# Variational Monte Carlo methods for quantum dots

Matteo Seclì <mattes@mail.uio.no>

July 8, 2015

# Preface

Write something in the preface. Who inspired this project, why, when....just think about something to write here.

# Introduction

COMPLETE THIS INTRODUCTION WITH A MORE DETAILED EXPLANATION AND INDICATING WHAT THERE'S IN THE FIRST CHAPTER!!!

The aim of this project is to use the Variational Monte Carlo (VMC) method to evaluate the ground state energy, one-body densities, expectation values of the kinetic and potential energies and single-particle energies of quantum dots with $N = 2$ and $N = 6$ electrons, so-called *closed shell systems*. We will begin with two electrons confined in a pure two-dimensional isotropic harmonic oscillator potential, with and without the repulsive interaction, developing a simple trial wave-function and a basic program that runs our calculations. Then we are going to improve the trial wave-function adding the so-called *Jastrow factor*, and finally we are going to extend the entire class of wave-functions and the program itself to a given number of electrons (obviously, for a closed-shell system). After that we are going to improve our sampling, switching from a brute-force approach to the clever *importance sampling*. In addition, we will present a method that implements analytical derivatives for calculating both the energy and the diffusive force of the importance sampling.

# Contents

**Preface** iii

**Introduction** v

**1 What are quantum dots?** 1

1.1 Definition and fabrication processes . . . 1

1.2 The shell-model . . . 2

1.3 Evidences of the shell-model . . . 5

1.4 The Coulomb blockade model . . . 7

1.5 The constant interaction model . . . 10

1.6 The parabolic confinement . . . 16

**2 Theoretical basis of the simulation** 19

2.1 The Variational Principle . . . 19

2.2 Considerations about the physical system . . . 22

2.3 The Metropolis algorithm . . . 24

2.3.1 Brute force sampling . . . 25

2.3.2 Importance sampling and Schödinger equation as a diffusion problem 26

2.3.3 Implementation . . . 30

**3 The 2-electrons system** 33

3.1 The unperturbed system . . . 33

3.2 The complete system . . . 36

3.2.1 Parallelization . . . 39

3.2.2 Importance sampling . . . 41

**4 The 6-electrons system** 45

4.1 $\omega = 0.28$ . . . 47

4.2 $\omega = 0.5$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 49
4.3 $\omega = 1$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 50

**5 The virial theorem** **53**
5.1 The 2-electrons system . . . . . . . . . . . . . . . . . . . . . . . . . . . 54
5.2 The 6-electrons system . . . . . . . . . . . . . . . . . . . . . . . . . . . 57

**6 Analytical derivatives** **59**
6.1 Execution times . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60

**7 Conclusions and perspectives** **63**

**Bibliography** **65**

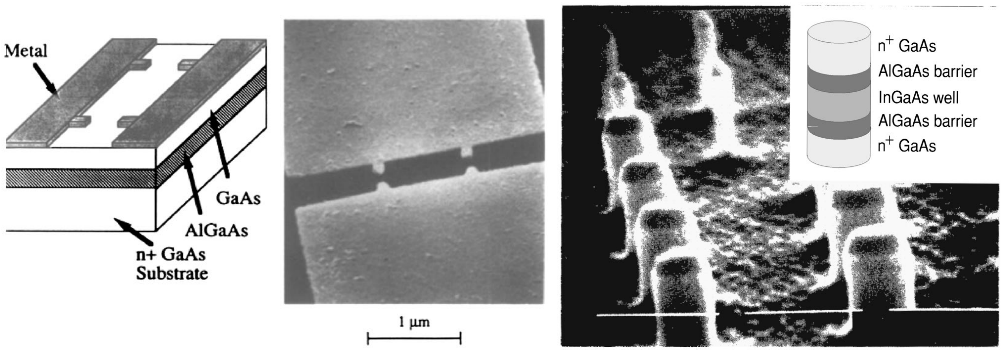
# Chapter 1

# What are quantum dots?

## 1.1 Definition and fabrication processes

The term “quantum dots” refers to finite fermion systems consisting of an artificial 3D confinement of a few electrons, that have a size of only a few hundred angstroms (Kastner, 1993; Reimann and Manninen, 2002).

Such confinement is usually achieved by restricting the two-dimensional electron gas that forms at the interface between two different semiconductor materials (or heterostructure), ether laterally or vertically. Let’s briefly examine these two methods.



*(a) Lateral device structure. From Meirav et al. (1990).*

*(b) Etched quantum dots. The white bars have a length of* 0.5 µm. *From Reed et al. (1988).*

*Figure 1.1: Lateral and vertical quantum dots structures.*

The lateral confinement is obtained with a device like the one in Figure 1.1a. The electron gas forms at the interface between GaAs and AlGaAs, and the confinement is obtained by applying a voltage to the top metal electrodes – called gates. The entire structure is only a few µm thick, so the gates are created by lithographic patterning.

Another common method to fabricate quantum dots is to build heterostructure pillars by etching techniques. Some of these structures are shown in Figure 1.1b the electrodes are at the top and at the bottom of the pillars.

Other fabricating processes include self-growth mechanisms, in which the growth conditions determine the form of the structure (that can be pyramidal, disk shaped or lens shaped), and cleaved-edge overgrowth, that consists in two separate MBE[1] growths on a specific substrate (see Reimann and Manninen, 2002).

Although the definition of quantum dots given above is quite precise, I prefer a more immediate one: "quantum dots are *artificial atoms*". Despite its simplicity, this definition contains a good amount of relevant information. Like natural atoms, in fact, quantum dots are made up of electrons confined in an attractive potential; and as one may guess, they show a similar shell-like structure with its relative *magic numbers* – the resemblance is striking. We begin our discussion by briefly illustrating this simple model.

## 1.2 The shell-model

The so-called *shell-model* is a particularly convenient and simple way to treat a system of interacting particles. This model is based on the *a priori* assumption that the *interacting* particles can be treated as *independent* particles subject to an average *effective potential*. This effective potential is created by the particles themselves, but one can also add an external component to the intrinsic one. Using this simplification it's possible to find a solution for the Schrödinger equation; the result is a certain distribution of single-particle energy levels. In many cases, this distribution it's nonuniform (look, for example, at the simple case of the infinite well) and shows a characteristic bunching of levels; these bunches are called *shells*.

---

[1] Molecular Beam Epitaxy

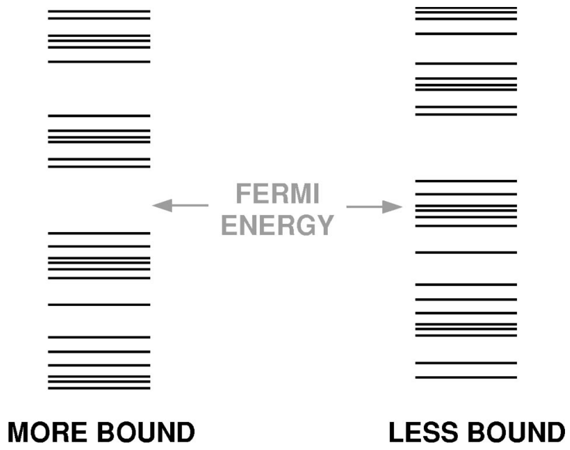


*Figure 1.2: Bunching of the single-particles levels and level density at the Fermi level. From Brack et al. (1972).*

It's very important to have a scheme of such energy levels, since it can immediately gives information about the stability of a system: if the level bunching at the Fermi surface has a minimum, then the system is more bound. In fact, this situation corresponds to particles occupying states with a lower energy (on average), thus minimizing the total energy. In terms of the shell-model, this means that all the shells below the Fermi level are filled (see Figure 1.2). A situation in which the level bunching doesn't have a minimum at the Fermi surface corresponds instead in a non-filled shell. In this case the system can rearrange itself in a non-symmetric shape, in order to reach a better stability.

This has been shown e.g. for atomic nuclei by means of nuclear spectroscopy. In that case, a measurement of the *quadrupole moment* can exploit the possibly non symmetric structure. As a simple example, the quadruple moment $Q$ for a single proton is

$$Q = e \int \psi^* (3z^2 - r^2) \psi \, d\tau, \tag{1.1}$$

where $\tau$ indicates the volume. If $|\psi|^2$ is spherically symmetric, then – on average – $x^2 = y^2 = z^2 = r^2/3$, and substituting $z^2 = r^2/3$ into (1.1) gives zero (see Krane, 1988). So, a non-zero quadrupole moment indicates a non-spherical shape; this shape can be determined by the analysis of the angular distribution of the radiation field.

But how can these deformations give a better stability to the system in the case of non-filled shells? The answer is simple: because they minimize the energy. To show this, we introduce the example of a two-dimensional anisotropic harmonic oscillator confinement of fermions (as in Reimann and Manninen, 2002), in which the effective potential is

$$V(x,y) = \frac{1}{2} m^* \omega^2 \left( \delta x^2 + \frac{1}{\delta} y^2 \right). \tag{1.2}$$

Here $\omega$ is the oscillator frequency, $m^*$ is the effective mass of the fermions and the deformation parameter $\delta \doteqdot \omega_x / \omega_y$ is defined such that $\omega_x = \omega\sqrt{\delta}$ and $\omega_y = \omega/\sqrt{\delta}$. This condition guarantees that the area is conserved with the deformation. By analogy with the simple harmonic oscillator, which has energy levels $E_n = \hbar\omega(n + 1/2)$, the energy levels of the 2D anisotropic case are just

$$E_{n_x,n_y}(\delta) = \hbar\omega \left[ \sqrt{\delta} \left( n_x + \frac{1}{2} \right) + \frac{1}{\sqrt{\delta}} \left( n_y + \frac{1}{2} \right) \right]. \tag{1.3}$$

These levels are plotted against the deformation in Figure 1.3. For $\delta = 1$ (that is the isotropic case) you see that there is a $(n + 1)$-fold degeneracy, where $n = n_x + n_y = 0, 1, 2, \ldots$ is the principal quantum number. Taking into account the spin degeneracy (that gives a factor 2), we obtain closed shells for $N = 2, 6, 12, 20, \ldots$ electrons. These numbers are called the *magic numbers*. As we increase the deformation parameter, the degenerate levels split. It is useful, in this case, to look at the total energies (on the right in Figure 1.3); you see that now the energy levels have cups and minima that significantly change the behavior of the system for non-closed shells. In fact the energies have minima at $\delta = 1$ for the closed-shell cases ($N = 2, 6, 12$), whereas their minima are shifted to different values of $\delta$ in the open-shell cases ($N = 4, 8, 10$). As a rough graphical estimate, for $N = 4$ the minimum energy (i.e., the stablest configuration) is reached about a deformation value $\delta = 2$ and for $N = 10$ about $\delta = 1.5$.

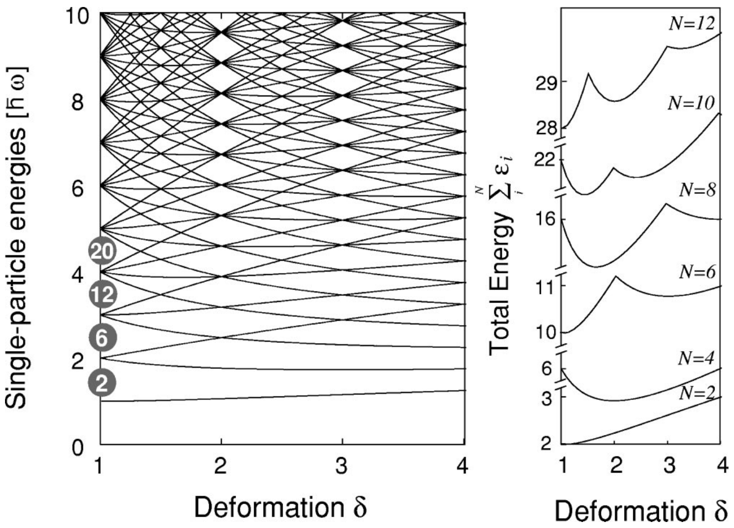


*Figure 1.3: Single particle energy levels (left) and total energies (right) as a function of the deformation in a two-dimensional harmonic oscillator. From Reimann and Manninen (2002).*

Even this simple example shows in an effective way how a deformation can guarantee a stabler configuration for non-closed shells. Further in our simulations, we will restrict only to closed-shell systems.

## 1.3 Evidences of the shell-model

We said before that the shell-structure seems to be a common property of finite fermion systems. Indeed, clear evidences of it had been shown for *atoms*, *atomic nuclei*, *clusters of atoms* and *quantum dots*. Let's briefly see which kind of measurements led to this results.

In the case of atoms, the shell-model is applied to the electrons "orbiting" around the nucleus, which is considered as an effective attractive Coulomb potential. A very natural quantity that one may think to measure in order to exploit the shell structure is the *ionization energy*, that is the energy required to remove an electron from a neutral atom. In fact we have shown that electrons in a closed-shell system are more bound, which means that we expect to need more energy to remove one of it (i.e., we expect to see a higher ionization energy). Indeed this happens pretty clearly, as you see in Figure 1.4, upper left, showing the sequence of magic numbers $N = 2, 10, 18, 36, 54, 86, \ldots$.

For atomic nuclei the idea is quite the same, with the only difference that now we have to measure the energy required to remove a nucleon from the nucleus (which is called the *separation energy*). The step-like behavior is also present in this case, giving

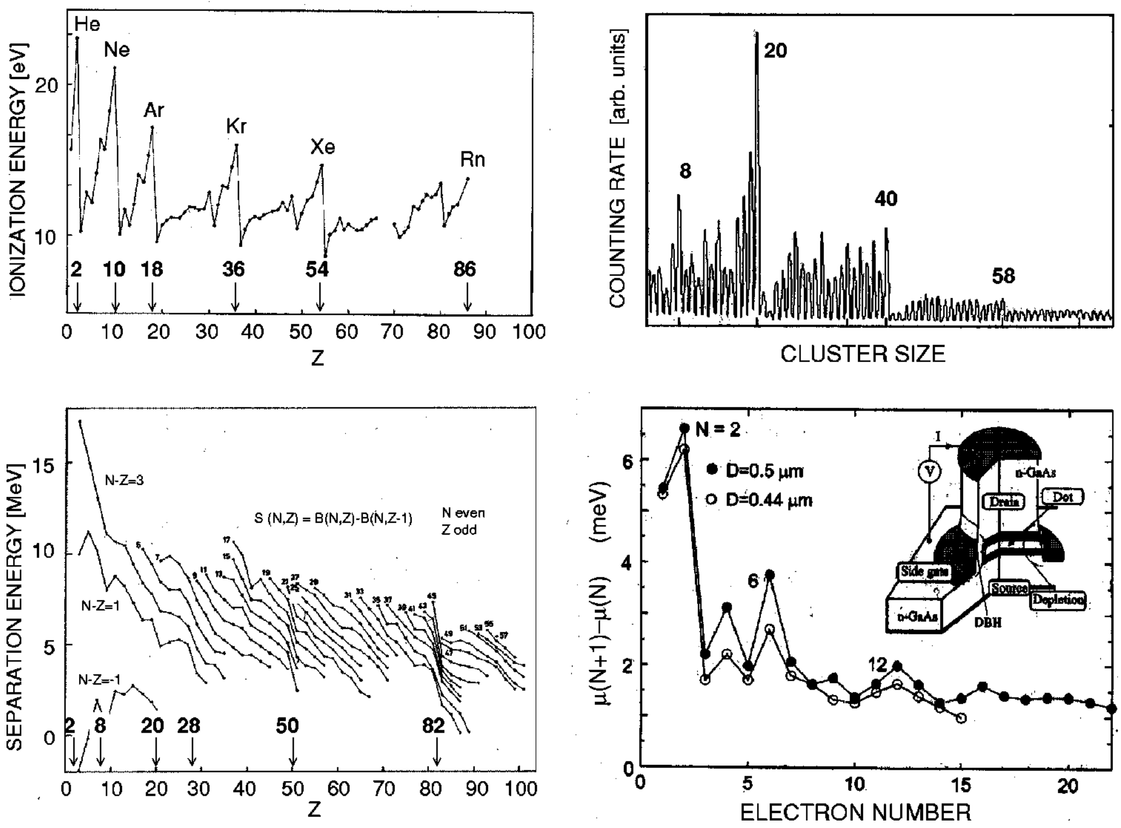


*Figure 1.4: Upper left, atomic ionization energies; lower left, separation energies of atomic nuclei; upper right, abundance spectra of metallic clusters; lower right, addition energies in a disk-shaped quantum dot (schematized in the inset). From Reimann and Manninen (2002).*

the sequence of magic numbers $N = 2, 8, 20, 28, 50, 82, \ldots$ (Figure 1.4, lower left).

Atomic clusters are also an interesting case, since they show anomalies in the mass abundance spectra: for certain cluster sizes, the clusters are more stable. This behavior has been explained with the “jellium” theory for metallic clusters; the valence electrons are treated as trapped in a homogeneous positive-charged background (the jellium), that stems from the atomic ions. Density functional calculations showed the magic number sequence $N = 2, 8, 20, 40, 58$ (see Figure 1.4, upper right) some years before it was discovered with experiments in the beginning of the 80’s (see Reimann and Manninen, 2002).

The sequence of magic numbers for quantum dots was discovered in 1996 by using an etched pillar of semiconducting material as shown in Figure 1.4, lower right (see also Tarucha et al., 1996). As we will see later, a quantum dot can be schematized as a one-electron transistor; measurements of the energy needed to add a single electron to a $N$-electrons quantum dot show large peaks for $N = 2, 6, 12$ (Figure 1.4). Noticeably, these numbers are the same quantum numbers obtained for a two-dimensional harmonic oscillator.

## 1.4 The Coulomb blockade model

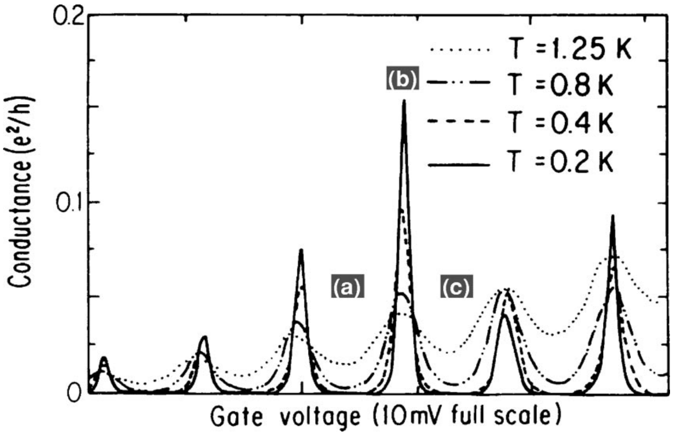


*Figure 1.5: Coulomb oscillations in a lateral quantum dot. The regions (a) and (c) have a fixed number of electrons since there is a Coulomb blockade; (b) is the conductance region, where the number of electrons oscillates. Plot based on the work of Meirav et al. (1990), by Meir et al. (1991).*

The behavior of a quantum dot can be analyzed by studying its electron transport mechanism. In order to do that, the dot is connected to a circuit that provides a certain gate voltage, and its conductance is measured. A plot like the one in Figure 1.5 is obtained; you see that there are conductance peaks at specific gate voltages, all of them evenly spaced. This behavior can be explained by the *Coulomb blockade model*.

Let's consider a *single-electron transistor* device, like the one in Figure 1.6. This name is due to the fact that the charge on the electron island is a multiple of the elementary charge. The electron island is weakly coupled with the source and drain contacts thanks to tunnel barriers, which are thick enough to guarantee that quantum resonances dominates the electron transport mechanism.; in other words, the conductance of the tunnel must be smaller than the quantum conductance $e^2/h$ (see Reimann and Manninen, 2002). By analogy with a capacitor, in which the energy stored is $U = Q^2/2C$, one can make a first estimate of the energy stored in a dot containing $N$ particles as

$$U(N) \simeq \frac{N(N-1)}{2C_\Sigma}. \tag{1.4}$$

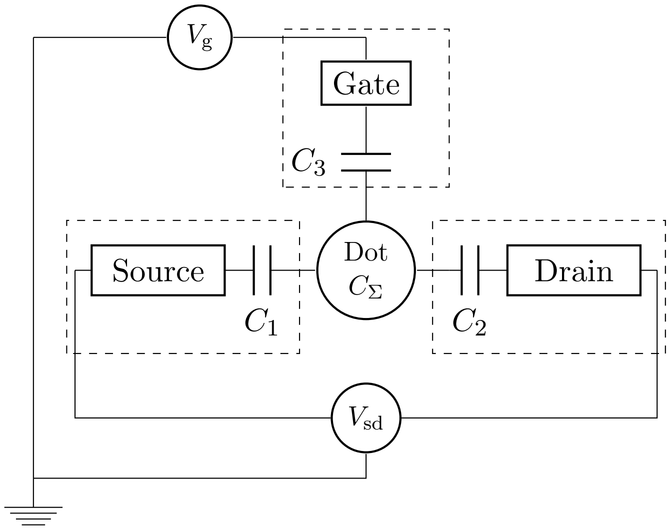


*Figure 1.6: Scheme of a single electron transistor. The island has index 0, the source has index 1, the drain index 2 and the gate index 3. The capacitances are meant to be intrinsic capacitances of the respective electrode (source, drain or gate). Adapted from Fasth and Fuhrer (2007).*

Here, the factor $(N-1)$ takes into account that each electron interacts with $N-1$ neighbors. Using this rough estimate, one can say that the energy required to add a single electron to a system already confining $N$ electrons, called the *addition energy*, is

$$\Delta U(N) = U(N+1) - U(N) = \frac{N(N+1)e^2}{2C_\Sigma} - \frac{N(N-1)e^2}{2C_\Sigma} = \frac{Ne^2}{C_\Sigma}. \qquad (1.5)$$

You see that, for successive $N$'s, the energy values are separated by the so-called *charging energy*, $e^2/C_\Sigma$, which is constant. At low enough temperatures and bias voltages, the charging energy might be larger than the thermal energy ($e^2/C_\Sigma \gg k_B T$), thus preventing the electrons from tunneling the barriers. In this case transport is blocked, and this effect is called the *Coulomb blockade.*

Let's now analyze the same situation in a more precise way (adapted from Fasth and Fuhrer, 2007). If we label the dot with the index $i = 0$, the source with 1, the drain with 2 and the gate with 3, the induced charge $\tilde{Q}_i$ on each of the electrodes can be written as

$$\tilde{Q}_i = \sum_{j=0}^{n} C_{ij} V_j, \qquad (1.6)$$

where $C_{ij}$ is the capacitance matrix, $V_j$ is the potential on the $j$'th element, and $n$ will be in general larger than 3 because of additional gate electrodes. If we refer to the dot

$(i = 0)$, we can then write

$$\tilde{Q}_0 = \sum_{j=0}^{n} C_{0j} V_j = C_{00} V_0 + \sum_{j=1}^{n} C_{0j} V_j, \tag{1.7}$$

from which

$$V_0 = \frac{1}{C_{00}} \left( \tilde{Q}_0 - \sum_{j=1}^{n} C_{0j} V_j \right). \tag{1.8}$$

In general, the induced charge $\tilde{Q}_0$ will be not equal to the measured one, $Q_0$, because of the presence of a background charge $Q_{\mathrm{bg}}$ (that can be measured if all the potentials are set to zero). This means that $\tilde{Q}_0 = Q_0 - Q_{\mathrm{bg}}$. Moreover, $C_\Sigma \equiv C_{00}$ by definition, and we can find an expression for it using the charge neutrality condition $\sum_{j=0}^{n} C_{0j} = 0$, which implies $C_\Sigma \equiv C_{00} = -\sum_{j=1}^{n} C_{0j}$. Using these two facts, we can rewrite equation (1.8) as

$$V_0(Q_0) = \frac{1}{C_\Sigma} \left( Q_0 - Q_{\mathrm{bg}} - \sum_{j=1}^{n} C_{0j} V_j \right). \tag{1.9}$$

At this point, we can calculate the energy $U(N)$ stored in a dot with $N$ electrons by performing an integration:

$$U(N) = \int_0^{-eN} V_0(Q_0)\, dQ_0 = \frac{e^2 N^2}{2 C_\Sigma} + eN \left( \frac{Q_{\mathrm{bg}}}{C_\Sigma} + \sum_{j=1}^{n} \frac{C_{0j}}{C_\Sigma} V_j \right). \tag{1.10}$$

Using this expression, we can finally calculate the addition energy as

$$\Delta U(N) = U(N+1) - U(N) = \frac{e^2}{C_\Sigma} \left( N + \frac{1}{2} \right) + e \left( \frac{Q_{\mathrm{bg}}}{C_\Sigma} + \sum_{j=1}^{n} \frac{C_{0j}}{C_\Sigma} V_j \right). \tag{1.11}$$

You see that, for small temperatures and bias voltages, we can approximate $\Delta U(N) \simeq e^2 N / C$ as we found previously by a rough estimate.

This calculation still lacks the fact that, in such a small device, quantization effects are also important and should be considered. A simple model that combines the Coulomb blockade effect and the quantum effects is the *constant interaction model.*

## 1.5 The constant interaction model

The basis assumption of the constant interaction model is that the total energy $E(N)$ of the electron island is the sum of the single-particle energies and the electrostatic energy $U(N)$:

$$E(N) = \sum_{i=1}^{N} \epsilon_i + U(N) = \sum_{i=1}^{N} \epsilon_i + \frac{e^2 N^2}{2C_\Sigma} + eN \left( \frac{Q_{\mathrm{bg}}}{C_\Sigma} + \sum_{j=1}^{n} \frac{C_{0j}}{C_\Sigma} V_j \right). \tag{1.12}$$

From this expression we can calculate the *electrochemical potential* $\mu_N$, that is defined as the necessary energy to add the $N$-th electron to a conductor, as

$$\mu_N = E(N) - E(N-1) = \epsilon_N + \frac{e^2}{C_\Sigma} \left( N - \frac{1}{2} \right) + e \left( \frac{Q_{\mathrm{bg}}}{C_\Sigma} - \sum_{j=1}^{n} \alpha_j V_j \right), \tag{1.13}$$

where $\alpha_j$, defined as

$$\alpha_j \doteqdot -\frac{C_{0j}}{C_\Sigma}, \tag{1.14}$$

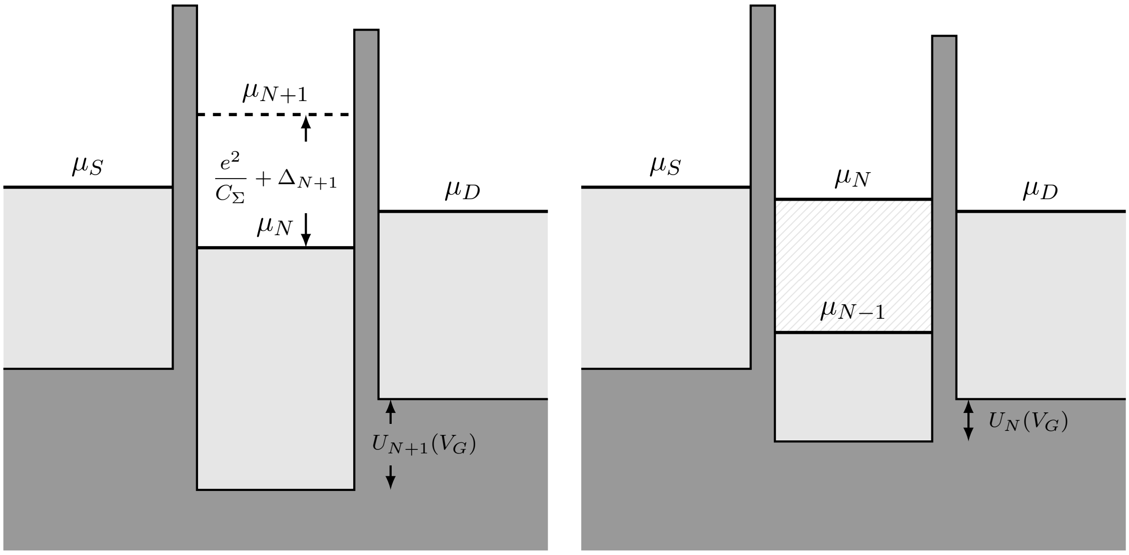
is called the *lever arm* of the gate $j$ and is always positive. Experiments show that the lever arm changes in a significant way only for very large gate voltage variations; so, in most cases the lever arm can be considered constant (Fasth and Fuhrer, 2007). This implies that the relation between the electrochemical potential and the gate voltage in equation (1.13) is linear; an increase in the gate voltage reflects in a decrease of the electrochemical potential, and vice-versa.

One can also calculate the spacing between different $\mu_N$'s, that is

$$\mu_{N+1} - \mu_N = \epsilon_{N+1} - \epsilon_N + \frac{e^2}{C_\Sigma} = \Delta_{N+1} + \frac{e^2}{C_\Sigma}, \tag{1.15}$$

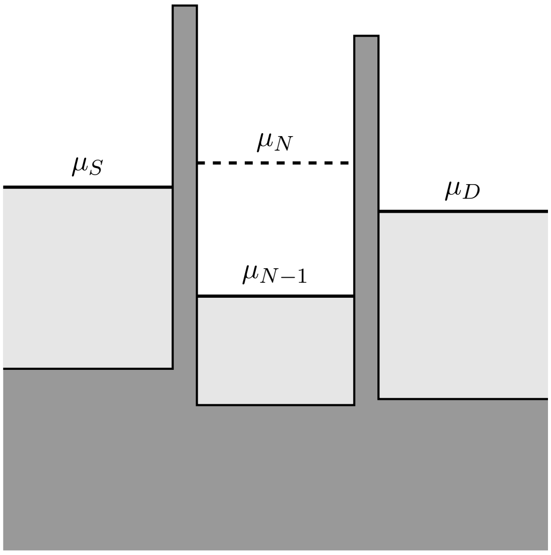
where $\Delta_{N+1} \doteqdot \epsilon_{N+1} - \epsilon_N$ is the spacing between two successive single-particles energy levels. You immediately note that, when $\Delta_{N+1} \ll k_B T \ll e^2/C_\Sigma$, the quantum effects can be neglected and the Coulomb blockade oscillations are periodic in $e^2/C_\Sigma$ (Reimann and Manninen, 2002).

Using this information, we can now analyze the transport mechanism in a quantum dot. Let's consider – as explained before – the low-temperature, low-bias voltage case ($eV_{\mathrm{bias}}, k_B T \ll e^2/C_\Sigma$). Let's also suppose that, initially, the electrochemical potential inside the dot – $\mu_N$ – is lower than the electrochemical potential of the drain – $\mu_D$: the transport is blocked due to the Coulomb blockade effect (Figure 1.7a). Now, we can


$\mu_S$
$\mu_{N+1}$
$\frac{e^2}{C_\Sigma} + \Delta_{N+1}$
$\mu_N$
$\mu_D$
$U_{N+1}(V_G)$
$\mu_S$
$\mu_N$
$\mu_{N-1}$
$\mu_D$
$U_N(V_G)$

*(a) $\mu_N < \mu_D$: the transport is blocked due to the Coulomb blockade. The number of electrons inside the dot is $N$.*

*(b) $\mu_S \gtrsim \mu_N \gtrsim \mu_D$: one electron can tunnel the barrier. The number of electrons inside the dot varies from $N-1$ to $N$. This configuration is obtain by lowering $V_G$, in order to increase $\mu_N$.*


$\mu_S$
$\mu_N$
$\mu_{N-1}$
$\mu_D$

*(c) $\mu_N > \mu_S$: the transport is blocked again due to the Coulomb blockade. The number of electrons inside the dot is $N-1$. This situation is obtained by further lowering $V_G$.*

*Figure 1.7: Energy diagram of a quantum dot. The sequence of images illustrates the transport process of the electrons (adapted from Fasth and Fuhrer, 2007).*

decrease the gate voltage: the overall effect, as you can see from equation (1.13) and from our previous considerations, is that we have a linear increase in $\mu_N$ wrt the gate voltage. Eventually, $\mu_N$ will align to $\mu_D$ ($\mu_N = \mu_D$), which means that an electron can leave the dot. At the same time, if $\mu_S \gtrsim \mu_D$, an electron from $\mu_S$ can enter the dot; the overall effect is that the number of the electrons inside the dot oscillates between $N$ and $N-1$, and an electric current flows though the dot from the drain to the source (Figure 1.7b). In terms of the conductance, this corresponds in a peak. If we keep lowering the gate voltage – i.e., increasing $\mu_N$ – until $\mu_N > \mu_S$, then the dot is left with only $N-1$ electrons. Now, the level to consider is $\mu_{N-1}$: you see that it's far below $\mu_D$, so the current is blocked again (see Figure 1.7c).

This qualitative analysis also permits to find the gate voltage values for which we have peaks in the conductance. The peak condition, in fact, is $\mu_S \approx \mu_N \approx \mu_D$; inserting this condition in (1.13) and solving for $V_3$ (the index 3 indicates the gate, so we will replace the subscript with "G" for simplicity) gives

$$V_G^{(N)} \approx \frac{1}{e\alpha_G}\left[\epsilon_N + \frac{e^2}{C_\Sigma}\left(N - \frac{1}{2}\right) + \frac{eQ_{\text{bg}}}{C_\Sigma} - \mu_S - e\sum_{j=4}^{n}\alpha_j V_j - e\sum_{j=1}^{2}\alpha_j V_j\right] \tag{1.16}$$

$$\approx \frac{1}{e\alpha_G}\left[\epsilon_N + \frac{e^2}{C_\Sigma}\left(N - \frac{1}{2}\right) + \frac{eQ_{\text{bg}}}{C_\Sigma} - \mu_S - e\sum_{j=4}^{n}\alpha_j V_j\right] \tag{1.17}$$

where the last term has been dropped because we are in the small-bias voltage case. The gate-voltage spacing between two peaks is finally found as (converted to energy)

$$e\Delta V_G(N) = e\Big[V_G^{(N)} - V_G^{(N-1)}\Big] \tag{1.18}$$

$$= \frac{1}{\alpha_G}\left(\epsilon_N - \epsilon_{N-1} + \frac{e^2}{C_\Sigma}\right) \tag{1.19}$$

$$= \frac{1}{\alpha_G}\left(\Delta_N + \frac{e^2}{C_\Sigma}\right). \tag{1.20}$$

As we already noticed before, for a lateral structure we can neglect the single-particle levels spacing (because the structure itself is quite large); this is not true for a vertical quantum dots structure, that has far smaller dimensions (as you can see in Figure 1.1). In that case the space quantization is evident and the conductance peaks are not evenly spaced, as you can see in Figure 1.8. It's also very nice that, by simply subtracting the quantity $1/\alpha_G(e^2/C_\Sigma)$ from the spacing of two successive peaks, we can reconstruct the spectrum of the dot.

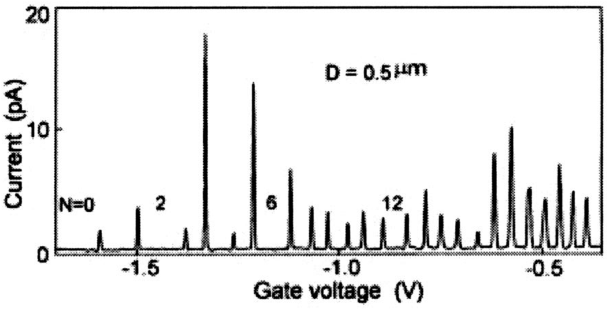


*Figure 1.8: Coulomb oscillations in a vertical quantum dot. Here, D is an estimate for the dot diameter. From Tarucha et al. (1996).*

Finally, we briefly discuss the case in which the bias voltage cannot be neglected. Let's put us in the simple case of a symmetric bias voltage between the drain and the source, i.e.

$$\begin{cases} V_S = V_{\text{bias}}/2 \\ V_D = -V_{\text{bias}}/2 \end{cases} \tag{1.21}$$

which means that

$$\begin{cases} \mu_S = \mu_0 + eV_{\text{bias}}/2 \\ \mu_D = \mu_0 - eV_{\text{bias}}/2 \end{cases} \tag{1.22}$$

Here, $\mu_0$ is the electrochemical potential of both drain and source when the bias voltage is zero.

We showed before that the dot does not conduct for

$$\mu_N < \mu_D < \mu_S < \mu_{N+1}, \qquad V_{\text{bias}} > 0, \tag{1.23}$$

and

$$\mu_N < \mu_S < \mu_D < \mu_{N+1}, \qquad V_{\text{bias}} < 0, \tag{1.24}$$

so the borders of the Coulomb blockade region are defined by

$$\mu_N = \begin{cases} \mu_D = \mu_0 - eV_{\text{bias}}/2, & V_{\text{bias}} > 0 \\ \mu_S = \mu_0 + eV_{\text{bias}}/2, & V_{\text{bias}} < 0 \end{cases} \tag{1.25}$$

and

$$\mu_{N+1} = \begin{cases} \mu_S = \mu_0 + eV_{\text{bias}}/2, & V_{\text{bias}} > 0 \\ \mu_D = \mu_0 - eV_{\text{bias}}/2, & V_{\text{bias}} < 0 \end{cases} \tag{1.26}$$

Substituting these values for $\mu_N$, $V_S$ and $V_D$ in (1.13), we finally find that the borders of the Coulomb blockade are the straight lines of equations

$$V_G^{(N)}(V_{\text{bias}}) = \frac{1}{e\alpha_G}\left[\epsilon_N + \frac{e^2}{C_\Sigma}\left(N - \frac{1}{2}\right)\right. \\ \left. + \frac{eQ_{\text{bg}}}{C_\Sigma} - \mu_0 - e(\alpha_S - \alpha_D - 1)\frac{V_{\text{bias}}}{2} - e\sum_{j=4}^{n}\alpha_j V_j\right] \tag{1.27}$$

and

$$V_G^{(N+1)}(V_{\text{bias}}) = \frac{1}{e\alpha_G}\left[\epsilon_{N+1} + \frac{e^2}{C_\Sigma}\left(N + \frac{1}{2}\right)\right. \\ \left. + \frac{eQ_{\text{bg}}}{C_\Sigma} - \mu_0 - e(\alpha_S - \alpha_D + 1)\frac{V_{\text{bias}}}{2} - e\sum_{j=4}^{n}\alpha_j V_j\right] \tag{1.28}$$

for $V_{\text{bias}} > 0$. The case $V_{\text{bias}} < 0$ is handled similarly.

By equating (1.27) and (1.28) we find that these two lines cross at $eV_{\text{bias}} = \Delta_{N+1} + e^2/C_\Sigma$, where $\Delta_{N+1} = \varepsilon_{N+1} - \varepsilon_N$ as usual. For $V_{\text{bias}} = 0$, instead, we find that the spacing between $V_G^{(N)}$ and $V_G^{(N+1)}$ is $(\Delta_{N+1} + e^2/C_\Sigma)/\alpha_G$ as we found previously for a negligible $V_{\text{bias}}$, equation (1.20). The situation is schematically pictured in Figure 1.9.

The diamond-shaped structure shown in Figure 1.9 have been shown in differential conductance measurements like the one shown in Figure 1.10. A 3D elaboration of the data obtained from this kind of experiment – that clarifies the ideas a little bit – is shown in Figure 1.11.

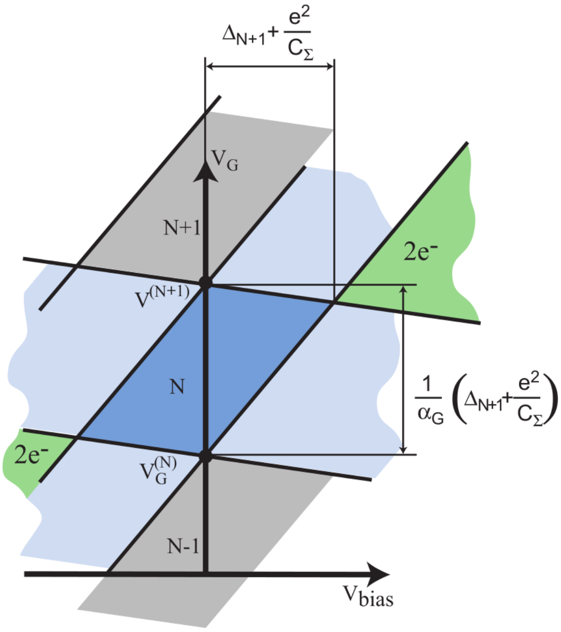

$\Delta_{N+1}+\frac{e^2}{C_\Sigma}$
$V_G$
N+1
$2e^-$
$V^{(N+1)}$
N
$\frac{1}{\alpha_G}\left(\Delta_{N+1}+\frac{e^2}{C_\Sigma}\right)$
$2e^-$
$V_G^{(N)}$
N-1
$V_{bias}$

*Figure 1.9: Coulomb blockade diamonds (gray and dark blue regions). In the dark blue region, the probability of finding N electrons in the dot is 1 (stable configuration); in the light blue regions, the probability varies from 0 to 1 and the electron number can change by one. In the green region, the bias voltage is so high that eventually $eV_{bias} > e^2/C_\Sigma$; due to this fact, two electrons can tunnel at the same time. From Fasth and Fuhrer (2007).*

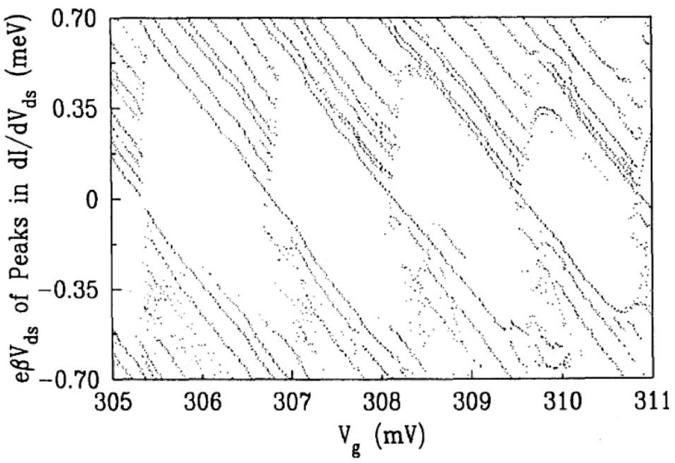


*Figure 1.10: Peaks in the differential conductance $\partial I/\partial V_{sd}$ (showed as solid lines) of a lateral quantum dot, as a function of $V_{sd}$ and $V_g$. The white regions are the Coulomb blockade regions; between successive regions, the number of electrons in the dot is increased by one. From Reimann and Manninen (2002).*

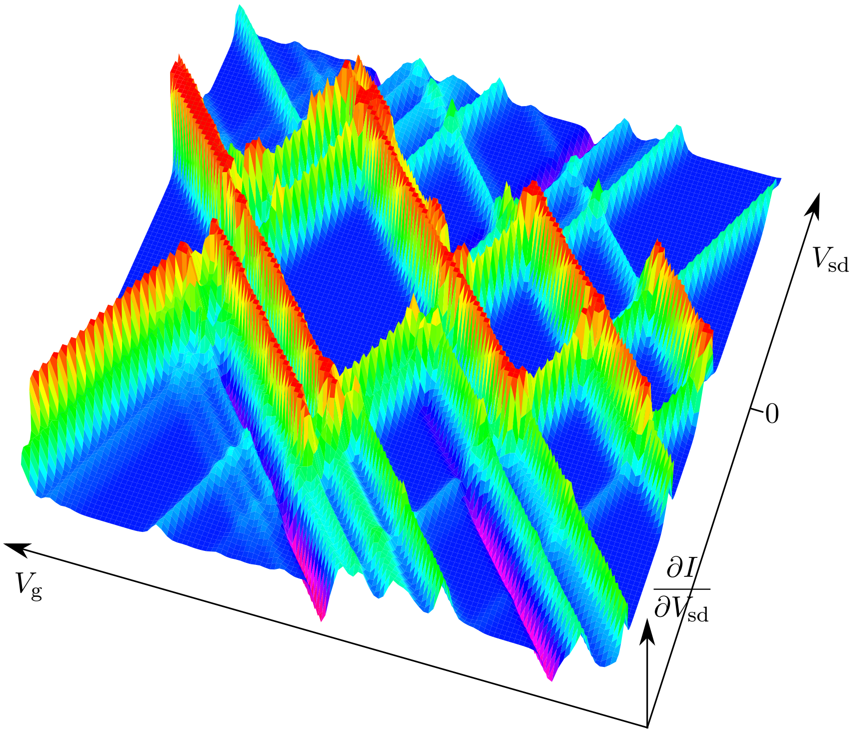


*Figure 1.11: Differential conductance $\partial I/\partial V_{sd}$ in a nanowire, plotted against $V_{sd}$ and $V_g$. Green and red indicate positive values; blue means close to zero, whereas pink stands for negative. From Weinmann (1994).*

## 1.6 The parabolic confinement

In our simulation, we will model the confining (single-particle) potential as a *two-dimensional harmonic trap*. This form for the confining potential is widely accepted as a standard for both exact and effective-field calculations, and is the result of experimental observations and numerical calculations.

The basis for this assumption can be found in the work of Kumar et al. (1990). They performed numerical calculations on the Poisson and Schrödinger equations to obtain the electron states in a GaAs/AlGaAs quantum dot – like the one sketched in Figure 1.1a. The calculations were done in the Hartree approximation. The results were quite interesting, as you can see in Figure 1.12; despite the square shape of the GaAs cap, the contours of the effective confinement potential were circular.

An experimental support for this model comes from the observation of far-infrared absorbption spectra, in combination with the Kohn theorem. This theorem (see Kohn,

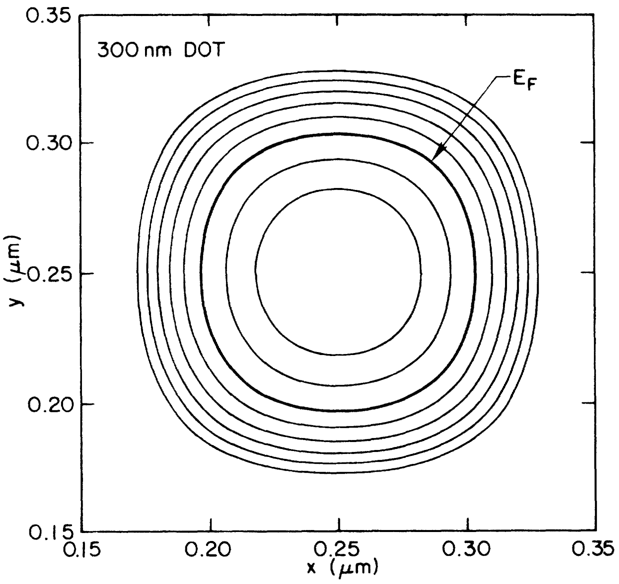


*Figure 1.12: Contours of the lateral potential in a GaAs/AlGaAs quantum dot. The measurements were taken* 8 mm *below the contact surface of the two materials, and the contours are evenly separated by* 10 meV *intervals. The Fermi level (thick line) is indicated by* $E_F$*. From Kumar et al. (1990).*

1961) states that, for an electron gas with short-range interactions, the latter don't change the cyclotron frequency. For far-infrared absorbption measurements in a quantum dot, this has the consequence that the effects of the electron-electron interaction show up only for a strong enough anharmonicity of the confining potential (see Reimann and Manninen, 2002). Since in many measurements the effects of the electron-electron interaction were not seen, the model of a harmonic confinement gained ground.

We also showed, in sections 1.2-1.3, that a 2D harmonic oscillator gives the same magic numbers observed experimentally. An improved model, that takes into account the non-zero thickness of a lateral dot (neglected for a pure 2D potential), would be to add a $z$-contribution to the potential, i.e. $V = V(x, y, z) = V(x, y) + V(z)$. Then, it is assumed that the only state occupying the $z$-direction is the ground state, so that the solution can be restricted to the $(x, y)$-plane (see Reimann and Manninen, 2002). $V(x, y)$ is modeled as the usual two-dimensional parabolic confinement:

$$V(x, y) = \frac{1}{2} m\omega^2 (x^2 + y^2) \tag{1.29}$$

It's worth noting that, in experimental observations, the effective confinement strength $\omega$ is not constant; it tends to lower as the number $N$ of confined electrons increases. The lowering of the effective strength translates into an increase of the “confinement volume”, so that the electron density remains constant. A model for $\omega$ that takes into account these effects was given by Koskinen et al. (1997), who modeled $\omega^2$ as

$$\omega^2 = \frac{e^2}{4\pi\varepsilon m^* r_s^3 \sqrt{N}}, \tag{1.30}$$

where $\varepsilon$ is the dielectric constant of the material, $m^*$ is the effective mass of the electron and $r_s$ is the average electron-density parameter[2], quite close to the equilibrium value for a two-dimensional electron gas (see Reimann and Manninen, 2002).

---

[2] This parameter is also called the *Wigner-Seitz* parameter. It's defined such that

$$\Omega_d(r_s) = n,$$

where $\Omega_d(r_s)$ is the spherical volume of radius $r_s$ in $d$ dimensions (e.g., for $d = 2$ is a circle with radius $r_s$) and $n$ is the mean volume per unit electron (i.e., the mean electron density). So, $r_s$ can be expressed as

$$r_s = \frac{1}{\sqrt{\pi n}}$$

for the two-dimensional case, or as

$$r_s = \left(\frac{3}{4\pi n}\right)^{1/3}$$

for the three-dimensional case.

# Chapter 2

# Theoretical basis of the simulation

## 2.1 The Variational Principle

The Variational Principle is a method of general validity that can be used to gather information about a system with a Hamiltonian that we are unable to diagonalize (i.e., we can't solve the Schrödinger equation for that system). Specifically, this principle gives us an *upper bound* for the energy of the ground state, that we will call here $E_{\text{gs}}$. The formulation is astonishingly simple: if you pick *any state* $|\psi\rangle$ *whatsoever*, then

$$E_{\text{gs}} \leq \frac{\left\langle \psi \,\middle|\, \hat{H} \,\middle|\, \psi \right\rangle}{\langle \psi \,|\, \psi \rangle}. \tag{2.1}$$

In this project we are going to use just this formula, because you see that the normalization of the state is not necessary. But if you have a normalized state, then equation (2.1) is further simplified in

$$E_{\text{gs}} \leq \left\langle \psi \,\middle|\, \hat{H} \,\middle|\, \psi \right\rangle \equiv \left\langle \hat{H} \right\rangle. \tag{2.2}$$

This equation seems a kind of magic but actually the proof of this fact is really simple, and we are going to sketch here a proof to show the power of this method (roughly, as it appears in Griffiths (2005)).

Proof: Since the (unknown) eigenstates $|\psi_n\rangle$ of $\hat{H}$ form a complete set, we can expand our random state $|\psi\rangle$ on this basis:

$$|\psi\rangle = \sum_n C_n |\psi_n\rangle, \qquad C_n \in \mathbb{C},$$

where the $|\psi_n\rangle$'s are such that

$$\hat{H} |\psi_n\rangle = E_n |\psi_n\rangle .$$

Then, we have

$$\begin{aligned} \left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle &= \sum_{nm} C_m^* C_n E_n \underbrace{\langle \psi_m \,|\, \psi_n \rangle}_{\delta_{mn}} \\ &= \sum_n |C_n|^2 E_n, \\ \langle \psi \,|\, \psi \rangle &= \sum_n |C_n|^2. \end{aligned}$$

Since $E_n \geq E_{\text{gs}}\ \forall n$, it follows immediately that

$$E_T \div \frac{\left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle}{\langle \psi \,|\, \psi \rangle} = \frac{\sum_n |C_n|^2 E_n}{\sum_n |C_n|^2} \geq \frac{\cancel{\sum_n |C_n|^2} E_{\text{gs}}}{\cancel{\sum_n |C_n|^2}} = E_{\text{gs}}. \tag{2.3}$$

■

In practice, one chooses a class of states for $|\psi\rangle$ parametrized by one or more parameters – the so-called *variational parameters*, and then calculates the quantity $E_T$ for multiple sets of values of the variational parameters $\alpha_1 \ldots, \alpha_n$. The lower value of $E_T$ obtained in this way is the required upper bound. If one manages to find an even lower value for $E_T$ with a more clever wave-function, then he has found a better upper bound for the ground state energy.

The only trouble with this method is that we never know for sure how close we are to the *actual* ground state energy; all we get for sure is just an upper bound. However, if one manages to guess a *realistic* $|\psi\rangle$, he often gets values for the ground state energy that miraculously match the actual ones.

In the following, we represent $|\psi\rangle$ in its function-representation as

$$|\psi\rangle \simeq \psi_T,$$

where "$\simeq$" means "represented by" and the subscript $T$ is just to remember that $\psi_T$ is our *trial* wave-function.

In our case the Hamiltonian lies in an infinite-dimension Hilbert space, so we have to replace the sums in (2.3) with integrals over all the space. Indicating with $d\vec{\tau}$ a space element, we can write

$$E_T = \frac{\int d\vec{\tau}\, \psi_T^* \hat{H} \psi_T}{\int d\vec{\tau}\, |\psi_T|^2}. \tag{2.4}$$

Calculating $E_T$ as it appears in (2.4) is not a piece of cake. However, we can recast that equation in a simpler form if we multiply and divide by $\psi_T$ at the left of $\hat{H}$. In fact,

$$\begin{aligned}
E_T &= \frac{\int d\vec{\tau}\, \psi_T^* \dfrac{\psi_T}{\psi_T} \hat{H} \psi_T}{\int d\vec{\tau}\, |\psi_T|^2} && (2.5)\\
&= \frac{\int d\vec{\tau}\, |\psi_T|^2 \dfrac{1}{\psi_T} \hat{H} \psi_T}{\int d\vec{\tau}\, |\psi_T|^2} && (2.6)\\
&= \int d\vec{\tau}\, \frac{|\psi_T|^2}{\int d\vec{\tau}\, |\psi_T|^2} \frac{1}{\psi_T} \hat{H} \psi_T && (2.7)\\
&= \int d\vec{\tau}\, \mathcal{P}(\vec{\tau}) E_L(\vec{\tau}) && (2.8)\\
&\simeq \frac{1}{n} \sum_{i=1}^{n} E_L(\vec{\tau}) && (2.9)
\end{aligned}$$

where

$$\mathcal{P}(\vec{\tau}) = \frac{|\psi_T|^2}{\int d\vec{\tau}\, |\psi_T|^2} \tag{2.10}$$

and

$$E_L(\vec{\tau}) = \frac{1}{\psi_T} \hat{H} \psi_T. \tag{2.11}$$

The dependence of $\psi_T$ on the position in space and on the variational parameters has been omitted just for better readability. The quantity $E_L$ is called the *local energy*, and it's what we are going to calculate in this project. Obviously, for bigger $n$ we obtain better results.

## 2.2 Considerations about the physical system

Our system is represented by a number $N$ of electrons that have total energy [1]

$$\hat{H} = \sum_{i=1}^{N} \left( -\frac{1}{2}\nabla_i^2 + \frac{1}{2}\omega^2 r_i^2 \right) + \sum_{i<j} \frac{1}{r_{ij}}, \tag{2.12}$$

where $r_{ij} = |\vec{r_i} - \vec{r_j}|$ and natural units ($\hbar = c = e = m_e = 1$) are used in order to have the energy in atomic units. You see that the Hamiltonian includes a standard part (a harmonic oscillator) plus a repulsion potential, that is the one that gives troubles in analytical calculations.

An important feature of our system is that it is made up of *identical particles*. Let's try to exploit this feature to simplify our trial wave-function.

Suppose that we have a system made up of two particles, let's call them 1 and 2. Then, the state $|\psi\rangle$ of the system can be expressed (for example) as

$$|\psi\rangle = |1\rangle \otimes |2\rangle \,. \tag{2.13}$$

Now we can introduce the *permutation operator* $\hat{P}$, defined as

$$\hat{P}\,|1\rangle \otimes |2\rangle \doteq |2\rangle \otimes |1\rangle \,. \tag{2.14}$$

One immediately sees that $\hat{P} = \hat{P}^\dagger$ and $\hat{P}^2 = \hat{I}$ ($\hat{I}$ is the identity operator), which means that $\hat{P}$ is hermitian and unitary. These two properties allow us to say that the eigenvalues of $\hat{P}$ are $\pm 1$. In general, this operator $\hat{P}$ does not commute with $\hat{H}$; but – and here is the trick – it *does* commute with $\hat{H}$ for a system of *identical particles*. Since in that case $[\hat{H}, \hat{P}] = 0$, the eigenstates of $\hat{H}$ must also be eigenstates of $\hat{P}$. Recalling that the eigenvalue of $\hat{P}$ are $\pm 1$, one can rewrite the state $\psi$ as

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big( |1\rangle \otimes |2\rangle \pm |2\rangle \otimes |1\rangle \Big) \tag{2.15}$$

to make it also eigenstate of $\hat{P}$. You see that, in this form, the state is either symmetric or antisymmetric; particles with symmetric wave-function are called *bosons*, and particles with antisymmetric wave-function are called *fermions*. It turns out that electrons are fermions, so we are interested only in the antisymmetric case. Rewriting the fermions wave-function in coordinate representation, we obtain:

---

[1] Calling the Hamiltonian the total energy is improper. Rather it should be called the *generalized energy*, since in general it is *not* the total energy because it also contains terms of the 1st and 0th order in the generalized velocities. However in this case the system does not contain non-conservative forces (from which the extra terms stem), so we can safely identify the Hamiltonian with the total energy.

$$\psi(\vec{r}_1, \vec{r}_2, \sigma_1, \sigma_2) = \Big( \langle \vec{r}_1 | \otimes \langle \vec{r}_2 | \Big) |\psi\rangle = \frac{1}{\sqrt{2}} \Big( \langle \vec{r}_1 \,|\, 1 \rangle \langle \vec{r}_2 \,|\, 2 \rangle - \langle \vec{r}_1 \,|\, 2 \rangle \langle \vec{r}_2 \,|\, 1 \rangle \Big) \tag{2.16}$$

$$= \frac{1}{\sqrt{2}} \Big( \phi_1(\vec{r}_1, \sigma_1) \phi_2(\vec{r}_2, \sigma_2) - \phi_2(\vec{r}_1, \sigma_2) \phi_1(\vec{r}_2, \sigma_1) \Big) = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_1(\vec{r}_1, \sigma_1) & \phi_1(\vec{r}_2, \sigma_1) \\ \phi_2(\vec{r}_1, \sigma_2) & \phi_2(\vec{r}_2, \sigma_2) \end{vmatrix} \tag{2.17}$$

where the $\phi_i$'s are the single-particle wave-functions, the $\sigma_i$'s just indicate the explicit dependence on the spin, and the determinant in (2.17) is called the *Slater determinant.* For $N$ particles the determinant has exactly the same shape (is just a $N \times N$ matrix) and the factor $\frac{1}{\sqrt{2}}$ is replaced with $\frac{1}{\sqrt{N!}}$. However, we don't care about the normalization factor because it's not required to apply the variational principle.

We can recast this equation in a simpler form. In fact, it's possible to rewrite $|\phi\rangle$ as

$$|\phi\rangle = |\eta\rangle \otimes |\chi\rangle \,, \tag{2.18}$$

where $|\eta\rangle$ is the *spatial part* and $|\chi\rangle$ is the *spin part.* Note that – since the Hamiltonian is spin-independent – the spin part just takes into account the spin configuration of the system, and it has nothing to do with the positions of the particles themselves! In other words, while the spatial part depends on the positions of the particles, the spin part does not at all and we can "forget" about it if we are smart.

In fact, the best way to minimize the energy is to put the two electrons in the same orbital. This means that the spatial part is *symmetric* under the interchange of two particles; but we have also said that this wave-function has to be antisymmetric. So, the spin part has to be *antisymmetric.* If you look at the Clebsch-Gordan table for combining two spin-1/2 (Figure 2.1), you see that the only antisymmetric combination is the total spin state $|0,0\rangle$, that can be written as

$$|0,0\rangle = \frac{1}{\sqrt{2}} \left( |\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle \right) \tag{2.19}$$

This fact permits a simplification of the Slater determinant; in fact, if we suppose – for example – that particle 1 has spin-up and particle 2 has spin-down, we can write

$$\Psi_T(\vec{r}_1, \vec{r}_2, \sigma_1, \sigma_2) = \begin{vmatrix} \phi_1(\vec{r}_1, \sigma_1) & \phi_1(\vec{r}_2, \sigma_1) \\ \phi_2(\vec{r}_1, \sigma_2) & \phi_2(\vec{r}_2, \sigma_2) \end{vmatrix} = \phi_1(\vec{r}_1)\phi_2(\vec{r}_2). \tag{2.20}$$

where the $\phi_i$'s are now independent on the spin. In other words, they are just the spatial parts $\eta_i$'s of the single-particle wave-functions. You also see that, since now the single-particle wave-functions are independent on the spin, the are *exactly* the same function because they occupy the same energy level!

The generalization to $N$ particles is straightforward, and will be given later.

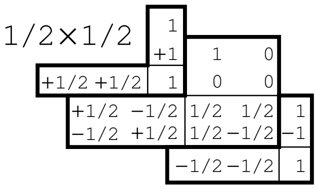

1/2×1/2
1
+1
1 0
0 0
+1/2 +1/2 1
+1/2 −1/2 1/2 1/2
−1/2 +1/2 1/2 −1/2
1
−1
−1/2 −1/2 1

*Figure 2.1: The Clebsch-Gordan table for combining two spin-1/2. Source:* `http://pdg.lbl.gov/2002/clebrpp.pdf`.

## 2.3 The Metropolis algorithm

Let's first briefly introduce what a *Markov chain* is. We call a Markov chain a stochastic process where the transition probability that characterizes the transition between states only depends on the previous step and not from the path to get to such step. We can then write (see Hjorth-Jensen, 2014)

$$w_i(t+\delta t) = \sum_i W(j \to i) w_j(t), \tag{2.21}$$

where $w_i(t+\delta t)$ is our PDF – in our case $|\psi_T(R)|^2$ and $W(j \to i)$ is the probability of jumping from $j$ to $i$, also called the *transition probability*. We also have

$$\sum_i w_i(t) = 1. \tag{2.22}$$

due to the normalization condition. We can see from equation (2.21) the probability density function at $t+\delta t$ only depends on its previous state $t$.

In the case of infinite dimensionality, representing $W$ as a matrix, we have

$$\hat{w}(t+\delta t) = \hat{W}\hat{w}(t). \tag{2.23}$$

from which

$$\hat{w}(\infty) = \hat{W}\hat{w}(\infty). \tag{2.24}$$

Then, the problem is just an eigenvalue problem with eigenvalue 1.

Now, let's consider an *ergodic* system, where the average over time of a process is the same as the ensemble average. Or in other words, as one increases the steps, there exists a positive probability measurement at step $n$ that is independent on the probability distribution at the initial step 0. It holds

$$W(i \to j)w_i = W(j \to i)w_j, \tag{2.25}$$

from which

$$R \doteqdot \frac{W(i \to j)}{W(j \to i)} = \frac{w_j}{w_i} = \frac{|\psi_T(r^{\text{new}})|}{|\psi_T(r^{\text{old}})|}. \tag{2.26}$$

Since $W(i \to j)$ is unknown, we model it as the product of two probabilities:

$$W(i \to j) = A(i \to j)\,T(i \to j), \tag{2.27}$$

where $A(i \to j)$ is the probability of accepting the move and $T(i \to j)$ is the probability to transition from the state $i$ to the state $j$. In a similar way we can write:

$$W(j \to i) = A(j \to i)\,T(j \to i). \tag{2.28}$$

### 2.3.1 Brute force sampling

The *brute force* Metropolis algorithm simply consists in the assumption that the transition probability is symmetric, that is

$$T(i \to j) = T(j \to i), \tag{2.29}$$

which leads to

$$\frac{W(i \to j)}{W(j \to i)} = \frac{A(i \to j)}{A(j \to i)}. \tag{2.30}$$

Then, we can distinguish between two cases.

- If $\dfrac{w_j}{w_i} \geq 1$, then
$$A(i \to j) \geq A(j \to i). \tag{2.31}$$
  So, if we manage to move to such a state, we accept this move because we are moving to a higher probability region.
- If $\dfrac{w_j}{w_i} < 1$, then
$$A(j \to i) > A(i \to j). \tag{2.32}$$

> We can't blindly reject this move just because we are moving to a lower probability region, but we can't also accept all of this kind of moves. So, what one does is to compute a random number $r \in (0,1)$ and to accept the move if $r \leq w_j/w_i$.

To summarize, one sets

$$A(i \to j) = \min\left(1, \frac{w_j}{w_i}\right) = \min\left(1, \frac{|\psi_T(r^{\text{new}})|^2}{|\psi_T(r^{\text{old}})|^2}\right) \tag{2.33}$$

and then compares with a random number $r \in (0,1)$. If $A(i \to j) \geq r$ the step is accepted, otherwise it's rejected.

### 2.3.2 Importance sampling and Schödinger equation as a diffusion problem

*Importance sampling* consists in doing more clever assumptions on the shape of $T(i \to j)$. The discussion is the same as the brute force case, but this time we have to set

$$A(i \to j) = \min\left(1, \frac{w_j T(j \to i)}{w_i T(i \to j)}\right) = \min\left(1, \frac{|\psi_T(r^{\text{new}})|^2 T(j \to i)}{|\psi_T(r^{\text{old}})|^2 T(i \to j)}\right). \tag{2.34}$$

If we recast the Schrödinger equation as a diffusion problem, $T(i \to j)$ is going to be a Gaussian distribution (or a modified Gaussian); let's do that.

The main idea is to model $T(i \to j)$ in order to reflect the shape of the trial wavefunction itself and have in this way a more "intelligent" walk, that wastes less points. The walk is generated using the Fokker-Planck and the Langevin equations; the former – for a one-dimension diffusion problem – is

$$\frac{\partial P}{\partial t} = D\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} - F\right) P(x,t), \tag{2.35}$$

where $P(x,t)$ is the time-dependent probability density, $D$ is the *diffusion coefficient* (it's 0.5 for the Schrödinger equation, see Hø gberget (2013)) and $F$ is the *drift term.*

The Langevin equation, on the other hand, reads as

$$\frac{\partial x(t)}{\partial t} = DF(x(t)) + \eta, \tag{2.36}$$

being $\eta$ a *random variable.* By solving equation (2.36) with Euler's method, it follows that the new position is given by

$$y = x + DF(x)\Delta t + \xi, \tag{2.37}$$

being $\xi$ a *Gaussian* random variable and $\Delta t$ a chosen time step.

For two or more dimensions, the Fokker-Planck equation is just

$$\frac{\partial P}{\partial t} = \sum_i D \frac{\partial}{\partial \vec{x}_i} \left( \frac{\partial}{\partial \vec{x}_i} - \vec{F}_i \right) P(\vec{x}_i, t) \tag{2.38}$$

We can obtain a convergence to a stationary probability density setting the left hand side to zero, that means all the terms of the sum must be equal to zero. Expanding the product, we obtain

$$\frac{\partial^2 P}{\partial \vec{x}_i^2} = P \frac{\partial}{\partial \vec{x}_i} \vec{F}_i + \vec{F}_i \frac{\partial}{\partial \vec{x}_i} P. \tag{2.39}$$

Since the drift vector has the form $\vec{F} = g(\vec{x}) \dfrac{\partial P}{\partial \vec{x}}$, we get

$$\frac{\partial^2 P}{\partial \vec{x}_i^2} = P \frac{\partial g}{\partial P} \left( \frac{\partial P}{\partial \vec{x}_i} \right)^2 + Pg \frac{\partial^2 P}{\partial \vec{x}_i^2} + g \left( \frac{\partial P}{\partial \vec{x}_i} \right)^2. \tag{2.40}$$

If we want a stationary density the left hand side of equation (2.40) must be zero, and this is possible only if $g = \dfrac{1}{P}$. With this substitution, the drift vector reads as

$$\vec{F} = \frac{1}{P} \nabla P. \tag{2.41}$$

Since – as we said many times before – our probability density is the modulus squared of the trial wave-function, that is

$$P = |\psi_T|^2, \tag{2.42}$$

the drift vector in equation (2.41) finally reads as

$$\vec{F} = 2 \frac{1}{\Psi_T} \nabla \Psi_T. \tag{2.43}$$

Written in this way, $\vec{F}$ is also called the *quantum force*. This force pushes the walker in those regions where the trial wave function is larger, increasing the efficiency of the simulation.

Solving the Fokker-Plank equation finally gives an expression for the transition probability $T(x \to y)$, that turns out can be expressed as the Green function

$$G(y, x, \Delta t) = \frac{1}{(4\pi D \Delta t)^{3N/2}} \exp(-(y - x - D\Delta t F(x))^2 / 4D\Delta t). \tag{2.44}$$

With this expression, the acceptance ratio in equation (2.34) becomes

$$R = \frac{|\psi_T(y)|^2 T(y \to x)}{|\psi_T(x)|^2 T(x \to y)} = \frac{|\psi_T(y)|^2 G(x, y, \Delta t)}{|\psi_T(x)|^2 G(y, x, \Delta t)}. \tag{2.45}$$

### 2.3.3 Implementation

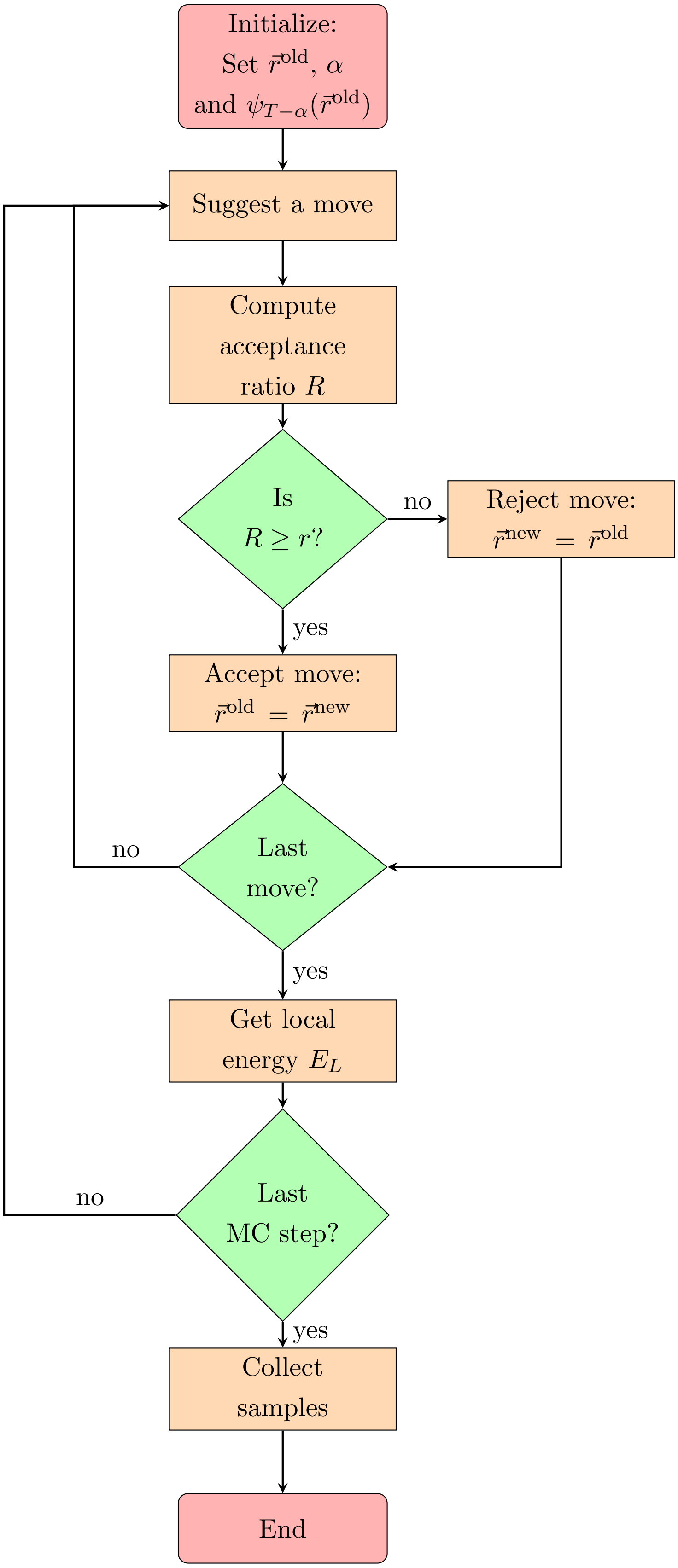

Initialize:
Set $\vec{r}^{\text{old}}$, $\alpha$
and $\psi_{T-\alpha}(\vec{r}^{\text{old}})$
Suggest a move
Compute acceptance ratio $R$
Is $R \geq r$?
no
Reject move: $\vec{r}^{\text{new}} = \vec{r}^{\text{old}}$
yes
Accept move: $\vec{r}^{\text{old}} = \vec{r}^{\text{new}}$
no
Last move?
yes
Get local energy $E_L$
no
Last MC step?
yes
Collect samples
End

*Figure 2.2: Chart flow for the Quantum Variational Monte Carlo algorithm.*

We can sum up the algorithm as follows. Given a random number generator that returns aleatory values in the interval $(0,1)$ with a uniform (or Gaussian for the importance sampling) distribution, then the Metropolis algorithm consists in these steps.

- Chosen a starting point $x_0$ and a step length $l$ (or a time-step $\Delta t$ for the importance sampling), a random number $\varepsilon$ is generated in the interval $(0,1)$.
- The proposed step is $x^{\text{new}} = x^{\text{old}} + l\varepsilon$ (or $x^{\text{new}} = x^{\text{old}} + D\vec{F}(x^{\text{old}})\Delta t + \varepsilon$ for the importance sampling).
- The probability ratio $R = \frac{|\psi_T(r^{\text{new}})|^2 T(j \to i)}{|\psi_T(r^{\text{old}})|^2 T(i \to j)}$ is computed.
- A new random number $r$ is generated in the interval $(0,1)$.
- If $R \geq r$, the step is accepted and the new position is stored by letting $x^{\text{old}} = x^{\text{new}}$.
- If $R < r$, the step is rejected and the new position is discarded by letting $x^{\text{new}} = x^{\text{old}}$.

A flow chart of this process is sketched in Figure 2.2.

# Chapter 3

# The 2-electrons system

## 3.1 The unperturbed system

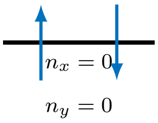


*Figure 3.1: The 2-electrons system configuration.*

We will first consider the unperturbed system (without the electron-electron repulsion) with $\omega = 1$, that has Hamiltonian

$$\hat{H}_0 = \sum_{i=1}^{N} \left( -\frac{1}{2}\nabla_i^2 + \frac{1}{2}\omega^2 r_i^2 \right), \tag{3.1}$$

with $N = 2$ and in two dimensions. For a single electron in a two-dimensional (isotropic) harmonic oscillator, the energy is $E_s = \hbar\omega(n_x + n_y + 1)$, where $n_x$ and $n_y$ are the quantum numbers for the $x$ and $y$ dimensions. The ground state (See Figure 3.1) is obtained for $(n_x, n_y) = (0, 0)$ which gives $E_s = 1\,\text{a.u.}$. Since we have two particles, the ground state energy of our system is in total

$$E_{\text{gs}} = 2\,\text{a.u.}.$$

In general, for $\omega \neq 1$, the energy will just be $E_{\text{gs}} = 2\omega$.

Now, we have to choose a trial wave-function. For a single electron, the exact wave-function of the ground state is

$$\phi_{n_x,n_y}(x, y) = A H_{n_x}(\sqrt{\omega}x) H_{n_y}(\sqrt{\omega}y) \exp\left(-\omega\left(x^2 + y^2\right)/2\right), \tag{3.2}$$

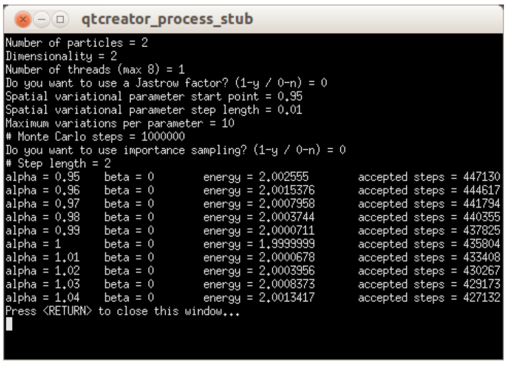
being $A$ a normalization constant and $H_n$ is the (physical) Hermite polynomial of order $n$. Since the 0-th order Hermite polynomials are just 1 and the two-particles wave-function can be seen as a product of the single-particle wave-functions, our total wave-function will just be

$$\phi(\vec{r}_1, \vec{r}_2) = C \exp\left(-\omega \left(r_1^2 + r_2^2\right)/2\right). \tag{3.3}$$

Adding a variational parameter $\alpha$ and forgetting about the normalization constant $C$, we obtain our trial wave-function for this case:

$$\psi_T(\vec{r}_1, \vec{r}_2) = \exp\left(-\alpha\omega \left(r_1^2 + r_2^2\right)/2\right). \tag{3.4}$$

Note that, since the correct solution of this problem is equation (3.3), we expect to obtain a variational energy *exactly* equal to 2 for $\alpha = 1$.



```
Number of particles = 2
Dimensionality = 2
Number of threads (max 8) = 1
Do you want to use a Jastrow factor? (1-y / 0-n) = 0
Spatial variational parameter start point = 0.95
Spatial variational parameter step length = 0.01
Maximum variations per parameter = 10
# Monte Carlo steps = 1000000
Do you want to use importance sampling? (1-y / 0-n) = 0
# Step length = 2
alpha = 0.95    beta = 0        energy = 2.002555       accepted steps = 447130
alpha = 0.96    beta = 0        energy = 2.0015376      accepted steps = 444617
alpha = 0.97    beta = 0        energy = 2.0007958      accepted steps = 441794
alpha = 0.98    beta = 0        energy = 2.0003744      accepted steps = 440355
alpha = 0.99    beta = 0        energy = 2.0000711      accepted steps = 437825
alpha = 1       beta = 0        energy = 1.9999999      accepted steps = 435804
alpha = 1.01    beta = 0        energy = 2.0000678      accepted steps = 433408
alpha = 1.02    beta = 0        energy = 2.0003956      accepted steps = 430267
alpha = 1.03    beta = 0        energy = 2.0008373      accepted steps = 429173
alpha = 1.04    beta = 0        energy = 2.0013417      accepted steps = 427132
Press <RETURN> to close this window...
```

*Figure 3.2: The program running for a two-electron system without repulsion.*

The electron-electron repulsion can be simply turned off by commenting out the line

```
    MySystem.add_potential(new eRepulsion(gP));
```

in the code. Running the program in Qt Creator shows the window in Figure 3.2, where some parameters can be specified before the actual calculation begins.

Just using a finer step for $\alpha$, we obtain the energy graph shown in Figure 3.3.

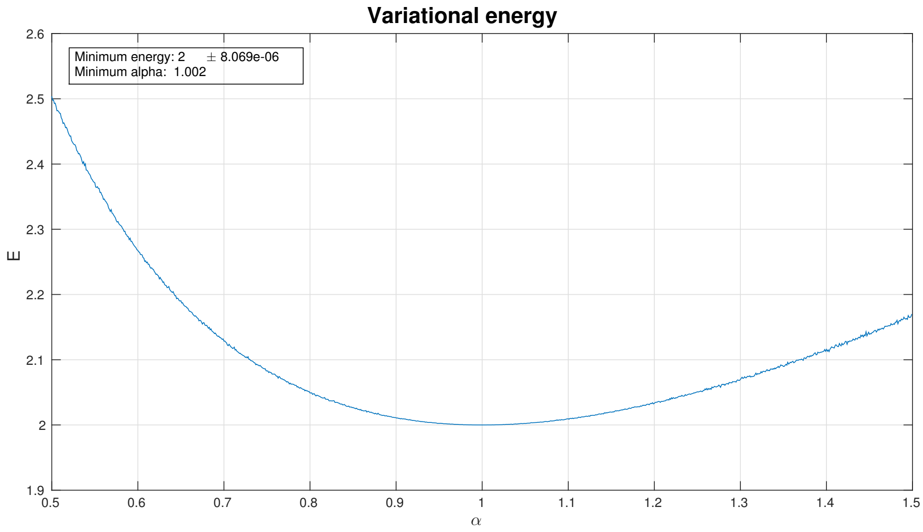


*Figure 3.3: The variational energy versus the variational parameter $\alpha$. The settings used are: brute force sampling with step length 2, no Jastrow factor, no parallelization,* 1000 *variations of* $\alpha$ *around* 1 *with step* 0.001, $1 \times 10^7$ *Monte Carlo steps. Acceptance ratio varies from* 40 *to* 60 %.

As you can see in Figure 3.3, we get $E_{\text{gs}} = 2$ with a high degree of precision for $\alpha = 1$[1].

We can also state that the total spin (better: its $z$-component) of the system is 0, as shown in equation (2.19). If you want a more "physical" reason than just reading off a table of coefficients, this is a direct consequence of the shape of the wave-function for a fermionic system, as shown in equation (2.17). In fact, in our case the spatial part of the two single-particle wave-functions is equal (they occupy the same energy level), yielding to

$$|1\rangle = |\phi\rangle \otimes |\chi_1\rangle \,, \qquad |2\rangle = |\phi\rangle \otimes |\chi_2\rangle \,. \tag{3.5}$$

Comparing with equation (2.15) one sees that $|\psi\rangle = 0$ unless $|1\rangle \neq |2\rangle$, and since the spatial parts are equal one has to conclude that $|\chi_1\rangle \neq |\chi_2\rangle$[2]. Now,

$$|\chi_1\rangle = |s, m_s\rangle \tag{3.6}$$

with $s = 1/2$, so we have to say that – for example – particle 1 has spin-up ($m_s = 1/2$) and particle 2 has spin-down ($m_s = -1/2$). This fact is also known as the *Pauli exclusion principle*, and applies for a generic system of fermions.

---

[1] actually the minimum is for $\alpha = 1.002$, but in this case we are not using so many steps to trust the third decimal digit. The smaller $\alpha$-step was just used to have a nicer plot.

[2] We don't want $|\psi\rangle = 0$ because it's not a physical state and it is not normalizable.

## 3.2 The complete system

We want now to analyse the system with the full Hamiltonian, namely the Hamiltonian in equation (2.12):

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + \frac{1}{2}\omega r_1^2 + \frac{1}{2}\omega r_2^2 + \frac{1}{|\vec{r_1} - \vec{r_2}|}, \tag{3.7}$$

where $\vec{r_1}$ and $\vec{r_2}$ are the position operators of the two particles.

You see that this time we have a problem when two particles are very near to each other, since the term

$$\sum_{i<j} \frac{1}{|\vec{r_1} - \vec{r_2}|}$$

tends to be a division by zero. Introducing the substitution

$$\begin{cases} \vec{r} = \vec{r_1} - \vec{r_2} \\ \vec{R} = \frac{1}{2}(\vec{r_1} + \vec{r_2}) \end{cases} \tag{3.8}$$

the Hamiltonian can be rewritten in a decoupled way as[3]

$$\hat{H} = -\nabla_{\vec{r}}^2 - \frac{1}{4}\nabla_{\vec{R}}^2 - \frac{1}{4}\omega^2 \vec{r}^2 + \omega^2 \vec{R}^2 + \frac{1}{r}. \tag{3.9}$$

Since the Hamiltonian is independent on spin, we can factorize the wave-function as

$$\psi(\vec{r}, \vec{R}, m_{s_1}, m_{s_2}) = \varphi(\vec{r})\xi(\vec{R})\chi(m_{s_1}, m_{s_2}) \tag{3.10}$$

which separates the Schrödinger equation into

$$\left(-\frac{1}{2}\nabla_{\vec{r}}^2 + \frac{1}{2}\omega_r^2 \vec{r}^2 + \frac{1}{2}\frac{1}{r}\right)\varphi(\vec{r}) = E_r \varphi(\vec{r}) \tag{3.11}$$

and

$$\left(-\frac{1}{2}\nabla_{\vec{R}}^2 + \frac{1}{2}\omega_R^2 \vec{R}^2\right)\xi(\vec{R}) = E_R \xi(\vec{R}), \tag{3.12}$$

where $\omega_r = \frac{1}{2}\omega$ and $\omega_R = 2\omega$.

We are not interested in the $R$-part of the solution, since it gives no troubles; we are more interested in equation (3.11), because we can cancel the divergence for $r \to 0$ working on the solution. We can also discard about the oscillator potential because it gives no divergence. Since we are interested only in the radial part of the solution, the relevant contributions from the laplacian in spherical coordinates are

---

[3] See Taut (1993).

$$\nabla_r^2 = \frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr}. \tag{3.13}$$

We also notice that $\frac{d^2}{dr^2}$ gives a finite contribution (is just the momentum part) and so does $\frac{d}{dr}$. The only piece that is not finite is the $\frac{1}{r}$ in front of $\frac{d}{dr}$, that we can use to cancel the divergence in the interaction potential. Using these facts, equation (3.11) can be rewritten in a way that is called the *cusp condition* for this system, namely

$$\lim_{r \to 0} \frac{1}{\varphi(r)} \left( -\frac{1}{\cancel{2}}\frac{\cancel{1}}{r}\frac{d}{dr} + \frac{1}{\cancel{2}}\frac{\cancel{1}}{r} \right) \varphi(r) = 0. \tag{3.14}$$

Solving for $\varphi$ by partial integration gives

$$\varphi(r) = \exp(Cr), \qquad\qquad r \to 0 \tag{3.15}$$

where $C$ is a constant. For these types of Monte Carlo calculations, a so-called *Padé-Jastrow* factor that has the property (3.15) is often used. This factor has the form (for two electrons)

$$\exp\left(\frac{ar}{(1+\beta r)}\right), \tag{3.16}$$

where (in 2 dimensions)

$$\beta = \begin{cases} 1 & \text{anti-parallel spin} \\ 1/3 & \text{parallel spin} \end{cases} \tag{3.17}$$

and $\beta$ is a variational parameter. With this factor, our trial wave-function reads as

$$\psi_T(\vec{r_1}, \vec{r_2}) = \exp\left(-\alpha\omega\left(r_1^2 + r_2^2\right)/2\right)\exp\left(\frac{ar}{(1+\beta r)}\right). \tag{3.18}$$

Using the ansatz in equation (3.18) in our program, we obtain a preliminary result shown in Figure 3.4.

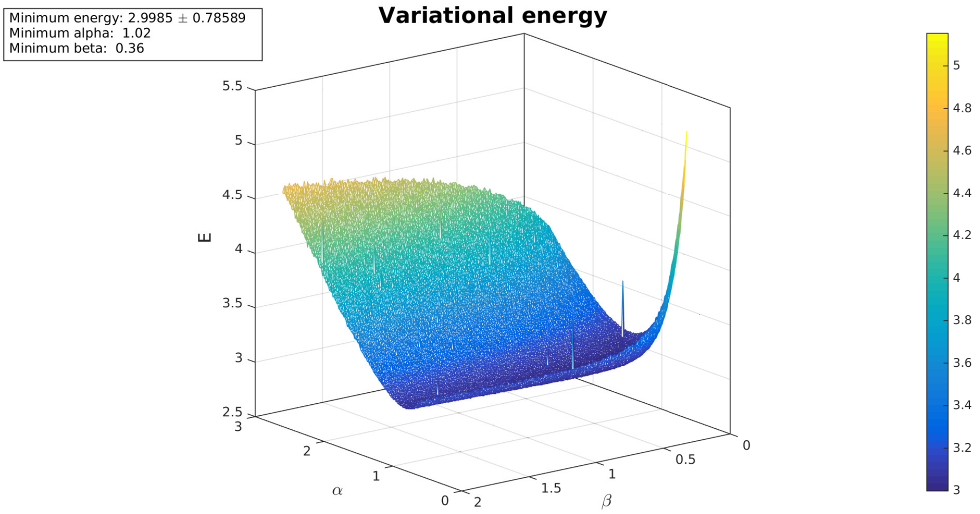


*Figure 3.4: The variational energy versus the variational parameters $\alpha$ and $\beta$. The settings used are: brute force sampling with step length 2, Jastrow factor, no parallelization,* 200 *variations of $\alpha$ and $\beta$ with step* 0.01*,* $1 \times 10^5$ *Monte Carlo steps. Acceptance ratio varies from* 45 *to* 55 %*.*

Refining around the minimum value to have better results, we finally obtain the plot in Figure 3.5.

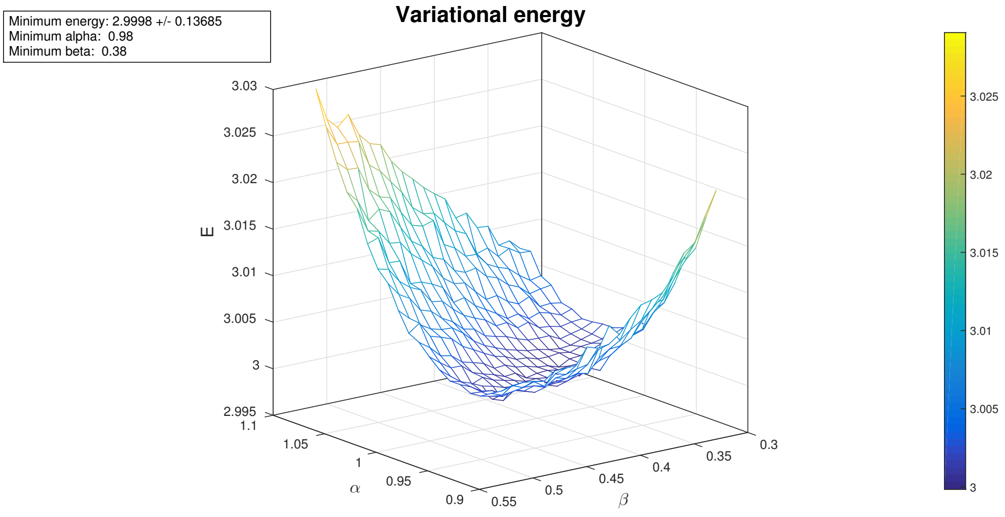


*Figure 3.5: The variational energy versus the variational parameters $\alpha$ and $\beta$. The settings used are: brute force sampling with step length 2, Jastrow factor, no parallelization,* 20 *variations of $\alpha$ and $\beta$ with step* 0.01*,* $1 \times 10^6$ *Monte Carlo steps. Acceptance ratio varies from* 45 *to* 55 %*. The relative distance is about* 1.64 *in natural units.*

You see that we are very near to the exact result $E_{\mathrm{gs}} = 3\,\mathrm{a.u.}$ calculated from Taut (1993) and reported in Pedersen Lohne et al. (2011). However, the energy error is still quite big; a way to improve the precision of our program is to *parallelize* it.

### 3.2.1 Parallelization

In this project, parallelization is done by using OpenMP[4]. The main idea is to run multiple computations of $E_T$ for the same pair of variational parameters $(\alpha, \beta)$ and then take the mean. Since the "measurements" of $E_T$ are not correlated, the error on its mean value falls like $1/\sqrt{n}$, where $n$ is the number of measurements. To be sure that the measurements are not correlated we have to make *private* local copies of the variables used in the computation, one for each thread used. Then, we have to sync al the threads to be sure that all the measurements are complete, and finally we can take the mean. An example of how this is implemented in our program is given below:

```
// Begin parallelization
omp_set_num_threads(gP.num_threads);
#pragma omp parallel shared(cumulative_e, cumulative_e2)
{
    // Make a private copy of all the stuff
    mat cumulative_e_local = zeros<mat>(vmcParams.max_variations+1,
        vmcParams.max_variations+1);
    mat cumulative_e2_local = zeros<mat>(vmcParams.max_variations+1,
        vmcParams.max_variations+1);
    struct GeneralParams gP_local = gP;
    struct VariationalParams vP_local = vP;
    struct VMCparams vmcParams_local = vmcParams;

    // Build the orbitals
    Orbitals* MyWaveFunction_local = new AlphaHarmonicOscillator(
        gP_local, vP_local);

    // Set up the Jastrow factor
    Jastrow* MyJastrowFactor_local;
    if (vmcParams_local.jF_active) MyJastrowFactor_local = new
        Pade_Jastrow(gP,vP);
    else MyJastrowFactor_local = new No_Jastrow();

    // Make a local copy of the system for convenience
    System MySystem_local = MySystem;

```

---

[4] http://openmp.org/.

```
        // Do the mc sampling
        mc_sampling(gP_local, vP_local, vmcParams_local,
                    cumulative_e_local, cumulative_e2_local,
                    MySystem_local, MyWaveFunction_local,
                        MyJastrowFactor_local);

        // Be sure to have all the contributions
        #pragma omp barrier
        #pragma omp critical
        {
            // Add the contributions
            cumulative_e += cumulative_e_local;
            cumulative_e2 += cumulative_e2_local;
        }
    }

    // Normalize to the number of threads
    cumulative_e = cumulative_e/((double) gP.num_threads);
    cumulative_e2 = cumulative_e2/((double) gP.num_threads);
```

Re-doing the calculations using parallelization gives the result in Figure 3.6.

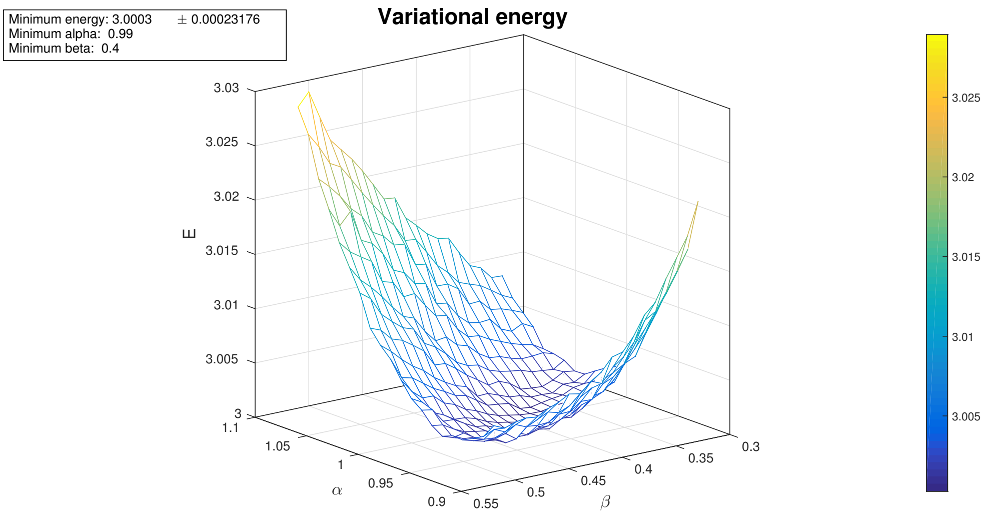


*Figure 3.6: The variational energy versus the variational parameters $\alpha$ and $\beta$. The settings used are: brute force sampling with step length 2, Jastrow factor, parallelization (8 threads), 20 variations of $\alpha$ and $\beta$ with step 0.01, $2 \times 10^5$ Monte Carlo steps. Acceptance ratio varies from 45 to 55 %.*

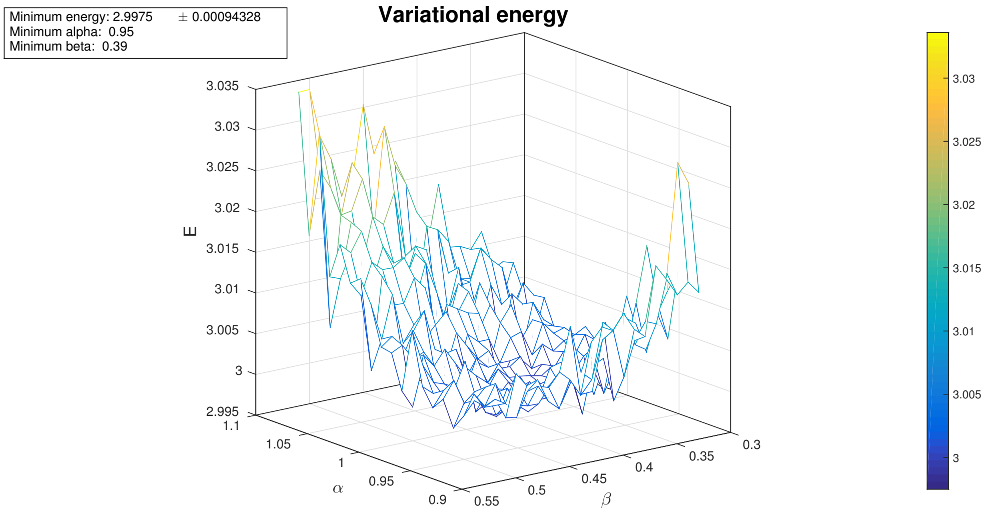


*Figure 3.7: The variational energy versus the variational parameters* $\alpha$ *and* $\beta$*. The settings used are: importance sampling with* $\Delta t = 0.001$*, Jastrow factor, parallelization (8 threads),* 20 *variations of* $\alpha$ *and* $\beta$ *with step* 0.01*,* $2 \times 10^5$ *Monte Carlo steps. Acceptance ratio is about* 99.999 %.

Note that this time, even using a lower number of Monte Carlo steps, we have an amazingly precise estimate for our energy, comparable with the one obtained in Pedersen Lohne et al. (2011) using Diffusion Monte Carlo.

### 3.2.2 Importance sampling

Let's now try to implement importance sampling, as explained in Section 2.3.2. In order to show the differences with the brute force algorithm, we will repeat the calculations with exactly the same settings used in Figure 3.6, this time with importance sampling active. We will also try to do the calculations for different time-steps $\Delta t = 0.001, 0.01, 0.1$.

The results are shown in Figures 3.7, 3.8 and 3.9.

As you see, there is a huge difference between the result obtained with $\Delta t = 0.001$ and the one obtained with $\Delta t = 0.1$. This is because, if the time step is too small and we don't have enough Monte Carlo steps, our walker cannot span completely the integration space. It will remain near its starting point and all the points reached will be inside the integration space, close to each other. In fact, the acceptance ratio for the case $\Delta t = 0.001$ and $\Delta t = 0.01$ is *practically* 100 %. The third case, for $\Delta t = 0.1$, has a little lower acceptance ratio, that is about 99.79 %. This fact guarantees that our walker is reaching all the borders of our integration space, and in fact it also reaches

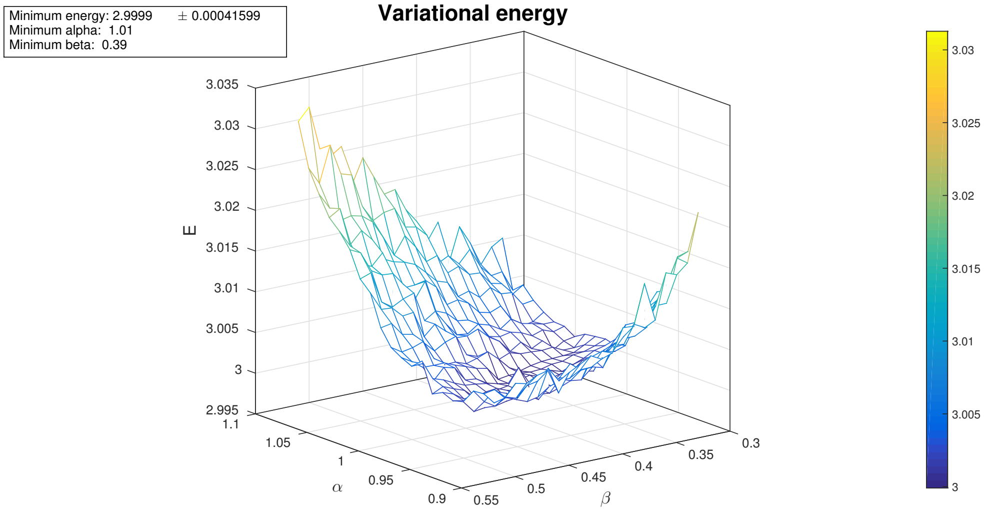


*Figure 3.8: The variational energy versus the variational parameters $\alpha$ and $\beta$. The settings used are: importance sampling with $\Delta t = 0.01$, Jastrow factor, parallelization (8 threads), 20 variations of $\alpha$ and $\beta$ with step 0.01, $2 \times 10^5$ Monte Carlo steps. Acceptance ratio is about 99.999 %.*

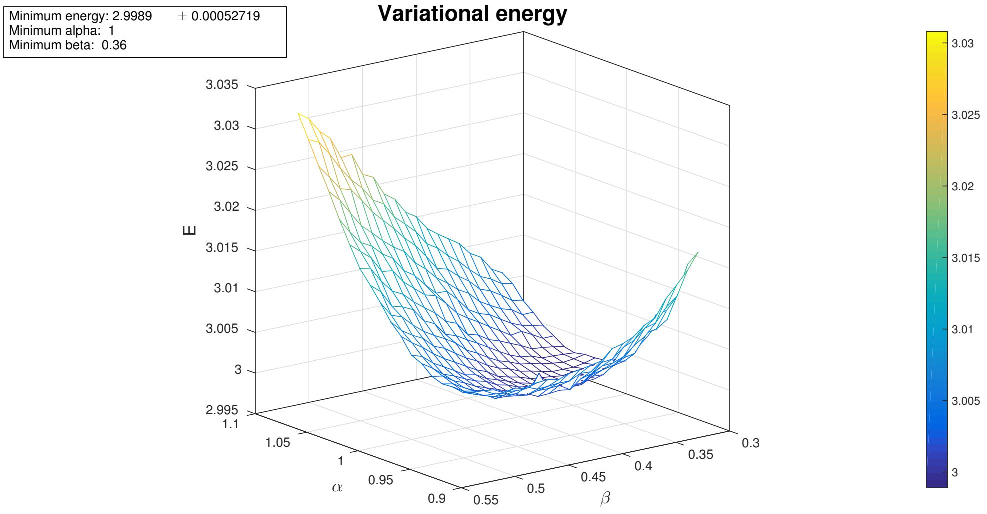


*Figure 3.9: The variational energy versus the variational parameters $\alpha$ and $\beta$. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 20 variations of $\alpha$ and $\beta$ with step 0.01, $2 \times 10^5$ Monte Carlo steps. Acceptance ratio is about 99.793 %.*

some points outside it. In other words, for the given number of Monte Carlo iterations, we have a large enough time-step to cover all the space. If one wants to use a lower time-step to have better precision, he has to increase the number of Monte Carlo cycles; as an example, for $\Delta t = 0.001$, one has to use at least $1 \times 10^9$ Monte Carlo loops.

# Chapter 4

# The 6-electrons system

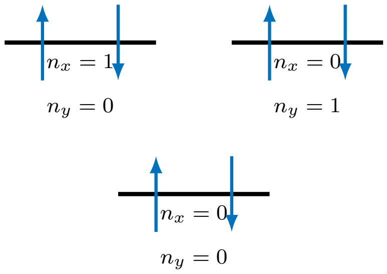


*Figure 4.1: The 6-electrons system configuration.*

The formulae discussed in the previous sections can be generalized for a $N$ electrons case. The trial wave function can be written as

$$\psi_T(\vec{r}_1 \ldots \vec{r}_N) = A\,\mathrm{Det}(S) \prod_{i<j}^{N} e^{\frac{r_{ij}}{1+\beta r_{ij}}}, \tag{4.1}$$

where Det$(S)$ is the Slater determinant

$$\mathrm{Det}(S) = \begin{vmatrix} \phi_1(\vec{r}_1) & \phi_2(\vec{r}_1) & \ldots & \phi_N(\vec{r}_1) \\ \phi_1(\vec{r}_2) & \phi_2(\vec{r}_2) & \ldots & \phi_N(\vec{r}_1) \\ \vdots & & & \vdots \\ \phi_1(\vec{r}_N) & \phi_2(\vec{r}_N) & \ldots & \phi_N(\vec{r}_N) \end{vmatrix} \tag{4.2}$$

As we did in Section 2.2, this determinant can be further simplified if we label our electrons in a smart way, as shown in Hjorth-Jensen (2014).

Let's consider a spin-independent quantum mechanical operator $\hat{O}(\vec{r})$ acting on our trial wave function. The latter depends on $\vec{x} = (\vec{r}, \sigma)$ where $\vec{r}$ includes the space coordinates and $\sigma$ the spin coordinates. We have that:

$$\left\langle \hat{O} \right\rangle = \frac{\left\langle \psi_T(\vec{x}) \middle| \hat{O}(\vec{r}) \middle| \psi_T(\vec{x}) \right\rangle}{\langle \psi_T(\vec{x}) \,|\, \psi_T(\vec{x}) \rangle} \tag{4.3}$$

Now we can replace the total antisymmetric wave function with one with permuted arguments and we can arbitrarily choose that the first $N/2$ arguments are spin up and the other half are spin down. We get:

$$\begin{aligned} \psi_T(\vec{x_1}, \ldots, \vec{x_N}) \longrightarrow & \, \psi_t(\vec{x_{i1}}, \ldots, \vec{x_{iN}}) && (4.4) \\ = & \, \psi_T(\{\vec{r_{i1}}, \uparrow\}, \ldots, \{r_{\vec{iN/2}}, \uparrow\}, \{\vec{r_{i1}}, \downarrow\}, \ldots, \{r_{\vec{iN/2}}, \downarrow\}) && (4.5) \\ = & \, \psi_T(\{\vec{r_1}, \uparrow\}, \ldots, \{r_{\vec{N/2}}, \uparrow\}, \{r_{\vec{N/2+1}}, \downarrow\}, \ldots, \{\vec{r_N}, \downarrow\}) && (4.6) \end{aligned}$$

The operator $\hat{O}$ is symmetric with the respect to the exchange of labels in a pair of particles, so:

$$\left\langle \hat{O} \right\rangle = \frac{\left\langle \psi_T(\vec{r}) \middle| \hat{O}(\vec{r}) \middle| \psi_T(\vec{r}) \right\rangle}{\langle \psi_T(\vec{r}) \,|\, \psi_T(\vec{r}) \rangle} \tag{4.7}$$

The wave function is now antisymmetric with respect to exchange of spatial coordinates of pairs of spin-up or spin-down electrons. Therefore, for spin-independent Hamiltonians, the Slater determinant can be splitted in a product of two Slater determinants, one for the single particle orbitals with spin up and the other with single particle orbitals with spin down. We have then:

$$\psi_S = S^{\uparrow} S^{\downarrow} \tag{4.8}$$

where

$$S^{\uparrow} = \begin{vmatrix} \phi_1(\vec{r}_1) & \phi_2(\vec{r}_1) & \ldots & \phi_{N/2}(\vec{r}_1) \\ \phi_1(\vec{r}_2) & \phi_2(\vec{r}_2) & \ldots & \phi_{N/2}(\vec{r}_1) \\ \vdots & & & \vdots \\ \phi_1(\vec{r}_{N/2}) & \phi_2(\vec{r}_{N/2}) & \ldots & \phi_{N/2}(\vec{r}_{N/2}) \end{vmatrix} \tag{4.9}$$

and $S^{\downarrow}$ is obviously defined in a similar way. The $\phi_i$'s functions are just as the one defined in equation (3.2).

So, our final trial wave-function is just

$$\psi_T = |S^{\uparrow}||S^{\downarrow}|J \tag{4.10}$$

being $J$ the Padé-Jastrow factor.

The C++ implementation of $\psi_S$ is therefore straightforward and it is shown below.

```
double Orbitals::SlaterD(mat& r) {
    mat D_up(n_half,n_half), D_down(n_half,n_half);

    for (int i = 0; i < n_half; i++) {
        this->set_qnum_indie_terms(r, i);
        for (int j = 0; j < n_half; j++) {
            D_up(i,j) = this->phi(r, i, j);
        }
    }

    for (int i = 0; i < n_half; i++) {
        this->set_qnum_indie_terms(r, i+n_half);
        for (int j = 0; j < n_half; j++) {
            D_down(i,j) = this->phi(r, i+n_half, j);
        }
    }

    double slater = det(D_up)*det(D_down);

    D_up.reset();
    D_down.reset();

    return slater;
}
```

## 4.1 $\omega = 0.28$

We start by studying our system for $\omega = 0.28$. A preliminary and very rough result is shown in Figure 4.2.

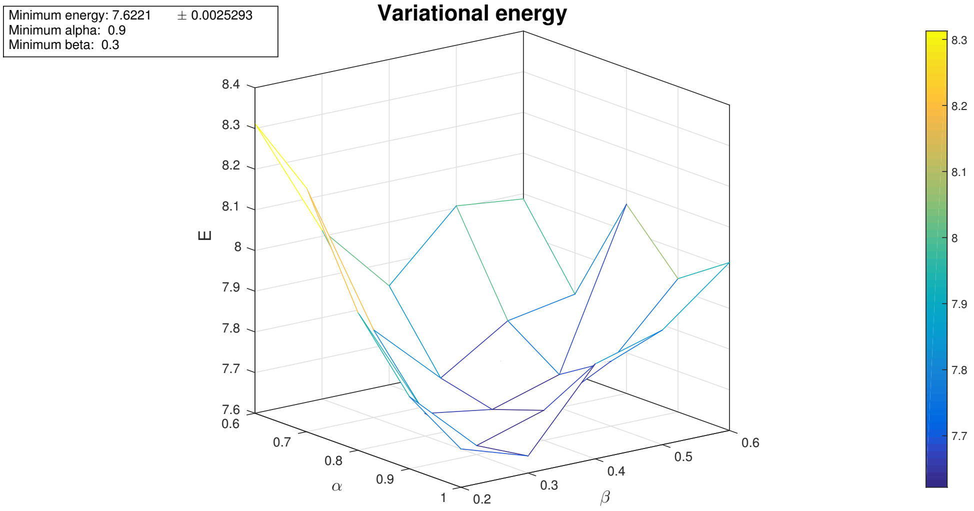


*Figure 4.3: The variational energy versus the variational parameters $\alpha$ and $\beta$. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 5 variations of $\alpha$ and $\beta$ with step 0.1, $2 \times 10^5$ Monte Carlo steps. $\omega = 0.28$.*

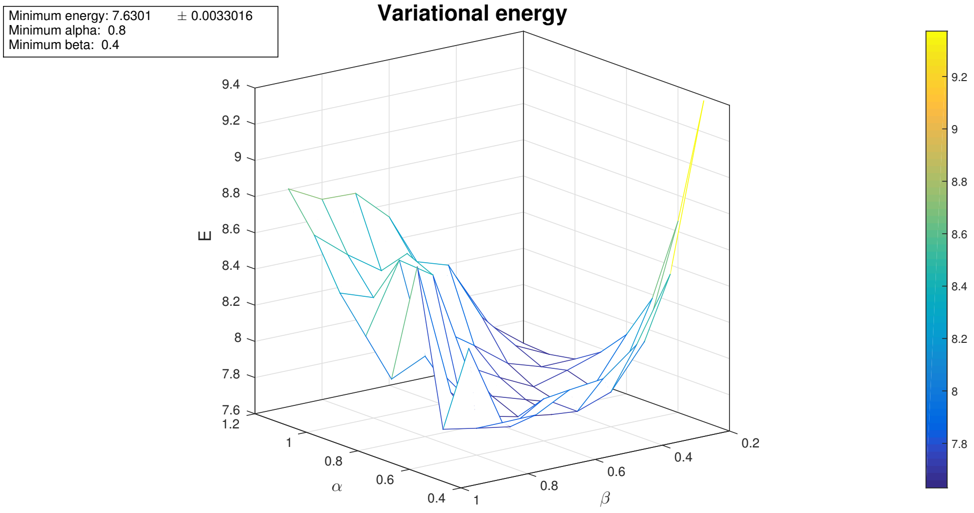


*Figure 4.2: The variational energy versus the variational parameters $\alpha$ and $\beta$. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 8 variations of $\alpha$ and $\beta$ with step 0.1, $1 \times 10^5$ Monte Carlo steps. $\omega = 0.28$.*

Restricting to the significant region, we get the result in Figure 4.3.

You see that the plot is no more nice like the 2-electrons case, mainly because we didn't used enough Monte Carlo steps. The problem is that now the program has to calculate $3 \times 3$ determinants every time, so it's much slower than the previous case. To refine the result, we just doubled the number of Monte Carlo loops and restricted a little bit the variational parameters grid; taking a finer grid with more Monte Carlo cycles would have required ages, literally.

Anyway, still with this configuration, we get a result that is really near the one calculated by DMC in Pedersen Lohne et al. (2011), that turns out to be 7.6001(1) a.u..

A nice way to test the Slater determinant code is to turn off the electron-electron repulsion. In that case, $\psi_S$ is the *exact* solution (for $\alpha = 1$). We already showed in Section 3.1 that the ground-state energy for the single electron is $E_s = \hbar\omega(n_x + n_y + 1)$. Since – in the 6 electrons case – we have 2 electrons with $n_x + n_y = 0$ and 4 electrons with $n_x + n_y = 1$ (See Figure 4.1), the total energy for the ground state is

$$E_{\text{gs}} = \hbar\omega\left[2(0+1) + 4(1+1)\right] = 10\omega. \tag{4.11}$$

Turning off the repulsion results in a minimum $E_{\text{gs}} = 2.799\,999\,999\,999\,9(2)$ a.u. for $\alpha = 1$, that confirms the prediction.

## 4.2 $\omega = 0.5$

We repeat the same calculations with $\omega = 0.5$, obtaining the results in Figure 4.4.

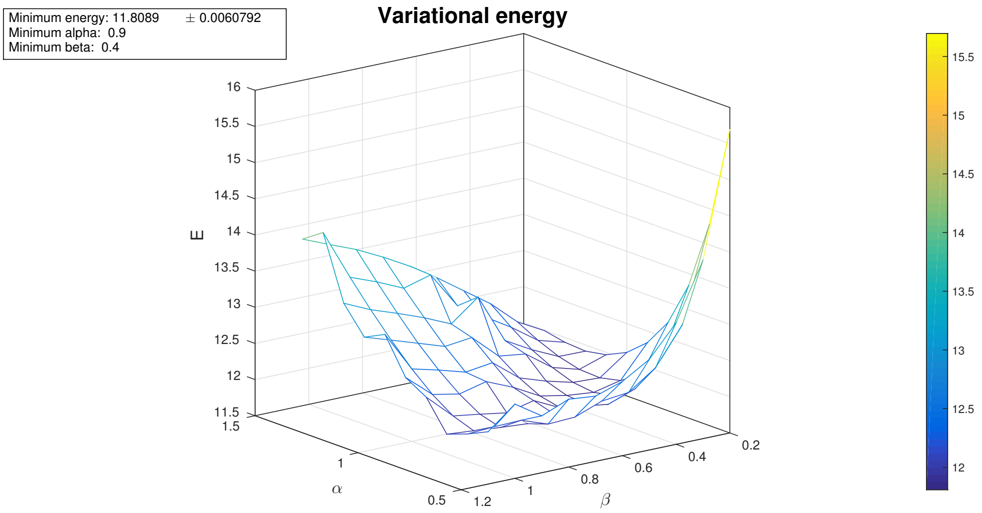


*Figure 4.4: The variational energy versus the variational parameters $\alpha$ and $\beta$. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 10 variations of $\alpha$ and $\beta$ with step 0.1, $2 \times 10^5$ Monte Carlo steps. $\omega = 0.50$.*

Again, our result is pretty consistent with the one calculated by DMC in Pedersen Lohne et al. (2011), that is 11.7888(2) a.u..

## 4.3 $\omega = 1$

Just with the same parameters used in Figure 4.4 but changing the frequency to $\omega = 1.00$, we have the results in Figure 4.5.

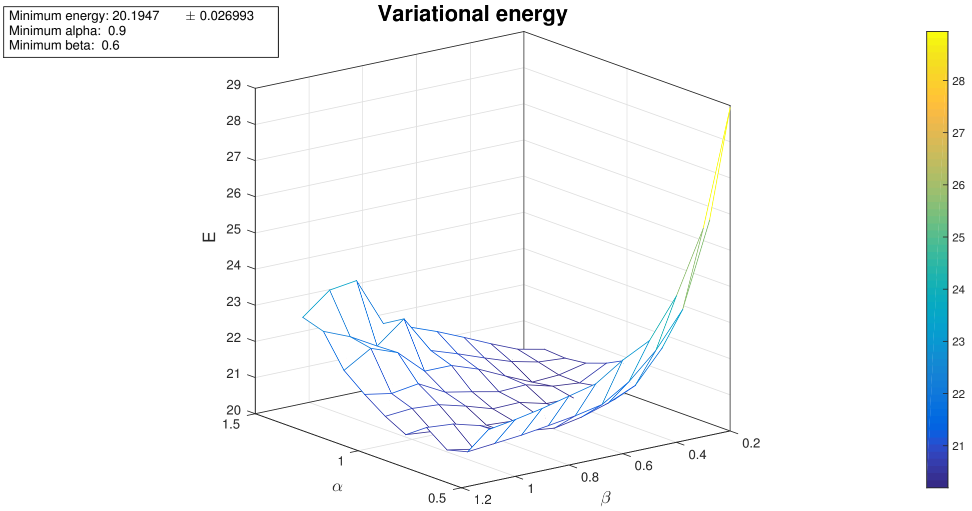


*Figure 4.5: The variational energy versus the variational parameters $\alpha$ and $\beta$. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads),* 10 *variations of $\alpha$ and $\beta$ with step* 0.1, $2 \times 10^5$ *Monte Carlo steps. $\omega = 1.00$.*

Once again, our result is in very good accordance with the one calculated by DMC in Pedersen Lohne et al. (2011), that is 20.1597(2) a.u..

# Chapter 5

# The virial theorem

The virial theorem states that the expectation value of the total kinetic energy $\langle T \rangle$ is proportional to the expectation value of the total potential energy $\langle V \rangle$. For a pure harmonic oscillator, this theorem is simply the theorem of the *equipartition theorem*:

$$\langle T \rangle = \langle V \rangle \,. \tag{5.1}$$

This is quite easy to prove. Let's introduce the ladder operators

$$\hat{a} = \frac{1}{\sqrt{2\hbar\omega m}} \left( m\omega\hat{x} + i\hat{p} \right) \qquad \text{and} \qquad \hat{a}^{\dagger} = \frac{1}{\sqrt{2\hbar\omega m}} \left( m\omega\hat{x} - i\hat{p} \right) \tag{5.2}$$

It can be shown that these operators act on the harmonic oscillator eigenstates such that

$$\hat{a} \left| n \right\rangle = \sqrt{n} \left| n-1 \right\rangle \quad \text{and} \quad \hat{a}^{\dagger} \left| n \right\rangle = \sqrt{n+1} \left| n+1 \right\rangle , \tag{5.3}$$

and we also have

$$N \left| n \right\rangle \doteqdot \hat{a}^{\dagger}\hat{a} \left| n \right\rangle = n \left| n \right\rangle . \tag{5.4}$$

Using these properties, the calculation of the expected value of the potential energy is straightforward:

$$\begin{aligned}
\left\langle n \left| \frac{1}{2} m\omega^2 \hat{x}^2 \right| n \right\rangle &= \frac{1}{2} m\omega^2 \left\langle n \,|\, \hat{x}^2 \,|\, n \right\rangle \qquad (5.5)\\
&= \frac{1}{2} \cancel{m\omega}^{\cancel{2}} \frac{\hbar}{2\cancel{m\omega}} \left\langle n \left| \left(\hat{a}^\dagger + \hat{a}\right)\left(\hat{a} + \hat{a}^\dagger\right) \right| n \right\rangle \qquad (5.6)\\
&= \frac{\hbar\omega}{4} \left( \left\langle n \left| \hat{a}^\dagger \hat{a}^\dagger \right| n \right\rangle + \left\langle n \left| \hat{a}\hat{a}^\dagger \right| n \right\rangle + \left\langle n \left| \hat{a}^\dagger \hat{a} \right| n \right\rangle + \left\langle n \,|\, \hat{a}\hat{a} \,|\, n \right\rangle \right) \qquad (5.7)\\
&= \frac{\hbar\omega}{4} \Big( \sqrt{n+2}\sqrt{n+1} \cancelto{0}{\langle n | n+2 \rangle} \qquad (5.8)\\
&+ \left\langle n \left| \hat{N} + 1 \right| n \right\rangle + \left\langle n \left| \hat{N} \right| n \right\rangle + \sqrt{n}\sqrt{n-1} \cancelto{0}{\langle n | n-2 \rangle} \Big) \qquad (5.9)\\
&= \frac{\hbar\omega}{4}((n+1)+n) = \frac{\hbar\omega}{4}(2n+1) \qquad (5.10)\\
&= \frac{\hbar\omega}{2}\left(n + \frac{1}{2}\right) \qquad (5.11)
\end{aligned}$$

If we perform the same calculation for $\left\langle n \left| \frac{\hat{p}^2}{2m} \right| n \right\rangle$ (the expected value of the kinetic energy) we will obtain the result, yielding

$$\langle T \rangle = \langle V \rangle \qquad (5.12)$$

and

$$\langle T \rangle + \langle V \rangle = \hbar\omega \left(n + \frac{1}{2}\right), \qquad (5.13)$$

that is the total energy of the harmonic oscillator, as expected.

Let's now graph the ratio $\langle T \rangle / \langle V \rangle$ for different values of $\omega$.

## 5.1 The 2-electrons system

We firstly computed $\langle T \rangle$ and $\langle V \rangle$ *without* the electron-electron repulsion, to check our code. We chose $\omega = 0.01$, 0.28, 0.50, 0.70, 1.00. The result – completely expected – is shown in Figure 5.1.

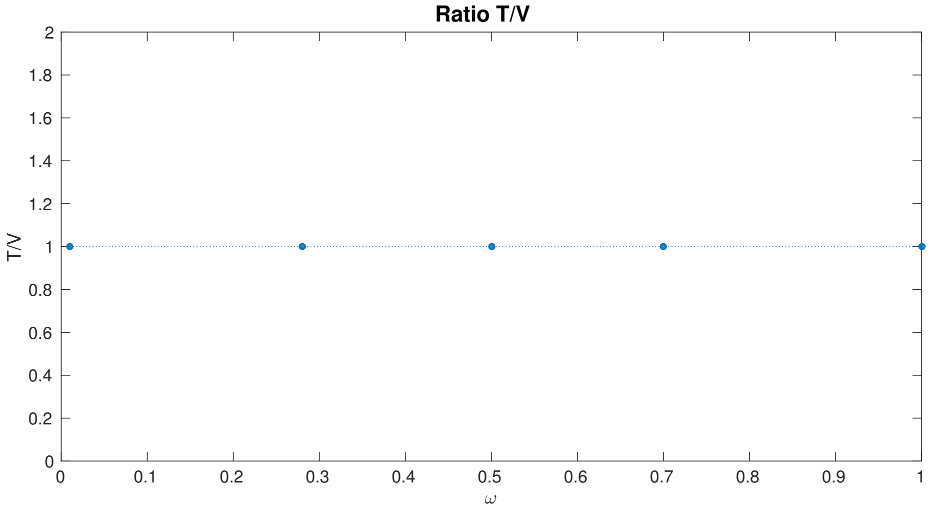


*Figure 5.1: The ratio* $\langle T \rangle / \langle V \rangle$ *for different values of* $\omega$ *and no electron-electron repulsion. The settings used are: importance sampling with* $\Delta t = 0.1$*, parallelization (8 threads),* $1 \times 10^7$ *Monte Carlo steps.*

After that, we added the electron-electron repulsion and we did again the calculations for the same $\omega$ values. The result, shown in Figure 5.2, is quite different.

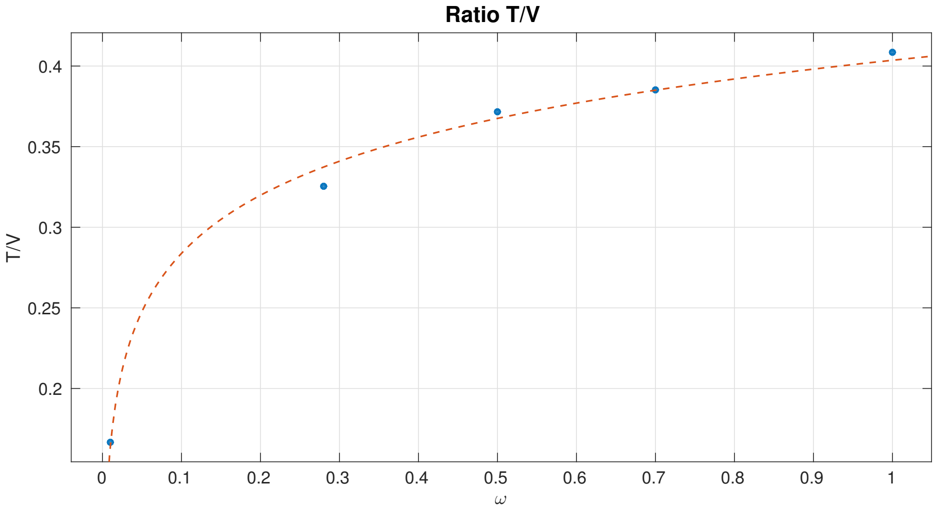


*Figure 5.2: The ratio* $\langle T \rangle / \langle V \rangle$ *for different values of* $\omega$ *with electron-electron repulsion. The settings used are: importance sampling with* $\Delta t = 0.1$*, Jastrow factor, parallelization (8 threads),* $1 \times 10^7$ *Monte Carlo steps. The fit is logarithmic.*

The dashed line is a fit of the type

$$\frac{\langle T\rangle}{\langle V\rangle} = b\log(c\omega), \tag{5.14}$$

which gives $b = 0.052(7)$ and $c = 2340(2021)$. We can't really say if the behaviour of our data is logarithmic without a theoretical basis and with just 5 points; the purpose of the dashed line in Figure 5.2 is just to show that our data is *no way* linear. This is due to the presence of the repulsive force, that can be dominant or not.

In a very intuitive way, we can at least say that what we obtain is coherent with the physics of the system. For $\omega \gg 1$ the oscillator potential is dominant and the repulsion can be considered as a small perturbation, which means that we have more kinetic energy than potential energy ($T > V$). For $\omega \ll 1$, we almost have only a repulsion potential plus a small harmonic perturbation; in other words, $T < V$.

If one really wants to find an analytical relation for the ratio $\langle T\rangle / \langle V\rangle$ of this system, he should compute the two terms separately in a way similar to the one shown in Section 5 and then take the ratio of them. Note that this is not always possible, since the necessary integrals could not be solved analytically.

## 5.2 The 6-electrons system

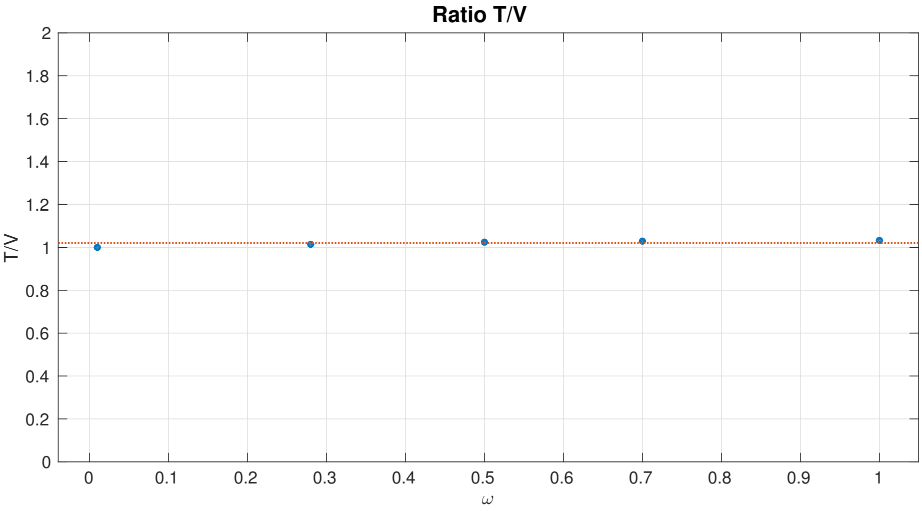


*Figure 5.3: The ratio* $\langle T \rangle / \langle V \rangle$ *for different values of* $\omega$ *and no electron-electron repulsion. The settings used are: importance sampling with* $\Delta t = 0.1$*, parallelization (8 threads),* $1 \times 10^6$ *Monte Carlo steps. A linear fit gives a ratio of* $1.02(2)$*, that is still compatible with the expected value* 1.

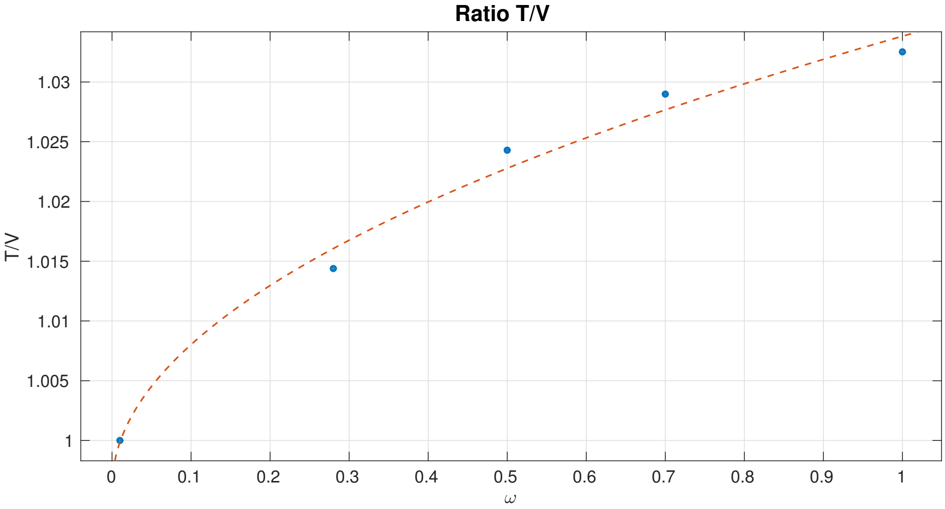


*Figure 5.4: The ratio* $\langle T \rangle / \langle V \rangle$ *for different values of* $\omega$ *with electron-electron repulsion. The settings used are: importance sampling with* $\Delta t = 0.1$*, Jastrow factor, parallelization (8 threads),* $1 \times 10^6$ *Monte Carlo steps. The fit is a square root.*

We did again the same calculations (with the same $\omega$ values) for 6 electrons. This time, we don't obtain a perfect 1 ratio for the non repulsive case (Figure 5.3); a linear fit gives 1.02(2), which is anyway still compatible with the expected value 1. A possible explanation for this behaviour is that we are not using enough Monte Carlo steps; they are ten times less than the previous case, when we are dealing with a number of electrons that is three times more the 2-electrons case.

The discussion of what we obtain adding the electron-electron repulsion, shown in Figure 5.4, is the same as the 2-electrons case in the previous section. The fit (that *is not* particularly indicative) is of the type

$$\frac{\langle T\rangle}{\langle V\rangle} = b\sqrt{c\omega} + d, \tag{5.15}$$

which gives $b = 0.06(1)$, $c = 0.416(1)$ and $d = 1.00(1)$.

# Chapter 6

# Analytical derivatives

The most time-consuming part of the Metropolis-Hastings algorithm is the computation of the acceptance ratio, the kinetic energy and the quantum force, that involve the calculation of several first and second derivatives. A way to – possibly – improve our program is to use analytical derivatives. Note that this is possible in very few cases – like this one, so we will take advantage of it.

Let's remember that the wave-function is

$$\psi_T = |\vec{S}^{\uparrow}||\vec{S}^{\downarrow}|J, \tag{6.1}$$

where $|\vec{S}^{\uparrow}|$ is the Slater determinant relative to the spin-up particles, $|\vec{S}^{\downarrow}|$ is the one relative to the spin-down particles and $J$ is the Jastrow factor.

For the quantum force $\vec{F}_i$ relative to particle $i$ we have that

$$\begin{align}\vec{F}_i &= 2\frac{1}{\psi_T}\nabla_i\Psi_T \tag{6.2}\\ &= 2\frac{\nabla_i(|\vec{S}^{\uparrow}||\vec{S}^{\downarrow}|J)}{|\vec{S}^{\uparrow}||\vec{S}^{\downarrow}|J} \tag{6.3}\\ &= 2\left(\frac{\nabla_i|\vec{S}^{\uparrow}|}{|\vec{S}^{\uparrow}|} + \frac{\nabla_i|\vec{S}^{\downarrow}|}{|\vec{S}^{\downarrow}|} + \frac{\nabla_i J}{J}\right). \tag{6.4}\end{align}$$

Since particle $i$ has either spin up or spin down, one of the two terms with the Slater determinant will vanish, because it's not dependant on that particle. So, labelling with $\alpha$ the spin of particle $i$, we have that

$$\vec{F}_i = 2\left(\frac{\nabla_i|\vec{S}^{\alpha}|}{|\vec{S}^{\alpha}|} + \frac{\nabla_i J}{J}\right) \tag{6.5}$$

Doing the same with the local energy, we end up with

$$E_L = \frac{1}{2}\sum_i \frac{1}{\psi_T}\nabla_i^2\psi_T + \sum_i V_i \tag{6.6}$$

Then,

$$\frac{1}{\psi_T}\nabla_i^2\psi_T = \frac{\nabla_i^2|\vec{S}^\alpha|}{|\vec{S}^\alpha|} + \frac{\nabla_i^2 J}{J} + 2\frac{\nabla_i|\vec{S}^\alpha|}{|\vec{S}^\alpha|}\frac{\nabla_i J}{J}, \tag{6.7}$$

where the components relative to the spin opposite to the one of the particle taken into account vanish for the reason explained before.

According to Høgberget (2013), the four terms that appear in (6.7) are

$$\frac{\nabla_i J}{J} = \sum_{k\neq i=1}^{N} \frac{a_{ik}}{r_{ik}}\frac{\vec{r}_i - \vec{r}_k}{(1+\beta\, r_{ik})^2} \tag{6.8}$$

$$\frac{\nabla_i^2 J}{J} = \left|\frac{\nabla_i J}{J}\right|^2 + \sum_{k\neq i=1}^{N} a_{ik}\frac{(d-3)(\beta\, r_{ik}+1)+2}{r_{ik}(\beta\, r_{ik}+1)^3} \tag{6.9}$$

$$\frac{\nabla_i|\vec{S}|}{|\vec{S}|} = \sum_k \nabla_i\phi_k(\vec{r}_i^{\,\text{new}})(\vec{S}_{ki}^{\text{new}})^{-1} \tag{6.10}$$

$$\frac{\nabla_i^2|\vec{S}|}{|\vec{S}|} = \sum_k \nabla_i^2\phi_k(\vec{r}_i^{\,\text{new}})(\vec{S}_{ki}^{\text{new}})^{-1} \tag{6.11}$$

with $k$ that spans the Slater matrix relative to the $i$th particle. The quantities $\nabla_i\phi_k$ and $\nabla_i^2\phi_k$ are also tabulated in Høgberget (2013). It's also possible to optimize the determinant of the inverse of the Slater matrix, but we opted here for a numerical computation of the inverse matrix using Armadillo. If one does that, it's also possible to optimize the acceptance ratio to have a even faster code.

## 6.1 Execution times

Let's see if we have a real gain in speed for the 2-electrons and the 6-electrons cases. Note that we have to run the program with a single thread to avoid extra waiting time on the CPU due to the line

```
        #pragma omp barrier
```

in the parallel cycles. Our benchmarks are shown in Figures 6.1 and 6.2.

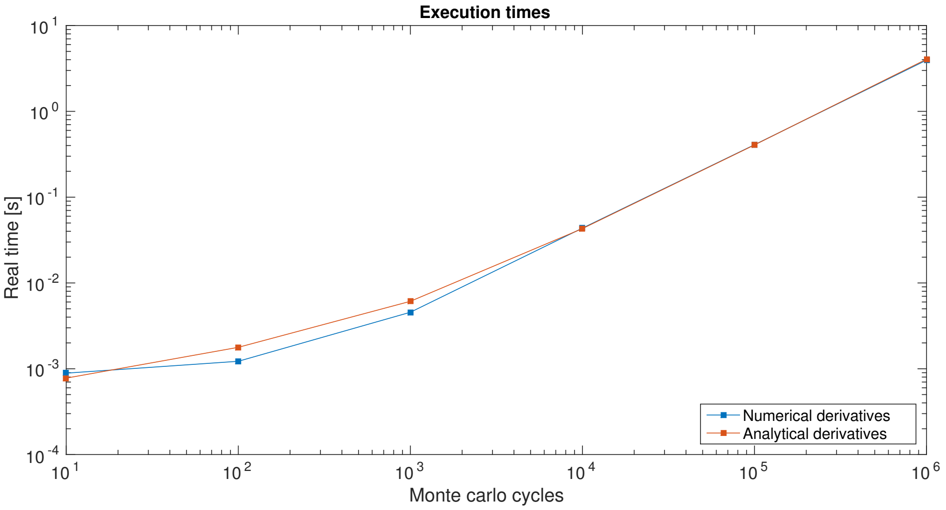


*Figure 6.1: Execution times for 2-electrons and a single pair of variational parameters $(\alpha, \beta)$. The GNU/Linux system tool `time` was used to take the measurements.*

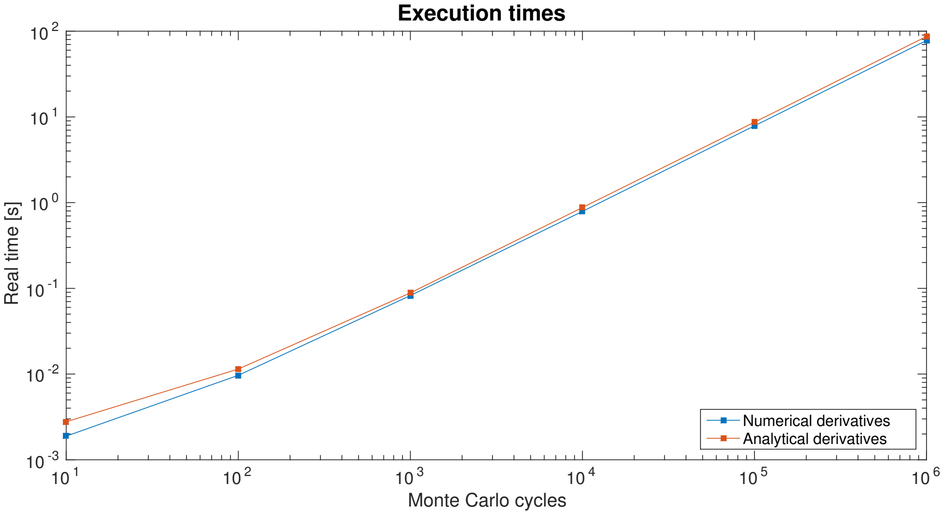


*Figure 6.2: Execution times for 6-electrons and a single pair of variational parameters $(\alpha, \beta)$. The GNU/Linux system tool `time` was used to take the measurements.*

As you see, we don't have great improvements; on the contrary, we obtain a code that is even a little bit slower than the numerical one for the 6-electrons case! However, there are some things that have to be pointed out. First of all we decided to perform a numerical inversion of a matrix, that is really time-consuming (it's a $\mathcal{O}(n^3)$ operation). Secondly there is no optimization of the acceptance ratio, that can be further simplified –

as shown again in Hø gberget (2013). Lastly, the benefits of numerical derivatives could not appear until one reaches 12 or even 20 particles; this is because the computational complexity of a determinant of a matrix is $\mathcal{O}(n^3)$, that for $3 \times 3$ matrices (as the 6-electrons case) is still acceptable. The struggle in calculating big determinants all the time is more evident for higher numbers of particles; for the 12-particles case the Slater determinant is 8 times slower than the 6-particles case, while for the 20-particles case is approximately 37 times slower.

Another thing that has to be noted is that – with the analytical derivatives – we obtain slightly lower values for the energy. Deactivating the electron-electron repulsion the discrepancy still persists, meaning that the reason of this fact has to be found in the Slater determinant derivatives implementation. A possible explanation is that we are (numerically) performing the inverse of a matrix that has very small elements, and this could cause a loss of precision.

# Chapter 7

# Conclusions and perspectives

In this project we really showed the power of the VMC method, that even if it doesn't give automatically the ground state energy of a system can be used to approximate it – to excess – with a considerable degree of precision. If we have no clues about the properties of the system we can use a brute force sampling, trying to manually set a length-step that gives about 50%. But if we are smarter, we can use the information we have about the expected result to implement a more efficient sampling.

This program has been implemented in a highly object-oriented way, following the suggestions contained in Hø gberget (2013), and can also fully handle the 12-electrons case (look at the implementation of the orbitals in the code). However, it is still *extremely slow* – practically unusable. So, I'm planning to improve this program using MPI and GPU parallelization, that would give me some extra speed.

I'm also planning to extend this program to handle a custom number of variational parameters, and to improve the trial wave-function with some extra terms as suggested by professor Francesco Pederiva[1] and/or using the Hartree-Fock method. This would require to abandon the analytical derivatives route and concentrate the efforts in the algorithm improvements in terms of efficiency.

Finally, for me this project has been a springboard into the quantum computational physics world; I enjoyed so much in doing this simulation that I've decided to goo deeper into this subject and start to explore of its full potential in advanced courses at UiO and NMBU.

---

[1] University of Trento, Department of Physics.

# Bibliography

M. Brack, Jens Damgaard, A. S. Jensen, H. C. Pauli, V. M. Strutinsky, and C. Y. Wong. Funny hills: The shell-correction approach to nuclear shell effects and its applications to the fission process. *Reviews of Modern Physics*, 44(2):320–405, April 1972. ISSN 00346861. doi: 10.1103/RevModPhys.44.320. URL `http://link.aps.org/doi/10.1103/RevModPhys.44.320`.

C. Fasth and A. Fuhrer. Coulomb blockade in Quantum Dots, April 2007. URL `http://www.ftf.lth.se/fileadmin/ftf/Course_pages/FFF042/cb_lecture07.pdf`.

D. J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Education, Inc., second ed. edition, 2005.

M. Hjorth-Jensen. Computational Physics – Lecture Notes Fall 2014, 2014. URL `http://www.uio.no/studier/emner/matnat/fys/FYS3150/h14/undervisningsmateriale/LectureNotes/lecture2014.pdf`.

Jø rgen Hø gberget. *Quantum Monte-Carlo Studies of Generalized Many-body Systems*. PhD thesis, University of Oslo, 2013. URL `http://urn.nb.no/URN:NBN:no-38645`.

M. A. Kastner. Artificial atoms. *Physics Today*, 46:24–31, 1993.

Walter Kohn. Cyclotron Resonance and de Haas-van Alphen Oscillations of an Interacting Electron Gas. *Phys. Rev.*, 123(4):1242–1244, 1961. ISSN 0031899X. doi: 10.1103/PhysRev.123.1242. URL `http://link.aps.org/doi/10.1103/PhysRev.123.1242`.

M. Koskinen, M. Manninen, and S. M Reimann. Hund's Rules and Spin Density Waves in Quantum Dots. *Phys. Rev. Lett.*, 79(7):1389–1392, 1997. ISSN 0031-9007. doi: 10.1103/PhysRevLett.79.1389. URL `http://link.aps.org/doi/10.1103/PhysRevLett.79.1389`.

K. S. Krane. *Introductory Nuclear Physics*. Wiley, 1988. ISBN 9780471859147. URL `https://books.google.no/books?id=rhVlQgAACAAJ`.

Arvind Kumar, Se Steven E Laux, and Frank Stern. Electron states in a GaAs quantum dot in a magnetic field. *Phys. Rev. B*, 42(8):5166–5175, 1990. doi: 10.1103/PhysRevB.42.5166. URL http://link.aps.org/doi/10.1103/PhysRevB.42.5166.

Yigal Meir, Ned S. Wingreen, and Patrick A. Lee. Transport through a strongly interacting electron system: Theory of periodic conductance oscillations. *Physical Review Letters*, 66(23):3048–3051, June 1991. ISSN 0031-9007. doi: 10.1103/PhysRevLett.66.3048. URL http://link.aps.org/doi/10.1103/PhysRevLett.66.3048.

U Meirav, M A Kastner, and S J Wind. Single-electron charging and periodic conductance resonances in GaAs nanostructures. *Phys. Rev. Lett.*, 65(6):771–774, 1990. doi: 10.1103/PhysRevLett.65.771. URL http://link.aps.org/doi/10.1103/PhysRevLett.65.771.

M. Pedersen Lohne, G. Hagen, M. Hjorth-Jensen, S. Kvaal, and F. Pederiva. *Ab initio* computation of the energies of circular quantum dots. *Phys. Rev. B*, 84(11):115302, 2011. doi: 10.1103/PhysRevB.84.115302. URL http://link.aps.org/doi/10.1103/PhysRevB.84.115302.

M. A. Reed, J. N. Randall, R. J. Aggarwal, R. J. Matyi, T. M. Moore, and A. E. Wetsel. Observation of discrete electronic states in a zero-dimensional semiconductor nanostructure. *Phys. Rev. Lett.*, 60(6):535–537, February 1988. doi: 10.1103/PhysRevLett.60.535. URL http://link.aps.org/doi/10.1103/PhysRevLett.60.535.

Stephanie M. Reimann and Matti Manninen. Electronic structure of quantum dots. *Rev. Mod. Phys.*, 74(4):1283–1342, November 2002. ISSN 0034-6861. doi: 10.1103/RevModPhys.74.1283. URL http://link.aps.org/doi/10.1103/RevModPhys.74.1283.

S. Tarucha, D. G. Austing, T. Honda, R. J. van der Hage, and L. P Kouwenhoven. Shell Filling and Spin Effects in a Few Electron Quantum Dot. *Phys. Rev. Lett.*, 77(17):3613–3616, 1996. ISSN 0031-9007. doi: 10.1103/PhysRevLett.77.3613. URL http://link.aps.org/doi/10.1103/PhysRevLett.77.3613.

M. Taut. Two electrons in an external oscillator potential: Particular analytic solutions of a Coulomb correlation problem. *Phys. Rev. A*, 48(5):3561–3566, November 1993. ISSN 10502947. doi: 10.1103/PhysRevA.48.3561. URL http://link.aps.org/doi/10.1103/PhysRevA.48.3561.

Dietmar Weinmann. *Quantum transport in nanostructures*. PhD thesis, Universitat Hamburg, 1994. URL http://www.ipcms.unistra.fr/?page_id=13869.