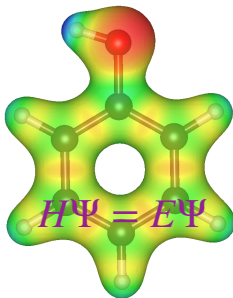


Computational chemistry methods in drug design: Quantum chemical approaches



CCNSB, IIITH, Shampa Raghunathan

MC-610, NIPER Hyderabad: Lecture 3 (27 May 2021)

Introduction to electronic structure theory I

Schrödinger equation

$$\hat{H}\psi = E\psi$$

- ▶ \hat{H} is the QM Hamiltonian of a system (an operator containing derivatives)
- ▶ E is the energy of the system
- ▶ ψ is the wavefunction of the system (contains every information we need to extract a certain property)
- ▶ $|\psi|^2$ is the probability density of finding the system over a certain region of space
- ▶ Eigenvalue-eigenfunction equation; E is scalar



Erwin
Schrödinger



Paul Dirac

One-electron system

The Hamiltonian operator \hat{H} acting upon Ψ of a system returns the system energy E ,

$$\hat{H}\Psi = E\Psi$$

can be expressed as time-independent Schrödinger equation,

$$\underbrace{\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) \right]}_{\hat{H}} \Psi(\mathbf{r}) = E\Psi(\mathbf{r})$$

can be interpreted as a system of one electron with a mass m under the influence of certain stationary potential $V(\mathbf{r})$ (which might be electrostatic potential due to the nucleus and the electron); \hbar is Planck's constant.

In Cartesian coordinates, the Laplacian operator has the form

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

Exercise 1:

The Schrödinger equation (SE) falls into a category of equations known as partial differential eigenvalue equation in which an operator acts on a function (the eigenfunction) and returns the function multiplied by a scalar (the eigenvalue).

A simple example of an eigenvalue equation is:

$$\frac{d}{dx}(y) = ry \quad (1)$$

The operator is $\frac{d}{dx}$, e.g., an eigenfunction can be $y = e^{kx}$, with the eigenvalue $r = k$

The SE involves second-order differential operator, therefore an example:

$$\frac{d^2}{dx^2}(y) = ry \quad (2)$$

with the eigenfunction $y = A e^{ikx} + B e^{-ikx}$ where A, B, k are constants.

Solve! $r = ?$ Ans: $r = -k^2$

One-dimensional SE, solve for an infinite well

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r})$$

Let's solve this equation for in 1-D:

$$V(\mathbf{x}) = \begin{cases} 0, & \text{if } 0 \leq x \leq a \\ \infty, & \text{if } x < 0, x > a \end{cases}$$

Inside well : $-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \Psi(x) = E \Psi(x)$

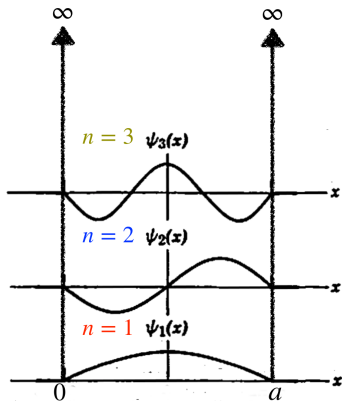
Solution : $\Psi(x) = A e^{ikx} + B e^{-ikx}$

with, $k = \frac{\sqrt{2mE}}{\hbar}$

With the boundary conditions, $\Psi(0) = \Psi(a) = 0$

Simplified Solution:

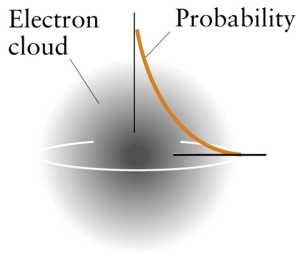
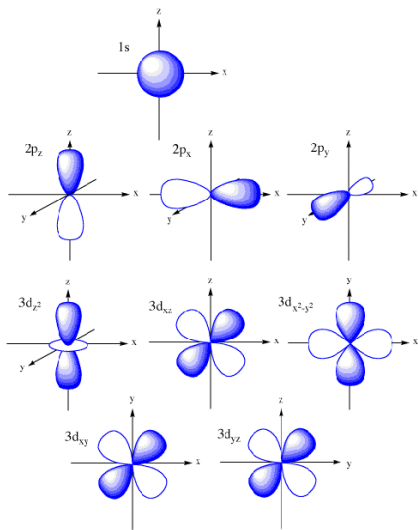
$$\Psi(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) \quad n = 1, 2, 3 \dots \quad \text{with, } E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2}$$



Related to probability of finding the electron

Constructing one-electron wavefunctions

- Atomic orbitals define electron distribution of an atom



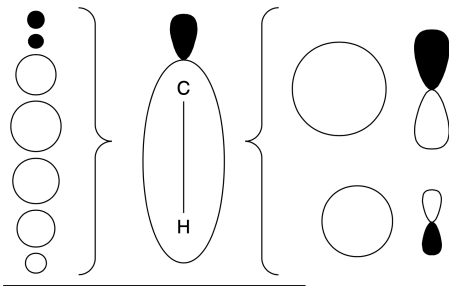
$$\Psi(r, \theta, \phi) = R(r)Y(\theta, \phi)$$

In spherical polar coordinates, the atomic wavefunction has two parts: **radial** and angular **$R(r)$** : a mathematical function varies with position and behaves like the probability of finding an electron at any point

Constructing one-electron wavefunctions

- **Linear combination of atomic orbitals (LCAO)**: To permit additional flexibility, we may take our (guess) wavefunction to be a linear combination of some set of common “basis” functions, e.g., atomic orbitals. Thus,

$$\psi_i = \sum_{i=1}^N a_i \phi_i.$$



For example, consider the wavefunction for an electron in a C–H bond. It could be represented by *s* and *p* functions on the atomic positions, or *s* functions along the bond axis, or any other convenient fashion.

Cramer, C. J. (2013). Essentials of Computational Chemistry: Theories and Models. John Wiley & Sons: Chapter 4

Molecular Hamiltonian

$$\hat{H}\Psi = E\Psi$$

$$\hat{H} = \hat{T} + \hat{V}$$

$$\hat{H} = \hat{T}_e + \hat{T}_N + \hat{V}_{e-N} + \hat{V}_{e-e} + \hat{V}_{N-N}$$

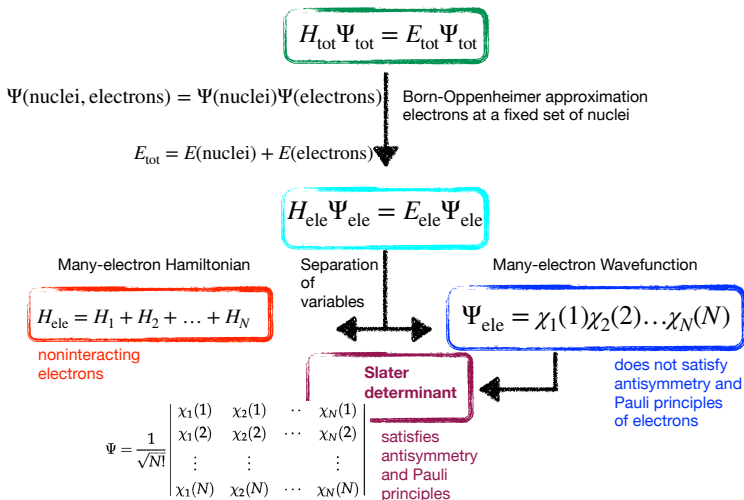
Hamiltonian describing N electrons with M nuclei system,

$$\hat{H} = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}$$

kinetic energy of the electrons; kinetic energy of the nuclei; electrostatic interaction between the electrons and the nuclei; electrostatic interaction between the electrons; electrostatic interaction between the nuclei

Approximations for solving Schrödinger equation

SE can only be solved for simple systems like, H atoms (with one electron). Therefore, approximations are made to be able to use it for molecules.



Expectation value

- ▶ For a measurable property, we can construct an operator \hat{O}
- ▶ Repeated measurements will give an average value of the operator
- ▶ The average value or expectation value of an operator can be calculated as

$$\frac{\int \Psi^* \hat{O} \Psi d\tau}{\int \Psi^* \Psi d\tau} = \langle O \rangle$$

Variational principle

The expectation value of the Hamiltonian is the variational energy:

$$\frac{\int \Psi^* \hat{H} \Psi d\tau}{\int \Psi^* \Psi d\tau} = E \geq E_{\text{exact}}$$

E which is obtained from a “trial” wavefunction will always be the upper bound to the exact (lowest) energy of the system

Born-Oppenheimer approximation

The Schrödinger equation of the total system is given by

$$\hat{\mathbf{H}}_{\text{tot}} \Psi_{\text{tot}}(\vec{r}, \vec{R}) = E_{\text{tot}} \Psi_{\text{tot}}(\vec{r}, \vec{R})$$

The total wave function Ψ_{tot} depends on both electronic \vec{r} and nuclear \vec{R} coordinates.

Uncoupling the electronic and nuclear motion is Born-Oppenheimer's approximation:

$$\hat{\mathbf{H}}_{\text{ele}} = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}}$$

Similarly, the electronic wavefunction and energy can be written as:

$$\hat{\mathbf{H}}_{\text{ele}} \Psi_{\text{ele}} = E_{\text{ele}} \Psi_{\text{ele}}, \quad \frac{\int \Psi_{\text{ele}}^* \hat{\mathbf{H}} \Psi_{\text{ele}} d\tau}{\int \Psi_{\text{ele}}^* \Psi_{\text{ele}} d\tau} = E_{\text{ele}}$$

Now, E depends on the nuclear positions through the **nuclear-electron attraction** and **nuclear-nuclear repulsion** (the last term in the total $\hat{\mathbf{H}}$)

Bra-ket notation

It is conventional to use the *bra-ket notation* for wave functions and multidimensional integrals in electronic structure theory in order to simplify the notation. The equivalences are defined as follows:

$$\begin{aligned} |\Psi\rangle &\equiv \Psi; & \langle\Psi| &\equiv \Psi^* \\ \int \Psi^* \Psi d\mathbf{r} &= \langle\Psi|\Psi\rangle \\ \int \Psi^* \mathbf{H} \Psi d\mathbf{r} &= \langle\Psi|\mathbf{H}|\Psi\rangle \end{aligned} \tag{3.2}$$

The *bra* $\langle n|$ denotes a *complex conjugate* wave function with quantum number n standing to the *left* of the operator, while the *ket* $|m\rangle$ denotes a wave function with quantum number m standing to the *right* of the operator, and the combined *bracket* denotes that the whole expression should be integrated over all coordinates. Such a bracket is often referred to as a *matrix element*, or as an *overlap element* when there is no operator involved.

Chapter 3 from Jensen, F. (2017). Introduction to Computational Chemistry. John Wiley & Sons.

Hamiltonian operator expressed into 1-electron operator

For H atom–2-body problem: one electron and one nucleus

The one-electron operator:

$$\hat{h}_1 = -\frac{1}{2}\nabla^2 - \frac{Z_A}{r_A}$$

$$\langle \chi_1(1) | \hat{h}_1 | \chi_1(1) \rangle = h_1$$

Hamiltonian operator expressed into 2-electron operator

For He atom–3-body problem: two electrons and one nucleus

Total Hamiltonian = $h_1 + h_2 + ?$ We are missing a term \hat{V}_{ee} , the repulsion between two electrons $\frac{1}{r_{12}}$, where r_{12} is the distance between two electrons. The Slater determinantal wavefunction:

$$\Psi = \begin{vmatrix} \chi_1(1) & \chi_2(1) \\ \chi_1(2) & \chi_2(2) \end{vmatrix} = \chi_1(1)\chi_2(2) - \chi_1(2)\chi_2(1)$$

electron coordinates are along ROWS

single electron wavefunctions are along COLUMNS

$$\left\langle \chi_1(1)\chi_2(2) \left| \frac{1}{r_{12}} \right| \chi_1(1)\chi_2(2) \right\rangle = J_{12} \quad \text{Coulomb integral}$$

$$\left\langle \chi_1(1)\chi_2(2) \left| \frac{1}{r_{12}} \right| \chi_2(1)\chi_1(2) \right\rangle = K_{12} \quad \text{Exchange integral}$$

Evaluation of the integral associated with the two electrons—the so-called two electron integrals (Coulomb and Exchange) cannot be exactly solved.

HF approximation

In order to derive the Hartree-Fock (HF) equations, we need an expression for the energy of a single Slater determinant.

$$E = \sum_{i=1}^N h_i + \sum_{i=1}^N \sum_{j>i}^N (J_{ij} - K_{ij})$$

HF approximation is the stepping stone on the way to solving the exact SE equation.

In the HF approximation the electron-electron interaction operator r_{ij}^{-1} is replaced by an *average* field in the form of operator \hat{V}_i^{HF} , seen by the i^{th} electron due to the presence of other electrons $\{j\}$. **Each electron is moving in the mean field of the other electrons.** The *Fock operator* \hat{F} for each i electron is defined as:

$$\hat{F}_i = \hat{h}_i + \hat{V}_i^{\text{HF}}\{j\} \quad \text{and} \quad \text{the Fock equation : } \hat{F}_i \chi_i = \epsilon_i \chi_i$$

where, ϵ is the orbital energy for each χ

HF equations (in practice): HF MOs represented with basis sets

Each molecular orbital (MO) ψ is expanded in terms of the basis functions ϕ , conventionally called atomic orbitals (MO = LCAO, linear combination of atomic orbitals),

$$\psi_i = \sum_{\mu=1}^K C_{\mu i} \phi_{\mu}.$$

Using the basis set representation of MOs Roothan-Hall formulations of HF equations for a closed-shell system, in matrix notation:

$$\begin{aligned} \mathbf{FC} &= \mathbf{SC}\epsilon \\ F_{\mu\nu} &= \left\langle \mu \left| -\frac{1}{2} \nabla^2 \right| \nu \right\rangle - \sum_k^{\text{nuclei}} Z_k \left\langle \mu \left| \frac{1}{r_k} \right| \nu \right\rangle \\ &\quad + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\lambda|\nu\sigma) \right] \end{aligned}$$

Solving these equations produce a total of M_{basis} no. of MOs, that is, if there are N_{elec} occupied and $M_{\text{basis}} - N_{\text{elec}}$ unoccupied, or virtual, MOs.

The Hartree-Fock self-consistent field method

The HF equations are solved using a self-consistent field (SCF) procedure applying the variational principle. This ensures that the determination of the best set of basis set coefficients will give the lowest energy of the system.

Basis sets: Slater type orbitals (STO)

Basis sets: mathematical functions chosen to give the maximum flexibility to the molecular orbitals AND then use LCAO
– each function is centered (has its origin) at some point in the molecule (usually on the nuclei)

► Slater type orbitals

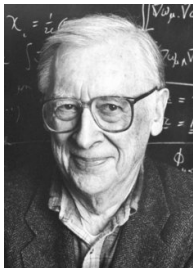
$$\phi_{1S}^{STO} = \left(\frac{\zeta^3}{\pi} \right)^{1/2} e^{-\zeta r}$$

ζ Slater orbital exponent; giving reasonable representations of atomic orbitals, near nucleus is well-represented

• Difficult to evaluate, especially when atomic orbitals are centered on different nuclei

$$(\mu\nu|\lambda\sigma) = \int \int \phi_\mu(1)\phi_\nu(1) \frac{1}{r_{12}} \phi_\lambda(2)\phi_\sigma(2) dr_1 dr_2 \quad (3)$$

Also 2-electron integrals scale as N^4



Sir John Pople

Revolutionized computational chemistry methods with a simple idea!

GAUSSIANS

Basis sets: Gaussian type orbitals (GTO)

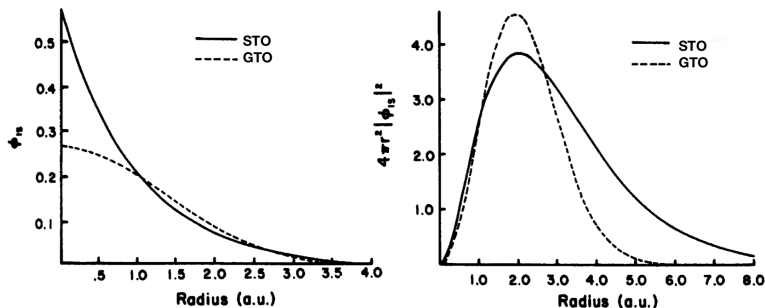
► Gaussian type orbitals

$$\phi_{1S}^{GTO} = \left(\frac{2\alpha}{\pi} \right)^{3/4} e^{-\alpha r^2}$$

α Gaussian orbital exponent

- Easy to evaluate 2-electron integrals, and product of two gaussians is a gaussian as well

STO vs. GTO



GTOs are less satisfactory than STOs in describing the AOs close to the nucleus. These two types of functions substantially differ for $r = 0$ and also, for very large values of r . **Cusp condition** At $r = 0$ STO has finite slope and GTO has a zero slope. $[d/dr e^{\zeta r}]_{r=0} \neq 0$; $[d/dr e^{\alpha r^2}]_{r=0} = 0$

Adopted from Szabo A., Ostlund N., (1996). Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory. Dover Pub. Inc., Mineola, NY.

Basis sets: Contractions

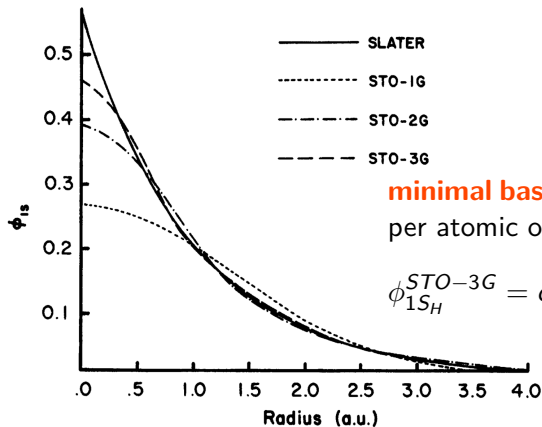
Gaussian functions are not optimum to what a functional behavior of MOs. Linear combination of the primitive gaussian functions – the so-called contractions lead to contracted Gaussian functions:

$$\phi_{1S}^{CGF} = \sum_i^N a_i \phi^{GTO}$$

N is the length of the contraction, and a is the contraction coefficient

Contracted basis sets

STO-#G



minimal basis: One basis function per atomic orbital

$$\phi_{1s}^{STO-3G} = c_1 e^{-\alpha_1 r^2} + c_2 e^{-\alpha_2 r^2} + c_3 e^{-\alpha_3 r^2}$$

Figure 3.3 Comparison of the quality of the least-squares fit of a 1s Slater function ($\zeta = 1.0$) obtained at the STO-1G, STO-2G, and STO-3G levels.














Szabo A., Ostlund N., (1996). Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory. Dover Pub. Inc., Mineola, NY.

Basis sets with multiple- ζ values



- ▶ How do you describe different bonding situations?
 - Try more than 1S orbital, in this case
- ▶ Multiple- ζ basis set
 - Multiple functions for the same atomic orbital – giving flexibility
- ▶ Only for valence orbitals
- ▶ Acronyms: single- ζ (SZ), double- ζ (DZ), triple- ζ (TZ), etc.

Pople basis sets (split-valence): 3-21G, 6-311G, etc.

	No. of basis functions	No. of primitive GTOs
3-21G		
Core 	1	3
Valence  	2	2+1
4-31G		
Core 	1	4
Valence  	2	3+1
6-31G		
Core 	1	6
Valence  	2	3+1
6-311G		
Core 	1	6
Valence   	3	3+1+1

What is the no. of basis functions for a NH_3 molecule using STO-3G & 3-21G basis sets?

STO-3G

Atoms	Core orbital	Valence orbital	No. of basis functions
H	0	1s	$1 \times 3 = 3$
N	1s	2s, 2px, 2py, 2pz	$5 \times 1 = 5$
Total			$3 + 5 = 8$

3-21G

Atoms	Core orbital	Valence orbital	No. of basis functions
H	0	1s	$2 \times 3 = 6$
N	1s	2s, 2px, 2py, 2pz	$1 + 4 \times 2 = 9$
Total			$6 + 9 = 15$

Increasing basis set sizes. . .

- ▶ Add more basis functions – allow orbitals to change *size*
- ▶ Add more polarization functions – allow orbitals to change *shape*
- ▶ Add more diffusion functions for loosely-bound electrons
- ▶ Add higher angular momentum functions

Polarization functions allow orbitals to change *shape*

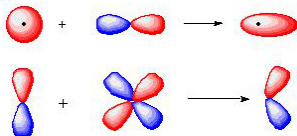
▶ Add p orbitals to H

▶ Add d orbitals to 2nd row atoms

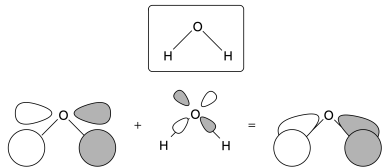
▶ Add f orbitals to transition metals

▶ 6-31G(d) – d functions per heavy atom; also known as 6-31G* (single *)

▶ 6-31G(d,p) – d functions per heavy atom and p functions for H atom; also known as 6-31G** (double **)



Ex.



Add diffusion functions to *diffuse electrons*

- ▶ Ex. Lone pairs, anions, excited states, etc.
- ▶ 6-31+G – add diffuse functions per heavy atom
- ▶ 6-31++G – add diffuse functions per heavy atom and for H atom as well

Exercise: What is the no. of basis functions for a NH_3 molecule 6-31G** basis sets?

6-31G**

Atoms	Core orbital	Valence orbital	No. of basis functions
H	0	1s	$(2+3) \times 3 = 9$
N	1s	2s, 2px, 2py, 2pz	$1+4 \times 2+5 = 14$
Total			$9+14 = 23$

6-31++G**

Atoms	Core orbital	Valence orbital	No. of basis functions
H	0	1s	$(2+1+3) \times 3 = 10$
N	1s	2s, 2px, 2py, 2pz	$1+4 \times 2+4+5 = 18$
Total			$10+18 = 28$

Dunning basis sets

- ▶ Correlation consistent basis sets designed for electron correlation methods, like, MP2, CI, CCSD, CCSD(T) etc.
- ▶ Ex. cc-pVXZ–Dunning's correlation consistent polarized valence X-zeta basis; X = D, T, Q, 5, 6, 7
- ▶ A prefix “aug” is included once diffuse functions added, like, aug-cc-pVXZ

Reading assignment

Section 2.5.5: A Simple Illustration of the Roothan-Hall Approach,
Chapter 2 from Leach, A. R. (2001). Molecular Modelling: Principles and
Applications. Pearson Education.