The Variational Monte Carlo Method

The theorem + restrictions

Upper and lower bounds: variance
How to do the variation.
McMillan’s calculation on liquid $^4\text{He}$
Fermion VMC
First Major QMC Calculation

- VMC calculation of ground state of liquid helium 4.
- Applied MC techniques from classical liquid theory.
- Ceperley, Chester and Kalos (1976) generalized to fermions.

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 \( \Pi_{\alpha \beta}(r_{ij}) \). The Lennard-Jones 12-6 potential is used with parameters determined from the gas data by deBoer and Michels. The configuration space integrals are performed by a Monte Carlo technique for 32 and 108 atoms in a cube with periodic boundary conditions. With \( f(r) = \exp[-(2.6 \text{ Å}/r)\,^6] \), the ground-state energy is found to be \(-0.78 \times 10^{-14} \text{ ergs/atom} \), 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
- Can be generalized to finite temperature by using "trial" density matrix instead of "trial" wavefunction.

Ceperley Variational Methods
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.

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

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

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

$$E_L(R; a) \equiv \frac{1}{\psi(R; a)}H\psi(R; a) =$"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.
Expand trial function in terms of the exact eigenfunctions:

$$\psi(R; a) = \sum_{\alpha} \phi_{\alpha}(R) \langle \alpha | \psi(a) \rangle$$

$$E_v(a) = \frac{\sum_{\alpha} E_{\alpha} |\langle \alpha | \psi(a) \rangle|^2}{\sum_{\alpha} |\langle \alpha | \psi(a) \rangle|^2} = \int dE \rho_a(E) E$$

$$\rho_a(E) = \sum_{\alpha} \delta(E - E_{\alpha}) \frac{|\langle \alpha | \psi(a) \rangle|^2}{\langle \psi(a) | \psi(a) \rangle} \geq 0 \quad \int_{-\infty}^{\infty} dE \rho(E) = 1$$

$$\sigma^2(a) = \int dE \rho_a(E) (E - E_v(a))^2$$

$$O(a) = \rho_a(E_0) = 1 - \int_{E_0}^{\infty} dE \rho_a(E) = \frac{|\langle 0 | \psi(a) \rangle|^2}{\langle \psi(a) | \psi(a) \rangle} = \text{overlap with ground state}$$

Energy and variance are second order in \((1-\text{overlap})\). Other properties are first order. Temple lower bound:

$$\frac{E_v - \sigma^2}{E_1 - E_0} \leq E_0 \leq E_v$$
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 $\psi$ to get the best upper bound

  \[ E_V = \frac{\int dR \langle \psi | H | \psi \rangle}{\int dR \langle \psi \psi \rangle} \geq E_0 \]

  \[ \sigma^2 = \frac{\int dR \langle \psi | H^2 | \psi \rangle}{\int dR \langle \psi \psi \rangle} - E_V^2 \]

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

  \[ E_V = \left\langle E_L(R) \right\rangle_{\psi}^2 \geq E_0 \]

  Better wavefunction, lower variance! “Zero variance” principle. (non-classical)
Liquid helium
the prototypic quantum fluid

- Interatomic potential is known more accurately than any other atom because electronic excitations are so high.
- A helium atom is an elementary particle. A weakly interacting hard sphere.

Two isotopes:
- $^3$He (fermion: antisymmetric trial function, spin 1/2)
- $^4$He (boson: symmetric trial function, spin zero)
Trial function for helium

- We want finite variance of the local energy.
- Whenever 2 atoms get close together wavefunction should vanish.
- The pseudopotential \( u(r) \) is similar to classical potential
- Local energy has the form:
  \[ G \text{ is the pseudoforce:} \]
  \[
  \psi(R) = \prod_{i<j} e^{-u(r_{ij})}
  \]
  \[
  E_{\psi}(R) = \sum_{i<j} v(r_{ij}) - 2\lambda \nabla^2 u(r_{ij}) - \lambda \sum_i G_i^2
  \]
  \[
  G_i = \sum \nabla_i u(r_{ij})
  \]

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

\[
 U(r) = \alpha r^{-m}
\]

Gives a cusp condition on \( u \).

- \[
  -e^2 r^{-1} = 2\lambda \nabla^2 u(r) = 2\lambda \left( u'' + \frac{D-1}{r} u' \right)
  \]
- \[
  u'(0) = -\frac{e^2}{2\lambda(D-1)}
  \]
- \[
  \epsilon r^{-n} = 2\lambda \left( \alpha mr^{-m-1} \right)^2 \text{ for } n > 2
  \]
  \[
  m = \frac{n}{2} - 1
  \]
  \[
  \alpha = \frac{1}{m} \sqrt{\frac{\epsilon}{2\lambda}}
  \]
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 \( a_0 \).

\[
E_v(a) = \frac{\int \psi(a) H \psi(a)}{\int |\psi(a)|^2} \sum_k w(R_i, a) E(R_i, a)
= \frac{\sum_k w(R_i, a) E(R_i, a)}{\sum_k w(R_i, a)}
\]

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

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

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

Ceperley Variational Methods
Other quantum properties

- Kinetic energy
- Potential energy
- Pair correlation function
- Structure function
- Pressure (virial relation)

Like properties from classical simulations

No upper bound property
Only first order in accuracy

- Momentum distribution
  - Non-classical showing effects of bose or fermi statistics
  - Fourier transform is the single particle off-diagonal density matrix
- Compute with McMillan Method.
- Condensate fraction ~10%

\[
n(r, r') = \int dr_2...dr_N \psi^*(r, r_2...)\psi(r', r_2...)
\]

\[
= \left\langle \frac{\psi^*(r', r_2,...)}{\psi(r, r_2,...)} \right\rangle
\]
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 pseudopotential. 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 $O(N^2)$.

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

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

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

*Slater-Jastrow trial function.*

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

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

Ceperley Variational Methods 17
The electron gas

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

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

\[
\Gamma = \frac{e^2}{Ta}
\]

\[
r_s = \frac{a}{a_0}
\]

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

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

\[
\log(r_s)
\]

\[
\Gamma < r_s \quad \text{classical OCP}
\]

\[
\Gamma = 175 \quad \text{classical melting}
\]

\[
\log(\Gamma)
\]
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
- 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).

\[ 2 \rho u_k = -\frac{1}{S_k} + \sqrt{\frac{1}{S_k^2} + \frac{V_k}{\lambda k^2}} \]

- 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.

Long range properties important
- Gives rise to dielectric properties
- Energy is insensitive to \( u_k \)

Long range properties important
- Gives rise to dielectric properties
- Energy is insensitive to \( u_k \)
Comparison of Trial functions

• What do we choose for the trial function in VMC and DMC?
• Slater-Jastrow (SJ) with plane wave orbitals:
  \[ \Psi_2(R) = \text{Det}\{\phi_k(r_j)\}e^{-\sum_{i<j}u_{ij}(r_{ij})} \]

• For higher accuracy we need to go beyond this form.
• Include backflow-three body.

Example of incorrect physics within SJ
Wavefunctions beyond Jastrow

\[ \phi_{n+1}(R) \approx \phi_n(R)e^{-\tau<\phi_n^{-1}H\phi_n>} \]

• Use method of residuals construct a sequence of increasingly better trial wave functions.
• 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.

\[ \Psi_2(R) \exp\{ \sum_i \sum_j [\xi_{ij}(r_{ij})(r_i - r_j)]^2 \} \]
Backflow- 3B Wave functions

- Backflow means change the coordinates to quasi-coordinates.

\[ \text{Det}\{e^{ik_ir_j}\} \Rightarrow \text{Det}\{e^{ik_ix_j}\} \]

\[ x_i = r_i + \sum_j \eta_{ij} (r_{ij})(r_i - r_j) \]

- Leads to a much improved energy

\[ \text{Kwon PRB 58, 6800 (1998).} \]

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 $\langle \psi | H | \psi \rangle$ converges to ground state
- Variance $\langle \psi [H-E]^2 \psi \rangle$ to zero.
- Using 3B-BF gains a factor of 4.
- Using DMC gains a factor of 4.
Wigner Crystal Trial Function

• Jastrow trial function does not “freeze” at appropriate density.
• Solution is to break spatial symmetry “by hand.”
• Introduce a bcc lattice \{Z_i\}
• bcc has the lowest Madelung energy, but others may have lower zero point energy.
• Introduce localized one-body terms (Wannier functions).
• Non-symmetric but provides a very good description of a quantum crystal.
• More complicated trial functions and methods are also possible.

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

\[
\psi(R) = \prod_i \phi(r_i - Z_i) \prod_{i<j} e^{-u(r_{ij})}
\]

\[
\phi(r) = e^{-Cr^2}
\]

“C” is a variational parameter to be optimized.
Twist averaged boundary conditions

- In periodic boundary conditions ($\Gamma$ point), the wavefunction is periodic → **Large finite size effects for metals because of shell effects.**
- Fermi liquid theory can be used to correct the properties.
- In twist averaged BC we use an arbitrary phase $\theta$ as $r \rightarrow r+L$
- If one integrates over all phases the momentum distribution changes from a lattice of k-vectors to a fermi sea.
- Smaller finite size effects

\[
\varphi = e^{ikr} \\
kL = 2\pi n + \theta
\]

\[
\Psi(x + L) = e^{i\theta} \Psi(x)
\]

\[
\bar{A} = \frac{1}{(2\pi)^3} \int_{-\pi}^{\pi} \! \! \! \! \! \! \! \! d^3\theta \langle \Psi_\theta A \Psi_\theta \rangle
\]

Error with PBC
Error with TABC
Twist averaged MC

- Make twist vector dynamical by changing during the random walk. 
  \[ -\pi < \theta_i \leq \pi \quad i = (1, 2, 3) \]

- Within GCE, change the number of electrons
- Within TA-VMC
  - Initialize twist vector.
  - Run usual VMC (with warmup)
  - Resample twist angle within cube
  - (iterate)
- Or do in parallel.
• Momentum distribution
  – Non-classical showing effects of bose or fermi statistics
  – Fourier transform is the single particle off-diagonal density matrix
• Compute with McMillan Method.

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

\[
= \left< \frac{\psi^*(r, r_2...)}{\psi(r', r_2...)} \right>
\]
Single particle size effects

- Exact single particle properties with TA within HF
- Implies momentum distribution is a continuous curve with a sharp feature at $k_F$.
- With PBC only 5 points on curve

\[ r_s = 4 \quad N = 33 \text{ polarized} \]

- No size effect within single particle theory!
- Kinetic energy will have much smaller size effects.

\[ T = \int d^3k \frac{\hbar^2}{2m} k^2 n(k) \]
Potential energy

- Write potential as integral over structure function:
  \[ V = \int d^3 k \frac{4\pi}{k^2} S(k) \quad S(k) = \langle \rho_{-k} \rho_k \rangle = 1 + (N - 1) \langle e^{i(r_1 - r_2)k} \rangle \]

- Error comes from 2 effects.
  - Approximating integral by sum
  - Finite size effects in S(k) at a given k.
- Within HF we get exact S(k) with TABC.
- Discretization errors come only from non-analytic points of S(k).
  - The absence of the k=0 term in sum. We can put it in by hand since we know the limit S(k) at small k (plasmon regime)
  - Remaining size effects are smaller, coming from the non-analytic behavior of S(k) at 2k_F.

\[ S(k) = \frac{k^2}{2m\omega_p} \Rightarrow \Delta E = \frac{\omega_p}{N} \]
Summary of T=0 methods:

Variational (VMC), Fixed-node (FN), Released-node (RN)

![Graph showing error vs. computer time for Simple trial function and Better trial function]