Abstract

The multiparticle Schrödinger equation is the basic governing equation in Quantum Mechanics. Its solution, called a wavefunction, is a function of many variables and is constrained to be antisymmetric under exchange of these variables. I will show how to use antisymmetric inner products to incorporate the antisymmetry condition into a representation of the wavefunction as a sum of separable functions. We will then consider how to incorporate the potential operators into such inner products, and finally construct a Green's function iteration to produce an approximate wavefunction.

Approximating the Wavefunction of the Multiparticle Schrödinger equation.

Martin J. Mohlenkamp

Department of Mathematics

OHIO UNIVERSITY

Goals

For people working in Quantum Mechanics:

Present a new strategy for approximating the wavefunction.

For others:

Give a case study on using sums of separable functions.

Outline

- The multiparticle Schrödinger equation
- An iteration based on the Green's function.
- Sums of Slater determinants
  - Unconstrained!!!
- Algorithm overview
- Selected algorithm details

What are we talking about?

The N-particle electronic Schrödinger equation is the basic governing equation in Quantum Mechanics.

It serves as the foundation for chemistry, physics, and the physical world as we understand it.

It misses some things, such as relativity, but we won't worry about those today.

Functions and Variables

The wavefunction for one electron has one 3D spatial variable $\mathbf{r} = (x, y, z)$ and one discrete spin variable $\sigma \in \{-\frac{1}{2}, \frac{1}{2}\}$.
We combine these into $\gamma = (\mathbf{r}, \sigma)$.

A system with N electrons does **not** correspond to a set of functions $\{\psi_1(\gamma), \ldots, \psi_N(\gamma)\}$ but rather a single function

$$\psi(\gamma_1, \gamma_2, \ldots, \gamma_N).$$

Don't even think about using a regular grid.

Orbitals are a (useful) myth.

The Kinetic Energy Operator

The kinetic energy operator is

$$\mathcal{T} = -\frac{1}{2}\sum_{i=1}^{N} \Delta_i$$

where the Laplacian in $\mathbf{r}_i$ is

$$\Delta_i = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}\,.$$

The Nuclear Potential Operator

A nucleus at position $\mathbf{R}_a$ with charge z_a creates a potential

$$\frac{-z_a}{\|\mathbf{r}-\mathbf{R}_a\|},$$

which we add over all nuclei to give the potential $V(\mathbf{r})$.

The nuclear potential operator is the multiplication operator

$$\mathcal{V} = \sum_{i=1}^{N} V(\mathbf{r}_i)\,.$$

The Electron-Electron Interaction Operator

Two electrons affect each other by the potential

$$\frac{1}{\|\mathbf{r}_1 - \mathbf{r}_2\|}\,.$$

The electron-electron interaction operator is the multiplication operator

$$\mathcal{W} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}\frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|}$$

Find the Low(est) Eigenvalues to get Energies

The Hamiltonian is the sum of these three:

$$\mathcal{H} = \mathcal{T} + \mathcal{V} + \mathcal{W}\,.$$

The basic goal is to find the most negative (discrete) eigenvalues

$$\mathcal{H}\psi = \lambda\psi\,.$$

These eigenvalues are the possible energies of the system. Other operators applied to the eigenfunctions give other properties of the system.

. . . subject to an Antisymmetry Constraint

Electrons are (observed to be) fermions, which means physical wavefunctions must be antisymmetric under exchange of coordinates: e.g.

$$\psi(\gamma_1, \gamma_2, \ldots, \gamma_N) = -\psi(\gamma_2, \gamma_1, \ldots, \gamma_N).$$

We can (in theory) project a function onto its antisymmetric part using the *antisymmetrizer*

$$\mathcal{A} = \frac{1}{N!} \sum_{p \in S_N} (-1)^p \mathcal{P},$$

where S_N is the permutation group on N elements.

Most eigenfunctions project to zero, and those eigenvalues must be discarded.

A Green's Function Iteration

$$\begin{aligned}(\mathcal{T}+\mathcal{V}+\mathcal{W})\psi = \lambda\psi &\Leftrightarrow (\mathcal{T}-\lambda\mathcal{I})\psi = -(\mathcal{V}+\mathcal{W})\psi \\ &\Leftrightarrow \psi = -(\mathcal{T}-\lambda\mathcal{I})^{-1}(\mathcal{V}+\mathcal{W})\psi\end{aligned}$$

Define the Green's function

$$\mathcal{G}_\mu = (\mathcal{T}-\mu\mathcal{I})^{-1}\,,$$

for $\mu < 0$. The Green's function iteration (*a la* Kalos) is

$$\begin{aligned}\tilde{\psi}_n &= -\mathcal{G}_{\mu_n}[(\mathcal{V}+\mathcal{W})\psi_n] \\ \mu_{n+1} &= \mu_n - \langle(\mathcal{V}+\mathcal{W})\psi_n, \psi_n - \tilde{\psi}_n\rangle/\|\tilde{\psi}_n\|^2 \\ \psi_{n+1} &= \tilde{\psi}_n/\|\tilde{\psi}_n\|\end{aligned}$$

and has

$$\mu_n \longrightarrow \lambda\,.$$

An Unconstrained Sum of Slater Determinants

We consider approximations to the wavefunction of the form

$$\psi_{(r)} = \mathcal{A} \sum_{l=1}^{r} s_l \prod_{i=1}^{N} \phi_i^l(\gamma_i) = \frac{1}{N!} \sum_{l=1}^{r} s_l \begin{vmatrix} \phi_1^l(\gamma_1) & \phi_1^l(\gamma_2) & \cdots & \phi_1^l(\gamma_N) \\ \phi_2^l(\gamma_1) & \phi_2^l(\gamma_2) & \cdots & \phi_2^l(\gamma_N) \\ \vdots & \vdots & & \vdots \\ \phi_N^l(\gamma_1) & \phi_N^l(\gamma_2) & \cdots & \phi_N^l(\gamma_N) \end{vmatrix} .$$

We impose no constraints (such as orthogonality) on the “orbitals” $\phi_i^l(\gamma_i)$, and thus seek the best approximation of this form.

Do Constraints Really Hurt You?

Without constraints, we can have

$$\psi = \mathcal{A} \prod_{i=1}^{N} \phi_i(\gamma_i) + \mathcal{A} \prod_{i=1}^{N} \left(\phi_i(\gamma_i) + \phi_{i+N}(\gamma_i)\right)$$

where $\{\phi_j\}_{j=1}^{2N}$ form an orthonormal set.

To represent the same function while constraining the factors to come from on orthogonal set would force you to multiply out the second term and thus obtain 2^N terms.

A Modified Green's Function Iteration

Since

$$\tilde{\psi}_n = -\mathcal{G}_{\mu_n}[(\mathcal{V}+\mathcal{W})\psi_n]$$

does not preserve the form $\psi_{(r)}$, we define $\tilde{\psi}_n$ to be the function of the correct form that minimizes

$$\|\tilde{\psi}_n - (-\mathcal{G}_{\mu_n}[(\mathcal{V}+\mathcal{W})\psi_n])\|.$$

In order to assure convergence to an antisymmetric solution, we use the pseudo-norm induced by the pseudo inner product

$$\langle \cdot, \cdot \rangle_{\mathcal{A}} = \langle \mathcal{A}(\cdot), \mathcal{A}(\cdot) \rangle.$$

This makes non-antisymmetric subspaces "invisible".

Alternating Least Squares Fitting

To find $\tilde{\psi}_n$ we loop through the directions, optimizing the functions in that coordinate while fixing the functions in other coordinates.

For each direction the minimization is a linear problem, solving the normal equations

$$\mathbb{A}\mathbf{x} = \mathbf{b}\,.$$

Rather than optimize the coefficients in some basis, we optimize the point values, resulting in a linear system of integral equations, which we solve with the Conjugate Gradient iterative method.

Refresher on the Usual Linear Least Squares Fitting

To find the coefficients $\{c_i\}$ to minimize

$$\left\| f - \sum_i c_i g_i \right\|^2 = \left\langle f - \sum_i c_i g_i , f - \sum_i c_i g_i \right\rangle ,$$

take the gradient with respect to $\{c_i\}$ and set it equal to 0 to obtain the normal equations

$$\mathbb{A}\mathbf{x} = \mathbf{b} ,$$

with

$$A(k,i) = \langle g_k, g_i \rangle \quad \text{and} \quad b(k) = \langle g_k, f \rangle .$$

As long as $\{g_i\}$ is linearly independent, the system has a unique solution $\mathbf{x}$, and one has $c_i = x(i)$.

Construction of the Integral System

For direction $i = 1$, the kernels in $\mathbb{A}$ are defined by

$$A(l,l')(\gamma,\gamma') = \left\langle \delta(\gamma' - \gamma_1) s_{l'} \prod_{i=2}^{N} \tilde{\phi}_i^{l'}(\gamma_i), \delta(\gamma - \gamma_1) s_l \prod_{i=2}^{N} \tilde{\phi}_i^{l}(\gamma_i) \right\rangle_{\mathcal{A}} ,$$

where $\delta(\cdot)$ is the delta function.

The functions in $\mathbf{b}$ are defined by

$$b(l)(\gamma) = \left\langle \delta(\gamma - \gamma_1) s_l \prod_{i=2}^{N} \tilde{\phi}_i^{l}(\gamma_i), -\mathcal{G}_\mu[\mathcal{V} + \mathcal{W}] \sum_{m=1}^{r} s_m \prod_{i=1}^{N} \phi_i^{m}(\gamma_i) \right\rangle_{\mathcal{A}} .$$

The work is in actually constructing these objects.

Antisymmetric Inner Products

Define the matrix $\mathbb{L} = \mathbb{L}(\prod \tilde{\phi}_i, \prod \phi_i)$ with entries

$$L(i,j) = \langle \tilde{\phi}_i, \phi_j \rangle$$

A simple derivation gives us Löwdin's rule, that

$$\left\langle \prod \tilde{\phi}_i, \prod \phi_i \right\rangle_{\mathcal{A}} = \frac{|\mathbb{L}|}{N!}\,.$$

Entries in $\mathbb{A}$ are computed via Löwdin's rule and some trickery with the delta functions. Entries in $\mathbf{b}$ will also use this, but need additional formulae.

Handling the Green's function in $\mathbf{b}$

Ignoring some trickery with the delta functions, to compute entries in $\mathbf{b}$ we need need to compute things of the form

$$\left\langle \prod \tilde{\phi}_i, \mathcal{G}_\mu \left[\mathcal{V} + \mathcal{W}\right] \prod \phi_i \right\rangle_{\mathcal{A}} .$$

$\mathcal{G}_\mu$ is self-adjoint, so we can convert to

$$\left\langle \mathcal{G}_\mu \prod \tilde{\phi}_i, \left[\mathcal{V} + \mathcal{W}\right] \prod \phi_i \right\rangle_{\mathcal{A}} .$$

It appears that the Green's function couples all directions, and so does not allow separation of variables to be used.

Expansion of the Green's Function

We will approximate

$$\mathcal{G}_\mu = (\mathcal{T} - \mu\mathcal{I})^{-1} \approx \sum_{p=1}^{L} \bigotimes_{i=1}^{N} \mathcal{F}_{\mathbf{r}_i}^{\mu,p} ,$$

where $\mathcal{F}_{\mathbf{r}_i}^{\mu,p}$ are a set of convolution operators, so that

$$\begin{aligned} \left\langle \mathcal{G}_\mu \prod \tilde{\phi}_i, [\mathcal{V}+\mathcal{W}] \prod \phi_i \right\rangle_{\mathcal{A}} &\approx \left\langle \sum_{p=1}^{L} \bigotimes_{i=1}^{N} \mathcal{F}_{\mathbf{r}_i}^{\mu,p} \prod \tilde{\phi}_i, [\mathcal{V}+\mathcal{W}] \prod \phi_i \right\rangle_{\mathcal{A}} \\ &= \sum_{p=1}^{L} \left\langle \prod (\mathcal{F}^{\mu,p} \tilde{\phi}_i), [\mathcal{V}+\mathcal{W}] \prod \phi_i \right\rangle_{\mathcal{A}} . \end{aligned}$$

Since $\mathcal{F}^{\mu,p}\tilde{\phi}_i$ is just some other function, we have handled the Green's function, with a multiplicative cost L.

How Good is this Approximation?

Theorem: For any $\epsilon > 0$, $\mu < 0$, and N, the N-particle Green's function $\mathcal{G}_\mu$ has a separated representation with relative error in $L^2 \to L^2$ operator norm bounded by ϵ using $L = \mathcal{O}((\ln \epsilon)^2)$ terms, with L independent of μ and N.

Thus this representation can be used for large N.

The Construction of the Expansion of the Green's Function

Step 1: approximate $1/t$: Given the desired ϵ, find positive w_p and τ_p such that

$$\left| \frac{1}{t} - \sum_{p=1}^{L} w_p \exp(-\tau_p t) \right| < \epsilon \,,$$

on the interval $t \in [1, \infty)$. Translating results of Braess and Hackbusch (2005), we can use $L = \mathcal{O}((\ln \epsilon)^2)$.

Step 2: stretch the interval: Substituting $t = s/(-\mu)$ for $\mu < 0$, and then dividing by $-\mu$, on $[-\mu, \infty)$ we get

$$\left| \frac{1}{s} - \sum_{p=1}^{L} \frac{w_p}{-\mu} \exp\left(-\frac{\tau_p}{-\mu} s \right) \right| < \frac{\epsilon}{-\mu} \,.$$

Step 3: approximate $\mathcal{G}_\mu$ in Fourier:

$$\mathcal{G}_\mu = (\mathcal{T} - \mu)^{-1} = \frac{1}{2\pi^2 \sum_i \xi_i^2 - \mu},$$

so we can substitute into the $1/s$ approximation and obtain

$$\left| \mathcal{G}_\mu - \sum_{p=1}^{L} \frac{w_p}{-\mu} e^{-\tau_p} \prod_{i=1}^{N} \exp\left(-\frac{2\pi^2 \tau_p}{-\mu} \xi_i^2 \right) \right| < \frac{\epsilon}{-\mu} = \epsilon \|\mathcal{G}_\mu\| .$$

Step 4: convert to space: $\mathcal{G}_\mu \approx \sum_{p=1}^{L} \bigotimes_{i=1}^{N} \mathcal{F}_{\mathbf{r}_i}^{\mu,p}$ with

$$\mathcal{F}_{\mathbf{r}_i}^{\mu,p} f(\gamma_1, \ldots, \gamma_N) = \left(\frac{w_p}{-\mu e^{\tau_p}} \right)^{1/N} \left(\frac{-\mu}{2\pi\tau_p} \right)^{3/2} \times$$

$$\int \exp\left(-\frac{-\mu}{2\tau_p} \|\mathbf{r}_i - \mathbf{r}\|^2 \right) f(\gamma_1, \ldots, \gamma_{i-1}, (\mathbf{r}, \sigma_i), \gamma_{i+1}, \ldots, \gamma_N) d\mathbf{r} .$$

Antisymmetric Inner Products with Operators

To compute entries in $\mathbf{b}$, we still need formulae to compute

$$\left\langle \prod \tilde{\phi}_i, \mathcal{W} \prod \phi_i \right\rangle_{\mathcal{A}} \quad \text{and} \quad \left\langle \prod \tilde{\phi}_i, \mathcal{V} \prod \phi_i \right\rangle_{\mathcal{A}} .$$

Define

$$\mathcal{W}\,[f](\mathbf{r}) = \int \frac{1}{\|\mathbf{r}-\mathbf{r}'\|} f(\gamma')d\gamma'$$

and let

$$\mathbf{\Phi} = [\phi_1, \phi_2, \ldots, \phi_d]^*$$

denote the column vector containing the functions in the product $\prod_{i=1}^{N} \phi_i(\gamma_i)$.
(Sometimes $\mathbf{\Phi}$ is just shorthand for the product.)

Maximum Coincidence

Using the matrix $\mathbb{L} = \mathbb{L}(\tilde{\mathbf{\Phi}}, \mathbf{\Phi})$, define

$$\mathbf{\Theta} = \mathbb{L}^{-1} \tilde{\mathbf{\Phi}} \, .$$

The antisymmetrization only differs by a constant:

$$|\mathbb{L}| \mathcal{A} \mathbf{\Theta} = \mathcal{A} \tilde{\mathbf{\Phi}} \, ,$$

but the functions in $\mathbf{\Theta}$ are aligned with those in $\mathbf{\Phi}$, in the sense that

$$\mathbb{L}(\mathbf{\Theta}, \mathbf{\Phi}) = \mathbb{I} \, ,$$

or equivalently

$$\langle \theta_i, \phi_j \rangle = \delta_{ij} \, .$$

Determinant of a Low-Rank Perturbation of the Identity

Let $\{\mathbf{u}_q\}_{q=1}^{Q}$ and $\{\mathbf{v}_q\}_{q=1}^{Q}$ be two sets of vectors. Then

$$\left|\mathbb{I} + \sum_{q=1}^{Q} \mathbf{u}_q \mathbf{v}_q^*\right| = \begin{vmatrix} 1 + \mathbf{v}_1^*\mathbf{u}_1 & \mathbf{v}_1^*\mathbf{u}_2 & \cdots & \mathbf{v}_1^*\mathbf{u}_Q \\ \mathbf{v}_2^*\mathbf{u}_1 & 1 + \mathbf{v}_2^*\mathbf{u}_2 & \cdots & \mathbf{v}_2^*\mathbf{u}_Q \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_Q^*\mathbf{u}_1 & \mathbf{v}_Q^*\mathbf{u}_2 & \cdots & 1 + \mathbf{v}_Q^*\mathbf{u}_Q \end{vmatrix}.$$

This lemma can be proven using Schur complements.

Inner Product with $\mathcal{W}$:

$$2\frac{N!}{|\mathbb{L}|}\left\langle \tilde{\mathbf{\Phi}}, \mathcal{W}\mathbf{\Phi}\right\rangle_{\mathcal{A}} = 2N!\,\langle\mathbf{\Theta},\mathcal{W}\mathbf{\Phi}\rangle_{\mathcal{A}} =$$

$$\int\left(\sum_{i\neq j}\frac{1}{\|\mathbf{r}_i-\mathbf{r}_j\|}\right)\prod_{j=1}^{N} \overline{\phi}_j(\gamma_j)\begin{vmatrix}\theta_1(\gamma_1)&\cdots&\theta_1(\gamma_N)\\ \vdots&\ddots&\vdots\\ \theta_N(\gamma_1)&\cdots&\theta_N(\gamma_N)\end{vmatrix} d\gamma_1\cdots d\gamma_N =$$

$$\sum_{i\neq j}\int\frac{1}{\|\mathbf{r}-\mathbf{r}'\|}\left|\mathbb{I}+\left(\overline{\phi}_i(\gamma)\mathbf{\Theta}(\gamma)-\mathbf{e}_i\right)\mathbf{e}_i^*+\left(\overline{\phi}_j(\gamma')\mathbf{\Theta}(\gamma')-\mathbf{e}_j\right)\mathbf{e}_j^*\right| d\gamma d\gamma'$$

$$= \sum_{i\neq j}\int \frac{1}{\|\mathbf{r}-\mathbf{r}'\|}\overline{\phi}_i(\gamma)\overline{\phi}_j(\gamma')\begin{vmatrix}\theta_i(\gamma)&\theta_i(\gamma')\\ \theta_j(\gamma)&\theta_j(\gamma')\end{vmatrix} d\gamma d\gamma' =$$

$$\int\left(\sum_i \overline{\phi}_i(\gamma)\theta_i(\gamma)\right)\left[\int\frac{1}{\|\mathbf{r}-\mathbf{r}'\|}\left(\sum_j\overline{\phi}_j(\gamma')\theta_j(\gamma')\right)d\gamma'\right]d\gamma$$

$$-\int\sum_i\sum_j\overline{\phi}_i(\gamma)\theta_j(\gamma)\left[\int\frac{1}{\|\mathbf{r}-\mathbf{r}'\|}\overline{\phi}_j(\gamma')\theta_i(\gamma')d\gamma'\right]d\gamma$$

$$= \int \mathbf{\Phi}^*\mathbf{\Theta}\mathcal{W}\left[\mathbf{\Phi}^*\mathbf{\Theta}\right] - \mathbf{\Phi}^*\mathcal{W}\left[\mathbf{\Theta}\mathbf{\Phi}^*\right]\mathbf{\Theta}d\gamma\,.$$

Inner Product with $\mathcal{V}$

$$\left\langle \tilde{\mathbf{\Phi}}, \mathcal{V}\mathbf{\Phi} \right\rangle_{\mathcal{A}} = \frac{|\mathbb{L}|}{N!} \int V(\mathbf{r})\mathbf{\Phi}^{*}\mathbf{\Theta}\, d\gamma \,.$$

(left as an exercise)

Recap: Sketch of the Method

Within the Green’s function iteration:

1. Modify the iteration to a least-squares fitting problem.
2. Collapse that to a set of one-electron least-squares fitting problems.
3. Update the one-electron functions using:
   - a numerical method for operating on one-electron functions,
   - formulas involving the nuclear potential and the Poisson kernel, and
   - an expansion of the Green’s function into Gaussian convolutions.

Features of the Algorithm

The integral equations formulation means we do not have to work in a fixed basis, but can adapt as necessary. We use adaptive polynomial multiwavelets, which also allow us to compute $\mathcal{W}[\cdot]$ efficiently.

With respect to the number of terms r, the number of electrons N, and the cost M to represent a function of γ, the computational cost is

$$\mathcal{O}(r^2 N^2 (N + M \log M)).$$

For comparison, the cost to evaluate a single instance of Löwdin's rules is

$$\mathcal{O}(N^2 (N + M)).$$

Status

- The multiparticle method that I presented here is fully developed and clean.
- The numerical method for the single-electron functions is giving us a lot of trouble.
- Work is in progress to develop extensions that would be size-consistent and capture the inter-electron cusp.
- We have no idea how to prove that small r is is sufficient to give good approximations.