The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

1929-6

Advanced School on Quantum Monte Carlo Methods in Physics and Chemistry

21 January - 1 February, 2008

DMC 1

C. Filippi
Leiden University

Strada Costiera 11, 34014 Trieste, Italy - Tel. +39 040 2240 111; Fax +39 040 224 163 - sci_info@ictp.it, www.ictp.it

Introduction to quantum Monte Carlo methods: Diffusion Monte Carlo

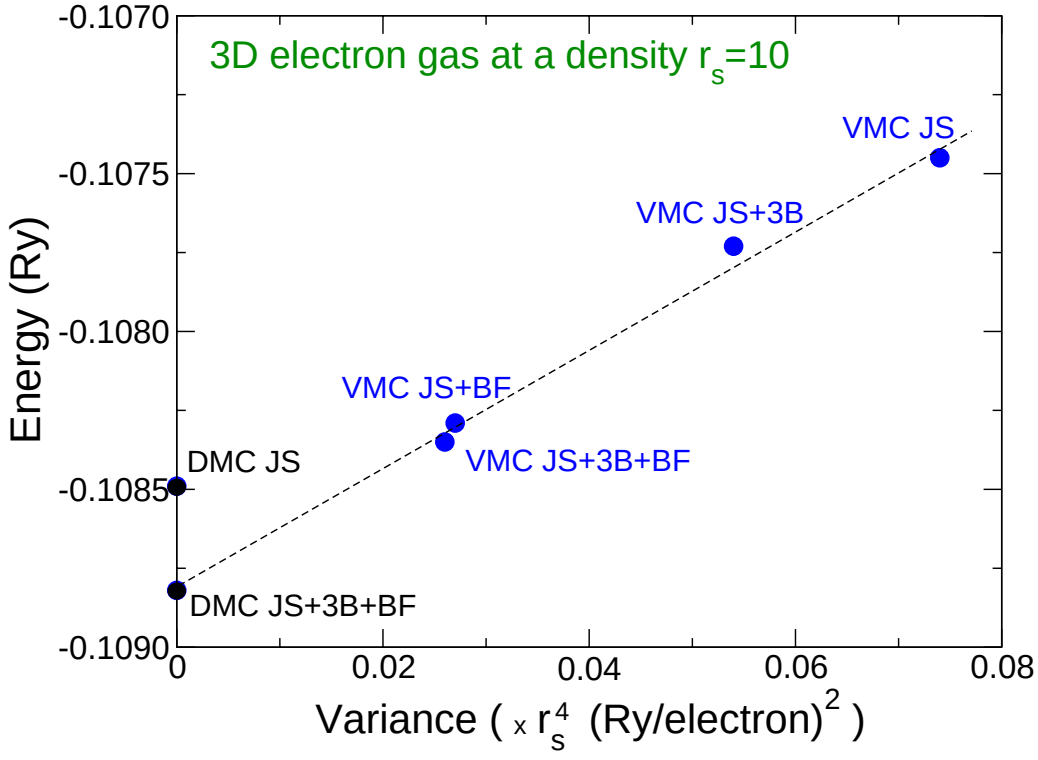
Claudia Filippi

Instituut-Lorentz, Universiteit Leiden, The Netherlands

Advanced School on QMC Methods in Physics and Chemistry
Jan 21-Feb 2, 2008, ICTP, Trieste

Why going beyond VMC?

Dependence of VMC from wave function Ψ



Kwon, Ceperley, Martin, Phys. Rev. B **58**, 6800 (1998)

Why going beyond VMC?

- Dependence on wave function: What goes in, comes out!
- No automatic way of constructing wave function Ψ

  Choices must be made about functional form (human time)
- Hard to ensure good error cancelation on energy differences

  e.g. easier to construct good Ψ for closed than open shells

Can we remove wave function bias?

Projector Monte Carlo methods

- ▷ Construct an operator which inverts spectrum of $\mathcal{H}$
- ▷ Use it to stochastically project the ground state of $\mathcal{H}$

Diffusion Monte Carlo	$\exp[-\tau(\mathcal{H} - E_{\mathrm{T}})]$
Green's function Monte Carlo	$1/(\mathcal{H} - E_{\mathrm{T}})$
Power Monte Carlo	$E_{\mathrm{T}} - \mathcal{H}$

Diffusion Monte Carlo

Consider initial guess $\Psi^{(0)}$ and repeatedly apply projection operator

$$\boxed{\Psi^{(n)} = e^{-\tau(\mathcal{H}-E_{\mathrm{T}})}\Psi^{(n-1)}}$$

Expand $\Psi^{(0)}$ on the eigenstates Ψ_i with energies E_i of $\mathcal{H}$

$$\Psi^{(n)} = e^{-n\tau(\mathcal{H}-E_{\mathrm{T}})}\Psi^{(0)} = \sum_i \Psi_i \left\langle \Psi^{(0)} | \Psi_i \right\rangle e^{-n\tau(E_i-E_{\mathrm{T}})}$$

and obtain in the limit of $n \to \infty$

$$\lim_{n\to\infty} \Psi^{(n)} = \Psi_0 \left\langle \Psi^{(0)} | \Psi_0 \right\rangle e^{-n\tau(E_0-E_{\mathrm{T}})}$$

If we choose $E_{\mathrm{T}} \approx E_0$, we obtain $\boxed{\lim_{n\to\infty} \Psi^{(n)} = \Psi_0}$

How do we perform the projection?

Rewrite projection equation in integral form

$$\Psi(\mathbf{R}', t+\tau) = \int \mathrm{d}\mathbf{R}\, G(\mathbf{R}', \mathbf{R}, \tau)\Psi(\mathbf{R}, t)$$

where $G(\mathbf{R}', \mathbf{R}, \tau) = \langle \mathbf{R}' | e^{-\tau(\mathcal{H}-E_{\mathrm{T}})} | \mathbf{R} \rangle$

- Can we sample the wave function?

  For the moment, assume we are dealing with bosons, so $\Psi > 0$

- Can we interpret $G(\mathbf{R}', \mathbf{R}, \tau)$ as a transition probability?

  If yes, we can perform this integral by Monte Carlo integration

VMC and DMC as power methods

VMC Distribution function is given $\rho(\mathbf{R}) = \dfrac{|\Psi(\mathbf{R})|^2}{\int \mathrm{d}\mathbf{R}|\Psi(\mathbf{R})|^2}$

Construct P which satisfies stationarity condition $P\rho = \rho$

$\rightarrow$ ρ is eigenvector of P with eigenvalue 1

$\rightarrow$ ρ is the dominant eigenvector $\Rightarrow \lim_{n\rightarrow\infty} P^n \rho_{\mathrm{initial}} = \rho$

DMC Opposite procedure!

The matrix P is given $\rightarrow$ $P = \langle \mathbf{R}' | e^{-\tau(\mathcal{H}-E_{\mathrm{T}})} | \mathbf{R} \rangle$

We want to find the dominant eigenvector $\rho = \Psi_0$

What can we say about the Green's function?

$$G(\mathbf{R}', \mathbf{R}, \tau) = \langle \mathbf{R}' | e^{-\tau(\mathcal{H} - E_{\mathrm{T}})} | \mathbf{R} \rangle$$

$G(\mathbf{R}', \mathbf{R}, \tau)$ satisfies the imaginary-time Schrödinger equation

$$\boxed{(\mathcal{H} - E_{\mathrm{T}}) G(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial G(\mathbf{R}, \mathbf{R}_0, t)}{\partial t}}$$

with $G(\mathbf{R}', \mathbf{R}, 0) = \delta(\mathbf{R}' - \mathbf{R})$

Can we interpret $G(\mathbf{R}', \mathbf{R}, \tau)$ as a transition probability? (1)

$\mathcal{H} = \mathcal{T}$

Imaginary-time Schrödinger equation is a diffusion equation

$$-\frac{1}{2}\nabla^2 G(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial G(\mathbf{R}, \mathbf{R}_0, t)}{\partial t}$$

The Green's function is given by a Gaussian

$$G(\mathbf{R}', \mathbf{R}, \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R})^2}{2\tau}\right]$$

Positive and can be sampled

Can we interpret $G(\mathbf{R}', \mathbf{R}, \tau)$ as a transition probability? (2)

$\mathcal{H} = \mathcal{V}$

$$(\mathcal{V}(\mathbf{R}) - E_{\mathrm{T}}) G(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial G(\mathbf{R}, \mathbf{R}_0, t)}{\partial t} ,$$

The Green's function is given by

$$G(\mathbf{R}', \mathbf{R}, \tau) = \exp\left[-\tau \left(\mathcal{V}(\mathbf{R}) - E_{\mathrm{T}}\right)\right] \delta(\mathbf{R} - \mathbf{R}'),$$

Positive but does not preserve the normalization

It is a factor by which we multiply the distribution $\Psi(\mathbf{R}, t)$

$\mathcal{H} = \mathcal{T} + \mathcal{V}$ and a combination of diffusion and branching

Trotter's theorem $\longrightarrow$ $e^{(A+B)\tau} = e^{A\tau} e^{B\tau} + \mathcal{O}(\tau^2)$

$$\begin{aligned} \langle \mathbf{R}' | e^{-\mathcal{H}\tau} | \mathbf{R}_0 \rangle &\approx \langle \mathbf{R}' | e^{-\mathcal{T}\tau} e^{-\mathcal{V}\tau} | \mathbf{R}_0 \rangle \\ &= \int \mathrm{d}\mathbf{R}'' \langle \mathbf{R}' | e^{-\mathcal{T}\tau} | \mathbf{R}'' \rangle \langle \mathbf{R}'' | e^{-\mathcal{V}\tau} | \mathbf{R}_0 \rangle \\ &= \langle \mathbf{R}' | e^{-\mathcal{T}\tau} | \mathbf{R}_0 \rangle e^{-\mathcal{V}(\mathbf{R}_0)\tau} \end{aligned}$$

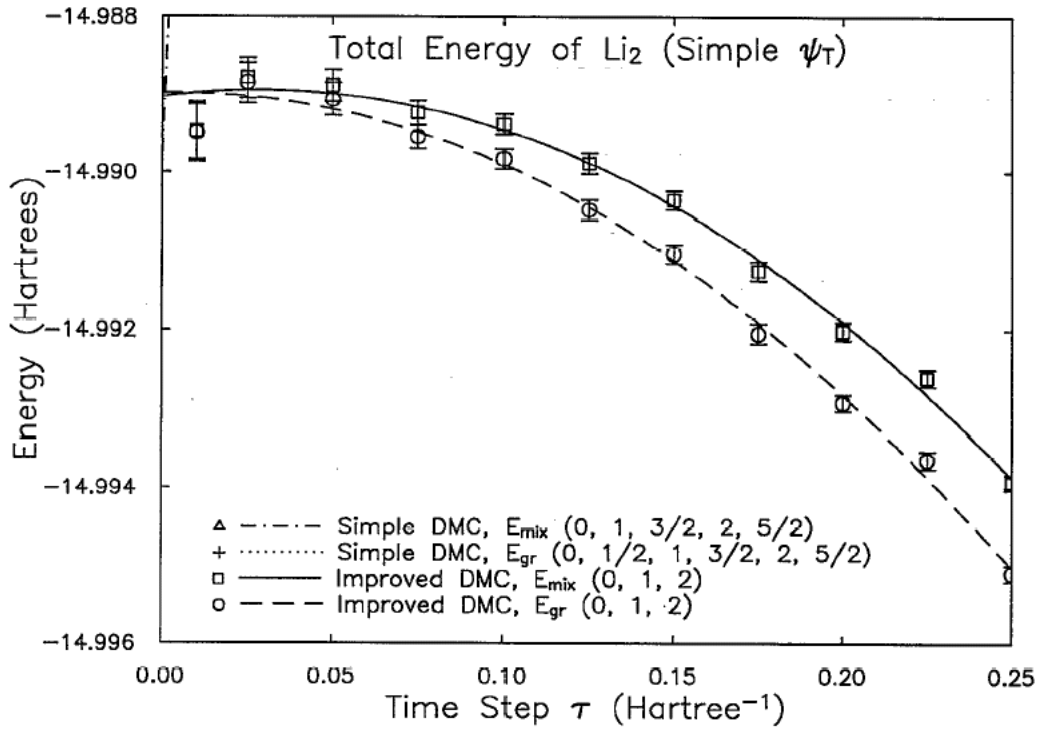
The Green's function in the short-time approximation to $\mathcal{O}(\tau^2)$ is

$$G(\mathbf{R}', \mathbf{R}, \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R})^2}{2\tau}\right] \exp\left[-\tau \left(\mathcal{V}(\mathbf{R}) - E_{\mathrm{T}}\right)\right]$$

DMC results must be extrapolated at short time-steps ($\tau \longrightarrow 0$)

Time-step extrapolation

Example: Energy of Li_2 versus time-step τ



Umrigar, Nightingale, Runge, J. Chem. Phys. **94**, 2865 (1993)

Diffusion Monte Carlo as a branching random walk (1)

The basic DMC algorithm is rather simple:

1. Sample $\Psi^{(0)}(\mathbf{R})$ with the Metropolis algorithm

   Generate M_0 walkers $\mathbf{R}_1, \ldots, \mathbf{R}_{M_0}$ (zeroth generation)

2. Diffuse each walker as $\boxed{\mathbf{R}' = \mathbf{R} + \xi}$

   where ξ is sampled from $g(\xi) = (2\pi\tau)^{-3N/2} \exp\left(-\xi^2/2\tau\right)$

3. For each walker, compute the factor

   $$\boxed{p = \exp\left[-\tau(\mathcal{V}(\mathbf{R}) - E_{\mathrm{T}})\right]}$$

   Branch the walker with p the probability to survive

Continue →

Diffusion Monte Carlo as a branching random walk (2)

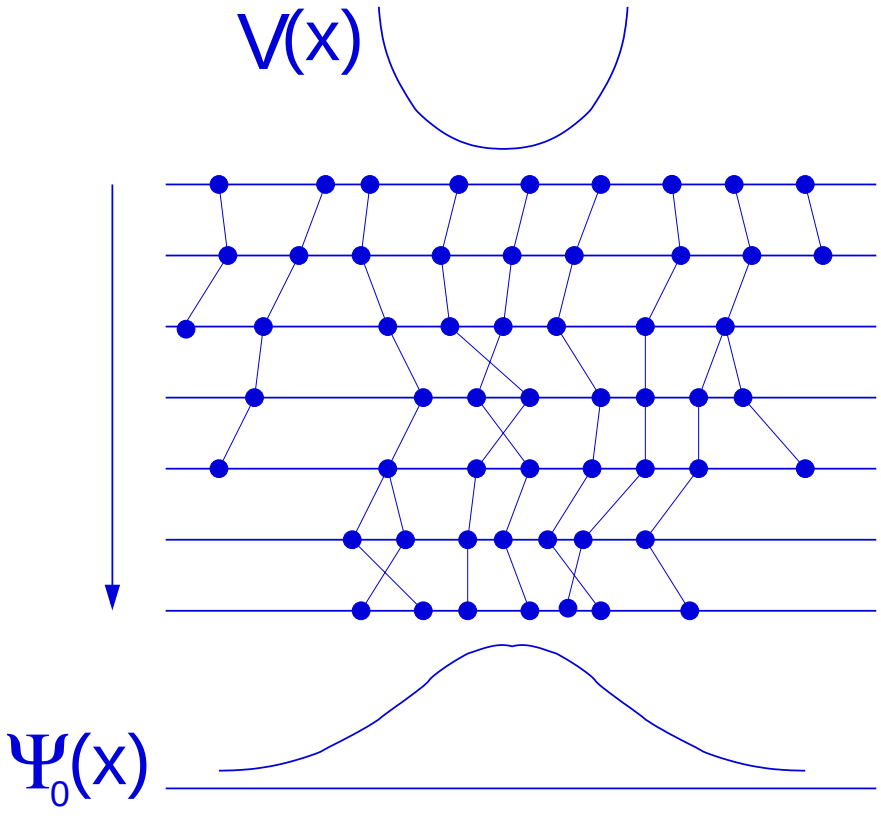
4. Branch the walker with p the probability to survive
   - ▷ If $p < 1$, the walker survives with probablity p
   - ▷ If $p > 1$, the walker continues and new walkers with the same coordinates are created with probability $p - 1$

   $\Rightarrow$ Number of copies of the current walker equal to $\mathrm{int}(p + \eta)$ where η is a random number between (0,1)

5. Adjust E_{T} so that population fluctuates around target M_0

$\rightarrow$ After many iterations, walkers distributed as $\Psi_0(\mathbf{R})$

Diffusion and branching in a harmonic potential



Walkers proliferate/die in regions of lower/higher potential than E_{T}

Some comments on the simple DMC algorithm

- E_{T} is adjusted to keep population stable

  IF $M(t)$ is the current and M_0 the desired population

$$M(t+T) = M(t)\, e^{-T(-\delta E_{\mathrm{T}})} = M_0 \quad \Rightarrow \quad \delta E_{\mathrm{T}} = \frac{1}{T} \ln \left[\frac{M_0}{M(t)} \right]$$

  If $E_{\mathrm{est}}(t)$ is current best estimate of the ground state

$$\boxed{E_{\mathrm{T}}(t+\tau) = E_{\mathrm{est}}(t) + \frac{1}{g\tau} \ln \left[M_0 / M(t) \right]}$$

  $\Rightarrow$ Feedback on E_{T} introduces population control bias

- Symmetric branching $\exp[-\tau(\mathcal{V}(\mathbf{R}) + \mathcal{V}(\mathbf{R}'))/2]$ starting from

$$e^{(A+B)\tau} = e^{A\tau/2} e^{B\tau} e^{A\tau/2} + \mathcal{O}(\tau^3)$$

Problems with simple algorithm

The simple algorithm is **inefficient and unstable**

- Potential can vary a lot and be unbounded

  e.g. electron-nucleus interaction $\longrightarrow$ Exploding population

- Branching factor grows with system size

Importance sampling

Start from integral equation

$$\Psi(\mathbf{R}', t+\tau) = \int \mathrm{d}\mathbf{R}\, G(\mathbf{R}', \mathbf{R}, \tau)\Psi(\mathbf{R}, t)$$

Multiply each side by trial Ψ and define $\boxed{f(\mathbf{R}, t) = \Psi(\mathbf{R})\Psi(\mathbf{R}, t)}$

$$f(\mathbf{R}', t+\tau) = \int \mathrm{d}\mathbf{R}\, \tilde{G}(\mathbf{R}', \mathbf{R}, \tau) f(\mathbf{R}, t)$$

where the importance sampled Green's function is

$$\boxed{\tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = \Psi(\mathbf{R}')\langle \mathbf{R}'|e^{-\tau(\mathcal{H}-E_{\mathrm{T}})}|\mathbf{R}\rangle/\Psi(\mathbf{R})}$$

We obtain $\boxed{\lim_{n\to\infty} f(\mathbf{R}) = \Psi(\mathbf{R})\Psi_0(\mathbf{R})}$

Importance sampled Green's function

The importance sampled $\tilde{G}(\mathbf{R}, \mathbf{R}_0, \tau)$ satisfies

$$\boxed{-\frac{1}{2}\nabla^2 \tilde{G} + \nabla \cdot [\tilde{G}\, \mathbf{V}(\mathbf{R})] + [E_{\mathrm{L}}(\mathbf{R}) - E_{\mathrm{T}}]\, \tilde{G} = -\frac{\partial \tilde{G}}{\partial \tau}}$$

with the quantum velocity $\mathbf{V}(\mathbf{R}) = \dfrac{\nabla \Psi(\mathbf{R})}{\Psi(\mathbf{R})}$

We now have drift in addition to diffusion and branching terms

Trotter's theorem $\Rightarrow$ Consider them separately for small enough τ

The drift-branching components: Reminder

Diffusion term

$$-\frac{1}{2}\nabla^2 \tilde{G}(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial \tilde{G}(\mathbf{R}, \mathbf{R}_0, t)}{\partial t}$$

$$\Rightarrow \tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R})^2}{2\tau}\right]$$

Branching term

$$(E_{\mathrm{L}}(\mathbf{R}) - E_{\mathrm{T}})\tilde{G}(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial \tilde{G}(\mathbf{R}, \mathbf{R}_0, t)}{\partial t}$$

$$\Rightarrow \tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = \exp\left[-\tau\left(E_{\mathrm{L}}(\mathbf{R}) - E_{\mathrm{T}}\right)\right] \delta(\mathbf{R} - \mathbf{R}')$$

The drift-diffusion-branching Green's function

$$-\frac{1}{2}\nabla^2\tilde{G} + \nabla\cdot[\tilde{G}\,\mathbf{V}(\mathbf{R})] + [E_{\mathrm{L}}(\mathbf{R}) - E_{\mathrm{T}}]\,\tilde{G} = -\frac{\partial\tilde{G}}{\partial\tau}$$

Drift term

Assume $\mathbf{V}(\mathbf{R}) = \dfrac{\nabla\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$ constant over the move (true as $\tau \to 0$)

The drift operator becomes $\boxed{\mathbf{V}\cdot\nabla + \nabla\cdot\mathbf{V} \approx \mathbf{V}\cdot\nabla}$ so that

$$\mathbf{V}\cdot\nabla\tilde{G}(\mathbf{R},\mathbf{R}_0,t) = -\frac{\partial\tilde{G}(\mathbf{R},\mathbf{R}_0,t)}{\partial t}$$

with solution $\boxed{\tilde{G}(\mathbf{R},\mathbf{R}_0,t) = \delta(\mathbf{R}-\mathbf{R}_0-\mathbf{V}t)}$

The drift-diffusion-branching Green's function

Drift-diffusion-branching short-time Green's function is

$$\tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R} - \tau\mathbf{V}(\mathbf{R}))^2}{2\tau}\right] \times$$
$$\times \exp\left\{-\tau\left[(E_{\mathrm{L}}(\mathbf{R}) + E_{\mathrm{L}}(\mathbf{R}'))/2 - E_{\mathrm{T}}\right]\right\} + \mathcal{O}(\tau^2)$$

What is new in the drift-diffusion-branching expression?

- $\mathbf{V}(\mathbf{R})$ pushes walkers where Ψ is large
- $E_{\mathrm{L}}(\mathbf{R})$ is better behaved than the potential $\mathcal{V}(\mathbf{R})$

  Cusp conditions $\Rightarrow$ No divergences when particles approach

  As $\Psi \rightarrow \Psi_0$, $E_{\mathrm{L}} \rightarrow E_0$ and branching factor is smaller

DMC algorithm with importance sampling

1. Sample initial walkers from $|\Psi(\mathbf{R})|^2$
2. Drift and diffuse the walkers as $\mathbf{R}' = \mathbf{R} + \tau \mathbf{V}(\mathbf{R}) + \xi$
   where ξ is sampled from $g(\xi) = (2\pi\tau)^{-3N/2} \exp\left(-\xi^2/2\tau\right)$
3. Branching step as in the simple algorithm but with the factor

$$p = \exp\left\{-\tau[(E_{\mathrm{L}}(\mathbf{R}) + E_{\mathrm{L}}(\mathbf{R}'))/2 - E_{\mathrm{T}}]\right\}$$

4. Adjust the trial energy to keep the population stable

$\rightarrow$ After many iterations, walkers distributed as $\Psi(\mathbf{R})\Psi_0(\mathbf{R})$

An important and simple improvement

If $\Psi = \Psi_0$, $E_{\rm L}(\mathbf{R}) = E_0 \rightarrow$ No branching term $\rightarrow$ Sample Ψ^2

Due to time-step approximation, we only sample Ψ^2 as $\tau \rightarrow 0$!

Solution Introduce accept/reject step like in Metropolis algorithm

$$\tilde{G}(\mathbf{R}', \mathbf{R}, \tau) \approx \underbrace{\mathcal{N} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R} - \mathbf{V}(\mathbf{R})\tau)^2}{2\tau}\right]}_{T(\mathbf{R}', \mathbf{R}, \tau)} \exp\left[-(E_{\rm L}(\mathbf{R}) + E_{\rm L}(\mathbf{R}'))\frac{\tau}{2}\right]$$

Walker drifts, diffuses and the move is accepted with probability

$$p = \min\left\{1, \frac{|\Psi(\mathbf{R}')|^2 \; T(\mathbf{R}, \mathbf{R}', \tau)}{|\Psi(\mathbf{R})|^2 \; T(\mathbf{R}', \mathbf{R}, \tau)}\right\}$$

$\rightarrow$ Improved algorithm with smaller time-step error

Evolution equation of the probability distribution

▷ $$\Psi(\mathbf{R}', t+\tau) = \int \mathrm{d}\mathbf{R}\, G(\mathbf{R}', \mathbf{R}, \tau)\Psi(\mathbf{R}, t)$$

where $G(\mathbf{R}', \mathbf{R}, \tau) = \langle \mathbf{R}' | e^{-\tau(\mathcal{H} - E_\mathrm{T})} | \mathbf{R} \rangle$

$$\boxed{(\mathcal{H} - E_\mathrm{T}) G(\mathbf{R}, \mathbf{R}_0, t) = -\frac{\partial G(\mathbf{R}, \mathbf{R}_0, t)}{\partial t}}$$

▷ $$\Psi(\mathbf{R}, t) = \int \mathrm{d}\mathbf{R}_0\, G(\mathbf{R}, \mathbf{R}_0, t)\Psi^{(0)}(\mathbf{R}_0)$$

satisfies the imaginary-time Schrödinger equation

$$\boxed{(\mathcal{H} - E_\mathrm{T})\Psi(\mathbf{R}, t) = -\frac{\partial \Psi(\mathbf{R}, t)}{\partial t}}$$

Electrons are fermions!

We assumed that $\Psi_0 > 0$ and that we are dealing with bosons

Fermions $\longrightarrow$ Ψ is antisymmetric and changes sign!

How can we impose antisymmetry in DMC method?

Stay tuned for second part of the lecture by Matthew Foulkes