
    for (i = 0; i < number_particles; i++) {
        for (j=0; j < dimension; j++) {
            r_plus[i][j] = r_minus[i][j] = r[i][j];
        }
    }
}

```

```

}
// compute the kinetic energy
e_kinetic = 0;
for (i = 0; i < number_particles; i++) {
    for (j = 0; j < dimension; j++) {
        r_plus[i][j] = r[i][j]+h;
        r_minus[i][j] = r[i][j]-h;
        wfminus = wave_function(r_minus, alpha, dimension, number_particles);
        wfplus = wave_function(r_plus, alpha, dimension, number_particles);
        e_kinetic -= (wfminus+wfplus-2*wfold);
        r_plus[i][j] = r[i][j];
        r_minus[i][j] = r[i][j];
    }
}
// include electron mass and hbar squared and divide by wave function
e_kinetic = 0.5*h2*e_kinetic/wfold;
// compute the potential energy
e_potential = 0;
// contribution from electron-proton potential
for (i = 0; i < number_particles; i++) {
    r_single_particle = 0;
    for (j = 0; j < dimension; j++) {
        r_single_particle += r[i][j]*r[i][j];
    }
    e_potential -= charge/sqrt(r_single_particle);
}
// contribution from electron-electron potential
for (i = 0; i < number_particles-1; i++) {
    for (j = i+1; j < number_particles; j++) {
        r_12 = 0;
        for (k = 0; k < dimension; k++) {
            r_12 += (r[i][k]-r[j][k])*(r[i][k]-r[j][k]);
        }
        e_potential += 1/sqrt(r_12);
    }
}
}
free_matrix((void **) r_plus); // free memory
free_matrix((void **) r_minus);
e_local = e_potential+e_kinetic;
return e_local;
}

```

The remaining part of the program consists of the output and initialize functions and is not listed here.

The way we have rewritten Schrödinger's equation results in energies given in atomic units. If we wish to convert these energies into more familiar units like electronvolt (eV), we have to multiply our results with $2E_0$ where $E_0 = 13.6$ eV, the binding energy of the hydrogen atom. Using Eq. (11.105) for the trial wave function, we obtain an energy minimum at $\alpha \approx 1.75$. The ground state is $E = -2.85$ in atomic units or $E = -77.5$ eV. The experimental value is -78.8 eV. Obviously, improvements to the wave function such as including the 'cusp'-condition for the two electrons as well, see Eq. (11.106), could improve our agreement with experiment. Such an implementation is the topic for the next project.

We note that the effective charge is less than the charge of the nucleus. We can interpret this reduction as an effective way of incorporating the repulsive electron-electron interaction. Finally, since we do not

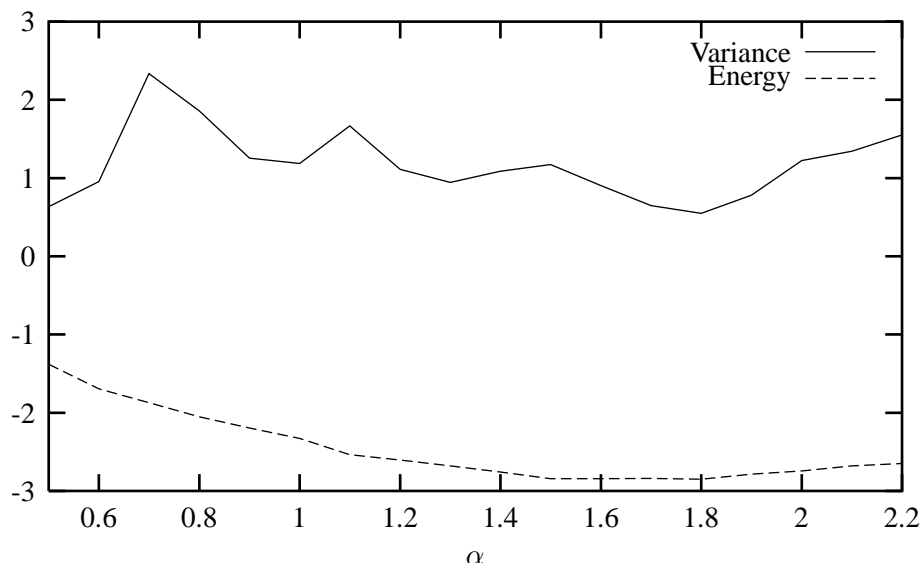


Figure 11.3: Result for ground state energy of the helium atom using Eq. (11.105) for the trial wave function. The variance is also plotted. A total of 100000 Monte Carlo moves were used with a step length of 2 Bohr radii.

have the exact wave function, we see from Fig. 11.3 that the variance is not zero at the energy minimum. Techniques such as importance sampling, to be contrasted to the brute force Metropolis sampling used here, and various optimization techniques of the variance and the energy, will be discussed under advanced topics, see chapter 18.

11.5.6 Helium and beyond

We need to establish some rules regarding the construction of physically reliable wave-functions for systems with more than one electron. The *Pauli principle*, after Wolfgang Pauli states that *The total wave function must be antisymmetric under the interchange of any pair of identical fermions and symmetric under the interchange of any pair of identical bosons.*

A result of the Pauli principle is the so-called Pauli exclusion principle which that *no two electrons can occupy the same state.*

Overall wave functions that satisfy the Pauli principle are often written as *Slater Determinants.*

The Slater Determinant

For the helium atom we assumed that the two electrons were both in the $1s$ state. This fulfills the Pauli exclusion principle as the two electrons in the ground state have different intrinsic spin. However, the wave function we discussed above was not antisymmetric with respect to an interchange of the different electrons. This is not totally true as we only included the spatial part of the wave function. For the helium ground state the spatial part of the wave function is symmetric and the spin part is antisymmetric. The product is therefore antisymmetric as well. The Slater-determinant consists of single-particle *spin-orbitals*; joint spin-space states of the electrons

$$\Psi_{1s}^\uparrow(1) = \Psi_{1s}(1) \uparrow(1),$$

and similarly

$$\Psi_{1s}^\downarrow(2) = \Psi_{1s}(2) \downarrow(2).$$

Here the two spin functions are given by

$$\uparrow(I) = \begin{cases} 1 & \text{if } m_s(I) = \frac{1}{2} \\ 0 & \text{if } m_s(I) = -\frac{1}{2} \end{cases},$$

and

$$\downarrow(I) = \begin{cases} 0 & \text{if } m_s(I) = \frac{1}{2} \\ 1 & \text{if } m_s(I) = -\frac{1}{2} \end{cases}, \quad (11.109)$$

with $I = 1, 2$. The ground state can then be expressed by the following determinant

$$\Psi(1, 2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \Psi_{1s}(1) \uparrow(1) & \Psi_{1s}(2) \uparrow(2) \\ \Psi_{1s}(1) \downarrow(1) & \Psi_{1s}(2) \downarrow(2) \end{vmatrix}.$$

This is an example of a *Slater determinant*. This determinant is antisymmetric since particle interchange is identical to an interchange of the two columns. For the ground state the spatial wave-function is symmetric. Therefore we simply get

$$\Psi(1, 2) = \Psi_{1s}(1)\Psi_{1s}(2) [\uparrow(1) \downarrow(2) - \uparrow(2) \downarrow(1)].$$

The spin part of the wave-function is here antisymmetric. This has no effect when calculating physical observables because the sign of the wave function is squared in all expectation values.

The general form of a Slater determinant composed of n orthonormal orbitals $\{\phi_i\}$ is

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \dots & \phi_1(N) \\ \phi_2(1) & \phi_2(2) & \dots & \phi_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(1) & \phi_N(2) & \dots & \phi_N(N) \end{vmatrix}. \quad (11.110)$$

The introduction of the Slater determinant is very important for the treatment of many-body systems, and is the principal building block for various variational wave functions. As long as we express the wave-function in terms of either one Slater determinant or a linear combination of several Slater determinants, the Pauli principle is satisfied. When constructing many-electron wave functions this picture provides an easy way to include many of the physical features. One problem with the Slater matrix is that it is computationally demanding. Limiting the number of calculations will be one of the most important issues concerning the implementation of the Slater determinant. This will be discussed in detail in chapter 18. Chapter 18 is dedicated to the discussion of advanced many-body methods for solving Schrödinger's equation.

The $1s$ hydrogen like wave function

$$R_{10}(r) = 2 \left(\frac{Z}{a_0} \right)^{3/2} \exp(-Zr/a_0) = u_{10}/r$$

The total energy for helium (not the Hartree or Fock terms) from the direct and the exchange term should give $5Z/8$.

The single-particle energy with no interactions should give $-Z^2/2n^2$.
 The $2s$ hydrogen-like wave function is

$$R_{20}(r) = 2 \left(\frac{Z}{2a_0} \right)^{3/2} \left(1 - \frac{Zr}{2a_0} \right) \exp(-Zr/2a_0) = u_{20}/r$$

and the $2p$ hydrogen-like wave function is

$$R_{21}(r) = \frac{1}{\sqrt{3}} \left(\frac{Z}{2a_0} \right)^{3/2} \frac{Zr}{a_0} \exp(-Zr/2a_0) = u_{21}/r$$

We use $a_0 = 1$.

If we compute the total energy of the helium atom with the function

$$R_{10}(r) = 2 \left(\frac{Z}{a_0} \right)^{3/2} \exp(-Zr/a_0) = u_{10}/r,$$

as a trial single-particle wave function, we obtain a total energy (one-body and two-body)

$$E[Z] = Z^2 - 4Z + \frac{5}{8}Z.$$

The minimum is not at $Z = 2$. Take the derivative wrt Z and we find that the minimum is at

$$Z = 2 - \frac{5}{16} = 1.6875$$

and represents an optimal effective charge.

11.6 The H_2^+ molecule

The H_2^+ molecule consists of two protons and one electron, with binding energy $E_B = -2.8$ eV and an equilibrium position $r_0 = 0.106$ nm between the two protons.

We define our system through the following variables. The electron is at a distance \mathbf{r} from a chosen origin, one of the protons is at the distance $-\mathbf{R}/2$ while the other one is placed at $\mathbf{R}/2$ from origin, resulting in a distance to the electron of $\mathbf{r} - \mathbf{R}/2$ and $\mathbf{r} + \mathbf{R}/2$, respectively.

In our solution of Schrödinger's equation for this system we are going to neglect the kinetic energies of the protons, since they are 2000 times heavier than the electron. We assume thus that their velocities are negligible compared to the velocity of the electron. In addition we omit contributions from nuclear forces, since they act at distances of several orders of magnitude smaller than the equilibrium position.

We can then write Schrödinger's equation as follows

$$\left\{ -\frac{\hbar^2 \nabla_{\mathbf{r}}^2}{2m_e} - \frac{ke^2}{|\mathbf{r} - \mathbf{R}/2|} - \frac{ke^2}{|\mathbf{r} + \mathbf{R}/2|} + \frac{ke^2}{R} \right\} \psi(\mathbf{r}, \mathbf{R}) = E\psi(\mathbf{r}, \mathbf{R}), \quad (11.111)$$

where the first term is the kinetic energy of the electron, the second term is the potential energy the electron feels from the proton at $-\mathbf{R}/2$ while the third term arises from the potential energy contribution from the proton at $\mathbf{R}/2$. The last term arises due to the repulsion between the two protons. In Fig. 11.4 we show a plot of the potential energy

$$V(\mathbf{r}, \mathbf{R}) = -\frac{ke^2}{|\mathbf{r} - \mathbf{R}/2|} - \frac{ke^2}{|\mathbf{r} + \mathbf{R}/2|} + \frac{ke^2}{R}. \quad (11.112)$$

Here we have fixed $|\mathbf{R}| = 2a_0$ og $|\mathbf{R}| = 8a_0$, being 2 and 8 Bohr radii, respectively. Note that in the region between $|\mathbf{r}| = -|\mathbf{R}|/2$ (units are r/a_0 in this figure, with $a_0 = 0.0529$) and $|\mathbf{r}| = |\mathbf{R}|/2$ the electron can tunnel through the potential barrier. Recall that $-\mathbf{R}/2$ og $\mathbf{R}/2$ correspond to the positions of the two protons. We note also that if R is increased, the potential becomes less attractive. This has consequences for the binding energy of the molecule. The binding energy decreases as the distance \mathbf{R} increases. Since the potential is symmetric with respect to the interchange of $\mathbf{R} \rightarrow -\mathbf{R}$ and $\mathbf{r} \rightarrow -\mathbf{r}$ it

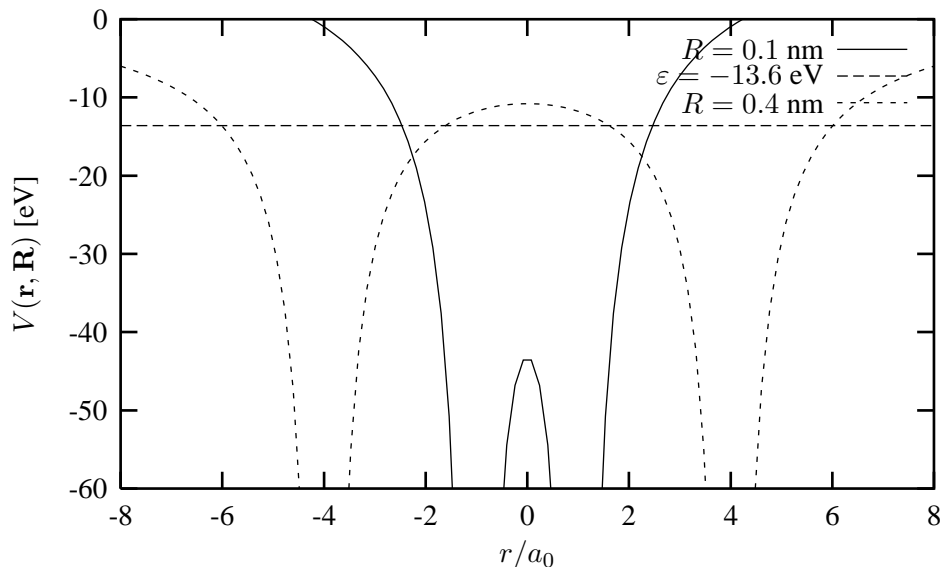


Figure 11.4: Plot of $V(r, R)$ for $|\mathbf{R}|=0.1$ and 0.4 nm. Units along the x -axis are r/a_0 . The straight line is the binding energy of the hydrogen atom, $\varepsilon = -13.6$ eV.

means that the probability for the electron to move from one proton to the other must be equal in both directions. We can say that the electron shares it's time between both protons.

With this caveat, we can now construct a model for simulating this molecule. Since we have only one electron, we could assume that in the limit $R \rightarrow \infty$, i.e., when the distance between the two protons is large, the electron is essentially bound to only one of the protons. This should correspond to a hydrogen atom. As a trial wave function, we could therefore use the electronic wave function for the ground state of hydrogen, namely

$$\psi_{100}(r) = \left(\frac{1}{\pi a_0^3} \right)^{1/2} e^{-r/a_0}. \quad (11.113)$$

Since we do not know exactly where the electron is, we have to allow for the possibility that the electron can be coupled to one of the two protons. This form includes the 'cusp'-condition discussed in the previous section. We define thence two hydrogen wave functions

$$\psi_1(\mathbf{r}, \mathbf{R}) = \left(\frac{1}{\pi a_0^3} \right)^{1/2} e^{-|\mathbf{r}-\mathbf{R}/2|/a_0}, \quad (11.114)$$

and

$$\psi_2(\mathbf{r}, \mathbf{R}) = \left(\frac{1}{\pi a_0^3} \right)^{1/2} e^{-|\mathbf{r}+\mathbf{R}/2|/a_0}. \quad (11.115)$$

Based on these two wave functions, which represent where the electron can be, we attempt at the following linear combination

$$\psi_{\pm}(\mathbf{r}, \mathbf{R}) = C_{\pm} (\psi_1(\mathbf{r}, \mathbf{R}) \pm \psi_2(\mathbf{r}, \mathbf{R})), \quad (11.116)$$

with C_{\pm} a constant. Based on this discussion, we add a second electron in order to simulate the H_2 molecule. That is the topic for project 11.3.

11.7 Improved variational calculations

To be finished for spring 2010.

11.7.1 Importance sampling

As mentioned in connection with the generation of random numbers, sequential correlation must be given thorough attention as it may lead to bad error estimates of our numerical results.

There are several things we need to keep in mind in order to keep the correlation low. First of all, the transition acceptance must be kept as high as possible. Otherwise, a walker will dwell at the same spot in state space for several iterations at a time, which will clearly lead to high correlation between nearby succeeding measurements.

Secondly, when using the simple symmetric form of $\omega(\mathbf{x}_{\text{old}}, \mathbf{x}_{\text{new}})$, one has to keep in mind the random walk nature of the algorithm. Transitions will be made between points that are relatively close to each other in state space, which also clearly contributes to increase correlation. The seemingly obvious way to deal with this would be just to increase the step size, allowing the walkers to cover more of the state space in fewer steps (thus requiring fewer steps to reach ergodicity). But unfortunately, long before the step length becomes desirably large, the algorithm breaks down. When proposing moves symmetrically and uniformly around \mathbf{x}_{old} , the step acceptance becomes directly dependent on the step length in such a way that a too large step length reduces the acceptance. The reason for this is very simple. As the step length increases, a walker will more likely be given a move proposition to areas of very low probability, particularly if the governing trial wave function describes a localized system. In effect, the effective movement of the walkers again becomes too small, resulting in large correlation. For optimal results we therefore have to balance the step length with the acceptance.

With a transition suggestion rule ω as simple as the uniform symmetrical one emphasized so far, the usual rule of thumb is to keep the acceptance around 0.5. But the optimal interval varies a lot from case to case. We therefore have to treat each numerical experiment with care.

By choosing a better ω , we can still improve the efficiency of the step length versus acceptance. Recall that ω may be chosen arbitrarily as long as it fulfills ergodicity, meaning that it has to allow the walker to reach any point of the state space in a finite number of steps. What we basically want is an ω that pushes the ratio towards unity, increasing the acceptance. The theoretical situation of ω exactly equal to p itself:

$$\omega(\mathbf{x}_{\text{new}}, \mathbf{x}_{\text{old}}) = \omega(\mathbf{x}_{\text{new}}) = p(\mathbf{x}_{\text{new}})$$

would give the maximal acceptance of 1. But then we would already have solved the problem of producing points distributed according to p . One typically settles on modifying the symmetrical ω so that the walkers move more towards areas of the state space where the distribution is large. One such procedure is the Fokker-Planck formalism where the walkers are moved according to the gradient of the distribution. The formalism “pushes” the walkers in a “desirable” direction. The idea is to propose moves similarly to an isotropic diffusion process with a drift. A new position \mathbf{x}_{new} is calculated from the old one, \mathbf{x}_{old} , as

follows:

$$\mathbf{x}_{\text{new}} = \mathbf{x}_{\text{old}} + \chi + D\mathbf{F}(\mathbf{x}_{\text{old}})\delta t \quad (11.117)$$

Here χ is a Gaussian pseudo-random number with mean equal zero and variance equal $2D\delta t$. It accounts for the diffusion part of the transition. The third term on the left hand side accounts for the drift. \mathbf{F} is a drift velocity dependent on the position of the walker and is derived from the quantum mechanical wave function ψ . The constant D , being the diffusion constant of χ , also adjusts the size of the drift. δt is a time step parameter whose presence will be clarified shortly.

It can be shown that the ω corresponding to the move proposition rule in Eq. (11.117) becomes (in non-normalized form):

$$\omega(\mathbf{x}_{\text{old}}, \mathbf{x}_{\text{new}}) = \exp\left(-\frac{(\mathbf{x}_{\text{new}} - \mathbf{x}_{\text{old}} - D\delta t\mathbf{F}(\mathbf{x}_{\text{old}}))^2}{4D\delta t}\right) \quad (11.118)$$

which, as expected, is a Gaussian with variance $2D\delta t$ centered slightly off \mathbf{x}_{old} due to the drift term $D\mathbf{F}(\mathbf{x}_{\text{old}})\delta t$.

What is the optimal choice for the drift term? From statistical mechanics we know that a simple isotropic drift diffusion process obeys a Fokker-Planck equation of the form:

$$\frac{\partial f}{\partial t} = \sum_i D \frac{\partial}{\partial x_i} \left(\frac{\partial}{\partial x_i} - F_i(\mathbf{F}) \right) f \quad (11.119)$$

where f is the continuous distribution of walkers. Equation (11.117) is a discretized realization of such a process where δt is the discretized time step. In order for the solution f to converge to the desired distribution p , it can be shown that the drift velocity has to be chosen as follows:

$$\mathbf{F} = \frac{1}{f} \nabla f$$

where the operator ∇ is the vector of first derivatives of all spatial coordinates. Convergence for such a diffusion process is only guaranteed when the time step approaches zero. But in the Metropolis algorithm, where drift diffusion is used just as a transition proposition rule, this bias is corrected automatically by the rejection mechanism. In our application, the desired PDF being the square absolute of the wave function, $f = |\psi|^2$, the drift velocity becomes:

$$\mathbf{F} = 2 \frac{1}{\psi} \nabla \psi \quad (11.120)$$

As expected, the walker is “pushed” along the gradient of the wave function.

When dealing with many-particle systems, as those that will be studied in this thesis, we should also consider whether to move only one particle at a time at each transition or all at once. The former method may often be more efficient. A movement of only one particle will restrict the accessible space a walker can move to in a single transition even more, thus introducing correlation. But on the other hand, the acceptance is increased so that each particle can be moved further than it could in a standard all-particle move. It is also computationally far more efficient to do one-particle transitions particularly when dealing with complicated distributions governing many-dimensional anti-symmetrical fermionic systems.

Alternatively, we can treat the sequence of all one-particle transitions as one total transition of all particles. This gives a larger effective step length thus reducing the correlation. From a computational point of view, we may not gain any speed by summing up the individual one-particle transitions as opposed to doing an all-particle transition. But the reduced correlation increases the total efficiency. We are able to do fewer calculations in order to reach the same numerical accuracy.

Another way to acquire some control over the correlation is to do a so called blocking procedure on our set of numerical measurements.

11.7.2 Guiding Functions

Evaluation of the energy functional can be further improved by a transformation similar to the one in the importance sampling scheme. Consider the mean of a function f over a PDF p and let p_g be a function that resembles p but is computationally less expensive. We can then estimate the mean of f as follows:

$$\langle f \rangle = \frac{\int p(\mathbf{x})f(\mathbf{x})d\mathbf{x}}{\int p(\mathbf{x})d\mathbf{x}} = \frac{\int p_g(\mathbf{x})\omega(\mathbf{x})f(\mathbf{x})d\mathbf{x}}{\int p_g(\mathbf{x})\omega(\mathbf{x})d\mathbf{x}} \approx \frac{\sum_{\mathbf{x} \in p_g} \omega(\mathbf{x})f(\mathbf{x})}{\sum_{\mathbf{x} \in p_g} \omega(\mathbf{x})}$$

where

$$\omega(\mathbf{x}) \equiv \frac{p(\mathbf{x})}{p_g(\mathbf{x})}$$

is a weight converting the distribution from p_g to p . Similarly as with regular importance sampling described above, we need to accumulate f reweighted by ω . Note that the normalization factor has to be reweighted as well so that we actually need to evaluate two integrals. Fortunately this does not pose any significant efficiency loss. The only extra thing we need to do is to accumulate the value of ω itself, which we have to evaluate anyway.

We might worry that the estimate (the rightmost part of the above expression) actually represents a normalized integration while the exact expressions do not contain any normalization constants. This is needless to worry about because the normalization constant of the numerator is the same as that of the denominator, namely $\int p_g(\mathbf{x})d\mathbf{x}$. It would cancel from the exact expressions anyway.

Guiding functions become especially useful in conjunction with the Metropolis algorithm. There we wished to keep the sequential correlation as low as possible by increasing the effective step size. This can be done by keeping only every n th point that is produced in the random walk. But such a method is most efficient when a guiding function Ψ_g mimicking the trial wave function Ψ is used to generate the actual walk. A reweighted local energy ωE_L is then used to obtain the correct energy average. If Ψ_g is sufficiently less expensive to evaluate than Ψ , then the extra time used on the rejected points is compensated for by the increased evaluation speed.

To see the benefit of using a guiding function, the correlated sample error is

$$\text{err}^2 = \frac{\text{Var}(x)}{n} \tau = \frac{\text{Var}(x)}{n} \left(1 + 2 \sum_{d=1}^m \kappa_d \right)$$

where τ is the correlation time. Assuming m equals at least a correlation length, we can cut off the sum at m , since correlations exceeding this distance are expected to give little contribution to τ . Now using a guiding function and keeping only every m th point we reduce the number of points from n to n/m but then we also reduce τ to 1, so that:

$$\text{err}_g^2 = \frac{m \text{Var}(x)}{n}$$

The ratio of these two error estimates is:

$$\frac{\text{err}_g^2}{\text{err}^2} = \frac{m}{1 + 2 \sum_{d=1}^m \kappa_d}$$

Because κ_d decreases exponentially starting at $\kappa_0 = 1$, the sum in the denominator must yield a value less than m . A reasonable estimate of $\kappa_d \approx 1/2$ for all d up to m gives that the ratio $\text{err}_g^2/\text{err}^2 \approx 1$. It appears that we have not gained any efficiency. But if p_g can be evaluated d times as fast as p , then we can produce d times as many points in the same time, thus reducing the error by a factor \sqrt{d} .

11.7.3 Energy and Variance Minimization

There are two main approaches for optimizing a trial wave function Ψ_T . By the variational principle we can do a straight forward minimization of the energy functional on the space of the parameters of Ψ_T . This means minimizing the estimate of the mean of the local energy. Because of its simplicity, such a scheme seems to be extremely powerful. The energy minimum is sought out by generating a grid on the space of variational parameters and estimating the functional on each point of this grid. A suitable minimization algorithm, like the steepest descent or the Newton method, can then be applied. Unfortunately this approach is often haunted by problems of instability complicating the minimization algorithms that usually rely on very accurate estimates of energy differences.

If the number of integration points is kept constant, then the further away we move from the optimal point on the parameter grid, the greater the variance of E_L becomes. This greatly hampers a proper estimation of the energy minimum since the local energy means fluctuate. One possible way to work around the instabilities is by the method of correlated sampling, to be described shortly.

The second approach for optimizing the trial wave function is by minimizing the variance. Such an approach is often preferred to the minimization of the energy since the variance has an a priori known lower bound, namely zero, as opposed to the energy itself which is unknown. Also, the variance vanishes for all eigenstates, not just the ground state, which is advantageous when one wishes to optimize excited states. In addition, the variance generally shows higher stability, thus being more easily applicable to minimization algorithms.

The usual strategy is to combine both energy and variance minimization. What we often notice is that the minimum of the variance does not coincide exactly with the minimum of the energy, unless we are actually able to reproduce the exact solution, in which case the two minima must coincide.

This discrepancy may in part be explained by realizing that the sample variance that we attempt to minimize is only an estimate of the true theoretical variance in that we are unable to obtain an exact value of the mean of the local energy. Let δE be the small difference between $\langle E_L \rangle_{\{\mathbf{x}_i\}}$ and $\langle E_L \rangle$. We can then write the estimate of the variance as follows:

$$\begin{aligned} \frac{1}{n} \sum_{\mathbf{x}_i \in p} (E_L - \langle E_L \rangle_{\{\mathbf{x}_i\}})^2 &= \frac{1}{n} \sum_{\mathbf{x}_i \in p} (E_L - \langle E_L \rangle - \delta E)^2 \\ &= \frac{1}{n} \sum_{\mathbf{x}_i \in p} \left((E_L - \langle E_L \rangle)^2 - 2\delta E (E_L - \langle E_L \rangle) + \delta E^2 \right) \\ &= \frac{1}{n} \sum_{\mathbf{x}_i \in p} (E_L - \langle E_L \rangle)^2 - 2\delta E (\langle E_L \rangle_{\{\mathbf{x}_i\}} - \langle E_L \rangle) + \delta E^2 \end{aligned}$$

Focusing our attention on the last line, we see that the first term is the one that we intend to minimize. The last term is just a constant that does not influence the position of the minimum. The middle term, however, is linearly dependent on the local energy thus shifting the minimum of the estimate from the true minimum. Of course, increasing the computational effort by sampling more points should make $\langle E_L \rangle_{\{\mathbf{x}_i\}}$ approach $\langle E_L \rangle$, reducing this effect.

Instead of the typical estimate of the variance one often substitutes $\langle E_L \rangle_{\{\mathbf{x}_i\}}$ by a reference energy E_{ref} :

$$\frac{1}{n} \sum_{\mathbf{x}_i \in p} (E_L - E_{\text{ref}})^2$$

E_{ref} being as close to the best optimized value of $\langle E_L \rangle$ as possible. E_{ref} can be updated continuously as the optimization converges.

Finally, one always has to be aware of the risk of getting stuck on local minima. Since the energy functional estimate, as noted, is never necessarily well behaved on the parameter space, it is difficult to construct efficient automated minimization algorithms. Most often we need to study the behavior of the energy carefully “by hand” to determine whether our results are trustworthy.

11.7.4 Correlated Sampling

Minimization algorithms are, as mentioned, easily corrupted by the statistical fluctuations of the mean energy estimates. The problem may in part be overcome by applying so called correlated sampling, a procedure involving guiding functions. The method smooths out the mean energy estimates on the variational parameter grid.

The main idea is to choose a particular point on the grid, preferably one that is as close to the true minimum as is possible to achieve. Let this point be represented by the vector α . Now use the trial wave function $|\Psi_\alpha(\mathbf{x})|^2$ for that particular point as a guiding function for the mean energy estimates of all the other points α' (instead of writing the trial wave function as Ψ_T we use $\Psi_{\alpha'}$ to explicitly state which point α' on the parameter grid it belongs to). Estimating the energy mean for all points α' now becomes:

$$\langle E_L \rangle_{\alpha'} = \frac{\int |\Psi_{\alpha'}(\mathbf{x})|^2 E_L(\mathbf{x}) d\mathbf{x}}{\int |\Psi_{\alpha'}(\mathbf{x})|^2 d\mathbf{x}} = \frac{\int |\Psi_\alpha(\mathbf{x})|^2 \omega(\mathbf{x}) E_L(\mathbf{x}) d\mathbf{x}}{\int |\Psi_\alpha(\mathbf{x})|^2 \omega(\mathbf{x}) d\mathbf{x}} \approx \frac{\sum \omega(\mathbf{x}) E_L(\mathbf{x})}{\sum \omega(\mathbf{x})}$$

where the points \mathbf{x} are sampled from the guiding function $|\Psi_\alpha(\mathbf{x})|^2$ and the weighting factor is:

$$\omega(\mathbf{x}) = \frac{|\Psi_{\alpha'}(\mathbf{x})|^2}{|\Psi_\alpha(\mathbf{x})|^2}$$

The trick is now to generate the set $\{\mathbf{x}_i\}$ of integration points sampled from $|\Psi_\alpha(\mathbf{x})|^2$ *only once* and use this set to estimate the means and variances for all the other points as well. By imposing such a statistical dependency we hope that the fluctuations of the different estimates become synchronized so that their graph is smoothed out. We can indeed show that correlated sampling reduces the variance when estimating energy *differences*. These are typically needed by minimization algorithms. We also save computation time since we only need to generate one set of integration points for all the estimates. But we still need to evaluate the weight ω individually for all points on the grid of variational parameters.

11.8 Exercises and projects

Project 11.1: Studies of light Atoms

The aim of this project is to test the variational Monte Carlo applied to light atoms. We will test different trial wave function Ψ_T . The systems we study are atoms consisting of two electrons only, such as the helium atom, Li_{II} and Be_{III} . The atom Li_{II} has two electrons and $Z = 3$ while Be_{III} has $Z = 4$ but still two electrons only. A general ansatz for the trial wave function is

$$\psi_T(\mathbf{R}) = \phi(\mathbf{r}_1)\phi(\mathbf{r}_2)f(r_{12}). \quad (11.121)$$

For all systems we assume that the one-electron wave functions $\phi(\mathbf{r}_i)$ are described by the an electron in the lowest hydrogen orbital $1s$.

The specific trial functions we study are

$$\psi_{T1}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12}) = \exp(-\alpha(r_1 + r_2)), \quad (11.122)$$

where α is the variational parameter,

$$\psi_{T2}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12}) = \exp(-\alpha(r_1 + r_2))(1 + \beta r_{12}), \quad (11.123)$$

with β as a new variational parameter and

$$\psi_{T3}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12}) = \exp(-\alpha(r_1 + r_2)) \exp\left(\frac{r_{12}}{2(1 + \beta r_{12})}\right). \quad (11.124)$$

- a) Find the analytic expressions for the local energy for the above trial wave function for the helium atom. Study the behavior of the local energy with these functions in the limits $r_1 \rightarrow 0$, $r_2 \rightarrow 0$ and $r_{12} \rightarrow 0$.
- b) Compute

$$\langle \hat{\mathbf{H}} \rangle = \frac{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \hat{\mathbf{H}}(\mathbf{R}) \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R})}, \quad (11.125)$$

for the helium atom using the variational Monte Carlo method employing the Metropolis algorithm to sample the different states using the trial wave function $\psi_{T1}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12})$. Compare your results with the analytic expression

$$\langle \hat{\mathbf{H}} \rangle = \frac{\hbar^2}{m_e} \alpha^2 - \frac{27}{32} \frac{e^2}{\pi \epsilon_0} \alpha. \quad (11.126)$$

- c) Use the optimal value of α from the previous point to compute the ground state of the helium atom using the other two trial wave functions $\psi_{T2}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12})$ and $\psi_{T3}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12})$. In this case you have to vary both α and β . Explain briefly which function $\psi_{T1}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12})$, $\psi_{T2}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12})$ and $\psi_{T3}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12})$ is the best.
- d) Use the optimal value for all parameters and all wave functions to compute the expectation value of the mean distance $\langle r_{12} \rangle$ between the two electrons. Comment your results.
- e) We will now repeat point 1c), but we replace the helium atom with the ions Li_{II} and Be_{III} . Perform first a variational calculation using the first ansatz for the trial wave function $\psi_{T1}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12})$ in order to find an optimal value for α . Use then this value to start the variational calculation of the energy for the wave functions $\psi_{T2}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12})$ and $\psi_{T3}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12})$. Comment your results.

Project 11.2: Ground state of He, Be and Ne

The task here is to perform a variational Monte Carlo calculation of the ground state energy of the atoms He, Be and Ne.

- a) Here we limit the attention to He and employ the following trial wave function

$$\psi_T(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_{12}) = \exp(-\alpha(r_1 + r_2)) \exp\left(\frac{r_{12}}{2(1 + \beta r_{12})}\right), \quad (11.127)$$

with α and β as variational parameters. The interaction is

$$V(r_1, r_2) = -\frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}, \quad (11.128)$$

yielding the following hamiltonian for the helium atom

$$\hat{H} = -\frac{\nabla_1^2}{2} - \frac{\nabla_2^2}{2} - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}. \quad (11.129)$$

Your task is to perform a Variational Monte Carlo calculation using the Metropolis algorithm to compute the integral

$$\langle H \rangle = \frac{\int d\mathbf{R} \psi_T^*(\mathbf{R}) H(\mathbf{R}) \psi_T(\mathbf{R})}{\int d\mathbf{R} \psi_T^*(\mathbf{R}) \psi_T(\mathbf{R})}. \quad (11.130)$$

- b) We turn the attention to the ground state energy for the Be atom. In this case the trial wave function is given by

$$\psi_T(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = Det(\phi_1(\mathbf{r}_1), \phi_2(\mathbf{r}_2), \phi_3(\mathbf{r}_3), \phi_4(\mathbf{r}_4)) \prod_{i<j}^4 \exp\left(\frac{r_{ij}}{2(1+\beta r_{ij})}\right), \quad (11.131)$$

where the *Det* is a Slater determinant and the single-particle wave functions are the hydrogen wave functions for the 1s and 2s orbitals. Their form within the variational ansatz is given by

$$\phi_{1s}(\mathbf{r}_i) = e^{-\alpha r_i}, \quad (11.132)$$

and

$$\phi_{2s}(\mathbf{r}_i) = (2 - \alpha r_i) e^{-\alpha r_i/2}. \quad (11.133)$$

Set up the expression for the Slater determinant and perform a variational calculation with α and β as variational parameters.

- c) Now we compute the ground state energy for the Neon atom following the same steps as in a) and b) but with the trial wave function

$$\psi_T(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{10}) = Det(\phi_1(\mathbf{r}_1), \phi_2(\mathbf{r}_2), \dots, \phi_{10}(\mathbf{r}_{10})) \prod_{i<j}^{10} \exp\left(\frac{r_{ij}}{2(1+\beta r_{ij})}\right), \quad (11.134)$$

Set up the expression for the Slater determinant and repeat steps a) and b) including the Slater determinant. The variational parameters are still α and β only. In this case you need to include the 2p wave function as well. It is given as

$$\phi_{2p}(\mathbf{r}_i) = \alpha \mathbf{r}_i e^{-\alpha r_i/2}. \quad (11.135)$$

Observe that $r_i = \sqrt{r_{i_x}^2 + r_{i_y}^2 + r_{i_z}^2}$.

Project 11.3: the H₂ molecule

The H₂ molecule consists of two protons and two electrons with a ground state energy $E = -1.17460$ a.u. and equilibrium distance between the two hydrogen atoms of $r_0 = 1.40$ Bohr radii. We define our systems using the following variables. Origo is chosen to be halfway between the two protons. The distance from proton 1 is defined as $-\mathbf{R}/2$ whereas proton 2 has a distance $\mathbf{R}/2$. Calculations are performed for fixed distances \mathbf{R} between the two protons.

Electron 1 has a distance r_1 from the chose origo, while electron 2 has a distance r_2 . The kinetic energy operator becomes then

$$-\frac{\nabla_1^2}{2} - \frac{\nabla_2^2}{2}. \quad (11.136)$$

The distance between the two electrons is $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$. The repulsion between the two electrons results in a potential energy term given by

$$+ \frac{1}{r_{12}}. \quad (11.137)$$

In a similar way we obtain a repulsive contribution from the interaction between the two protons given by

$$+ \frac{1}{|\mathbf{R}|}, \quad (11.138)$$

where \mathbf{R} is the distance between the two protons. To obtain the final potential energy we need to include the attraction the electrons feel from the protons. To model this, we need to define the distance between the electrons and the two protons. If we model this along a chosen z -akse with electron 1 placed at a distance \mathbf{r}_1 from a chose origo, one proton at $-\mathbf{R}/2$ and the other at $\mathbf{R}/2$, the distance from proton 1 to electron 1 becomes

$$\mathbf{r}_{1p1} = \mathbf{r}_1 + \mathbf{R}/2, \quad (11.139)$$

and

$$\mathbf{r}_{1p2} = \mathbf{r}_1 - \mathbf{R}/2, \quad (11.140)$$

from proton 2. Similarly, for electron 2 we obtain

$$\mathbf{r}_{2p1} = \mathbf{r}_2 + \mathbf{R}/2, \quad (11.141)$$

and

$$\mathbf{r}_{2p2} = \mathbf{r}_2 - \mathbf{R}/2. \quad (11.142)$$

These four distances define the attractive contributions to the potential energy

$$-\frac{1}{r_{1p1}} - \frac{1}{r_{1p2}} - \frac{1}{r_{2p1}} - \frac{1}{r_{2p2}}. \quad (11.143)$$

We can then write the total Hamiltonian as

$$\hat{\mathbf{H}} = -\frac{\nabla_1^2}{2} - \frac{\nabla_2^2}{2} - \frac{1}{r_{1p1}} - \frac{1}{r_{1p2}} - \frac{1}{r_{2p1}} - \frac{1}{r_{2p2}} + \frac{1}{r_{12}} + \frac{1}{|\mathbf{R}|}, \quad (11.144)$$

and if we choose $\mathbf{R} = 0$ we obtain the helium atom.

In this project we will use a trial wave function of the form

$$\psi_T(\mathbf{r}_1, \mathbf{r}_2, \mathbf{R}) = \psi(\mathbf{r}_1, \mathbf{R})\psi(\mathbf{r}_2, \mathbf{R}) \exp\left(\frac{r_{12}}{2(1 + \beta r_{12})}\right), \quad (11.145)$$

with the following trial wave function

$$\psi(\mathbf{r}_1, \mathbf{R}) = (\exp(-\alpha r_{1p1}) + \exp(-\alpha r_{1p2})), \quad (11.146)$$

for electron 1 and

$$\psi(\mathbf{r}_2, \mathbf{R}) = (\exp(-\alpha r_{2p1}) + \exp(-\alpha r_{2p2})). \quad (11.147)$$

The variational parameters are α and β .

One can show that in the limit where all distances approach zero that

$$\alpha = 1 + \exp(-R/\alpha), \quad (11.148)$$

resulting in β as the only variational parameter. The last equation is a non-linear equation which we can solve with for example Newton's method discussed in chapter 5.

- a) Find the local energy as function of R .
- b) Set up an algorithm and write a program which computes the expectation value of $\langle \hat{\mathbf{H}} \rangle$ using the variational Monte Carlo method with a brute force Metropolis sampling. For each inter-proton distance R you must find the parameter β which minimizes the energy. Plot the corresponding energy as function of the distance R between the protons.
- c) Use thereafter the optimal parameter sets to compute the average distance $\langle r_{12} \rangle$ between the electrons where the energy as function of R exhibits its minimum. Comment your results.
- d) We modify now the approximation for the wave functions of electrons 1 and 2 by subtracting the two terms instead of adding up, viz

$$\psi(\mathbf{r}_1, \mathbf{R}) = (\exp(-\alpha r_{1p1}) - \exp(-\alpha r_{1p2})), \quad (11.149)$$

for electron 1

$$\psi(\mathbf{r}_2, \mathbf{R}) = (\exp(-\alpha r_{2p1}) - \exp(-\alpha r_{2p2})), \quad (11.150)$$

for electron 2. Mathematically, this approach is equally viable as the previous one. Repeat your calculations from point b) and see if you can obtain an energy minimum as function of R . Comment your results.