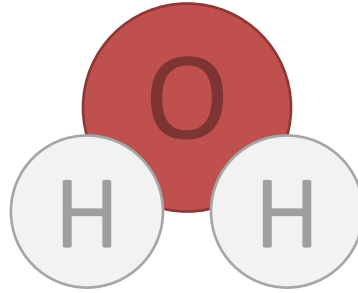


Electronic structure theory, *quickly*

Shiva Kaul (skkaul@cs.cmu.edu)

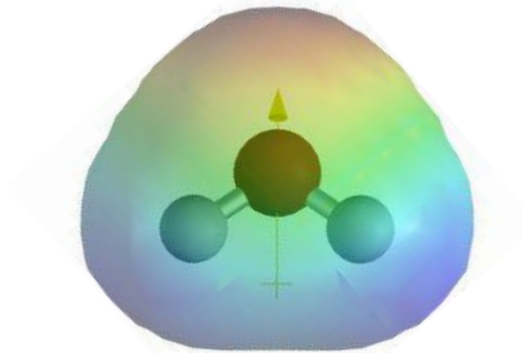
January 25th, 2011 – SELECT Lab

I have a molecule.



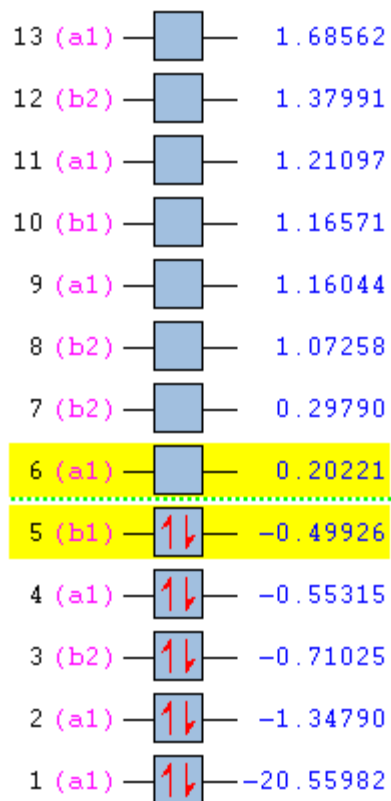
It's water. It has $n = 10$ electrons.

Where are the electrons going to go?



Knowing this is helpful for making medicines, fuels, materials, etc.

According to theory I will describe, electrons occupy **molecular orbitals**, and nature tends to assign electrons to orbitals with the **lowest energy**.



The 5 orbitals with the lowest energy are (doubly) occupied.

These numbers, though plausible, aren't experimentally observed...

...they were computed in about 3 seconds on my desktop PC.

Given:

$$\boldsymbol{\phi} = (\phi_1, \dots, \phi_m)$$

Precompute:

$$\mathbf{h} = \int \boldsymbol{\phi}^T \mathbf{h} \boldsymbol{\phi} d\mathbf{o}_1 \cdots d\mathbf{o}_m$$

$$\mathbf{S} = \int \boldsymbol{\phi}^T \boldsymbol{\phi} d\mathbf{o}_1 \cdots d\mathbf{o}_m$$

$$\mathbf{R}(r, s, t, u) = \int \int \frac{\phi_r(\mathbf{o}_1)\phi_s(\mathbf{o}_1)\phi_t(\mathbf{o}_2)\phi_u(\mathbf{o}_2)}{|\mathbf{o}_1 - \mathbf{o}_2|} d\mathbf{o}_1 d\mathbf{o}_2$$

Let:

$$\mathbf{G}_{r,s} = \sum_{t,u=1}^m (\mathbf{C}\mathbf{C}^T)_{t,u} (2\mathbf{R}(r, s, t, u) - \mathbf{R}(r, u, t, s))$$

$$\mathbf{h}^F = \mathbf{h} + \mathbf{G}$$

Solve for $\mathbf{C}, \boldsymbol{\epsilon}$ in:

$$\mathbf{h}^F \mathbf{C} = \mathbf{S} \mathbf{C} \boldsymbol{\epsilon}$$

closed-shell matrix Hartree-Fock equation

Today's goal: implement water.

Physically motivate aforementioned equation.

“Fill in the blanks” for water.

Use a simple iterative algorithm.

time-dependent Schrödinger equation

ignore relativity
focus on equilibrium



Schrödinger equation

fix nuclei



electronic Schrödinger equation

introduce antisymmetry

single determinant

determinants



single determinant approximation

solution as minimization



differential Hartree-Fock equation

finite basis set

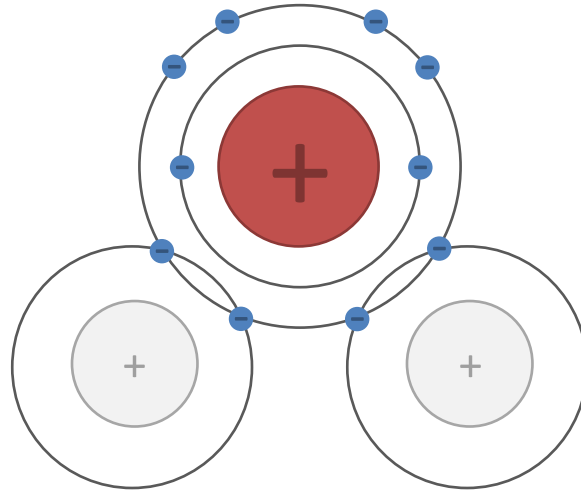
closed shell

replace differential equation by algebraic equation



closed-shell matrix Hartree-Fock equation

First we'll simplify the model without worrying about approximation quality.



Each of the n electrons has:

- position ($\mathbf{o}_i \in \mathbb{R}^3$) and spin (which we'll gloss over; watch for σ)
- low mass, so we can ignore relativistic effects
- high velocity, so we can pretend the nuclei are stationary

Pair these simplifications with Schrödinger's mechanics to get...

$$H(\Psi(\sigma_1, \dots, \sigma_n)) = E \cdot \Psi(\sigma_1, \dots, \sigma_n)$$

electronic Schrödinger equation

$E: \mathbb{R}$ is an energy (eigenvalue)

$\Psi: (\mathbb{R}^3 \rightarrow \dots \rightarrow \mathbb{R}^3) \rightarrow \mathbb{C}$ is a wavefunction (eigenfunction)

$\int_R \Psi^* \Psi dx_1 \dots dx_n$ is the probability that the electrons are in R .

H : wavefunction \rightarrow wavefunction is a Hamiltonian, a known operator representing observable aspects of the molecular setup.

$$H(o_1, \dots, o_n) = \sum_{i=1}^n h(o_i) + \sum_{i=1}^n \sum_{j<i} \frac{1}{|o_i - o_j|}$$

kinetic energy of electron

nuclei-electron attraction

electron-electron repulsion

The electrons are also unlabeled, so for any permutation operator π :

$$|\pi(\Psi)(\sigma_1, \dots, \sigma_n)|^2 = |\Psi(\sigma_1, \dots, \sigma_n)|^2$$

We want to solve for Ψ so there are two possibilities:

$$\pi(\Psi) = \Psi \text{ and } \pi(\Psi) = -\Psi$$

A permutation is just a sequence of transpositions, so consider:

$$\pi(\Psi)(\sigma_1, \sigma_2, \dots, \sigma_n) = s\Psi(\sigma_2, \sigma_1, \dots, \sigma_n)$$

Asserting (with physical motivation) that $s = -1$, we get:

$$\pi(\Psi) = (-1)^{\text{parity}(\pi)}\Psi$$

That is, the wave function needs to be **antisymmetric**.

We'd like to factorize (without yet making approximations) the n -electron Ψ into a product of 1-electron χ_i 's. But simple products aren't antisymmetric.

Completeness yields a product expansion which is antisymmetric, but the coefficients are redundant.

$$\Psi(\sigma_1, \dots, \sigma_n) = \sum_{i_1, \dots, i_n=1}^{\infty} c_{i_1, \dots, i_n} \chi_{i_1}(\sigma_1) \cdots \chi_{i_n}(\sigma_n)$$

$$\Psi(\cdot) = \sum_{K=1}^{\infty} D_K \Phi_K(\cdot) \quad \text{where} \quad \Phi_K(\sigma_1, \dots, \sigma_n) = \begin{vmatrix} \chi_{i_1}(\sigma_1) & \cdots & \chi_{i_1}(\sigma_n) \\ \vdots & \ddots & \vdots \\ \chi_{i_n}(\sigma_1) & \cdots & \chi_{i_n}(\sigma_n) \end{vmatrix}$$

is a determinant, and the χ_i are taken to be orthonormal WLOG.

Explicitly enforce antisymmetry upon product with operator:

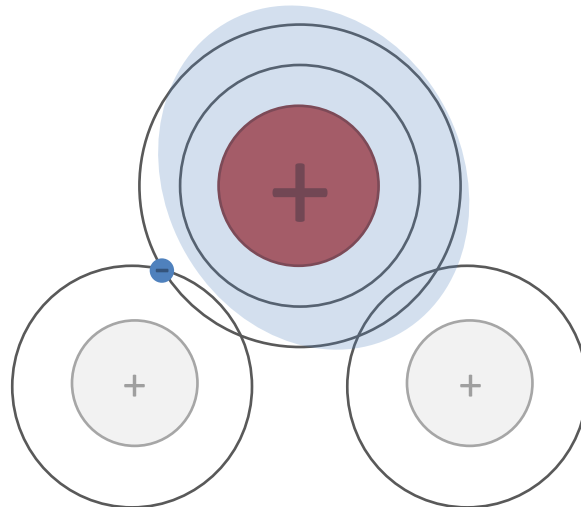
$$\Phi_K = \text{antisym}((\chi_{i_1}(\cdot_1) \cdots \chi_{i_n}(\cdot_n)))(\sigma_1, \dots, \sigma_n) \quad \text{where} \quad \text{antisym}(\cdot) = \sum_{\text{all } \pi} (-1)^{\text{parity}(\pi)} \pi(\cdot)$$

We will phrase a significant approximation as a truncation of Ψ 's expansion.

$$\Psi(\sigma) \approx \Phi(\sigma)$$

single determinant approximation

Another way to phrase it is as a simplification of H : each electron no longer individually interacts with the others; it feels only the *mean* of the others. (This is why product factorizations are sometimes called “mean-field” approximations.)



We've lost completeness; now is a good time to define approximation quality. As previously mentioned, we seek solutions $\tilde{\Psi}$ (defined by Φ and in turn by the χ_i) corresponding to low energy \tilde{E} .

$$\tilde{E} = \frac{\int \Phi^* H(\Phi) d\sigma_1 \dots d\sigma_n}{\int \Phi^* \Phi d\sigma_1 \dots d\sigma_n} \quad \text{a normalized mean}$$

$$= \sum_{i=1}^n \int \chi_i^*(\sigma) h(o) \chi_i(\sigma) d\sigma \quad \text{energy of electron } i \text{ in presence of nuclei but in absence of other electrons}$$

$$+ \frac{1}{2} \sum_{i,j=1}^n \iint \frac{\chi_i^*(\sigma_1) \chi_j^*(\sigma_2) \chi_i(\sigma_1) \chi_j(\sigma_2)}{|\sigma_1 - \sigma_2|} d\sigma_1 d\sigma_2 \quad \text{mean repulsions between electrons occupying } \chi_i \text{ and } \chi_j$$

$$- \frac{1}{2} \sum_{i,j=1}^n \iint \frac{\chi_i^*(\sigma_1) \chi_j^*(\sigma_2) \chi_j(\sigma_1) \chi_i(\sigma_2)}{|\sigma_1 - \sigma_2|} d\sigma_1 d\sigma_2 \quad \text{"exchange terms" from determinant}$$

It is easy to show that for all $\tilde{\Psi}$:

$$\tilde{E} \geq E_1$$

where E_1 is the true 'ground state' energy. Since the solution can be expressed using minimization, we can't get in trouble by performing it.

Take that expression for \tilde{E} and treat it as objective. Enforce orthonormality of χ_i with Lagrange multipliers. Use variational method (not just "take derivative, set to zero") to find stationary points. This shows that the optimal χ_i must satisfy:

$$h^F \chi_i = \epsilon_i \chi_i$$

differential Hartree-Fock equation

where h^F is similarly split up into three terms.

Now we will make a finite computational approximation which will yield matrix equations.

$$\chi_i(o) \approx \sum_{k=1}^m C_{k,i} \phi_k(o)$$

We will handcraft the basis from sums

$$\phi_k(o) = \sum_{j=1}^{b_k} d_{j,k} \eta_j(o)$$

of primitives which are Gaussians multiplied by some ‘angular momentum’ terms:

$$\eta_j(o) = N(o_1 - \gamma_1)^{a_1} (o_2 - \gamma_2)^{a_2} (o_3 - \gamma_3)^{a_3} \exp(-\alpha \|o - \gamma\|^2)$$

$N: \mathbb{R}^+$ is normalization.

$\gamma: \mathbb{R}^3$ is position of nucleus

$a_i: \mathbb{Z}^+$ describe angular momentum

$\alpha: \mathbb{R}$ contracts as usual

We’ll assume electrons occupy orbitals in pairs, which effectively halves the basis size.

Atom	Shell	Function	Width	Primitive coefficients
H	1s	ϕ_1	Narrow	0.187311D+020.214935D
				0.282539D+010.364571D
				0.640122D+000.415051D
		ϕ_2	Diffuse	0.161278D+000.181381D
H	1s	ϕ_3	Narrow	0.187311D+020.214935D
				0.282539D+010.364571D
				0.640122D+000.415051D
		ϕ_4	Diffuse	0.161278D+000.181381D
O	1s	ϕ_5	N/A	0.548467D+040.831724D+000.000000D
				0.825235D+030.153082D+010.000000D
				0.188047D+030.247715D+010.000000D
				0.529645D+020.325628D+010.000000D
				0.168976D+020.279289D+010.000000D
				0.579964D+010.954938D+000.000000D
	2s	ϕ_6	Narrow	0.155396D
				0.359993D
				0.101376D
		ϕ_7	Diffuse	0.270006D
	2px	ϕ_8	Narrow	0.155396D+02-.617934D
				0.359993D+01-.275721D
				0.101376D+010.814208D
		ϕ_9	Diffuse	0.270006D+000.266956D
	2py	ϕ_{10}	Narrow	0.155396D+02-.617934D+000.311694D
				0.359993D+01-.275721D+000.240144D
				0.101376D+010.814208D+000.105436D
		ϕ_{11}	Diffuse	0.270006D+000.266956D+000.277432D
	2pz	ϕ_{12}	Narrow	0.155396D+02-.617934D+000.311694D+010.000000D
				0.359993D+01-.275721D+000.240144D+010.000000D
				0.101376D+010.814208D+000.105436D+010.000000D
		ϕ_{13}	Diffuse	0.270006D+000.266956D+000.277432D+000.000000D

This is a basis set called 6-31G.

It encodes physical intuition, computational expedience, and deprecated ideas.

Given:

$$\boldsymbol{\phi} = (\phi_1, \dots, \phi_m)$$

Precompute:

$$\mathbf{h} = \int \boldsymbol{\phi}^T \mathbf{h} \boldsymbol{\phi} \, do_1 \cdots do_m$$

$$\mathbf{S} = \int \boldsymbol{\phi}^T \boldsymbol{\phi} \, do_1 \cdots do_m$$

m^4 of these

$$\mathbf{R}(r, s, t, u) = \int \int \frac{\phi_r(o_1) \phi_s(o_1) \phi_t(o_2) \phi_u(o_2)}{|o_1 - o_2|} \, do_1 do_2$$

Guess \mathbf{C} .

Let:

$$\mathbf{G}_{r,s} = \sum_{t,u=1}^m (\mathbf{C}\mathbf{C}^T)_{t,u} (2\mathbf{R}(r, s, t, u) - \mathbf{R}(r, u, t, s))$$

$$\mathbf{h}^F = \mathbf{h} + \mathbf{G}$$

Solve for $\mathbf{C}, \boldsymbol{\epsilon}$ in:

$$\mathbf{h}^F \mathbf{C} = \mathbf{S} \mathbf{C} \boldsymbol{\epsilon}$$

eigenvalue equation, but we need to get rid of \mathbf{S} by orthogonalizing basis functions. This isn't done at the outset because we don't want to change the basis of \mathbf{R} .

(unless converged)

- Modeling (local) interactions
- Choosing basis sets
- Evaluating integrals
- Iteratively solving matrix equations
- Oh, and the industry-standard software (used in this presentation) originated from CMU

It's a small world.