

Variational Monte Carlo

- Historically first quantum simulation method
- Slater-Jastrow trial function
- Calculations of properties: $g(r)$ $n(k)$.
- Example: electron gas.

The “Variational Theorem”

Assume $\psi(R; a)$ is a trial function where R are the quantum degrees of freedom (positions, spin) “ a ” are “variational” parameters.

$$E_V(a) = \frac{\langle \psi(a) H \psi(a) \rangle}{\langle \psi(a) \psi(a) \rangle} \geq E_0 = \text{exact ground state energy}$$

$$\langle \psi(a) H \psi(a) \rangle \equiv \int dR \psi^*(R; a) H \psi(R; a)$$

$$E_V(a) = E_0 \Leftrightarrow \psi(R; a) = \phi_0(R)$$

$$E_L(R; a) \equiv \frac{1}{\psi(R; a)} H \psi(R; a) = \text{"local energy" of trial function}$$

$$E_V(a) = \left\langle \left\langle E_L(R; a) \right\rangle \right\rangle_{\psi^2} \quad \text{where} \quad \left\langle \left\langle O \right\rangle \right\rangle_{\psi^2} \equiv \frac{\langle \psi(a) O \psi(a) \rangle}{\langle \psi(a) \psi(a) \rangle}$$

$$\sigma^2(a) \equiv \frac{\langle \psi(a) (H - E_V(a))^2 \psi(a) \rangle}{\langle \psi(a) \psi(a) \rangle} = \left\langle \left\langle (E_L(R; a) - E_V(a))^2 \right\rangle \right\rangle_{\psi^2} = \text{variance of the trial function}$$

$$\frac{dE_V(a)}{da} = 2 \left\langle \left\langle \frac{d \ln \psi(R; a)}{da} (E_L(R; a) - E_V(a)) \right\rangle \right\rangle = 0$$

Conditions: matrix elements exist, symmetries and boundary conditions are correct.

First Major QMC Calculation

- PhD thesis of W. McMillan (1964) University of Illinois.
- VMC calculation of ground state of liquid helium 4.
- Applied MC techniques from classical liquid theory.
- Ceperley, Chester and Kalos (1976) generalized to fermions.

PHYSICAL REVIEW

VOLUME 133, NUMBER 2A

10 APRIL 1965

Ground State of Liquid He⁴

W. L. McMILLAN*

Department of Physics, University of Illinois, Urbana, Illinois

(Received 16 November 1964)

The properties of the ground state of liquid He⁴ are studied using a variational wave function of the form $\prod_{i < j} f(r_{ij})$. The Lennard-Jones 12-6 potential is used with parameters determined from the gas data by deBoer and Michiels. The configuration space integrals are performed by a Monte Carlo technique for 32 and 108 atoms in a cube with periodic boundary conditions. With $f(r) = \exp[-(2.6 \text{ \AA}/r)^6]$, the ground-state energy is found to be -0.78×10^{-18} ergs/atom, which is 20% above the experimental value. The liquid structure factor and the two-particle correlation function are in reasonably good agreement with the x-ray and neutron scattering experiments.

- Zero temperature (single state) method

Variational Monte Carlo (VMC)

- Variational Principle. Given an appropriate trial function:
 - Continuous
 - Proper symmetry
 - Normalizable
 - **Finite variance**
- Quantum chemistry uses a product of single particle functions
- With MC we can use any “computable” function.

- Sample R from $|\psi|^2$ using MCMC.
- Take average of local energy:
- Optimize ψ to get the best upper bound
- Error in energy is 2nd order

- Better wavefunction, lower variance!
“Zero variance” principle. (non-classical)

$$E_V = \frac{\int dR \langle \psi | H | \psi \rangle}{\int dR \langle \psi \psi \rangle} \geq E_0$$
$$\sigma^2 = \frac{\int dR \langle \psi | H^2 | \psi \rangle}{\int dR \langle \psi \psi \rangle} - E_V^2$$

$$E_L(R) = \Re \left[\psi^{-1}(R) H \psi(R) \right]$$

$$E_V = \langle E_L(R) \rangle_{\psi^2} \geq E_0$$

VARIATIONAL MONTE CARLO CODE

```
call initstate (s_old) ← Initialize the state
p_old = psi2 (s_old) ← Evaluate psi_trial
LOOP {
  call sample (s_old,s_new,T_new,1) ← Sample new state
  p_new = psi2 (s_new) ← Evaluate psi_trial
  call sample (s_new,s_old,T_old,0) ← Find transition prob.
  A = (p_new/T_new)/(p_old/T_old) ← for going backward
  if(A > rand () ) { ← Acceptance prob.
    s_old=s_new
    p_old=p_new
    naccept = naccept +1} ← Accept the move
  call averages (s_old) ← } Collect statistics
```

Periodic boundary conditions

- Minimum Image Convention: take the closest distance

$$|r|_M = \min (r+nL)$$

Potential is cutoff so that $V(r)=0$ for $r>L/2$ since force needs to be continuous. Remember perturbation theory.

- Image potential

$$V_I = \sum_n v(r_i-r_j+nL)$$

For long range potential this leads to the Ewald image potential. You need a background and convergence method.

What do we use for the trial function?

- **Formal requirements** antisymmetry, continuity, finite variance.
- **Mean field approaches** Slater determinant,
- **Expand in basis sets** e.g. multideterminants, CI.
- **Local energy argument**
 - Control singularity at small r , or r_{12} . Leads to cusp condition.
 - Behavior at large r , plasmons, van der Waals (dispersion) interaction
- **Feynman-Kacs formula** gives connection to local energy
 - Jastrow, Backflow, three-body interactions

Two electron (bose) ground state.

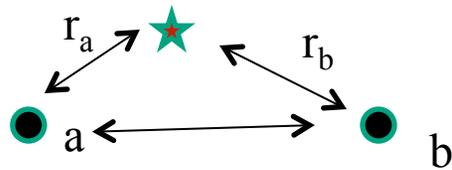
- Assume spin $\frac{1}{2}$ fermions (or bosons).
- Total wf is antisymmetric.
- Assume spin function is a singlet ($\uparrow\downarrow - \downarrow\uparrow$)
- Then spatial wavefunction is symmetric.

$$\Psi(r_1, r_2) = \Psi(r_2, r_1)$$

- If ground state is non-degenerate (true if space is “ergodic” or connected) then we can *assume* $\psi(r_1, r_2)$ is real, positive and symmetric by fixing the “gauge.”
- Why? Otherwise we could always lower the energy.

$$\therefore \Psi = e^{-u(r_1, r_2)} \quad u(r_1, r_2) = u(r_2, r_1)$$

H₂ molecule:



$$f(r) = \exp(-r)$$

$$\Psi_{RHF} = \phi(r_1)\phi(r_2) \quad \phi(r) = f_a(r) + f_b(r)$$

$$\Psi_{HL} = f_a(r_1)f_b(r_2) + f_b(r_1)f_a(r_2)$$

$$\Psi_c = \Psi_{HL} e^{-u(r_{12})}$$

$$\lim_{r_1 \rightarrow r_2} u(r_1, r_2) = u_0 - \frac{1}{2}|r_1 - r_2|$$

Cusp condition gives derivative of $u(r_1, r_2)$

- when two electrons approach each other
- when an electron gets close to a nuclei.

Correlation energy 1.11eV, spherical Jastrow 1.06eV

Generalized Feynman-Kacs

- We can find the correction to any wave function.

$$\phi(R_0) = \psi(R_0) \left\langle \exp \left[- \int_0^\infty dt E_\psi(R(t)) \right] \right\rangle_{RW}$$

$$\frac{dR}{dt} = \eta(t) - \nabla \ln \psi(R(t))$$



- The exact wavefunction is average over paths starting at R_0 .
- Gives intuition about how to improve it a given wavefunction.
- Can be used to compute the wavefunction (Next lecture).

Cumulant Approximation

- In FK formula take the average into the exponent
- It is possible to evaluate the average using fourier transforms.
- Very accurate for Coulomb problems
- Result: take functional form of local energy and smooth it out.
- Suppose we take for $\psi(R)$ an exact solution to the non-interacting SE. Then local energy is e-e interaction \rightarrow Jastrow

$$e^{-U(R_0)} \sim \exp \left[- \left\langle \int_0^\tau dt V(R(t)) \right\rangle_{RW} \right]$$

$$U_C(R_0) = \int_0^\infty dt V_s[R_0, t]$$

$$V_s(r, t) = \int dk e^{-ikr - \frac{tk^2}{2}} v_k$$



Trial function for bosons: “Jastrow” or pair product

- We want finite variance of the local energy.
- Whenever 2 atoms get close together wavefunction should vanish.
- The function $u(r)$ is similar to classical potential
- Local energy has the form:
G is the pseudoforce:

$$\psi(R) = \prod_{i < j} e^{-u(r_{ij})}$$

$$E_\psi(R) = \sum_{i < j} [v(r_{ij}) - 2\lambda \nabla^2 u(r_{ij})] - \lambda \sum_i G_i^2$$

$$G_i = \sum_j \nabla_i u(r_{ij})$$

If $v(r)$ diverges as r^{-1} how should $u(r)$ diverge? Assume:

$$u(r) = u(0) + u'(0)r + \dots$$

Keep $N-1$ electrons fixed and let 1 electron approach another and analyze the singular parts of the local energy.

Gives the **cusp condition** on u at small r .

$$r^{-1} = -u''(0) - \frac{2}{r}u'(0) + (u'(0))^2 + \dots$$

$$u'(0) = -\frac{1}{2}$$

Fermions: antisymmetric trial function

- At mean field level the wavefunction is a Slater determinant. Orbitals for homogenous systems are a filled set of plane waves.
- We can compute this energy analytically (HF).
- To include correlation we multiply by a “Jastrow”. We need MC to evaluate properties.
- New feature: how to compute the derivatives of a determinant and sample the determinant. Use tricks from linear algebra.
- Reduces complexity to $O(N^2)$.

$$\Psi_s(R) = \text{Det} \left\{ e^{ik_i r_j} \eta_i(\sigma_j) \right\}$$

$$\text{PBC: } k \cdot L = 2\pi n + \{\theta\}$$

$$\Psi_{SJ}(R) = \text{Det} \{ e^{ik_i r_j} \} e^{-\sum_{i<j} u(r_{ij})}$$

Slater-Jastrow trial function.

$$\det(\phi_k(r_j^T)) = \det(\phi_k(r_j)) \sum_k \phi_k(r_j^T) M_{k,i}^{-1}$$

$$\frac{1}{\det(M)} \frac{\partial \det(M)}{\partial a} = \text{Tr} \left\{ M^{-1} \frac{\partial M}{\partial a} \right\}$$

Spin & real vs. complex

- How do we treat spin in QMC?
- For extended systems we use the S_z representation.
- We have a fixed number of up and down electrons and we antisymmetrize among electrons with the same spin.
- This leads to 2 Slater determinants.
- For a given trial function, its real part is also a trial function (but it may have different symmetries), for example momentum

$$\left(e^{ikr}, e^{-ikr} \right) \quad \text{or} \quad \left(\cos(kr), \sin(kr) \right)$$

- For the ground state, without magnetic fields or spin-orbit interaction we can always work with real functions.
- However, it may be better to work with complex functions in some cases, e.g. for small molecules.

Preview of realistic trial functions

- Use LDA derived pseudopotentials
- Take orbitals from other methods:
 - “Gaussian” give orbitals for molecules
 - DFT-PW codes give orbitals for extended systems
- HF is slightly better because of self-interaction effects within DFT
- Multiply by a Jastrow function (electron gas or otherwise). Can include higher order e-e-n terms
- Must add a compensating e-n term in order to cancel out purely repulsive character of e-e correlation.
- Assuming LDA density is correct, this can be done by making sure VMC electron density=LDA electron density.

Spline Jastrow factor

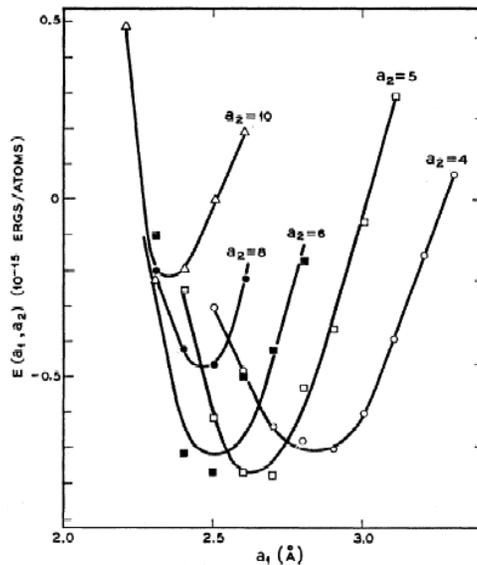
- For the HEG, the most general Jastrow factor has the form:

$$u(\vec{r}) = u_{sr}(r) + \sum_{\vec{k}} u_k e^{i\vec{k}\vec{r}}$$

$$u_{sr}(r) = 0 \quad \text{for } r > L/2$$

- $u(r)$ must be continuous, with a continuous derivative.
- We can impose the cusp condition at $r=0$, and BC at $r=L/2$.
- It is a smooth function: represent it as piecewise cubic polynomial in the region $0 < r < L/2$.
- M “Knots” at b_n . Total number of unknowns is $2M-1$
- Also u_k k -space Jastrows. Do we use RPA values?

Optimization of trial function



- Try to optimize $u(r)$ using reweighting (correlated sampling)
 - Sample R using $P(R)=\Psi^2(R,a_0)$
 - Now find minima of the analytic function $E_V(a)$
 - Or minimize the variance (more stable but wavefunctions less accurate).
- Statistical accuracy declines away from a_0 .

$$E_V(a) = \frac{\int \psi(a) H \psi(a)}{\int |\psi(a)|^2} = \frac{\sum_k w(R_i, a) E(R_i, a)}{\sum_k w(R_i, a)}$$

$$w(R_i, a) = \frac{|\psi(R, a)|^2}{P(R)}$$

$$E(R, a) = \psi^{-1}(R, a) H \psi(R, a)$$

$$N_{eff} = \frac{\left[\sum_i w_i \right]^2}{\sum_i w_i^2}$$

“modern” optimization

- With more computer time, we do a MC rw in both R and a (parameter space).
- Do usual VMC for a “block” and collect statistics on E , dE/da , $d^2E/(da_i da_j)$.
- Special estimators for these quantities.
- Then make a change in a: $a_{\text{new}} = a_{\text{old}} + c dE/da + \dots$
- Iterate until convergence.
- Lots more tricks to make it stable.
- Can do hundreds of parameters.

Scalar Properties, Static Correlations and Order Parameters

What do we get out of a simulation? Energy by itself doesn't tell you very much.

Other properties

- Do NOT have an upper bound property
- Only first order in accuracy

EXAMPLES

- Static properties: pressure,
- Density
- Pair correlation in real space and fourier space.
- Order parameters and broken symmetry: how to tell a liquid from a solid
- Specifically quantum: the momentum distribution

e-e Pair Correlation Function, $g(r)$

Primary quantity in a liquid is the probability distribution of pairs of particles. Given a particle at the origin what is the density of surrounding particles?

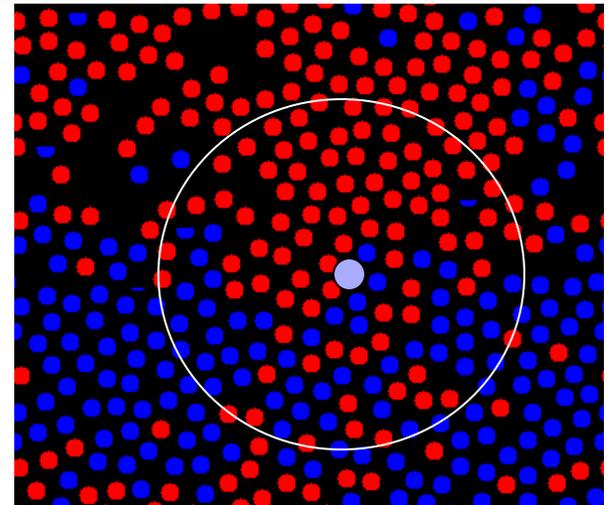
$$g(r) = \langle \sum_{i < j} \delta(r_i - r_j - r) \rangle$$

$$\rho(r, r') = \langle \sum_{i < j} \delta(r_i - r) \delta(r_j - r') \rangle$$

From $g(r)$ you can calculate all pair quantities (potential, pressure, ...)

$$V = \sum_{i < j} v(r_{ij}) = \frac{N\rho}{2} \int d^3r \frac{1}{r} g(r)$$

A function gives more information than a number!



(The static) structure factor $S(k)$

- The Fourier transform of the pair correlation function is the structure factor

$$S(k) = \frac{1}{N} \left\langle \left| \rho_k \right|^2 \right\rangle \quad \text{where } \rho_k = \sum_{i=1}^N e^{ik \cdot r}$$

$$S(k) = 1 + \rho \int dr e^{ik \cdot r} (g(r) - 1)$$

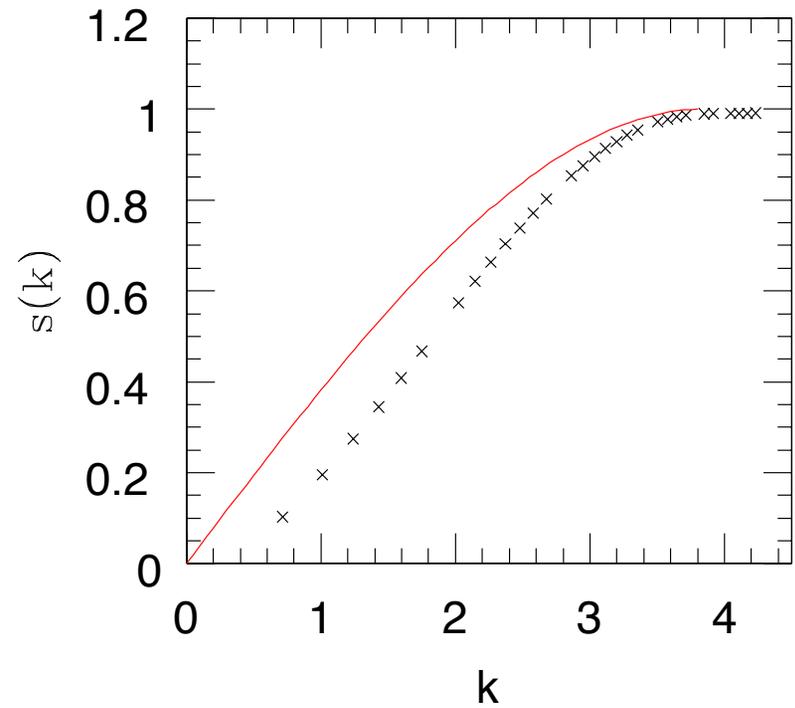
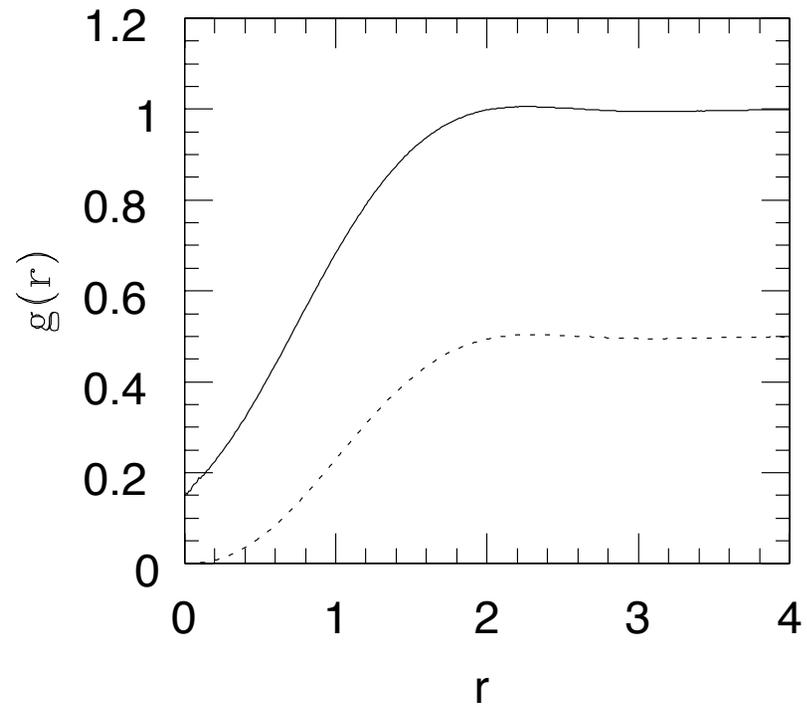
In PBC k lies on a lattice $k_n = (2\pi/L)$ (integer vector)
problem with (2) is to extend $g(r)$ to infinity

- $S(k)$ is measured in neutron and X-Ray scattering experiments.
- Can provide a direct test of the assumed potential.
- Used to see the state of a system:

liquid, solid, glass, gas? (much better than $g(r)$)

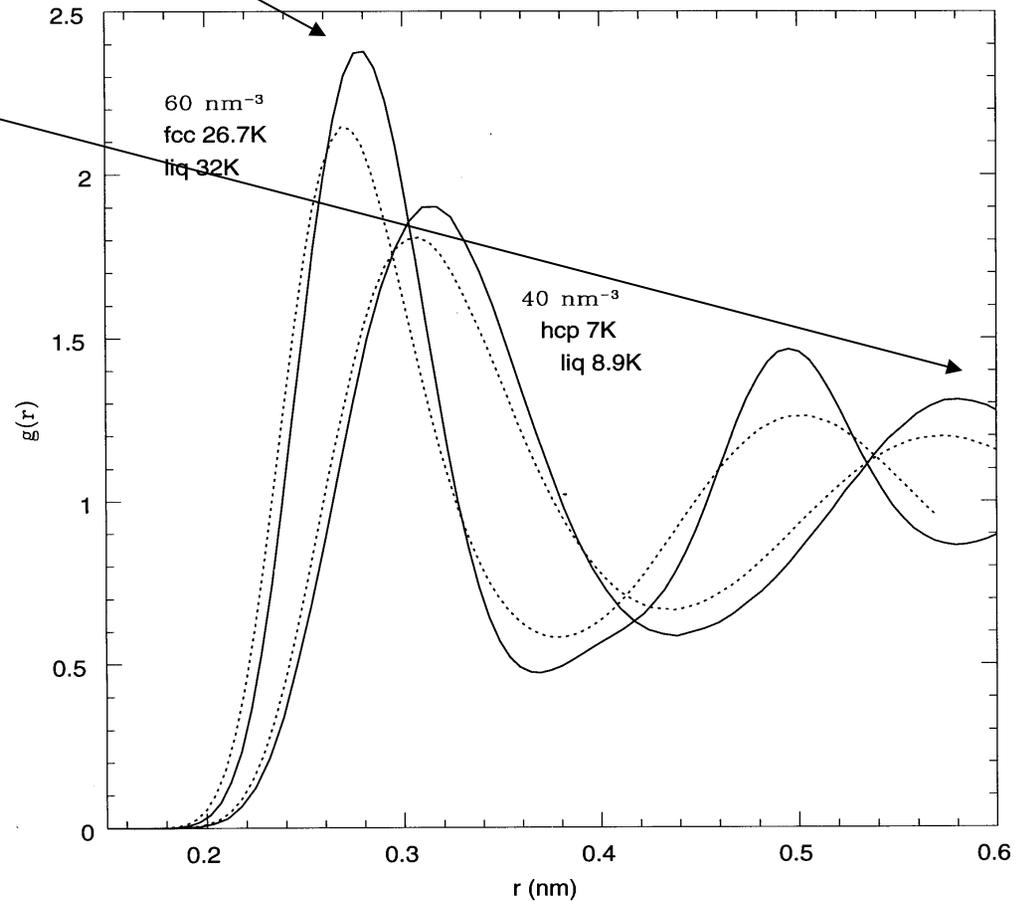
- Order parameter in solid is ρ_G where G is a particular wavevector (reciprocal lattice vector).

Electron gas $g(r)$ and $S(K)$



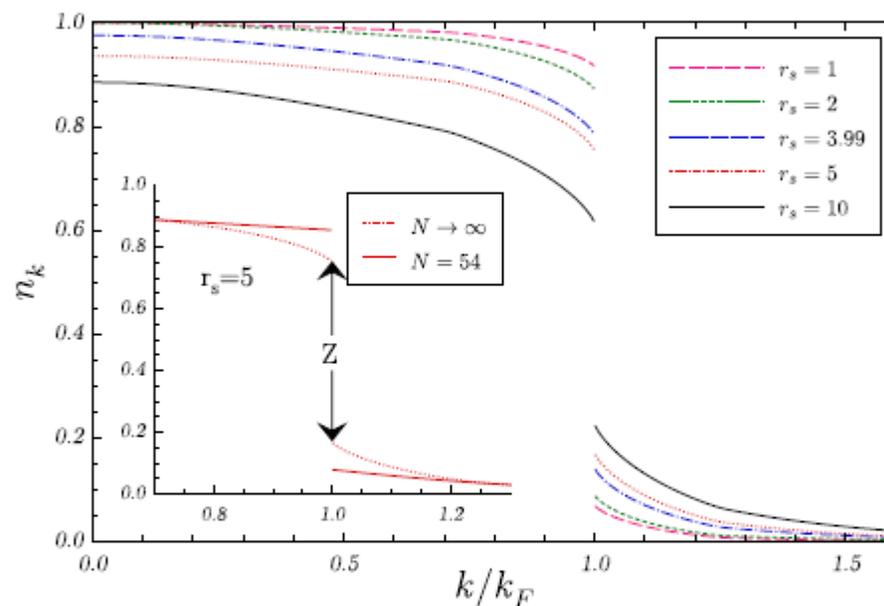
$g(r)$ in liquid and solid helium

- First peak is at inter-particle spacing. (shell around the particle)
- goes out to $r < L/2$ in periodic boundary conditions.



Momentum Distribution

- Momentum distribution
 - Classically momentum distribution is always a Gaussian
 - Non-classical showing effects of bose or fermi statistics
 - Fourier transform is the single particle off-diagonal density matrix
- Compute with McMillan Method.
- For fermions we need to use the determinant update formulas to find the effect of the movement of 1 electron.



$$n(r, r') = \frac{1}{Z} \int dr_2 \dots dr_N \psi^*(r, r_2 \dots) \psi(r', r_2 \dots)$$

$$= \left\langle \frac{\psi^*(r, r_2 \dots)}{\psi(r', r_2 \dots)} \right\rangle$$

Derivation of momentum formula

- Suppose we want the probability n_k that a given atom has momentum $\hbar k$.
- Find wavefunction in momentum space by FT wrt all the coordinates and integrating out all but one electron

$$\Pr(k_1, \dots, k_N) = \left| \int dR e^{-i(k_1 r_1 + \dots + k_N r_N)} \Psi(R) \right|^2$$

$$n_k = \int dk_2 \dots dk_N \Pr(k, k_2, \dots, k_N)$$

- Expanding out the square and performing the integrals we get.

$$n_k = \int \frac{d^3 r d^3 s}{(2\pi)^3 V} \exp(-ik(r-s)) n(r, s) = \int \frac{d^3 r}{(2\pi)^3} e^{-ikr} n(r)$$

Where: $n(r, s) = \frac{1}{Q} \int dr_2 \dots dr_N \psi^*(r, r_2 \dots r_N) \psi(s, r_2 \dots r_N)$

(states occupied with the Boltzmann distribution.)

For a homogeneous system, $n(r, s) = n(|r-s|)$

The electron gas

D. M. Ceperley, Phys. Rev. B 18, 3126 (1978)

$$H = -\frac{\hbar^2}{2m} \sum_i \nabla_i^2 + \sum_{i<j} \frac{1}{r_{ij}}$$

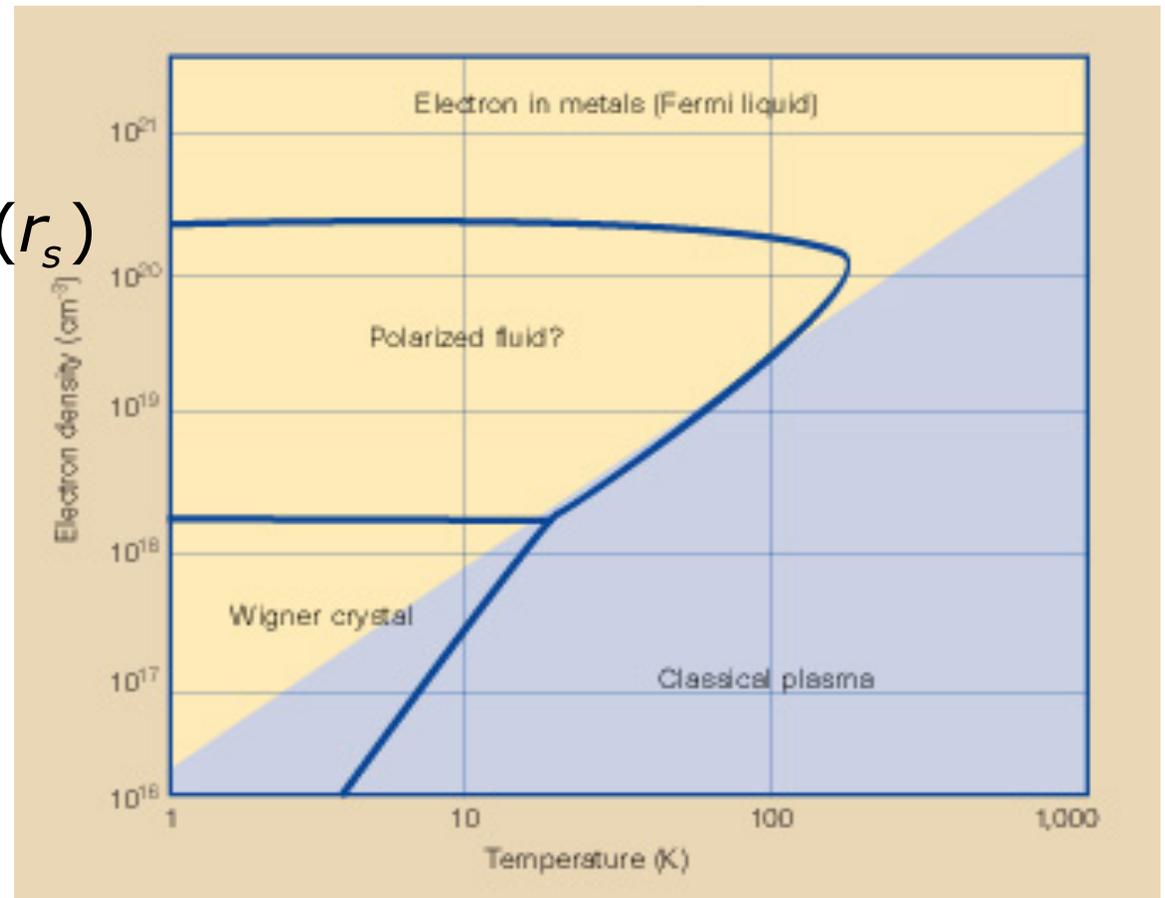
- Standard model for electrons in metals
- Basis of DFT.
- Characterized by 2 dimensionless parameters:
 - Density
 - Temperature

$$r_s = a / a_0$$

$$\Gamma = e^2 / Ta$$

- What is energy?
- When does it freeze?
- What is spin polarization?
- What are properties?

$\log(r_s)$



$\Gamma < r_s$ classical OCP

$\Gamma = 175$ classical melting

$\log(\Gamma)$

Jastrow factor for the e-gas

- Look at local energy either in r space or k-space:
- r-space as 2 electrons get close gives cusp condition: $du/dr|_0 = -1$
- K-space, charge-sloshing or plasmon modes.

$$2\rho u_k = \sqrt{\frac{V_k}{\lambda k^2}} \propto \frac{1}{k^2}$$

- Can combine 2 exact properties in the Gaskell form. Write E_V in terms structure factor making “random phase approximation.” (RPA).

$$2\rho u_k = -\frac{1}{S_k} + \sqrt{\frac{1}{S_k^2} + \frac{V_k}{\lambda k^2}} \quad S_k = \text{ideal structure factor}$$

- Optimization can hardly improve this form for the e-gas in either 2 or 3 dimensions. **RPA works better for trial function than for the energy.**
- NEED EWALD SUMS because potential trial function is long range, it also decays as $1/r$, but it is not a simple power.

$$\lim_{r \rightarrow \infty} u(r) = \begin{cases} r^{-1} & 3D \\ r^{-1/2} & 2D \\ \log(r) & 1D \end{cases}$$

Long range properties important

- Give rise to dielectric properties
- Energy is insensitive to u_k at small k
- Those modes converge $t \sim 1/k^2$

Wavefunctions beyond Jastrow

- Use method of residuals construct a sequence of increasingly better trial wave functions. Justify from the Importance sampled DMC.
- Zeroth order is Hartree-Fock wavefunction
- First order is Slater-Jastrow pair wavefunction (RPA for electrons gives an analytic formula)
- Second order is **3-body backflow** wavefunction
- Three-body form is like a squared force. It is a bosonic term that does not change the nodes.

$$\exp\left\{\sum_i \left[\sum_j \xi_{ij}(r_{ij})(\mathbf{r}_i - \mathbf{r}_j)\right]^2\right\}$$

$$\phi_{n+1}(R) \approx \phi_n(R) e^{-\tau \langle \phi_n^{-1} H \phi_n \rangle}$$

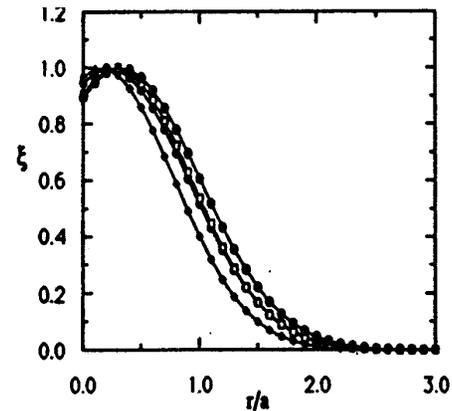
$$\phi_0 = e^{i \sum_j \mathbf{k}_j \cdot \mathbf{r}_j}$$

$$E_0 = V(R)$$

$$\phi_1 = \phi_0 e^{-U(R)}$$

$$E_1 = U(R) - \left[\nabla W(R) \right]^2 + i \sum_j \mathbf{k}_j \cdot \left(\mathbf{r}_j - \nabla_j Y(R) \right)$$

smoothing



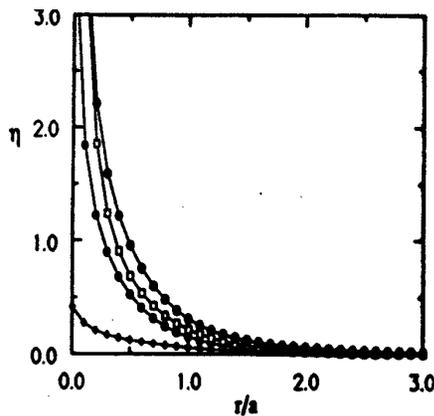
Backflow wave function

- Backflow means change the coordinates to quasi- coordinates.
- Leads to a much improved energy and to improvement in nodal surfaces. Couples nodal surfaces together.

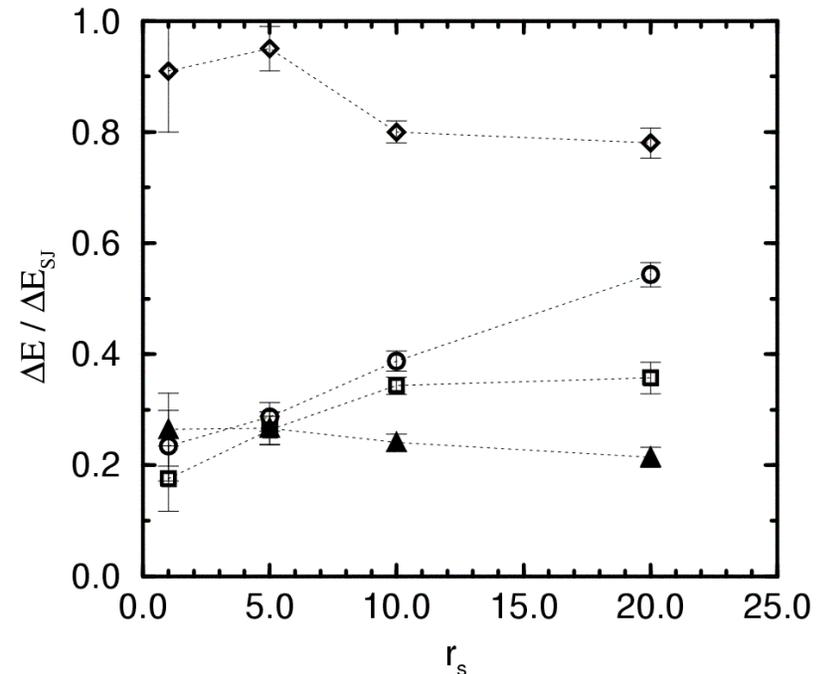
$$\text{Det}\{e^{i\mathbf{k}_i \mathbf{r}_j}\} \Rightarrow \text{Det}\{e^{i\mathbf{k}_i \mathbf{x}_j}\}$$

$$\mathbf{x}_i = \mathbf{r}_i + \sum_j \eta_{ij}(r_{ij})(\mathbf{r}_i - \mathbf{r}_j)$$

Kwon PRB 58, 6800 (1998).



3DEG

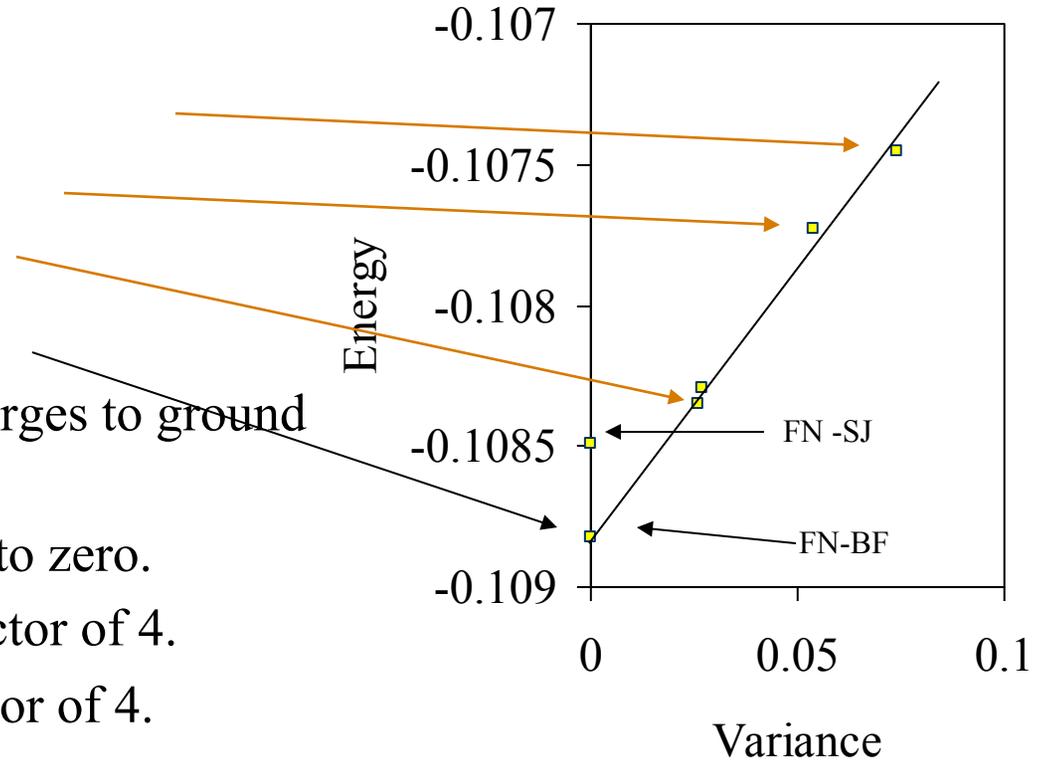


Dependence of energy on wavefunction

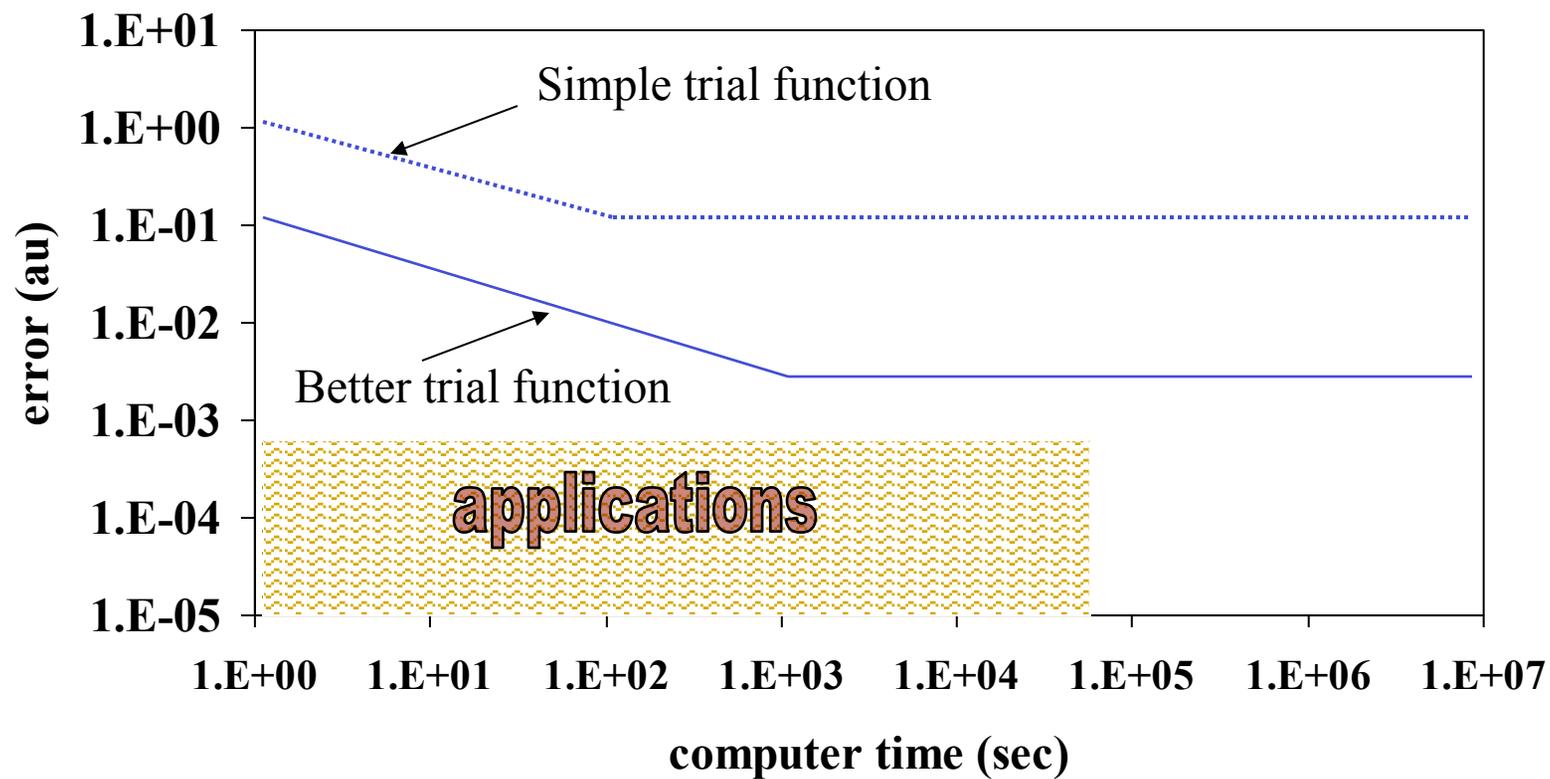
3d Electron fluid at a density $r_s=10$

Kwon, Ceperley, Martin, Phys. Rev. B58,6800, 1998

- Wavefunctions
 - Slater-Jastrow (SJ)
 - three-body (3)
 - backflow (BF)
 - fixed-node (FN)
- Energy $\langle \varphi | H | \varphi \rangle$ converges to ground state
- Variance $\langle \varphi | [H-E]^2 | \varphi \rangle$ to zero.
- Using 3B-BF gains a factor of 4.
- Using DMC gains a factor of 4.



Summary of Variational (VMC)



Summary and problems with variational methods

- Powerful method since you can use any trial function
- Scaling (computational effort vs. size) is almost classical
- Learn directly about what works in wavefunctions
- No sign problem
- Optimization is time consuming
- Energy is insensitive to order parameter
- Non-energetic properties are less accurate. $O(1)$ vs. $O(2)$ for energy.
- Difficult to find out how accurate results are.
- Favors simple states over more complicated states, e.g.
 - Solid over liquid
 - Polarized over unpolarized

What goes into the trial wave function comes out! “GIGO”

We need a more automatic method! Projector Monte Carlo