

# Introduction to the diffusion Monte Carlo method

Ioan Kosztin, Byron Faber, and Klaus Schulten<sup>a)</sup>

Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801

(Received 16 February 1995; accepted 25 September 1995)

A self-contained and tutorial presentation of the diffusion Monte Carlo method for determining the ground state energy and wave function of quantum systems is provided. First, the theoretical basis of the method is derived and then a numerical algorithm is formulated. The algorithm is applied to determine the ground state of the harmonic oscillator, the Morse oscillator, the hydrogen atom, and the electronic ground state of the  $H_2^+$  ion and of the  $H_2$  molecule. A computer program on which the sample calculations are based is available upon request. © 1996 American Association of Physics Teachers.

## I. INTRODUCTION

The Schrödinger equation provides the accepted description for microscopic phenomena at nonrelativistic energies. Many molecular and solid-state systems are governed by this equation. Unfortunately, the Schrödinger equation can be solved analytically only in a few highly idealized cases; for most realistic systems, one needs to resort to numerical descriptions. In this paper, we want to introduce the reader to a relatively recent numerical method of solving the Schrödinger equation, the diffusion Monte Carlo (DMC) method. This method is suitable to describe the ground state of many quantum systems.

The solution of the time-dependent Schrödinger equation can be written as a linear superposition of stationary states in which the time dependence is given by a phase factor  $\exp(-iE_n t/\hbar)$ , where  $E_n$  is the  $n$ th energy level of the quantum system in question. The energy scale can be chosen such that all energies are positive. In the DMC method, one actually considers the solution of the Schrödinger equation assuming imaginary time  $\tau$ , i.e., after replacing the time  $t$  by  $-i\tau$ . The solution is then given by a sum of transients of the form  $\exp(-E_n \tau/\hbar)$ ,  $n=0,1,\dots$ . The DMC method is based upon the observation that, as a quantum system evolves in imaginary time, the longest lasting transient corresponds to the ground state with energy  $E_0 < E_n$ ,  $n=1,2,\dots$ . Following the evolution of the wave function in imaginary time long enough, one can determine both the ground state energy  $E_0$  and the ground state wave function  $\phi_0$  of a quantum system, regardless of the initial state in which the system had been prepared. The DMC method provides a practical way of evolving in imaginary time the wave function of a quantum system and obtaining, ultimately, the ground state energy and wave function.

The DMC method can be formulated in two different ways. The first one is based on the similarity between the imaginary time Schrödinger equation and a generalized diffusion equation. The kinetic (potential) energy term of the Schrödinger equation corresponds to the diffusion (*source/sink* or *reaction*) term in the generalized diffusion equation. The diffusion-reaction equation arising can be solved by employing *stochastic calculus* as was first suggested by Fermi around 1945.<sup>1,2</sup> Indeed, the imaginary time Schrödinger equation can be solved by simulating random walks of particles which are subject to birth/death processes imposed by the source/sink term. The probability distribution of the random walks is identical to the wave function. This is

possible only for wave functions which are positive everywhere, a feature, which limits the range of applicability of the DMC method. Such a formulation of the DMC method was given for the first time by Anderson<sup>3</sup> who used this method to calculate the ground state energy of small molecules such as  $H_3^+$ .

A second formulation of the DMC method arises from the Feynman path integral solution of the time-dependent Schrödinger equation. By means of path integrals the wave function can be expressed as a multidimensional integral which can be evaluated by employing the Monte Carlo method. Algorithms to solve the diffusion-reaction equation obeyed by the wave function and algorithms to evaluate the path integral representation of the wave function yield essentially one and the same formulation of the DMC method. Which one of the two formulations of the DMC method one adopts depends on one's expertise: A formulation of the DMC method based on the diffusion-reaction equation requires basic knowledge of the theory of stochastic processes; a path integral formulation obviously requires familiarity with the corresponding formulation of quantum mechanics.

The purpose of this article is to provide a self-contained and tutorial presentation of the path-integral formulation of the DMC method. We also present a numerical algorithm and a computer program based on the DMC method, and we apply this program to calculate the ground state energy and wave function for some sample quantum systems.

The article is organized as follows: The formulation of the DMC method is presented in Sec. II. In Sec. III, a numerical algorithm for the DMC method is constructed. The results of the DMC simulation for some simple quantum mechanical systems are presented in Sec. IV. Finally, Sec. V provides suggestions for further numerical experiments and guides the reader to the literature on the DMC method.

## II. THEORY

The theoretical formulation of the DMC method, presented below, follows three steps. These steps will be outlined first, to provide the reader with an overview.

*First Step: Imaginary time Schrödinger equation.* In this step, the solution of the time-dependent Schrödinger equation of a quantum system is expressed as a formal series expansion in terms of the eigenfunctions of the Hamiltonian. One then performs a transformation from real time  $t$  to imaginary time  $\tau$ , replacing  $t \rightarrow -i\tau$ . The solution of the obtained imaginary time Schrödinger equation becomes a series

of transients which decay exponentially as  $\tau \rightarrow \infty$ . The longest lasting transient corresponds to the ground state (i.e., the state with the lowest possible energy) of the system.

*Second Step: Path integral formulation and Monte Carlo integration.* In this step, the imaginary time Schrödinger equation is investigated by means of the path integral method. By using path integrals, the solution of this equation can be reduced to quadrature, provided that an initial state wave function is given. Standard Monte Carlo methods<sup>4</sup> permit one to evaluate numerically the path integral to any desired accuracy, assuming that the initial state wave function and, therefore, the ground state wave function as well, is positive definite. In this case, the wave function itself can be interpreted as a probability density and the “classical” Monte Carlo method can be applied. According to the general principles of quantum mechanics, only the square of the absolute value of the wave function has the meaning of a probability density; the fact that the ground state wave function has to be a positive definite real quantity imposes severe limitations on the applicability of the Monte Carlo technique for solving the Schrödinger equation. An efficient implementation of the standard Monte Carlo algorithm for calculating the wave function as a large multidimensional integral is realized through an alternation of diffusive displacements and of so-called birth–death processes applied to a set of imaginary particles, termed “replicas,” distributed in the configuration space of the system. The spatial distribution of these replicas converges to a probability density which represents the ground state wave function. The diffusive displacements and birth–death processes can be simulated on a computer using random number generators.

*Third Step: Continuous estimate of the ground state energy and sampling of the ground state wave function.* In this step, the ground state energy and the ground state wave function are actually determined. As mentioned above, the Monte Carlo method samples the wave function after each time step. The spatial coordinate distribution of the replicas involved in the combined diffusion and birth–death processes, after each (finite) time step, provides an approximation to the wave function of the system at that given time. The wave function converges in (imaginary) time towards the (time-independent) ground state wave function, if and only if the origin of the energy scale is equal to the ground state energy. Since the ground state energy is initially unknown, one starts with a reasonable guess and, after each time step in which a diffusive displacement and birth–death process is applied to all particles once, one improves the estimate of the ground state energy. Ultimately, this estimate converges towards the desired ground state energy and the distribution of particles converges to the ground state wave function.

In the following, we shall provide a detailed account of the above steps.

## A. Imaginary time Schrödinger equation

For simplicity, let us consider a single particle of mass  $m$  which moves along the  $x$  axis in a potential  $V(x)$ . Its wave function  $\Psi(x, t)$  is governed by the time-dependent Schrödinger equation<sup>5</sup>

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi, \quad (2.1)$$

where the Hamiltonian has the form

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x). \quad (2.2)$$

Assuming that the potential for  $x \rightarrow \pm\infty$  becomes infinite, i.e., the particle motion is confined to a finite spatial domain, the formal solution of (2.1) can be written as a series expansion in terms of the eigenfunctions of  $\hat{H}$ :

$$\Psi(x, t) = \sum_{n=0}^{\infty} c_n \phi_n(x) e^{-(i/\hbar)E_n t}. \quad (2.3)$$

The eigenfunctions  $\phi_n(x)$ , which are square integrable in the present case, and the eigenvalues  $E_n$  are obtained from the time-independent Schrödinger equation

$$\hat{H}\phi_n(x) = E_n \phi_n(x), \quad (2.4)$$

subject to the boundary conditions  $\lim_{x \rightarrow \pm\infty} \phi_n(x) = 0$ . We label the energy eigenstates by  $n=0, 1, 2, \dots$  and order the energies

$$E_0 < E_1 \leq E_2 \leq \dots \quad (2.5)$$

The eigenfunctions  $\phi_n(x)$  are assumed to be orthonormal and real; i.e.,

$$\int_{-\infty}^{\infty} dx \phi_n(x) \phi_m(x) = \delta_{nm}. \quad (2.6)$$

The expansion coefficients  $c_n$  in Eq. (2.3) are then

$$c_n = \int_{-\infty}^{\infty} dx \phi_n(x) \Psi(x, 0), \quad n=0, 1, 2, \dots, \quad (2.7)$$

i.e., they describe the overlap of the initial state  $\Psi(x, 0)$ , else assumed real, with the eigenfunctions  $\phi_n(x)$  in (2.4).

### 1. Shift of energy scale

We perform now a trivial, but methodologically crucial shift of the energy scale introducing the replacements  $V(x) \rightarrow V(x) - E_R$  and  $E_n \rightarrow E_n - E_R$ . This leads to the Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial \tau} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + [V(x) - E_R]\Psi, \quad (2.8)$$

and to the expansion

$$\Psi(x, t) = \sum_{n=0}^{\infty} c_n \phi_n(x) e^{-i(E_n - E_R)t/\hbar}. \quad (2.9)$$

### 2. Wick rotation of time

Now let us perform a transformation from real time to imaginary time (also known as Wick rotation) by introducing the new variable  $\tau = it$ . The Schrödinger equation (2.8) becomes

$$\hbar \frac{\partial \Psi}{\partial \tau} = \frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} - [V(x) - E_R]\Psi, \quad (2.10)$$

and the expansion equation (2.9) reads

$$\Psi(x, \tau) = \sum_{n=0}^{\infty} c_n \phi_n(x) e^{-(E_n - E_R)\tau/\hbar}. \quad (2.11)$$

Noting the energy ordering equation (2.5), one can infer from (2.11) the following asymptotic behavior for  $\tau \rightarrow \infty$ :

- (i) if  $E_R > E_0$ ,  $\lim_{\tau \rightarrow \infty} \Psi(x, \tau) = \infty$ , the wave function diverges exponentially fast;
- (ii) if  $E_R < E_0$ ,  $\lim_{\tau \rightarrow \infty} \Psi(x, \tau) = 0$ , the wave function vanishes exponentially fast;
- (iii) if  $E_R = E_0$ ,  $\lim_{\tau \rightarrow \infty} \Psi(x, \tau) = c_0 \phi_0(x)$ , the wave function converges, up to a constant factor  $c_0$  defined through Eq. (2.7), to the ground state wave function.

This behavior provides the basis of the DMC method. For  $E_R = E_0$ , the function  $\Psi(x, \tau)$  converges to the ground state wave function  $\phi_0(x)$  regardless of the choice of the initial wave function  $\Psi(x, 0)$ , as long as there is a numerically significant overlap between  $\Psi(x, 0)$  and  $\phi_0(x)$ , i.e., as long as  $c_0$  is not too small. The ground state wave function, for a single particle, has no nodes (in case of many fermion systems this might not be true) and one can always fulfill the requirement of nonvanishing  $c_0$  by choosing a positive definite initial wave function centered in a region of space where  $\phi_0(x)$  is sufficiently large.

We now seek a practical way to integrate equation (2.10) for an arbitrary reference energy  $E_R$  and initial wave function  $\Psi(x, 0)$ . We shall accomplish this by using the path integral formalism.

## B. Path integral formalism

The solution of the imaginary time Schrödinger equation (2.10) can be written

$$\Psi(x, \tau) = \int_{-\infty}^{\infty} dx_0 K(x, \tau | x_0, 0) \Psi(x_0, 0), \quad (2.12)$$

where the propagator  $K(x, \tau | x_0, 0)$  is expressed in terms of the well-known path integral,<sup>6</sup> modified by the replacement  $t = -i\tau$

$$K(x, \tau | x_0, 0) = \lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_{N-1} \left( \frac{m}{2\pi\hbar\Delta\tau} \right)^{N/2} \times \exp \left\{ -\frac{\Delta\tau}{\hbar} \sum_{j=1}^N \left[ \frac{m}{2\Delta\tau^2} (x_j - x_{j-1})^2 + V(x_j) - E_R \right] \right\}. \quad (2.13)$$

Here,  $\Delta\tau = \tau/N$  is a small time step. One sets  $x_N \equiv x$ . The wave function  $\Psi(x, \tau)$  can be written in the form

$$\Psi(x, \tau) = \lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} \left( \prod_{j=0}^{N-1} dx_j \right) \prod_{n=1}^N W(x_n) \times P(x_n, x_{n-1}) \Psi(x_0, 0), \quad (2.14)$$

where we have defined

$$P(x_n, x_{n-1}) \equiv \left( \frac{m}{2\pi\hbar\Delta\tau} \right)^{1/2} \exp \left[ -\frac{m(x_n - x_{n-1})^2}{2\hbar\Delta\tau} \right], \quad (2.15)$$

and

$$W(x_n) \equiv \exp \left[ -\frac{[V(x_n) - E_R]\Delta\tau}{\hbar} \right]. \quad (2.16)$$

The function  $P(x_n, x_{n-1})$  is related to the kinetic energy term in (2.2). This function can be thought of as a Gaussian probability density for the random variable  $x_n$  with mean equal to  $x_{n-1}$  and variance

$$\sigma = \sqrt{\hbar\Delta\tau/m}. \quad (2.17)$$

The so-called *weight function*  $W(x_n)$  depends on both the potential energy in (2.2) and the reference energy  $E_R$ . The main difference between the functions  $P$  and  $W$  is that the former can be interpreted as a probability density since

$$\int_{-\infty}^{\infty} dy P(x, y) = 1, \quad (2.18)$$

while the latter cannot.

The path integral equation (2.14) can be evaluated analytically only for particular forms of the potential energy  $V(x)$ .<sup>7</sup> Fortunately, by choosing  $N$  sufficiently large, one can evaluate (2.14) numerically to any desired accuracy. However, since a suitable  $N$  is necessarily a large number, the standard algorithms of numerical integration cannot be employed directly;<sup>8</sup> instead, one uses the so-called Monte Carlo method.<sup>4</sup> According to this method, any (convergent)  $N$ -dimensional integral of the form

$$I = \int_{-\infty}^{\infty} \left( \prod_{j=0}^{N-1} dx_j \right) f(x_0, \dots, x_{N-1}) \mathcal{P}(x_0, \dots, x_{N-1}), \quad (2.19)$$

where  $\mathcal{P}$  is a probability density, i.e.,

$$\mathcal{P}(x_0, \dots, x_{N-1}) \geq 0$$

and

$$\int_{-\infty}^{\infty} \left( \prod_{j=0}^{N-1} dx_j \right) \mathcal{P}(x_0, \dots, x_{N-1}) = 1, \quad (2.20)$$

can be approximated by the expression

$$\mathcal{I} = \frac{1}{\mathcal{N}} \sum_{\substack{i=1 \\ x^{(i)} \in \mathcal{P}}}^{\mathcal{N}} f(x_0^{(i)}, \dots, x_{N-1}^{(i)}). \quad (2.21)$$

Here, the notation  $x^{(i)} \in \mathcal{P}$  means that the numbers  $x^{(i)}$ ,  $i=1, 2, \dots, \mathcal{N}$ ;  $j=0, \dots, N-1$ , are selected randomly with the probability density  $\mathcal{P}$ . The larger  $\mathcal{N}$ , the better is the approximation  $I \approx \mathcal{I}$ . In fact, according to the *central limit theorem*,<sup>4</sup> the values of  $\mathcal{I}$  obtained as a result of different simulations are distributed normally, i.e., according to a Gaussian distribution, around the exact value  $I$ , the standard deviation being proportional to  $1/\sqrt{\mathcal{N}}$ .

In order to evaluate  $\Psi(x, \tau)$  in Eq. (2.14), for given  $x$ ,  $\tau$ , and  $N$ , one defines

$$\mathcal{A}(x_0, \dots, x_{N-1}) = \prod_{n=1}^N P(x_n, x_{n-1}), \quad (2.22)$$

and

$$f(x_0, \dots, x_{N-1}) = \Psi(x_0, 0) \prod_{n=1}^N W(x_n), \quad (2.23)$$

such that one can apply (2.19) and (2.21). For this purpose, we note that due to

$$\int_{-\infty}^{\infty} dy P(x, y) P(y, z) = P(x, z), \quad (2.24)$$

and (2.18), the probability distribution equation (2.22) does indeed obey the property equation (2.20). The expression (2.21) can now be invoked to evaluate  $\Psi(x, \tau)$  by means of the path integral (2.14). This requires the generation of sets of coordinate vectors  $x^{(i)} \in \mathcal{P}$ ,  $x^{(i)} = (x_0^{(i)}, x_1^{(i)}, \dots, x_N^{(i)})$  for  $i=1, 2, \dots, \mathcal{N}$ , where  $x_N^{(i)} = x$ .

In order to obtain vectors  $x^{(i)}$  which sample the probability density  $\mathcal{P}$ , one proceeds as follows. In a first step one generates, for a fixed  $x = x_N^{(i)}$ , a Gaussian random number  $x_{N-1}^{(i)}$  with mean value  $x_N^{(i)}$  (i.e., a Gaussian random number distributed about  $x_N^{(i)}$ ) and variance  $\sigma$ , according to the probability density  $P(x_N^{(i)}, x_{N-1}^{(i)})$  given by (2.15). In a second step, a Gaussian random number  $x_{N-2}^{(i)}$ , with mean  $x_{N-1}^{(i)}$  and variance  $\sigma$ , is generated according to  $P(x_{N-1}^{(i)}, x_{N-2}^{(i)})$ . The steps are continued to produce random numbers  $x_n^{(i)}$  until one reaches  $x_0^{(i)}$ . Two consecutive random numbers  $x_n^{(i)}$  and  $x_{n-1}^{(i)}$  are related through the equation

$$x_n^{(i)} = x_{n-1}^{(i)} + \sigma \rho_n^{(i)}, \quad (2.25)$$

where  $\sigma$  is given by (2.17) and  $\rho_n^{(i)}$  is a Gaussian random number with zero mean and a variance equal to 1. The  $\rho_n^{(i)}$ 's can be generated numerically by means of algorithms referred to as random number generators. One can check, using (2.25) and (2.15), that the mean and the variance of  $x_n^{(i)}(x_{n-1}^{(i)})$  are equal to  $x_{n-1}^{(i)}(x_n^{(i)})$  and  $\sigma$ , respectively. Therefore, the coordinate vectors  $\{x_{N-1}^{(i)}, \dots, x_0^{(i)}\}$ , obtained through Eq. (2.25) for  $i=1, 2, \dots, \mathcal{N}$ , are distributed according to the probability density equation (2.22). We note in passing that the sequence of positions  $x_n$  given by (2.25) defines a stochastic process, namely, the well known Brownian diffusion process.

Repeating  $\mathcal{N}$  times the sampling of  $\mathcal{P}$ , the wave function  $\Psi(x, \tau)$  can be determined according to (2.21) and (2.14). Unfortunately, the algorithm outlined is impractical since it provides only a route to calculate  $\Psi(x, \tau)$  for a chosen time  $\tau$ , but no systematic method to obtain the ground state energy and wave function, which requires a description for  $\tau \rightarrow \infty$ .

Fortunately, the above so-called Monte Carlo algorithm can be improved and used to determine *simultaneously* both  $E_0$  and  $\phi_0$ . The basic idea is to consider the wave function itself a probability density. This implies that the wave function should be a positive definite function, a constraint which limits the applicability of the suggested method. By sampling the initial wave function  $\Psi(x, 0)$  at  $\mathcal{N}_0$  points, one generates as many Gaussian random walks which evolve in time according to Eq. (2.22) or, equivalently, according to (2.25); instead of tracing the motion of each random walk separately, one rather follows the motion of a whole ensemble of random walks simultaneously. The advantage of this procedure is that one can sample the wave function of the system, through the actual position of the random walks and the products of the weights  $W$  along the corresponding trajectories, after each time step  $\Delta\tau$ . This procedure, as we explain below, also provides the possibility to readjust the value of  $E_R$  after each time step and to follow the time evolution of the system for as many time steps as are needed to converge to the ground state wave function and energy.

The procedure, the so-called DMC method, interprets the integrand in (2.14), i.e.,

$$\underbrace{W(x_N)}_{\text{process } 2N} \underbrace{P(x_N, x_{N-1}) \cdots P(x_2, x_1)}_{\text{process } 2N-1} \underbrace{W(x_2)}_{\text{process } 4} \underbrace{P(x_2, x_1)}_{\text{process } 3} \underbrace{W(x_1)}_{\text{process } 2} \underbrace{P(x_1, x_0)}_{\text{process } 1} \underbrace{\Psi(x_0, 0)}_{\text{process } 0}, \quad (2.26)$$

as a product of probabilities and weights to be modeled by a series of sequential stochastic processes 0, 1, 2, ..., 2N. We will explain now how these processes are described numerically.

### 1. Initial state

The 0th process describes "particles" (random walks) distributed according to the initial wave function  $\Psi(x_0, 0)$ , which is typically chosen as a Dirac  $\delta$  function

$$\Psi(x_0, 0) = \delta(x - x_0), \quad (2.27)$$

where  $x_0$  is located in an area where the ground state of the quantum system is expected to be large. The initial distribution equation (2.27) is obtained by simply placing all particles initially at position  $x_0$ .

### 2. Diffusive displacement

As we explained earlier, the successive positions  $x_{n-1}, x_n$  in (2.26) can be generated through (2.25). The Monte Carlo algorithm produces then the positions  $x_1 = x_0 + \sigma\rho_1$ ,  $x_2 = x_1 + \sigma\rho_2$ , etc. by generating the series of random numbers  $\rho_n$ ,  $n=1, 2, \dots$ .

### 3. Birth-death process

Instead of accumulating the product of the weight factors  $W$  for each particle, it is more efficient numerically to replicate (see below) the particles after each time step with a probability proportional to  $W(x_n)$ . In this way, after each time step  $\Delta\tau$ , the (unnormalized) wave function is given by a histogram of the spatial distribution of the particles. The calculation of the wave function  $\Psi(x, \tau)$  can be regarded as a simulated diffusion-reaction process of some imaginary particles.

In the replication process each particle is replaced by a number of

$$m_n = \min[\text{int}[W(x_n) + u], 3] \quad (2.28)$$

particles, where  $\text{int}[x]$  denotes the integer part of  $x$  and where  $u$  represents a random number uniformly distributed in the interval  $[0, 1]$ . In case  $m_n=0$  the particle is deleted and one stops the diffusion process; this is referred to as a "death" of a particle. In case of  $m_n=1$ , the particle is unaffected and one continues with the next diffusion step. In case  $m_n=2, 3$ , one continues with the next diffusion step, but begins also a new series (in case of  $m_n=3$ , two new series) of diffusive displacements starting at the present location  $x_n$ .

This latter case is referred to as the “birth” of a particle (of two particles for  $m_n=3$ ). From (2.28), one can see that at most two new particles can be generated whereas one would expect that for  $\text{int}[W(x_n)+u] \geq 4$  three or more new series would be started. This limitation on the birth rate of the particles is necessary in order to avoid numerical instabilities, especially at the beginning of the Monte Carlo simulation, when  $E_R$  may differ significantly from  $E_0$ . The error resulting from the limitation  $m_n \leq 3$  is expected to be small since for sufficiently small  $\Delta\tau$  holds

$$W(x) \approx 1 - \frac{V(x) - E_R}{\hbar} \Delta\tau, \quad (2.29)$$

which, evidently, assumes values around unity.

### C. The diffusion Monte Carlo method

We want to summarize now the algorithmic steps actually taken in a straightforward application of the DMC method. To realize the suggested algorithm one starts with  $\mathcal{N}_0$  “particles,” referred to as “replicas,” which are placed according to a distribution  $\Psi(x_0, 0)$ . As pointed out above, the actual numbers of replicas will vary as replicas “die” and new ones are “born.” The replicas are characterized through a position  $x_n^{(j)}$  where the suffix  $n$  counts the diffusive displacements and where  $(j)$  counts the replicas. The number of replicas after  $n$  diffusive displacements will be denoted by  $\mathcal{N}_n$ . In the initial step, one samples  $\mathcal{N}_0$  replicas assigning positions  $x_0^{(i)}$ , ( $i=1, \dots, \mathcal{N}_0$ ) according to the distribution  $\Psi(x_0, 0)$ . As mentioned, we actually choose all replicas to begin at the same point  $x_0$ .

Rather than following the fate of each replica and of its descendants through many diffusive displacements, one follows all replicas simultaneously. Accordingly, one determines the positions  $x_1^{(j)}$  following Eq. (2.25); i.e., one sets

$$x_1^{(j)} = x_0^{(j)} + \sqrt{\hbar \Delta\tau / m} \cdot \rho_1^{(j)} \quad (2.30)$$

for replicas  $j=1, 2, \dots, \mathcal{N}_0$  generating appropriate random numbers  $\rho_1^{(j)}$ . This can be regarded as a one-step diffusion process of the system of replicas. Once the new positions equation (2.30) have been determined, one evaluates the weight  $W(x_1^{(i)})$  given by (2.16) and, according to (2.28), one determines the set of integers  $m_1^{(j)}$  for  $j=1, 2, \dots, \mathcal{N}_0$ . Replicas  $j$  with  $m_1^{(j)}=0$  are terminated, replicas  $m_1^{(j)}=1, 2, 3$  are left unaffected, except, that in case  $m_1^{(j)}=2, 3$  one or two new replicas  $j'$  are added to the system and their position is set to  $x_1^{(j')} = x_1^{(j)}$ . The number of replicas is counted and, thus,  $\mathcal{N}_1$  is determined. During the combined diffusion and birth-death process the distribution of replicas changes in such a way that the corresponding coordinates  $x_1^{(j)}$  will now be distributed according to a probability density which will be identified with  $\Psi(x, \Delta\tau)$ .

#### 1. Adaptation of the reference energy $E_R$

As a result of the birth-death process the total number  $\mathcal{N}_1$  of replicas changes from its original  $\mathcal{N}_0$  value. As discussed above, for  $E_R$  values less than the ground state energy the distribution  $\Psi(x, \tau)$  decays asymptotically to zero, i.e., the replicas eventually all die; for  $E_R$  values larger than the ground state energy, the distribution will increase indefinitely, i.e., the number  $\mathcal{N}_n$  will exceed all bounds. Only for  $E_R = E_0$  can one expect a stable asymptotic distribution such that the number of replicas fluctuates around an average value  $\mathcal{N}_0$ . The spatial distribution and increases/decreases

of the number of replicas allow one to adjust the value of  $E_R$  as to keep the total number of replicas approximately constant. To this end, one proceeds from (2.29). Averaged over all replicas this equation reads

$$\langle W \rangle_1 \approx 1 - \frac{\langle V \rangle_1 - E_R}{\hbar} \Delta\tau, \quad (2.31)$$

where the average potential energy is

$$\langle V \rangle_1 = \frac{1}{\mathcal{N}_1} \sum_{j=1}^{\mathcal{N}_1} V(x_1^{(j)}). \quad (2.32)$$

The reader should note that this average varies with the number of diffusion steps taken for all replicas; in the present case, we consider the average after the first diffusive displacement. One would like the value of  $\langle W \rangle$  to be eventually always unity such that the number of replicas remains constant. This leads to the proper choice of  $E_R$

$$E_R^{(1)} = \langle V \rangle_1. \quad (2.33)$$

However, if the distribution of replicas deviates too strongly from  $\phi_0(x)$ , the number of replicas can experience strong changes which need to be repaired by “overcompensating” through a suitable choice of  $E_R$ . In fact, in case  $\mathcal{N}_1/\mathcal{N}_0 < 1$ , one would like to increase subsequently the number of replicas in order to restore the initial number of replicas, and, hence, choose an  $E_R$  value larger than Eq. (2.33); in case  $\mathcal{N}_1/\mathcal{N}_0 > 1$ , one would like to decrease subsequently the number of replicas and, hence, choose an  $E_R$  value smaller than (2.33). A suitable choice is

$$E_R^{(2)} = \langle V \rangle_1 + \frac{\hbar}{\Delta\tau} \left( 1 - \frac{\mathcal{N}_1}{\mathcal{N}_0} \right). \quad (2.34)$$

In order to justify that (2.34) serves our purpose one notes that  $\exp[-\Delta\tau(\langle V \rangle - E_R)/\hbar]$ , for sufficiently small  $\Delta\tau$ , is  $1 - \Delta\tau(\langle V \rangle - E_R)/\hbar$ . In the subsequent birth-death process, this would lead to an  $\mathcal{N}_1$  value (recall that during the diffusion process the number of replicas does not change) related to  $E_R$  by

$$E_R = \langle V \rangle_1 - \frac{\hbar}{\Delta\tau} \left( 1 - \frac{\mathcal{N}_1}{\mathcal{N}_0} \right). \quad (2.35)$$

However, one would like the total number of replicas during the next birth-death process to return to the initial value  $\mathcal{N}_0$  and  $E_R$  to be given by an expression similar to (2.33). This can be achieved by adding to both sides of Eq. (2.35) the same quantity  $\hbar(1 - \mathcal{N}_1/\mathcal{N}_0)/\Delta\tau$ . Hence, the redefined value for the reference energy is

$$E_R^{(2)} = E_R^{(1)} + \frac{\hbar}{\Delta\tau} \left( 1 - \frac{\mathcal{N}_1}{\mathcal{N}_0} \right), \quad (2.36)$$

which, by taking (2.33) into account, is identical to (2.34).

Equation (2.34) should be regarded as an empirical result rather than an exact one. In fact, any expression of the form  $E_R = \langle V \rangle + \alpha(1 - \mathcal{N}_1/\mathcal{N}_0)$ , with arbitrary positive  $\alpha$ , can be used equally well in place of (2.34).<sup>9</sup> Usually, the actual value of the “feedback” parameter  $\alpha$  is chosen empirically for each individual problem so as to reduce as much as possible the statistical fluctuations in  $\mathcal{N}_0$  and, at the same time, to diminish unwanted correlations between the successive generation of replicas. Equation (2.34) suggests that a good starting value for  $\alpha$  is  $\hbar/\Delta\tau$ , a value used in our DMC program.

## 2. 2nd, 3rd,... displacements

The diffusive displacements, birth–death processes and estimates for new  $E_R$  values are repeated until, after a sufficiently large number of steps, the energy  $E_R$  and the distribution of replicas becomes stationary. The ground state energy is then

$$E_0 = \lim_{n \rightarrow \infty} \langle V \rangle_n \quad (2.37)$$

since  $\mathcal{N}_n/\mathcal{N}_{n-1} \rightarrow 1$ . The distribution of replicas provides the ground state function  $\phi_0(x)$ .

## D. Systems with several degrees of freedom

The DMC method described above is valid for a quantum system with one degree of freedom  $x$ . However, the method can be easily extended to quantum systems with several degrees of freedom. Such systems arise, for example, in case of a particle moving in two or three spatial dimensions or in case of several interacting particles. For these two cases the DMC method can be readily generalized.

### 1. Particle in $d$ dimensions

The Hamiltonian of a particle of mass  $m$  moving in a potential  $V(x_1, \dots, x_d)$  can be written

$$\hat{H} = -\frac{\hbar^2}{2m} \sum_{\alpha=1}^d \frac{\partial^2}{\partial x_\alpha^2} + V(x_1, \dots, x_d), \quad (2.38)$$

where  $x_\alpha$ ,  $\alpha=1, \dots, d$ , denotes the Cartesian coordinates of the particle and  $d(=2,3)$  represents the spatial dimension. For the Hamiltonian equation (2.38), the DMC algorithm can be devised in a similar way as for (2.2). The only difference to the one degree of freedom case is that during each time step  $\Delta\tau$  one needs to execute  $d$  random walks for each replica. Indeed, the kinetic energy term in (2.38) can be formally regarded as the sum of kinetic energies of  $d$  ‘‘particles’’ having the same mass  $m$  and moving along the  $x_\alpha$ ,  $\alpha=1, \dots, d$ , directions. Consequently, the diffusive displacements are governed by the distribution

$$P(x_{n,1}, x_{n-1,1}; \dots; x_{n,d}, x_{n-1,d}) = \prod_{\alpha=1}^d \left( \frac{m}{2\pi\hbar\Delta\tau} \right)^{1/2} \exp \left[ -\frac{m(x_{n,\alpha} - x_{n-1,\alpha})^2}{2\hbar\Delta\tau} \right]. \quad (2.39)$$

The product of probabilities can be described through independent random processes, i.e., one can reproduce the probability equation (2.39) through  $d$  independent diffusive displacements applied for each replica to the  $d$  spatial directions.

### 2. $S$ particles

In the case of  $S$  interacting particles, which move in  $d$  spatial dimensions, the most general form of the Hamiltonian is given by (assuming that no internal, e.g., spin, degrees of freedom are involved)

$$\hat{H} = \sum_{j=1}^S \left[ -\frac{\hbar^2}{2m_j} \sum_{\alpha=1}^d \frac{\partial^2}{\partial x_{j\alpha}^2} + V(\{x_{j\alpha}\}) \right], \quad (2.40)$$

where  $V(\{x_{j\alpha}\})$  accounts for both an interaction between particles and for an interaction due to an external field;  $\{x_{j\alpha}\}$  denotes the dependence on the coordinates of all particles. By rescaling the coordinates in (2.40),

$$x_{j\alpha} = \sqrt{\frac{m}{m_j}} x'_{j\alpha}, \quad j=1, \dots, S, \quad \alpha=1, \dots, d \quad (2.41)$$

where  $m$  is an arbitrary mass, one can make the Hamiltonian equation (2.40) look formally the same as the Hamiltonian equation (2.38) describing a single particle of mass  $m$  which moves in  $d'=S \times d$  spatial dimensions. Hence, the generalization of the DMC algorithm for this case is again apparent.

### 3. Sign problem

The case in which (2.40) actually describes a system of *identical* particles cannot be treated like that of a single particle moving in  $d'=S \times d$  spatial dimensions. For such systems, a prescribed *boson* or *fermion* symmetry of the wave function with respect to an interchange of particles must be obeyed. For bosons (particles with integer spin), the total wave function (i.e., the product of the orbital and the spin wave functions) is symmetric with respect to any permutation of the particles while for fermions (particles with half-integer spin) the total wave function is antisymmetric with respect to such permutation. This constraint determines the symmetry of the orbital part of the ground state wave function for fermionic systems (but not for bosonic ones): in many cases of interest, the (orbital) ground state wave function of fermionic systems will have nodes, i.e., regions with different signs, which make the DMC method, as presented here, inapplicable. We shall not address this issue in further detail and shall consider only cases where the so-called ‘‘sign problem’’ of the ground state wave function does not arise.

To summarize, a DMC algorithm for a system with one degree of freedom can be adopted to a system of  $S$  interacting particles moving in  $d$  spatial dimensions with an effective dimension of  $d'=S \times d$ . Exactly this feature of the DMC method makes it so attractive for the evaluation of the ground state of a quantum system. However, in the case of identical fermions, one needs to obey the actual symmetry of the ground state wave function and the method often is not applicable.

## III. ALGORITHM

Our goal is to provide an algorithm for the DMC method presented in the previous section and to apply this algorithm to obtain the ground state energy and wave function for sample quantum systems. Some of the examples chosen below, e.g., the harmonic oscillator, have an analytical solution and, therefore, allow one to test the diffusion Monte Carlo method. Other examples, e.g., the hydrogen molecule, can not be solved analytically and, hence, the DMC method provides a convenient way of solving the problem. The obtained results turn out to be in good agreement with results obtained by means of other numerical methods.<sup>10</sup>

### A. Dimensionless units

In order to implement the DMC method into a numerical algorithm, one needs to rewrite all the relevant equations in dimensionless units. One can go from conventional (e.g., SI) units to dimensionless units by explicitly writing each physical quantity as its magnitude times the corresponding unit. In mechanics the unit of any physical quantity can be expressed as a proper combination of three independent units, such as,  $L$ ,  $T$ , and  $\mathcal{E}$ , which denote the unit of length, time and energy, respectively. By denoting the value of a given physical

quantity with the same symbol as the quantity itself (e.g.,  $xL \rightarrow x$  in the case of coordinate  $x$ ), the Schrödinger equation (2.10) can be recast

$$\frac{\partial \Psi}{\partial \tau} = \frac{\hbar T}{2mL^2} \frac{\partial^2 \Psi}{\partial x^2} - \frac{T\mathcal{E}}{\hbar} [V(x) - E_R] \Psi. \quad (3.1)$$

It is convenient to choose  $L$ ,  $T$ , and  $\mathcal{E}$  such that

$$\frac{\hbar T}{2mL^2} = \frac{1}{2}, \quad \text{and} \quad \frac{T\mathcal{E}}{\hbar} = 1 \quad (3.2)$$

holds. Since there are three unknown units and only two relationships between them, one has the freedom to specify the actual value of either  $L$ ,  $T$ ,  $\mathcal{E}$  while the value of the other two units follows from (3.2).

In dimensionless units obeying (3.2), the original imaginary-time Schrödinger equation reads

$$\frac{\partial \Psi}{\partial \tau} = \frac{1}{2} \frac{\partial^2 \Psi}{\partial x^2} - [V(x) - E_R] \Psi. \quad (3.3)$$

It is not difficult to transcribe also the other relevant equations above in dimensionless units; for example, the functions (2.15) and (2.16) become

$$P(x_n, x_{n-1}) = \sqrt{\frac{1}{2\pi\Delta\tau}} \exp\left[-\frac{(x_n - x_{n-1})^2}{2\Delta\tau}\right]$$

and

$$W(x_n) = \exp\{-[V(x_n) - E_R]\},$$

respectively. As a consequence, the diffusive displacements (2.25) are described by

$$x_n = x_{n-1} + \sqrt{\Delta\tau} \rho_n \quad (3.4)$$

## B. Computer program

The flow diagram of the computer program<sup>11</sup> implementing the DMC method is shown in Fig. 1. Each block in the diagram performs specific tasks which are explained now.

All external data required for the calculation are collected through a menu driven, interactive interface in the `input` block. First, one has to select the quantum system on which the calculation is performed. At the program level, this means to define the right spatial dimensionality  $d$  and the potential energy  $V$  (see Sec. IV) which corresponds to the selected system. The quantum systems covered by our program are the ground state of the harmonic oscillator, of the Morse oscillator, of the hydrogen atom, of the  $H_2^+$  ion (electronic state) and of the hydrogen molecule (electronic state). The results of the simulation for all these cases are presented in Sec. IV below. The other input parameters are: the initial number of replicas ( $\mathcal{N}_0$ ), the maximum number of replicas ( $\mathcal{N}_{\max}$ ), the `seed` value for the random number generators, the number of time steps to run the simulation ( $\tau_0$ ), the value of the time step ( $\Delta\tau$ ), the limits of the coordinates for the spatial sampling of the replicas ( $x_{\min}$ ,  $x_{\max}$ ) and, finally, the number of spatial “boxes” ( $n_b$ ) for sorting the replicas during their sampling. Suggested values for these parameters are (in dimensionless units):  $\mathcal{N}_0=500$ ,  $\mathcal{N}_{\max}=2000$ ,  $\tau_0=1000$ ,  $\Delta\tau=0.1$ ,  $x_{\min}=-20$ ,  $x_{\max}=20$  and  $n_b=200$ .

In the next block, `initialize replicas`, a two-dimensional matrix, called `psips`, is defined with the following structure: The first (row) index identifies the replicas and takes integer values between one and  $\mathcal{N}_{\max}$ . The second

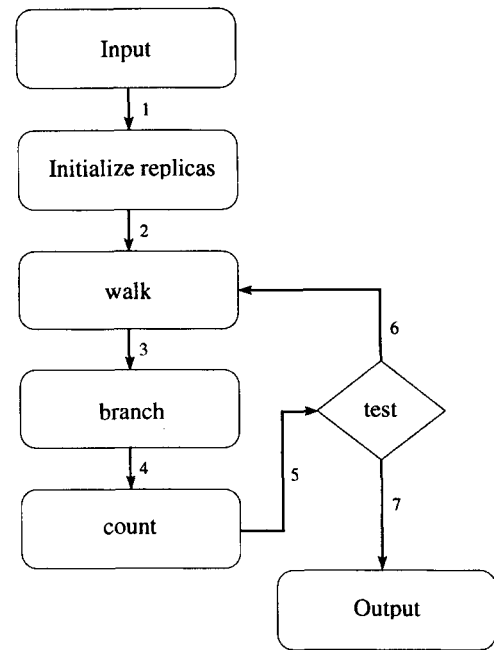


Fig. 1. Flow diagram of the DMC algorithm.

(column) index points to information regarding the replica identified by the first index, and is a positive integer less or equal to the number of degrees of freedom of the system (i.e.,  $d'$ , see Sec. II D) plus one. For a given replica, say  $i$ , `psips[i][1]` is used as an *existence flag*: it is zero (one) if the replica is *dead* (*alive*). The other elements `psips[i][j]`, ( $j=2, \dots, d'+1$ ) are used to store the coordinates of the replica (i.e.,  $x_{j,\alpha}$  according to the notations of Sec. II D) during the simulation.

To initialize the matrix `psips` one sets equal to one the value of the existence flag `psips[j][1]` for  $j=1, \dots, \mathcal{N}_0$  (replicas  $j \geq \mathcal{N}_0$  are not born yet and, accordingly, their existence flag is set to zero), and then one assigns the same coordinates for all these replicas (for the actual values of these coordinates, see Sec. IV). Such a choice for the initial distribution of the replicas corresponds to a  $\delta$  function for  $\Psi(x,0)$ . For a suitable choice of  $x_0$ , one can be certain that there is always a significant overlap of  $\Psi(x,0)$  and the ground state wave function  $\phi_0(x)$ . The initial value of the reference energy  $E_R$  is simply given by the value of the potential energy at the initial position of the replicas [c.f. Eq. (2.33)].

After initializing the replicas the program enters into a loop which essentially consists of three routines: `walk`, `branch` and `count`. One loop corresponds to taking one time step  $\Delta\tau$ . `walk`: This routine performs the diffusion process of the replicas by adding to the coordinates of the active (alive) replicas the value  $\sqrt{\Delta\tau} \rho$ , where  $\rho$  is a random number drawn from a Gaussian distribution with mean zero and standard deviation one [c.f. Eq. (3.4)]. The program uses for this purpose the Gaussian random number generator `gasdev`.<sup>4</sup> `branch`: The birth-death (branching) processes, which follow the diffusion steps of the replicas, are performed by the routine `branch`. For each alive replica, the number  $m_n$ , given by Eq. (2.28), is calculated. For the generation of  $u$  in Eq. (2.28), a uniformly distributed random number in the interval  $[0,1]$ , the program employs the function `rans3`.<sup>4</sup> If a value  $m_n=0$  results, the replica is killed by

setting the corresponding existence flag to zero. If a value  $m_n=1$  results, the replica is left as it is. If a value  $n=2$  results, the replica is duplicated (the first inactive replica in `psips` is set active with the same coordinates as the original replica). If a value  $n=3$  results, two identical copies of the replica are generated in `psips`. The reader may note that never more than two copies are born; this limitation is necessary to prevent the uncontrolled growth of the number of replicas which `psips`, due to its finite size, might not be able to accommodate. Such growth might occur when all the replicas are located in (almost) the same place and when one can expect that large values  $m_n$  can arise. Since (by choice) the replicas initially are located in one and the same point of the configuration space the first diffusion process does not spread the replicas far enough and, for certain initial positions,  $m_n$  could then become large for all replicas. The algorithm also terminates if all the  $m_n$ 's assume values zero. To avoid this possibility, one needs to choose the initial location of the replicas with care. In general, any point where the ground state wave function is large is a good choice.

**count:** The role of this routine is to return the ground state wave function of the system (i.e., the spatial distribution of the replicas) at the end of the simulation. To this end, the spatial interval ( $x_{\min}$ – $x_{\max}$ ) is divided equally into  $n_b$  “boxes” (subintervals) for each degree of freedom and then, by employing standard numerical methods,<sup>8</sup> one counts the distribution of replicas among these “boxes.” The counting process starts after  $\tau_0$  (in units of  $\Delta\tau$ ) time steps when the system has already reached its stationary state (identified through a converged  $\langle V \rangle_\tau$ ) and is performed in a cumulative way for another  $\tau_0$  time steps. This strategy can be justified as follows: once stationarity is reached, the wave function  $\Psi(x, \tau) \propto \phi_0(x)$  will practically not change in time; hence, the replicas will essentially sample one and the same wave function at any subsequent time and the cumulative counting of the replicas in the “boxes” can be used; this procedure yields better statistics for  $\phi_0(x)$  than sampling of replicas at only one instant in time; by cumulative counting, the effective number of replicas used to sample the wave function is enhanced by a factor of  $\tau_0$  (number of time steps the counting is done).

Once the spatial distribution of replicas is known, one can normalize the distribution to obtain the ground state wave function using

$$\phi_0(x_i) \approx \frac{N_i}{\sqrt{\sum_{i=1}^{n_b} N_i^2}}, \quad i=1, \dots, n_b. \quad (3.5)$$

As should be apparent by now, until the completion of the algorithm the routines `walk` and `branch` are called  $2\tau_0$  times while the `count` routine is called only  $\tau_0$  times, namely, during the second half of the calculation. This is controlled by the block `test` shown in Fig. 1.

Finally, the `output` block returns the results of the simulation. These results are (i) the average value of the reference energy  $\langle E_R \rangle \approx E_0$ , which is calculated during the second part of the simulation (i.e., from  $\tau_0$  to  $2\tau_0$ ) when the system is already stabilized; (ii) the corresponding standard deviation  $\delta E_R$ ; (iii) the (imaginary) time evolution of  $\langle E_R \rangle$  for the first  $\tau_0$  time steps (used basically to check how fast stationarity is reached by the system during the simulation); (iv) the normalized spatial distribution of the replicas, i.e., the ground state wave function. Note that the average reference energy after  $n$  time steps is defined through

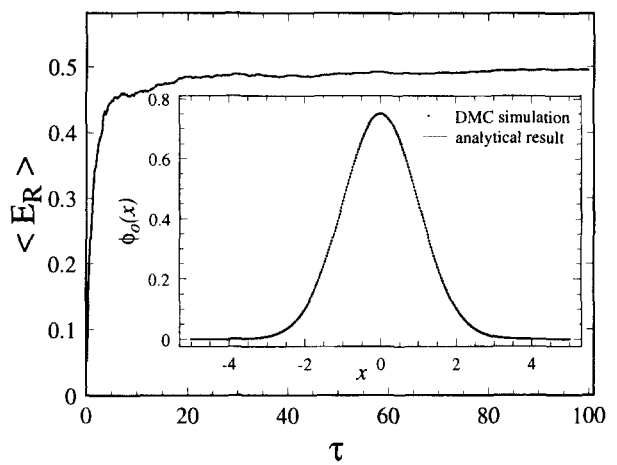


Fig. 2. Reference energy  $E_R$  (in dimensionless units) as a function of the imaginary time  $\tau$  (in units of time steps  $\Delta\tau$ ) obtained by the DMC method for the harmonic oscillator.  $E_R$  converges rapidly towards the exact ground state energy  $E_0=0.5$ . The inset shows the corresponding ground state wave function  $\phi_0(x)$ . The result of the simulation is represented by dots while the continuous line corresponds to the analytical solution equation (4.2)

$$\langle E_R(\tau = n\Delta\tau) \rangle = \frac{1}{n} \sum_{i=1}^n E_R(i\Delta\tau). \quad (3.6)$$

## IV. EXAMPLES

In this section, we report on the results obtained, by means of the DMC program, for the ground state energy and wave function of some quantum mechanical systems. The program was executed on an HP-9000 (series 700) workstation. In each case, we specify the units used and present numerical results of the simulation. For all simulations the values of the input parameters suggested in Sec. III B, have been employed.

### A. Harmonic oscillator

For a one-dimensional harmonic oscillator, characterized by the proper angular frequency  $\omega$ , one has (in SI units)

$$V(x) = \frac{1}{2}m\omega^2x^2. \quad (4.1)$$

Choosing  $T=1/\omega$ , one obtains from Eq. (3.2)  $L = \sqrt{\hbar/m\omega}$  and  $\mathcal{E} = \hbar\omega$ . In corresponding dimensionless units, the exact ground state energy is  $E_0=1/2$  and the ground state wave function is<sup>5</sup>

$$\phi_0(x) = \pi^{-1/4} \exp\left(-\frac{x^2}{2}\right). \quad (4.2)$$

The results of the Monte Carlo simulation for the harmonic oscillator are given in Fig. 2. At the beginning of the simulation all replicas are located at the origin. The main graph shows the rapid convergence of the reference energy  $\langle E_R(\tau) \rangle$  towards the exact ground state energy  $E_0$ . The inset contains the plot of the ground state wave function: the results of the simulation are represented by triangles, while the continuous line corresponds to the analytical result equation (4.2). The agreement between the diffusion Monte Carlo result and the analytical expressions is very good.



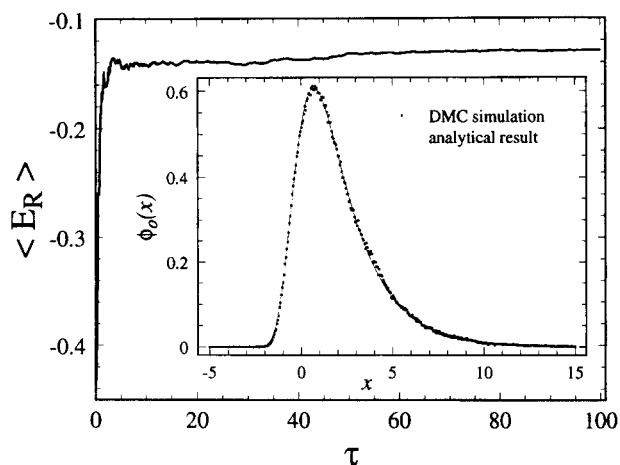


Fig. 3. DMC description of the Morse oscillator. The time evolution of  $\langle E_R \rangle$  is shown. The inset presents the ground state wave function; the numerical result is represented through dots, the wave function in the analytical form equation (4.4) is represented by a continuous line.

### B. Morse oscillator

The Morse potential is defined through

$$V(x) = D(e^{-2ax} - 2e^{-ax}). \quad (4.3)$$

In this case, one has a natural length scale through  $L = 1/a$  which can be used as the unit of length. As a result, from Eq. (3.2), one has  $T = m/\hbar a^2$  and  $\mathcal{E} = \hbar^2 a^2/m$ . For simplicity, we shall consider only the case when (in dimensionless units)  $D = 1/2$ ; the exact ground state energy is  $E_0 = -1/8$ , and the corresponding wave function is<sup>12</sup>

$$\phi_0(x) = \sqrt{2} \exp\left(-e^{-x} - \frac{x}{2}\right). \quad (4.4)$$

The results of the DMC description are presented in Fig. 3. Initially, all replicas were positioned in our simulation at the origin. The figure represents the time evolution of  $\langle E_R \rangle$  towards the exact ground state energy  $E_0 = -0.125$ . The figure demonstrates also that the resulting ground state wave function is in excellent agreement with Eq. (4.4).

### C. Hydrogen atom

In case of the hydrogen atom, it is customary to choose the unit of length equal to the Bohr radius  $a = \hbar^2/m_e^2 (\approx 0.53 \text{ \AA})$ . Thus, by setting  $L = a$ , one finds  $T = \hbar^3/m_e^4$  and  $\mathcal{E} = m_e^4/\hbar^2 (\approx 27.21 \text{ eV})$ . The well-known ground state energy (in dimensionless units) is  $E_0 = -1/2$  and the corresponding radial wave function is<sup>5</sup>

$$\phi_0(r) = 2e^{-r}. \quad (4.5)$$

We have carried out a DMC simulation for the hydrogen atom, generalizing the previous description to three spatial dimensions. At  $\tau = 0$  all the replicas were located at  $(0,0,1)$ , i.e., at a distance of one Bohr radius from the origin, along the positive  $z$  axis. Fig. 4 shows the convergence of  $\langle E_R \rangle$  to the exact ground state energy. This convergence, however, is not as rapid as in the case of the harmonic oscillator and the Morse oscillator. This is not surprising since the hydrogen atom has three degrees of freedom which require more sampling. The running time of the simulation has also increased (see Table I). The inset shows both the radial wave function  $\phi_0(r)$  (triangles) and the function  $\chi(r) = r\phi_0(r)$  (squares).

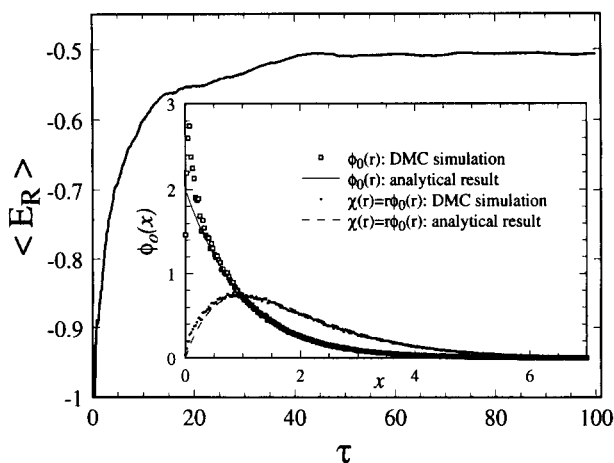


Fig. 4. DMC description of the hydrogen atom. The time evolution of  $\langle E_R \rangle$  is shown. The inset presents both the radial wave function  $\phi_0(r)$  and the function  $\chi(r) = r\phi_0(r)$ .

For comparison, the corresponding analytical solutions are also plotted with continuous lines. Again, the agreement between the analytical results and those obtained by our algorithm are good. The error of the radial wave function in the vicinity of the origin is due to insufficient sampling of the number of replicas in this region of the configuration space. To improve the wave function for  $r \sim 0$  one may decrease the size of the counting ‘‘boxes’’ in the vicinity of the origin, increase the number of replicas, or increase  $\tau$ .

### D. $H_2^+$ ion

The  $H_2^+$  ion is held stabilized through a single electron moving in the electric field of two protons separated, at equilibrium, by a distance  $R = 1.06 \text{ \AA}$ .<sup>10</sup> The ground state of this quantum system can not be determined analytically. The numerically determined ground state energy of  $H_2^+$  is  $E_0 = -16.252 \pm 0.002 \text{ eV}$ .<sup>10</sup> The DMC method can be applied in this case as for the hydrogen atom. The Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{|\mathbf{r} + \frac{1}{2}\mathbf{R}|} - \frac{e^2}{|\mathbf{r} - \frac{1}{2}\mathbf{R}|} \quad (4.6)$$

where  $\mathbf{R}$  denotes the separation between the two protons. The results of the DMC description are presented in Fig. 5. The same dimensionless units as in the case of the hydrogen atom were employed and replicas were located initially at the origin, the nuclei being located at  $(0,0,\pm R/2)$ , for  $R = 2$ . Figure 5 shows that the ground state energy obtained asymptotically is  $-16.75 \text{ eV}$  (see Table I) which differs from the more exact numerical value<sup>10</sup> by  $0.5 \text{ eV}$ , i.e., by about 3%. The inset in Fig. 5 shows a plot of the spatial distribution of the replicas, demonstrating that the electronic ground state wave function is nearly spherically symmetric.

### E. $H_2$ molecule

The  $H_2$  molecule is formed by two protons, at equilibrium separated by  $R = 0.74 \text{ \AA}$ ,<sup>10</sup> and by two electrons. In the Born approximation one considers the proton positions fixed and solves the corresponding stationary Schrödinger equation for the electrons. The wave function of the two electrons must be antisymmetric with respect to exchange of the electrons. In the ground state, the electrons are in a singlet spin state

Table I. Results obtained during two different simulations performed on four quantum systems listed in the first column. During Simulation I (II) the initial number of replicas was 500 (4000), the time step  $\Delta\tau=0.1$  (0.05), the length of simulation  $\tau=1000\Delta\tau$  ( $2000\Delta\tau$ ), the random number generator seed 1 and the number of boxes used to calculate the spatial distribution of the replicas 200 (400).  $\langle E_R \rangle$  and  $\delta E_R$  are defined in the text.  $\Delta t$  represents the actual running time of the simulation on a HP-9000 (series 700) workstation. For the hydrogen atom, the energies are given both in dimensionless units and in eV (in parentheses), respectively. For comparison,  $E_0$  in the last column represents the exact value (analytical or obtained through an alternative numerical method) of the corresponding ground state energy.

Quantum system	Simulation I			Simulation II			$E_0$
	$\langle E_R \rangle$	$\delta E_R$	$\Delta t$	$\langle E_R \rangle$	$\delta E_R$	$\Delta t$	
Harmonic oscillator	0.505	0.094	8 s	0.500	0.048	4 min	0.5
Morse oscillator	-0.1236	0.0749	10 s	-0.1245	0.0330	4 min	-0.125
Hydrogen atom	-0.495	0.080	18 s	-0.505	0.040	5 min	0.5
	(-13.477 eV)	(2.186 eV)		(-13.752 eV)	(1.093 eV)		(13.6 eV)
$H_2^+$ ion	-16.753 eV	2.869 eV	40 s	-16.476 eV	1.389 eV	11 min	-16.25(2) eV <sup>a</sup>
$H_2$ molecule	-30.973 eV	3.638 eV	55 s	-31.968 eV	1.754 eV	16 min	-31.6(87) eV <sup>a</sup>

<sup>a</sup>Numerical values calculated based upon the heat of dissociation given by Hertzberg (Ref. 10).

$$\chi(1,2) = \sqrt{\frac{1}{2}} [\chi_{1/2,+1/2}(1)\chi_{1/2,-1/2}(2) - \chi_{1/2,-1/2}(1)\chi_{1/2,+1/2}(2)] \quad (4.7)$$

which is antisymmetric. Here,  $\chi_{1/2,\pm 1/2}(1,2)$  denotes the wave function of electron 1,2 in the spin 1/2 state with magnetic quantum numbers  $\pm 1/2$ . Accordingly, in the wave function of the electronic ground state

$$\Psi_0 = \Phi(\mathbf{r}_1, \mathbf{r}_2)\chi(1,2), \quad (4.8)$$

the factor  $\Phi(\mathbf{r}_1, \mathbf{r}_2)$ , describing the spatial distribution of the electrons, must be symmetric. This factor can be determined through

$$\Phi(\mathbf{r}_1, \mathbf{r}_2) = \sqrt{\frac{1}{2}}(\phi(\mathbf{r}_1, \mathbf{r}_2) + \phi(\mathbf{r}_2, \mathbf{r}_1)), \quad (4.9)$$

where  $\phi(\mathbf{r}_1, \mathbf{r}_2)$  is a solution of

$$\left[ -\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) - \frac{e^2}{|\mathbf{r}_1 - \frac{1}{2}\mathbf{R}|} - \frac{e^2}{|\mathbf{r}_1 + \frac{1}{2}\mathbf{R}|} - \frac{e^2}{|\mathbf{r}_2 - \frac{1}{2}\mathbf{R}|} - \frac{e^2}{|\mathbf{r}_2 + \frac{1}{2}\mathbf{R}|} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right] \phi(\mathbf{r}_1, \mathbf{r}_2) = E_0 \phi(\mathbf{r}_1, \mathbf{r}_2). \quad (4.10)$$

The Schrödinger equation (4.10) cannot be solved analytically. The available numerical result for the ground state energy of  $H_2$  is  $E_0 = -31.688 \pm 0.013$  eV.<sup>10</sup> The results of the diffusion Monte Carlo simulation for  $H_2$  are shown in Fig. 6. The same dimensionless units as in the case of the hydrogen atom were employed and the replicas at  $\tau=0$  were located initially at  $\{(0,0,1)(0,0,-1)\}$ ; the position of the protons were  $(0,0,\pm R/2)$ , with  $R=1.398$ . The simulation took about 55 s to complete and the obtained ground state energy was  $E_0 = -30.973$  eV (see Table I). This result differs from the exact energy by 0.715 eV, i.e., by 2.3%. The inset of Fig. 6 shows the spatial distribution of the replicas which is nearly spherically symmetric. In Fig. 7, we present the results of a calculation in which an unphysically large  $R$  value of 8.398

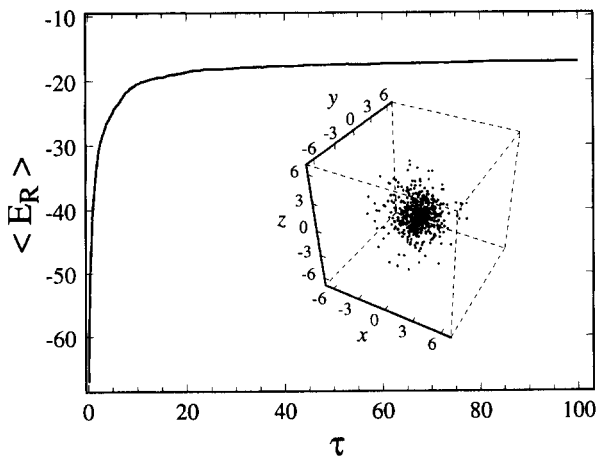


Fig. 5. DMC description of the  $H_2^+$  ion. The time evolution of  $E_R$  is shown. The inset presents the spatial distribution of the replicas, i.e., the electron cloud.

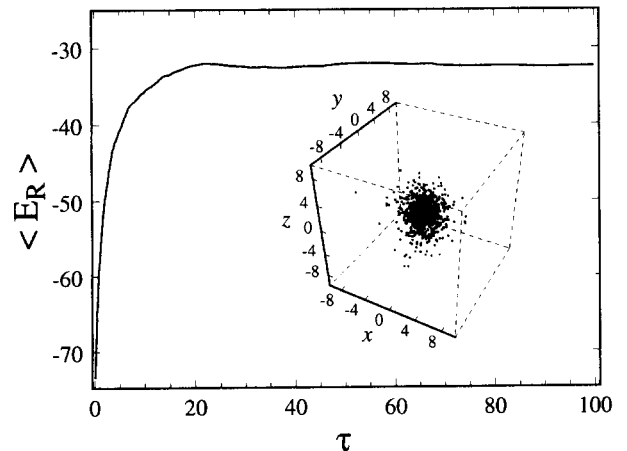


Fig. 6. DMC description of the electronic ground state of the  $H_2$  molecule. The time evolution of  $E_R$  is shown. The inset presents the spatial distribution of the replicas, i.e., the electron cloud.

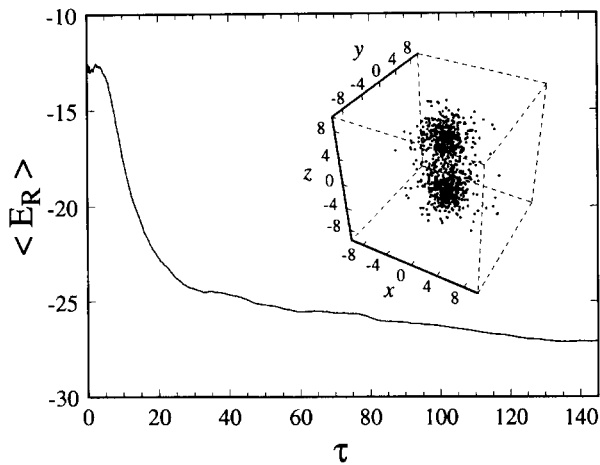


Fig. 7.  $E_R$  vs  $\tau$  and the spatial distribution of the replicas (inset) at the end of the simulation for the  $H_2$  molecule for an unphysically large separation between the two protons of 8.398 Bohr radius.

was assumed. In this case, the distribution of the electronic cloud around the protons is clearly anisotropic, the energy of the electrons is still negative.

## V. DISCUSSION

In this article, we have presented a detailed account of the path integral formulation of the DMC method devised for calculating simultaneously the ground state energy and wave function of an arbitrary quantum system. A simple numerical algorithm based on the DMC method has been formulated and a computer program, based on this algorithm, has been applied to determine numerically the ground state of a few quantum systems of pedagogical interest.

The DMC algorithm, as presented in this article, is quite unstable numerically. We want to demonstrate here the need for an improvement of the method. This need arises due to the strong fluctuations stemming from the birth-death processes employed in this method. These fluctuations affect the reference energy  $E_R(\tau)$  which is expected to converge to the ground state energy. This convergence is observed only for the value of  $\langle E_R \rangle$  defined through Eq. (3.6). The actual time dependence of  $E_R$  is presented in Fig. 8 for the case of the harmonic oscillator. The fluctuations in  $E_R$  are large, but symmetrically distributed around  $\langle E_R \rangle$  which explains the good agreement between the result of the simulation with the exact result. Attempts to use our DMC program to calculate the electronic ground state energy of the  $H_2^+$  ion and  $H_2$  molecule as a function of the separation between the two protons, which would allow one to determine the equilibrium separation from the minimum of this dependence, have failed due to large fluctuations in  $E_R$ .

The DMC algorithm can be significantly improved resorting to a method called ‘‘importance sampling’’.<sup>1,9</sup> The basic idea of this method is to change the probability distribution of the replicas in a controlled way. This can be achieved by reformulating Eq. (2.10) such that the resulting equation has a solution  $\Psi(x, \tau)$  multiplied by an approximation of the ground state wave function, the latter being obtained, for example, from a variational method. Application of the DMC method to this new equation yields then replicas which spend more time in ‘‘important’’ regions of the configuration

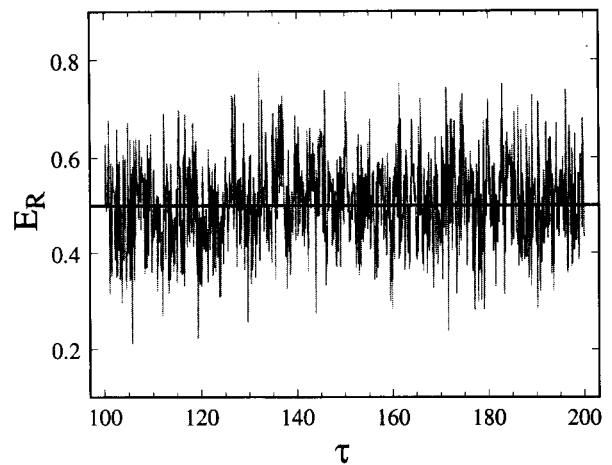


Fig. 8. Fluctuations of  $E_R$  about its average value  $\langle E_R \rangle \approx 0.5$  well after the distribution of replicas reached its stationary form. The relatively large fluctuations are due to the birth-death processes which persist even if the replicas are distributed according to the exact ground state wave function  $\phi_0(x)$  and the reference energy  $E_R$  has been adapted to the exact ground state energy.

space where the wave function  $\Psi(x, \tau)$  is expected to be large. For details regarding importance sampling the reader is referred to the literature cited.<sup>1,2,9</sup>

Any implementation of the DMC method leads to systematic errors due to the use of a finite time step  $\Delta\tau$ . Apparently, these errors can be reduced to zero by choosing  $\Delta\tau \rightarrow 0$ . Unfortunately, this is not the case. Even worse, making the time step shorter and shorter not only increases the needed computer time, but also renders the successive generations of replicas more and more correlated such that their distribution actually departs more and more from the ground state wave function. For each quantum system investigated, it is necessary to find the most convenient value of  $\Delta\tau$  which is short enough to produce small systematic errors and at the same time long enough to keep the successive distributions of replicas sufficiently uncorrelated.

In order to apply the DMC method to calculate the ground state properties of a system with interacting fermions one has to treat, as mentioned above, the ‘‘sign problem’’ due to the antisymmetry property of the many-fermion wave function. Two principle methods, the *fixed-node* method<sup>13,1</sup> and the *release-node* method,<sup>14,1</sup> have been proposed to deal with this problem.

Once some kind of importance sampling is implemented and the sign problem, in the case of many-fermion systems, is resolved, the DMC method can be used to compute ground state properties for molecules or molecular clusters<sup>9</sup> and for quantum spin, boson and fermion systems.<sup>15</sup> The method is also applicable to the study of ground-state phase transitions due to quantum fluctuations, a topic of modern condensed matter physics.<sup>15</sup>

Finally, we would like to mention that the DMC method has been extended and successfully applied to the study of the excited states of molecules and clusters<sup>9</sup> and also to the study of finite temperature properties of different condensed matter systems. Also, the DMC method has been successfully applied in quantum field theories.<sup>16</sup>

## ACKNOWLEDGMENTS

This work has been supported by a National Science Foundation REU fellowship to B.F., by funds of the University of Illinois at Urbana-Champaign, and by a grant from the National Science Foundation (BIR-9318159).

<sup>a)</sup>To whom correspondence should be addressed.

<sup>1</sup>D. Ceperley and B. Alder, "Quantum Monte Carlo," *Science* **231**, 555–560 (1986).

<sup>2</sup>M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods* (Wiley, New York, 1986).

<sup>3</sup>J. B. Anderson, "A random-walk simulation of the Schrödinger equation:  $H_3^+$ ," *J. Chem. Phys.* **63**, 1499–1503 (1975).

<sup>4</sup>W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge U.P., Cambridge, 1992).

<sup>5</sup>L. D. Landau and E. M. Lifshitz, *Quantum Mechanics, Vol. 3 of Course of Theoretical Physics* (Pergamon, Oxford, 1977).

<sup>6</sup>R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

<sup>7</sup>D. Khandekar, S. Lawande, and K. B. Hagwat, *Path Integral Methods and Their Applications* (World Scientific, London, 1993).

<sup>8</sup>S. E. Koonin, *Computational Physics* (Benjamin, Reading, MA, 1986).

<sup>9</sup>M. A. Suhm and R. O. Watts, "Quantum Monte Carlo studies of vibrational states in molecules and clusters," *Phys. Rep.* **204**, 293–329 (1991).

<sup>10</sup>G. Herzberg, *Molecular Spectra and Molecular Structure* (Van Nostrand, New York 1950).

<sup>11</sup>The program was written in the C programming language. A copy of the source code is freely available from one of the authors (KS).

<sup>12</sup>L. Infeld and T. E. Hull, "The factorization method," *Rev. Mod. Phys.* **23**, 21–68 (1951).

<sup>13</sup>P. J. Reynolds, D. M. Ceperley, B. J. Adler, and J. W. A. Lester, "Fixed-node quantum Monte Carlo for molecules," *J. Chem. Phys.* **77**, 5593–5603 (1982).

<sup>14</sup>D. M. Ceperley and B. J. Adler, "Quantum Monte Carlo for molecules: Green's function and nodal release," *J. Chem. Phys.* **81**, 5833–5844 (1984).

<sup>15</sup>M. Suzuki, *Quantum Monte Carlo Methods in Condensed Matter Physics* (World Scientific, Singapore, 1993).

<sup>16</sup>T. Barnes and G. J. Daniell, "Numerical Solution of Field Theories Using Random Walks," *Nucl. Phys. B* **257**, 173–198 (1985).

## Using model systems to demonstrate instances of mathematical inequalities

R. J. Tykodi

*Department of Chemistry, University of Massachusetts Dartmouth, North Dartmouth, Massachusetts 02747-2300*

(Received 5 December 1994; accepted 9 October 1995)

By considering certain properties of model systems, one can sometimes demonstrate instances of appropriate mathematical inequalities. For example, by considering the temperature-equalization process in a set of contiguous metal blocks, one can demonstrate instances of the mathematical inequality that states that the arithmetic average of a set of non-negative real numbers is never less than the geometric average of those same numbers. For each of six different mathematical inequalities, an appropriate model system is used to demonstrate instances of the inequality.

© 1996 American Association of Physics Teachers.

### I. INTRODUCTION

There is a curious fascination in finding that a model system can behave in such a way as to demonstrate instances of a general mathematical inequality.

As an illustration, consider a set of dc resistors having resistances  $\{R_i\}$  under certain conditions. Let there be  $n$  resistors in the set. If the resistors are connected in series, they have an effective resistance  $R_s$  such that

$$R_s = \sum_{i=1}^n R_i. \quad (1)$$

On the other hand, if the resistors are connected in parallel, they have an effective resistance  $R_p$  such that

$$R_p = \left( \sum_{i=1}^n 1/R_i \right)^{-1}. \quad (2)$$

The case  $n=1$  is a degenerate case in that the designations "series" and "parallel" do not properly apply to a single resistor. It will be convenient, nevertheless, to allow the case  $n=1$  and to accept as the definitions of "series" and "parallel" for  $n=1$  the outcomes of relations (1) and (2) for that case (i.e.,  $R_s = R_p = R_1$ ).

If we allow the case  $n=1$ , dc theory tells us that

$$R_s \geq R_p, \quad (3)$$

or, equivalently, that

$$\left( \sum_{i=1}^n R_i \right) \left( \sum_{i=1}^n 1/R_i \right) \geq 1. \quad (4)$$

If we examine the expression

$$\left( \sum_{i=1}^n R_i \right) \left( \sum_{i=1}^n 1/R_i \right) \quad (5)$$