

Quantum Mechanics

José R. Valverde
EMBnet/CNB

Introduction

- ~ A molecule is not a string/network of atoms
 - ~ It is a cloud of electrons in a field of nuclei
 - ~ Molecular orbitals
- ~ QM deals with electrons and nuclei
 - ~ May account for bond-breaking / formation
 - ~ Explore the fine structure of species
 - ~ Explores electronic behaviour (excited states)
- ~ Quantum Dynamics may explore system evolution in time

The wave equation

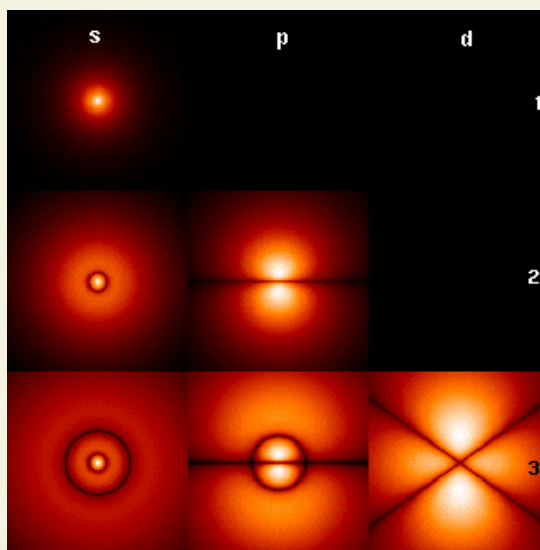
- ~ A particle is a wave is a particle is a wave is...
- ~ The core of QM is Schroedinger's equation

$$H\Psi = E\Psi$$

- ~ Ψ is a continuous natural function (3D [+ time]) in *all* the space
- ~ In practice we have a system of equations with many Ψ for many E levels
- ~ Given Ψ we can compute observables
 - ~ Observable = $\int \Psi^* \text{Operator } \Psi \, d\tau / \int \Psi^* \Psi \, d\tau$

Atomic orbitals: H

- ~ Schroedinger equation can be solved for H ($1s^1$)
- ~ Electrons are in orbitals
- ~ Orbitals are probabilistic *electron density clouds*
- ~ Higher probability regions have a higher *charge density*



Atomic orbitals: polielectronic

- ~ We consider each electron separately:
 - ~ With its own wave function X_i
 - ~ The total wave function is $\Psi = X_1 X_2 \cdots X_n$
 - ~ Electrons have spin ($1/2$ $-1/2$, α β)
 - ~ He: $1s^2 = 1s^\alpha 1s^\beta \rightarrow \Psi = X^\alpha X^\beta$
 - ~ Pauli's exclusion principle: if we exchange two electrons the function must change sign
 - ~ Rewrite the equation as *a determinant* to account for all combinations and normalize with $1/\sqrt{n!}$
 - ~ **Simplify notation** as $\Psi = X_1 X_2 \dots X_n$

Solving the equation

- ~ Assign a value to orbitals (e. g. STO functions defined by Hartree and simplified by Slater)
- ~ Apply the *variational principle*:
 - ~ As the function gets better, the energy will be lower
- ~ Orbitals should be orthonormal
 - ~ $\int X_i X_j d\tau = 0$ if $i \neq j$
 - ~ But as we use approximations we must account for overlap integrals:
 - ~ $\int X_i X_j d\tau = S_{ij}$

Molecular orbitals

- ~ The total wave function for the molecule will be again the product of individual electron wave functions

$$\Psi = \Phi_1 \Phi_2 \dots \Phi_n$$

- ~ Individual orbital wave functions are computed from atomic orbitals X_k as a LCAO

$$\Phi_i = \sum_k c_{ik} X_k$$

- ~ We know X_k so we need to find c_{ik}
 - ~ Use variational principle

Computing Ψ

~ We have

~ $\Psi = \Phi_1 \Phi_2 \dots \Phi_n$ (a normalized determinant) where
 $\Phi_i = \sum_k c_{ik} X_k$, with X_i known and the only
unknowns are c_{ik} .

~ Can be written as a system of equations

~ In matrix form

~ We can solve the system of equations to find c_{ik}

~ Born-Oppenheimer approximation

~ Nuclear movement is too slow w.r.t. Electrons

~ Nuclear movement effects can be ignored.

The Self Consistent Field

- ~ We can exploit the variational principle:
 - ~ Start with a guess for c_{ik}
 - ~ Substitute guessed values into the set of equations
 - ~ Compute energy
 - ~ Use this energy to define a new set of values for c_{ik}
 - ~ Continue until the set of values stabilizes (becomes self-consistent)
- ~ Problems
 - ~ There are factors we haven't considered
 - ~ We start from approximations to atomic orbitals

Relativity

- ~ In general we ignore it
 - ~ Its effects are negligible for most biological atoms/molecules
- ~ Relevant for heavy atoms
 - ~ External electrons repel internal electrons towards the nucleus
 - ~ Inner electrons speed increases and becomes relativistic
 - ~ Cr, Fe, Zn... owe relevant properties to relativity

Basis sets

~ LCAO: MOs as a linear combination of AOs

$$\sim \Phi_i = \sum_k c_{ik} X_k$$

~ X_k are represented by approximate functions:

~ Basis functions/basis set

~ STO: spherical harmonics

~ Good for diatomics and semiempirical methods

~ GTO: Gaussian curves

~ Bad close and far from nucleus

~ Can be enhanced by sums of gaussians

Basis sets

~ Minimal set

- ~ One STO function (or GTO set) for each AO

~ N-Zeta

- ~ DZ: double number of functions for each AO

- ~ TZ: three functions, QZ, 5Z...

- ~ Split valence: one function for core, two for valence orbitals

- ~ Polarization functions: DZP, QZP ...

~ Contracted sets

- ~ Compute core orbitals once from and fix them

Common basis sets

~STO-nG

- ~STOs approximated by n gaussian functions.
- ~Minimal basis set

~k-nlmG

- ~Split valence basis set
- ~k: number of PGTOs used to approximate core e-
- ~nlm: GTOs used for valence orbitals
- ~May add diffusion (+) and polarization functions (d, p orbitals)

Common Basis sets

~STO-3G

- ~Minimal basis set

~3-21G:

- ~core is a contraction of 3 PGTOs, valence is divided in two parts, one uses 2 PGTOs and the other 1 GTO

- ~ Similar to STO-3G but uses double PGTOs in the valence section

~6-31G

- ~Similar to 3-21G but higher precision

Common basis sets

~6-311G

~Splits valence in three levels, with 3 PGTOs for inner valence orbital and two levels for variation

~6-311++G(3d,2p)

~6 PGTOs for contracted (compute once) core

~3 GTOs for inner valence

~2 additional valence levels

~2 diffusion functions (++)

~3 polarization functions for d orbitals and 2 for p orbitals

Ab initio methods

Hartree-Fock

- ~ SCF or HF uses the variational principle as described. N^4
- ~ RHF: use same orbitals irrespective of spin
 - ~ Good for full occupancy (singlet)
- ~ UHF: different orbitals for different spins
 - ~ Good for unpaired electrons (doublet..)
 - ~ Spin contamination
- ~ ROHF: compute spins separately by force same orbitals
- ~ Recovers 99% of the energy

Electron correlation

- ~ Electron clouds do not consider correlation
- ~ Picture a single voxel:
 - ~ There is a probability that two electrons will be there simultaneously at any time
 - ~ But it is unlikely as they repel each other
 - ~ Electron movement is correlated
- ~ Including correlation increases precision
 - ~ But takes more time

Møller-Plesset Perturbation Theory

- ~ Uses many body perturbation theory
- ~ Count twice electron repulsion and apply a correction
 - ~ MP0, MP1, MP2... (depending on the correction order)
 - ~ MP2 is N^5 and recovers 80% of the correlation energy
 - ~ MP3 is N^6 , MP4 N^7 , MP5 N^{8-10} ...
- ~ MP is non-variational
 - ~ May give energies too low.

Configuration interaction

~ Add a new linear combination:

$$\Psi_{\text{enhanced}} = a\Psi_0 + b\Psi_1 + \dots$$

~ Where Ψ_0 is the function for a basic configuration, and Ψ_i are functions for excited states with the same geometry

~ We compute many times the wave function for different excited states.. how many? N^8

~ CIS (single excited state)

~ CISD (single and double)

~ CISDT (... and triple), CISDTQ (..and quadruple)

Multiconfiguration SCF

- ~ It is a kind of CI where we optimize
 - ~ the wave function coefficients
 - ~ and the MOs used
 - ~ May be more precise at a higher cost
 - ~ Run an HF method
 - ~ Run a CI method
 - ~ If $\text{HF} < 0.9 \text{ CI}$ (correlation effect $> 10\%$) then consider MCSCF
- ~ Requires manual intervention to select excited MOs to include

Coupled Cluster

- ~ Similar to CI
 - ~ Instead of including all corrections to a given order, compute one kind of corrections to infinite order.
 - ~ As for CI we need to truncate corrections
 - ~ CCS, CCSD, CCSDT...
 - ~ Precision is similar to CI
 - ~ Method is *extensive* in size

Localized methods

- ~ As we have seen we move in N^5 - N^8
- ~ Most interactions are mainly local
 - ~ Consider only local contributions of atomic orbitals to MOs
 - ~ Used for MP and CC
 - ~ May scale linearly

As a general rule

$$\text{HF} \ll \text{MP2} < \text{CISD} < \text{MP4}(\text{SDQ}) \sim \text{CCSD} < \text{MP4} \\ < \text{CCSD}(\text{T}) < \text{CCSDT}$$

Semiempirical methods

Semiempirical methods

- ~ Based on valence bond
 - ~ Substitute known values to speed up calculus
 - ~ From experimental data (includes correlation)
 - ~ From previous *ab initio* calculations
- ~ Hückel
 - ~ Very simple, imprecise, good start guess
- ~ NDO: neglect of differential overlap
 - ~ CNDO (complete neglect of differential overlap)
 - ~ MNDO (modified neglect)
 - ~ AM1 (Austin Model 1)
 - ~ PM1, PM2, PM3... PM6 (Parameter model n)

As a rule

CNDO < MNDO < AM1 < PM3 < PM6

Density Functional Theory

The basics

- ~ Energy may be computed from electron density (Hohenburg and Kohn) using a *functional*
 - ~ *A function that takes functions as arguments*
- ~ Density may be computed as a linear combination of *orbitals* (Kohn and Sham)
- ~ We need to identify the function of density that is to be feeded the functional to compute the energy

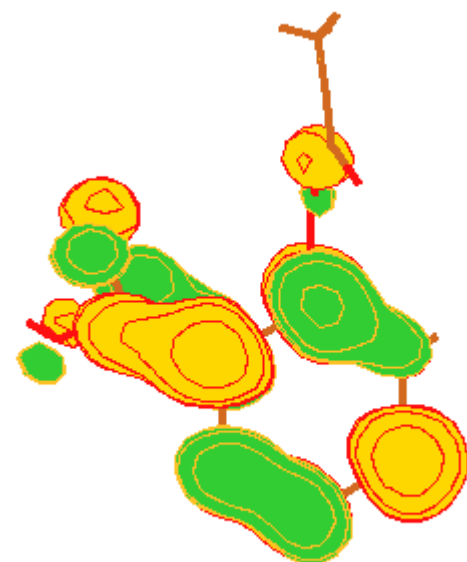
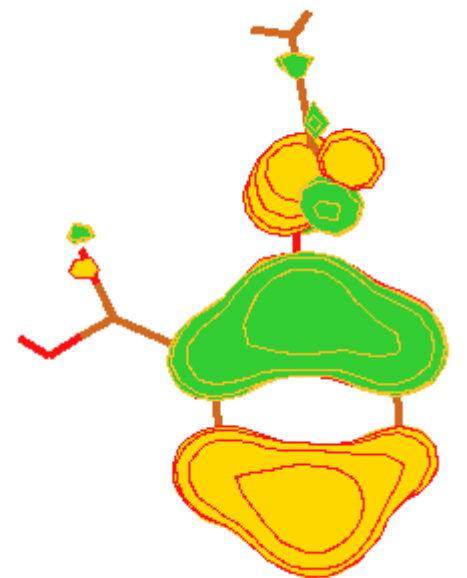
Computing DFT

- ~ Cost is N^3
 - ~ Assuming we can identify the function
 - ~ Which we do not know!
 - ~ Approximations
 - ~ LDA and LSDA depend on density
 - ~ Gradient corrected methods
 - ~ Hybrid methods (with HF): B3LYP
- ~ There are now linear scaling DFT methods
 - ~ FMM, CFMM...

Applications

Electronic structure

- ~ Orbitals
- ~ HOMO (highest occupied)
 - ~ nucleophiles
- ~ LUMO (lowest unoccupied)
 - ~ electrophiles
- ~ Frontier orbital theory



Population analysis

- ~ Charge distribution in space
- ~ Assignment to nuclei (*atoms*)
 - ~ Theory of atoms in molecules
 - ~ Partial charges
 - ~ Bond order
- ~ Mulliken
- ~ Lowdin
- ~ NBO

Charge distribution

- ~ Simplifications:
 - ~ Dipoles
 - ~ Multipoles
 - ~ Used in QSAR
- ~ Electrostatic field
 - ~ Map over VdW surface
 - ~ ESP
- ~ ChelpG method

Other properties

- ~ Structure
- ~ Molecular size and dimensions
- ~ Energy and tensions
 - ~ Vibration states (phonons)
 - ~ Forces for MD simulations
- ~ Electroaffinity
- ~ Additive properties (to build molecules)...

Reactions

- ~ Saddle point calculations
- ~ Transition state calculations
- ~ IRC (Internal reaction coordinates)
- ~ Reaction coordinate path

QM/MM QM/MD QD

~QM/MM

- ~Minimize complex structures

~QM/MD

- ~Model dynamical behavior of complex structures
 - ~Interest region modeled by QM
 - ~Rest modeled by MD
 - ~Border region needs special treatment (ghost atoms)
 - ~ONIOM, Carr-Parrinello

~QD

- ~Full quantum dynamical modeling

In practice

- ~ Select a basis set approximation to AOs
- ~ Select a method to compute MOs
 - ~ *Ab initio*
 - ~ Semiempirical
- ~ Build model
- ~ Launch program
- ~ Wait...

In biological practice

- ~ Usually use a semiempirical method
 - ~ MNDO, AM1, PM3, PM6
- ~ It is now feasible to use *ab initio*
 - ~ With STO-3G (minimal) or 6-311G or better
 - ~ RHF for singlet
 - ~ UHF/ROHF for unpaired electrons
 - ~ If correlation is relevant, use MP or CI
 - ~ If size is too big, use QM/MD
- ~ Consider DFT and Carr-Parrinello MD.

Method**Time**

Molecular Mechanics

 N^2

Molecular Dynamics

 N^2

Hartree-Fock (HF)

 $N^2 - N^4$ (depending on symmetry and cutoff use)

MP2

 N^5

MP3, MP4

 N^6

MP5

 N^8

CC2

 N^5

CCSD

 N^6

CISD

 N^6

CASSF

 $A!$ (depends on the number of orbitals)

CI complete

 $N!$

Semiempirical methods

 N^2 (for small systems)

Semiempirical methods

 N^3 (for large systems)

DFT

 N^3

DFT with linear scaling algorithms

 N

Note: adapted from *Computational Chemistry: A Practical Guide for applying Techniques to Real-World Problems*. David C. Young. Wiley & Sons. 2001.