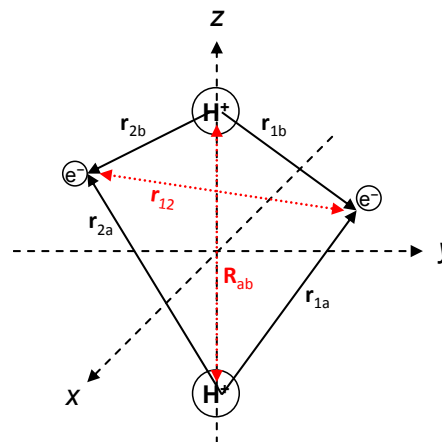


1. Born-Oppenheimer Approximation

Need to elaborate our models to handle molecules. More particles, more variables to keep track of. For H_2 , for instance, would have to solve for energy E and wavefunction

$\Upsilon(\mathbf{r}_1, \mathbf{r}_2, \mathbf{R}_a, \mathbf{R}_b)$ including all four particles. Here \mathbf{r} refer to the coordinates of the two electrons and \mathbf{R} to those of the two nuclei. All have kinetic energy and interact via Coulomb forces. Further, in principle we'd have to make sure the wavefunction obeys the Pauli principle for electrons and all identical nuclei.



Molecular Schrödinger equation

Electron KE	Nuclear KE	Nuclear- electron attraction	Electron- electron repulsion	Nuclear-nuclear repulsion
----------------	------------	------------------------------------	------------------------------------	------------------------------

$$\left\{ -\frac{\hbar^2}{2m_e} \sum_{i=1}^n \nabla_i^2 - \frac{\hbar^2}{2} \sum_{\alpha=1}^N \frac{1}{M_\alpha} \nabla_\alpha^2 - \sum_{i=1}^n \sum_{\alpha=1}^N \frac{Z_\alpha e^2}{r_{i\alpha}} + \sum_{i=1}^n \sum_{j=i+1}^n \frac{e^2}{r_{ij}} + \sum_{\alpha=1}^N \sum_{\beta=\alpha+1}^N \frac{Z_\alpha Z_\beta e^2}{R_{\alpha\beta}} \right\} \Upsilon(\mathbf{r}_1, \dots, \mathbf{r}_n, \mathbf{R}_1, \dots, \mathbf{R}_N) = E \Upsilon(\mathbf{r}_1, \dots, \mathbf{r}_n, \mathbf{R}_1, \dots, \mathbf{R}_N)$$

Quite a mess! Masses of nuclei M are always $\gg m_e$. Electrons will move much faster than nuclei, and nuclei instantaneously will appear immobile. The electron and nuclear motions can thus be approximately decoupled by separation of variables. Essence of *Born-Oppenheimer approximation*:

$$\Upsilon(\mathbf{r}_1, \dots, \mathbf{r}_n, \mathbf{R}_1, \dots, \mathbf{R}_N) \approx \Psi_{\text{elec}}(\mathbf{r}_1, \dots, \mathbf{r}_n; \mathbf{R}_1, \dots, \mathbf{R}_N) \Omega_{\text{nucl}}(\mathbf{R}_1, \dots, \mathbf{R}_N)$$

Substituting, we get *electronic Schrödinger equation*, which depends only parametrically on the locations of the nuclei:

$$\left\{ \sum_{i=1}^n \hat{h}_i + \sum_{i=1}^n \sum_{j=i+1}^n \frac{e^2}{r_{ij}} \right\} \Psi_{\text{elec}}(\mathbf{r}_1, \dots, \mathbf{r}_n; \mathbf{R}_1, \dots, \mathbf{R}_N) = E_{\text{elec}} \Psi_{\text{elec}}(\mathbf{r}_1, \dots, \mathbf{r}_n; \mathbf{R}_1, \dots, \mathbf{R}_N)$$

$$\hat{h}_i = -\frac{1}{2} \nabla_i^2 - \sum_{\alpha=1}^N \frac{Z_\alpha e^2}{r_{i\alpha}}$$

Same ideas about orbital approximation, $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_n) \approx |\psi_1(\mathbf{r}_1) \cdots \psi_n(\mathbf{r}_n)\rangle$ and applying variational principle to give effective

$$\left\{ \hat{h}_i + v_{\text{Coulomb}}[\rho] + v_{\text{exchange}}[\psi_j] + v_{\text{correlation}}[\psi_j] \right\} \psi_i = \hat{f}_i \psi_i = \epsilon_i \psi_i$$

In all approximations, $v_{\text{Coulomb}}[\rho] = \int \frac{\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2$, $\rho(\mathbf{r}) = \sum_i \psi_i^2(\mathbf{r})$

In Hartree-Fock, v_{exchange} is treated exactly and $v_{\text{correlation}}$ ignored. In Hartree-Fock-Slater, v_{exchange} is approximated and $v_{\text{correlation}}$ ignored.

In more advanced methods,

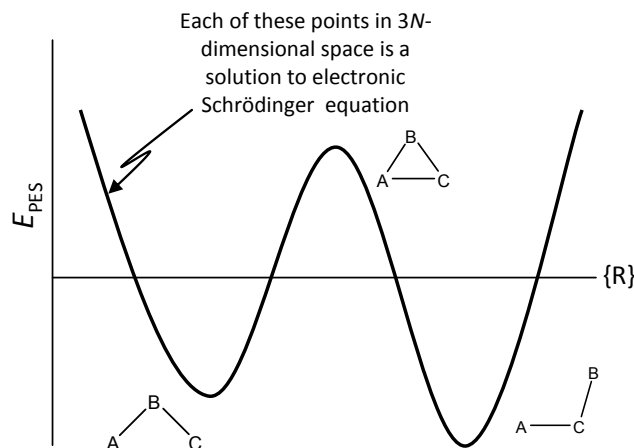
In all cases, we ultimately end up with equations that look *exactly* like our equations

for atoms, except that the one-electron part \hat{h}_i now includes Coulomb attractions to *all* the nuclei, instead of just one. Defines an adiabatic *potential energy surface*, PES

$$E_{\text{PES}}(\mathbf{R}_1, \dots, \mathbf{R}_N) = E_{\text{elec}} + \sum_{\alpha=1}^N \sum_{\beta=\alpha+1}^N \frac{Z_{\alpha} Z_{\beta} e^2}{R_{\alpha\beta}}$$

Nuclei can be thought to “travel” on this PES. Can be treated several ways:

1. Nuclear motion could be described quantum mechanically, e.g. to capture “tunneling.” Expensive and specialized.
2. Can be described as classical particles rolling along the PES. Essence of classical *ab initio* molecular dynamics.
3. Can focus on locating “critical points” along PES, like stable minima (“molecules”) and saddle points (“transition states”). Least expensive and most common in computational chemistry.



Born-Oppenheimer approximation is pretty robust, by and large.

2. Bring back the basis sets

The one-electron HF or HFS equations give us defining equations for the energy-optimal orbitals, but they aren't convenient to solve for anything more complicated than an atom. What to do? Reintroduce idea of a basis set.

$$\psi_i(\mathbf{r}) = \sum_{\mu} C_{\mu i} \phi_{\mu}(\mathbf{r})$$

This is perfect if the ϕ_{μ} form a complete basis, which of course is never possible. Trick again will be to find a convenient and sufficiently complete one. If basis functions are atomic-like, often called “LCAO-MO” approximation, or “linear combination of atomic orbitals-molecular orbital” approximation.

Substitute ψ into one electron equations, multiplying and integrating gives

$$\sum_{\nu} F_{\mu\nu} C_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu} C_{\nu i} \quad \text{or} \quad \mathbf{FC} = \mathbf{SC}\epsilon$$

where \mathbf{F} and \mathbf{S} define the Hamiltonian (“Fock”) and overlap matrices, just like we had when we developed the secular equations:

$$F_{\mu\nu} = \langle \phi_{\mu} | \hat{f} | \phi_{\nu} \rangle \quad S_{\mu\nu} = \langle \phi_{\mu} | \phi_{\nu} \rangle$$

Translates problem from integrodifferential equation into a secular matrix problem. Difficulty is that the Fock matrix elements depends on all the $\psi_i(\mathbf{r}) = \sum_{\mu} C_{\mu i} \phi_{\mu}(\mathbf{r})$.

In Hartree-Fock case, called Hartree-Fock-Roothaan equation. Substituting again

$$F_{\mu\nu} = \langle \phi_\mu | \hat{h} | \phi_\nu \rangle + \sum_{j=1}^{N/2} 2 \langle \phi_\mu | \hat{J}_j | \phi_\nu \rangle - \langle \phi_\mu | \hat{K}_j | \phi_\nu \rangle$$

$$= H_{\mu\nu} + \sum_{\lambda} \sum_{\sigma} P_{\lambda\sigma} \left[(\mu\nu | \lambda\sigma) - \frac{1}{2} (\mu\lambda | \sigma\nu) \right]$$

(Differs only in details for HFS or DFT.) This simplifies things down to a *charge density matrix*

$P_{\mu\nu} = 2 \sum_i^{N/2} C_{\mu i} C_{\nu i}$, which expresses the charge density in terms of the basis,

$$\rho(\mathbf{r}) = \sum_{\mu} \sum_{\nu} P_{\mu\nu} \phi_{\mu}(\mathbf{r}) \phi_{\nu}^*(\mathbf{r})$$

and a bunch of *two-electron integrals*

$$(\mu\nu | \lambda\sigma) = \iint \phi_{\mu}^*(1) \phi_{\nu}(1) \frac{1}{r_{12}} \phi_{\lambda}^*(2) \phi_{\sigma}(2) d\mathbf{r}_1 d\mathbf{r}_2$$

The two electron integrals are determined by the basis. The density matrix has to be guessed and then updated with new coefficients during SCF cycles. It is this matrix that is usually converged upon.

Here's our new *self-consistent field* algorithm:

1. Select a basis set
2. Calculate one-electron and two-electron integrals (only once!)
3. Guess some coefficients **C**/density matrix **P**/density ρ . (This guess matters. It determines what electronic configuration you are calculating!)
4. Construct and solve Fock matrix for coefficients **C** and eigenvalues ϵ
5. Construct new density matrix **P'** and compare to old
6. Difference greater than tolerance...update **P** and return to Step 4
7. Difference less than tolerance....all done! Check to see whether you arrived at the electronic configuration you wanted!)

3. Semi-empirical methods

A Hartree-Fock calculation roughly scales in computational expense as N^4 , where N is the number of electrons. Depending on your computer and the size of the system you want to model, this can be prohibitively expensive. Many methods developed to simplify the calculations. See Cramer for a nice description.

Extended Hückel Theory (EHT): Uses minimal basis of Slater orbitals, uses empirical rules to construct Hamiltonian matrix elements (by fitting to some experimental constants). Non-iterative, useful for molecular orbital analysis (won Robert Woodward and Roald Hoffmann a Nobel Prize!) Called *tight bonding* in the physics world.

Complete Neglect of Differential Overlap (CNDO) and *Intermediate Neglect of Differential Overlap* (INDO): Like EHT, but parameterizes one- and two-electron integrals rather than matrix elements. Sets $S_{\mu\nu} = \delta_{\mu\nu}$ in Fock matrix. Archaic.

Neglect of Diatomic Differential Overlap (NDDO): Includes more integrals. Many flavors, including MNDO, MINDO, AM1, PM3, All involve different parameterizations, often fit to experimental data.

Still actively developed, but superseded in many applications by DFT. Utility depends on particular problem.

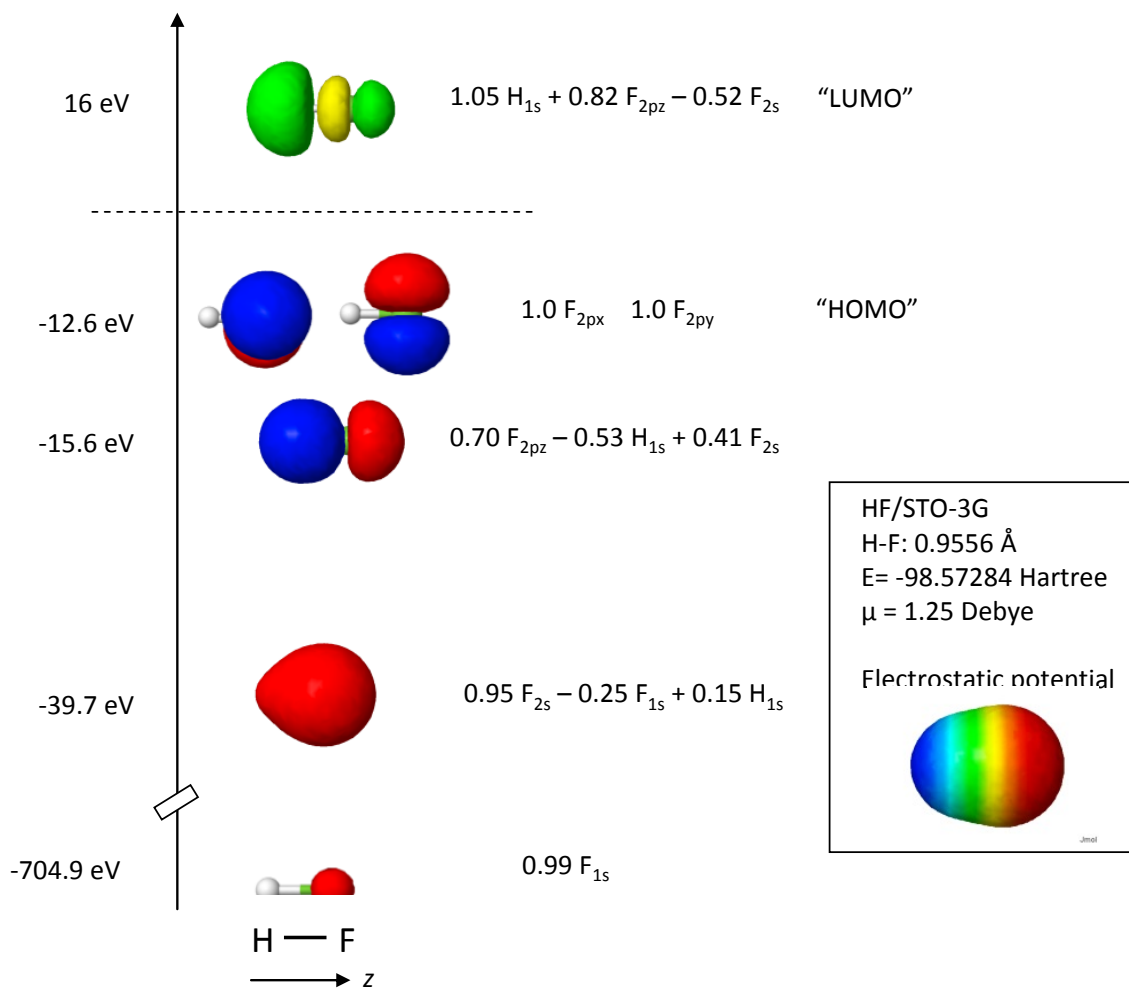
Some new variations on these ideas, like “Density Functional Tight Binding” (DFTB). Still active area of research.

4. H₂ example

Minimal basis set. See <http://www.webmo.net>. Walk through input and output files.

5. HF Example

Example....minimal basis set “HF/STO-3G” calculation on HF. Six basis functions input, six molecular orbitals output. Walk through Gaussian output. “Pop=Full GFInput” to see basis information



Exp't H-F: 0.917 Å

μ = 1.82 D

IP: 16.1 eV

Note where all the energy is...in the core! Note where all the chemistry is...in the valence! Basis sets have to balance treatment of both...good core lowers energy but may not help with chemical part.

6. Symmetry

In molecules that possess *symmetry*, Fock matrix is simplified because integrals of basis functions from different *symmetry classes* equal zero. Look at HF example above, F 2p_x and 2p_y orbitals have π symmetry and cannot interact directly with *s* basis functions on H. (Indirectly affected by changing charge density.) Other orbitals have σ symmetry. Fock matrix becomes block diagonal in these two symmetry classes. Good resource for this is Cotton's *Chemical Applications of Group Theory*.

Non-linear molecules, can construct *symmetry adapted* basis functions. Water example.

7. Population analysis

The molecular orbital solutions of the Hartree-Fock/molecular orbital model contain lots of useful information that can be helpful in understanding structure and bonding.

Perhaps most direct is the distribution of *charge* around the molecule,

$$\rho(\mathbf{r}) = \sum_{\text{occupied}} \psi^2(\mathbf{r}) = \sum_{\mu} \sum_{\nu} P_{\mu\nu} \phi_{\mu}(\mathbf{r}) \phi_{\nu}(\mathbf{r})$$

Can calculate moments of the charge (dipole, quadrupole, ...). Note these are only exactly defined for neutral species.

Another interesting quantity is the Coulomb, or often-termed “electrostatic” potential (ESP) created by the combination of electrons and nuclei. *CHELPG* works by trying to reproduce calculated electrostatic potential with charges placed at the atom centers.

Chemically it is intuitively nice to assign charge to individual atoms, in what is called a *population analysis*. There is no single “right” way to do this...the “charge” on an atom in a molecule is not uniquely defined!

Consider an occupied molecular orbital made up of two basis functions on two different atoms, α and β :

$$\begin{aligned} \psi &= c_{\alpha} \phi_{\alpha} + c_{\beta} \phi_{\beta} \\ \langle \psi | \psi \rangle &= c_{\alpha}^2 + c_{\beta}^2 + 2c_{\alpha} c_{\beta} \langle \phi_{\alpha} | \phi_{\beta} \rangle \end{aligned}$$

c_{α}^2 can be interpreted as the fraction of charge in ψ that can be assigned to α , and likewise c_{β}^2 for β . The remainder is the “overlap” population between α and β . Who gets these? In *Mulliken population analysis* they are divided equally between the two atoms. Summing over all the occupied orbitals and subtracting from the nuclear charge gives the *gross atomic populations*.

Mulliken populations are directionally very useful but their absolute magnitudes have no real meaning. They are sensitive to the choice of basis, and equal distribution of overlap density between atoms is rather arbitrary.

The overlap problem goes away if the basis is orthogonal, so the overlap terms go to zero. In the *Löwdin* approach the basis set is orthogonalized according to $\phi_{\mu}' = \sum_{\nu} (S^{-1/2})_{\nu\mu} \phi_{\nu}$, so that $\langle \phi_{\mu}' | \phi_{\nu}' \rangle = \delta_{\mu\nu}$. MO

coefficients in this transformed basis used to assign charge to individual atoms. More stable, less common than Mulliken approach. This orthogonalization scheme is not unique, so other implementations possible.

Natural orbitals are another, more sophisticated scheme for orthogonalizing and assigning charge. See <http://www.chem.wisc.edu/~nbo5/>. Again more stable than Mulliken, more information rich.

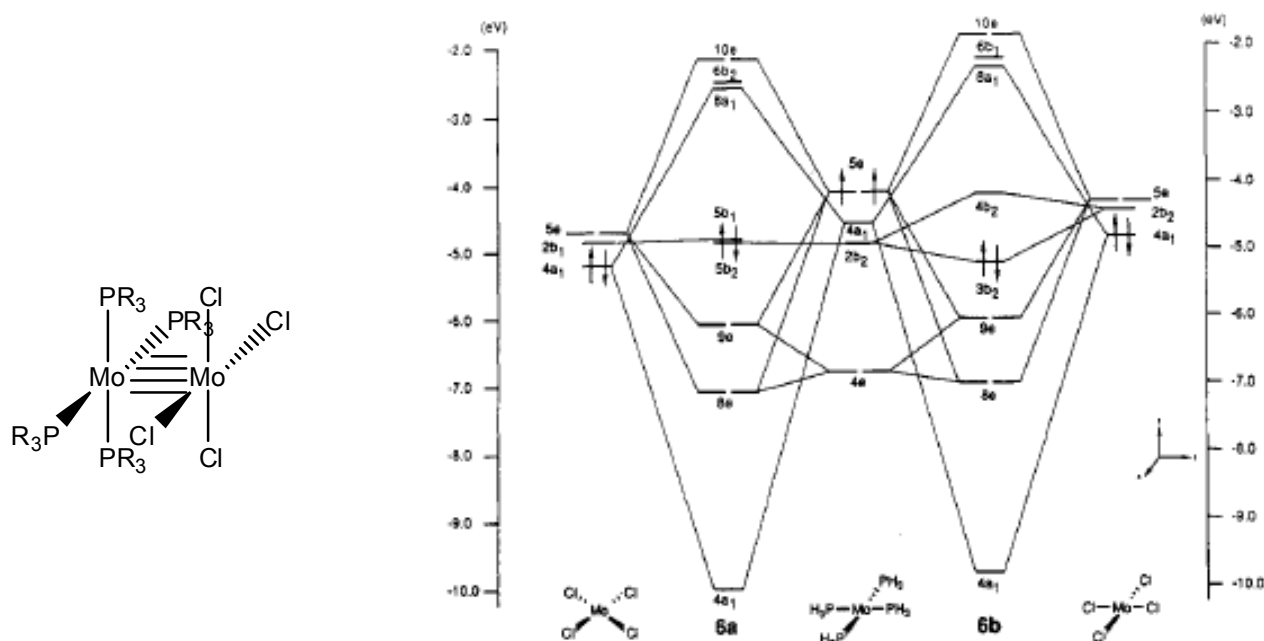
Bader analysis another more sophisticated method, based on a geometric analysis of the total charge density. See Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, 1990.

8. Molecular orbital diagrams

Direct analysis of the shapes and energies of molecular orbitals can often be useful for understanding molecular shapes and bonding. Molecular properties often dictated by *frontier molecular orbitals*, including *HOMO* (highest occupied molecular orbital) and *LUMO* (lowest unoccupied molecular orbital). Good resource is Albright, Burdett, and Wangbo, *Orbital Interactions in Chemistry*.

HF picture above is one good example of a molecular orbital diagram. Hartree-Fock creates virtual orbitals that are artificially high in energy and so is not the best for calculating these. Orbital analysis often applied within semi-empirical and DFT frameworks.

Here is an oldie but goodie, illustrating why $\text{Mo}_2\text{Cl}_4(\text{PR}_3)_4$ is eclipsed rather than staggered. Note formation of δ bond in eclipsed (**6b**) conformation that is not present in staggered.



9. Gaussian basis sets

Generally seek basis functions for which the necessary integrals can be calculated conveniently. One common approach, and the one we cover in the first part of this course, uses *atom-centered basis functions*.

Recall *Slater* functions ($e^{-\zeta r}$) have correct functional form near nucleus, but many-center integrals cannot be calculated analytically. *Gaussian* functions ($e^{-\zeta r^2}$) turn out to have functional form particularly convenient for calculating necessary integrals. Most common procedure for molecules is to write molecular orbitals in expansions in terms of Gaussian atomic-like functions centered on the atoms. Many such *Gaussian basis sets* have been created for various purposes, and considerable jargon associated with concept. One of John Pople's (1998 Nobel prize in Chemistry, w/Walter Kohn) major contributions.

The *Gaussian* commands GFPrint and GFInput print out the basis set; the Gen command can be used for specifying a basis set. See the manual.

Primitive Gaussian is one Gaussian basis function, defined by an *exponent* ζ and an angular part given by some product of x , y , and z :

$$g(\mathbf{r}, \xi, i, j, k) = \left(\frac{2\alpha}{\pi} \right)^{3/4} \left[\frac{(8\alpha)^{i+j+k} i! j! k!}{(2i)!(2j)!(2k)!} \right] x^i y^j z^k e^{-\xi r^2}$$

For instance $i = j = k = 0$ would define a single, spherical s -type primitive Gaussian.

The set $(i, j, k) = (1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ would define a set of three p -type primitive Gaussians.

Exponent ζ determines how fast or slow the basis function decays away from the atom. Big ζ = fast decay = function close to nucleus; small ζ = slow decay = function far from nucleus.

Because individual Gaussians are a pretty poor representation of the behavior of real atomic wavefunctions, especially near nuclei, common to bundle together, or *contract*, several Gaussians with different exponents into one basis function:

$$\phi(\mathbf{r}) = \sum_{a=1}^m g(\mathbf{r}, \xi_a)$$

m is typically a small number, maybe up to six. A *basis set* is a collection of exponents and contraction coefficients defining a basis for an atom or atoms. Constructed in many ways, be fitting to numerical results for atoms and for some set of properties of molecules. Common notation

primitive Gaussians \rightarrow contracted functions
e.g. (10s, 5p, 2d, 1f) \rightarrow [4s, 3p, 2d, 1f]

Note a "d" can be either 5 or 6 functions, depending on whether the symmetric component ($x^2 + y^2 + z^2$) is included ("cartesian") or removed (spherical). Similarly "f" can be 7 or 10.

Minimal basis sets

To describe an atom within a molecule, would want at a minimum one basis function for each occupied atomic orbital. Called a *minimal basis set*. For instance

Hydrogen	1s = one basis function
Fluorine	1s + 2s + 2p _x + 2p _y + 2p _z = five basis functions.
Iron	1s, 2s, 2p _{x,y,z} , 3s, 3p _{x,y,z} , 3d _{xy} , 3d _{xz} , 3d _{yz} , 3d _{x2} , 3d _{y2} , 3d _{z2} , 4s, 4p _{x,y,z} = 19 basis functions

Table. Some sample results for HF molecule (from CCCBDB, <http://cccbdb.nist.gov/>), Hartree-Fock calculations.

	STO-3G	6-31G	6-31G**	cc-pVDZ	aug-cc-pVDZ	6-311G**	cc-pVTZ
# basis functions	5/1	9/2	15/5	15/5	26/9	20/6	38/15
Energy (Hartree)	-98.5728	-99.9834	-100.0117	-100.0197	-100.0338	-100.0469	-100.0585
E HOMO (eV)	-12.6	-17.2	-17.1	-17.2	-17.8		-17.6
Dipole moment (D)	-1.252	-2.301	-1.944	-1.918	-1.899	-1.980	-1.905

For example, “STO-3G” is short-hand for a minimal basis set in which each basis function is a contraction of three Gaussians. For fluorine (6s 3p) \rightarrow [2s 1p]. STO-*n*G basis sets available for many atoms. Good for rough and ready calculations, but not very accurate.

Multi-zeta basis sets

Minimal basis set doesn't leave any real room for atomic orbitals to “breathe” (expand or contract) when forming bonds with other atoms. *Double zeta* basis set one with two basis functions per occupied atomic orbital. More common is *split valence*, which is double zeta for valence levels, single zeta for core.

Pople basis sets very common split valence type. For example, “6-31G” basis set for fluorine:

- 1s orbital described by 6 primitive Gaussians contracted to one basis function
- One set of 2s and 2p orbitals described by contraction of 3 primitive Gaussians
- One set of 2s and 2p orbitals described by 1 primitive Gaussian

Triple zeta and *triple split valence* similar, for instance “6-311G”

Polarization functions

To capture *polarization* of atoms that occurs when forming chemical bonds, it is necessary to include basis functions of higher angular momentum than the occupied atomic orbitals, e.g. *p* functions for H or *d* functions for C.

For example, 6-31G(d,p) (a.k.a. 6-31G**) augments 6-31G with a single set of polarization functions on each atom. Multiple polarization functions can be added, and are particularly important to capture electron correlation effects.

Dunning's “correlation-consistent” basis sets another family that includes polarization functions in a systematic way, e.g. cc-pVDZ (polarized valence double zeta) is similar in size to 6-31G(d,p), cc-pVTZ similar to 6-311G(2df,2pd), higher levels available.

Diffuse functions

Lastly, for anionic systems or loosely bound electrons, common to include “diffuse functions,” same angular momentum as occupied states, but small exponents. Denoted “6-31++G(d,p)” in Pople notation, “aug-cc-pVDZ” in Dunning.

As seen in the Table above, the total energy drops with increasing basis size (consequence of variational principle). “Hartree-Fock limit” is energy in limit of complete basis. Other quantities converge more quickly with basis size.

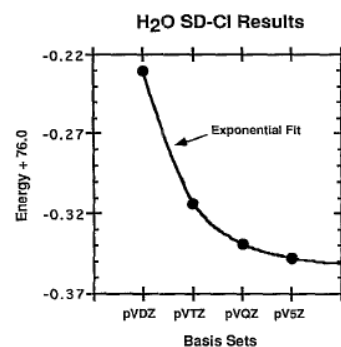


FIG. 1. Frozen core ground state (1A_1) HF SD-CI energies for the water molecule obtained with various correlation consistent basis sets at the experimental geometry. The solid line is a least squares exponential fit to the four data points.

Some families of basis sets are constructed to facilitate extrapolation to the complete basis limit (CBS). See for instance D. Feller, "The use of systematic sequences of wave functions for estimating the complete basis set, full configuration interaction limit in water," *J. Chem. Phys.* **1993**, 98, 7059.

The incompleteness of basis sets leads to a so-called "basis set superposition error" (BSSE) when calculating the interaction between two fragments. The counterpoise (CP) correction of Boys and Bernardi, *Mol Phys.* **1970**, 19, 553, is most common, but not entirely non-controversial, way of corrections.

Electron cores

Low energy "core" electrons typically don't participate much in chemical bonding, but can add a lot to computational cost of Hartree-Fock (and beyond) calculations. Generally seek approximations. Especially important for heavy elements/metals.

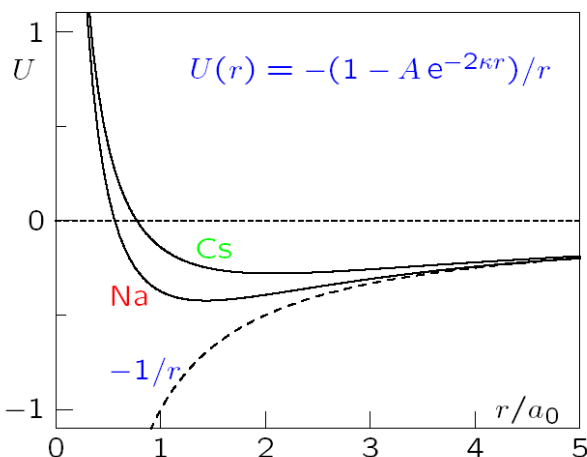
Basic assumption is so-called "frozen core" approximation, that the core electrons can be assumed to be constant in different chemical environments. Not strictly right, but for chemical bonding typically a good approximation. First, need to make (somewhat arbitrary) decision about what is *core* and what is *valence*.

Within Hartree-Fock, for instance, we could separate the Fock operator like this:

$$\left\{ \hat{h} + \sum_{core} (2\hat{J}_j - \hat{K}_j) + \sum_{valence} (2\hat{J}_j - \hat{K}_j) \right\} \psi_i(\mathbf{r}_1) = \hat{f} \psi_i(\mathbf{r}_1) = \varepsilon_i \psi_i(\mathbf{r}_1)$$

$$\hat{J}_j = \int |\psi_j(\mathbf{r}_2)|^2 \frac{1}{r_{12}} d\mathbf{r}_2 \quad \hat{K}_j \psi_i(\mathbf{r}_1) = \left[\int \psi_i(\mathbf{r}_2) \psi_j(\mathbf{r}_2) \frac{1}{r_{12}} d\mathbf{r}_2 \right] \psi_j(\mathbf{r}_1)$$

Goal is to replace core operators on ψ_i with some simpler-to-calculate operator(s) derived from accurate atomic calculations. Simplest way (not nearly good enough, though) would be to say that valence electrons see nucleus classically screened by core electrons. Instead of valence seeing $1/r$, see diminished $1/r$. Valence wavefunctions won't have the right nodal behavior near the core, but the valence part will (can) be ok.



Effective Core Potentials – 70 Years of Development –

Dirk Andrae
Theoretische Chemie
Fakultät für Chemie, Universität Bielefeld

Hellmann Memorial Meeting
Bonn, 2003-07-26

Effective core potential $U(r)$, in atomic units, for **sodium** ($A = 1.672$; $\kappa = 0.333$), used in this form by Hellmann & in alkali metals [H. Hellmann, W. Kassatotschkin: Di Näherungsverfahren. *Acta Physicochim. U. R. S. S.* **5** (19 term energies)].

To do right, have to take care of exchange and the different potentials seen by electrons with different angular momenta. Provides convenient opportunity, also, to take care of *relativistic effects*, which are most important near the core.

Relativistic kinetic energy is relativistic total energy minus the rest energy:

$$T = \sqrt{p^2 c^2 + m_0^2 c^4} - m_0 c^2$$

Taylor expanding about $p^2 = 0$ gives the first-order *mass-velocity correction*:

$$T \approx \frac{p^2}{2m} - \frac{p^4}{8m^3 c^2}$$

Reduces to non-relativistic result in limit that c goes to infinity. Electrons near core move at speeds close to c , second term becomes non-negligible and diminishes their energy. Most important for s states that penetrate closest to nucleus; they shield nucleus better and other states rise up in energy.

Magnetic field of electron coupling with magnetic field induced by nucleus that appears to be circulating around it causes *spin-orbit coupling*: orbitals split into $j = l \pm s$ states. Only manifest for $l > 0$.

Darwin correction: corrects s orbitals for electron and nucleus being at the same point; comes from solution of full Dirac relativistic equation for the atom.

Non-relativistic and relativistic *effective core potentials (ECPs)* available for many elements. These specify the potential felt by the valence electrons due to the core in terms of radial potential functions and angular projection operators. Typically these have to be combined with basis functions designed to work with them.

Most common are Hay-Wadt (LANL) and Stevens-Basch-Krause (SBK). Other more modern ones also available, like Stuttgart. See *Reviews in Computational Chemistry* 1996. Gaussian allows these common ones to be specified along with the basis set in the job line. Psuedo command allows arbitrary ECPs to be input.

10. Choosing basis sets/ECPs

Of course, would like a *complete* basis set, but computational cost almost always limits basis set size. Further, there is no one unique way to construct atom-centered basis functions *or* core potentials, so there is always some arbitrariness to these basis sets. Much research has gone into constructing basis sets.

Rules of thumb

1. More basis functions are generally better than fewer. Trick is to add basis flexibility in regions you care about. Also, generally need to have balance in flexibility amongst atoms.

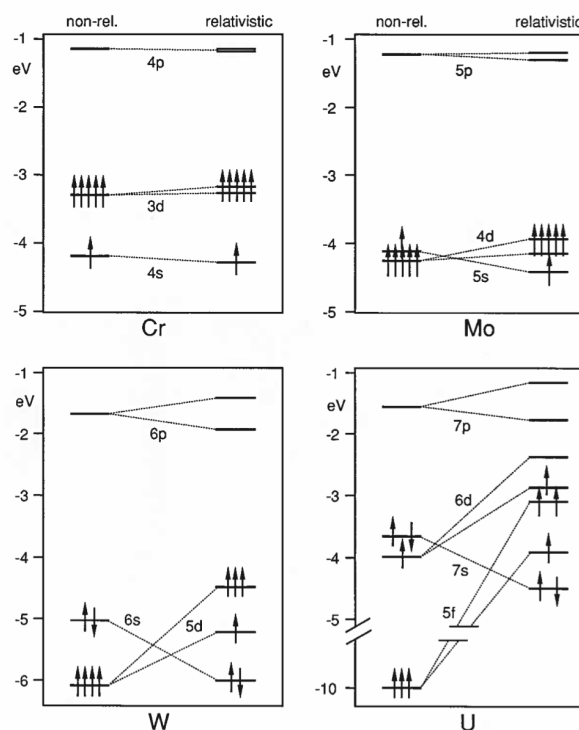


Figure 29. Comparison of non-relativistic HFS and relativistic DFS atomic calculations on four metals with six valence electrons: chromium ($Z = 24$), molybdenum ($Z = 42$), tungsten ($Z = 74$), and uranium ($Z = 92$).

2. Consider nature of problem at hand. Structures, dipole moments, vibrational frequencies can generally be captured with modest (double zeta + polarization) basis. Accurate absolute bond or excited state energies may require larger bases, but relative energies could be achieved more modestly. Anions require variational flexibility far from the atom, so diffuse functions. NMR requires good cover near the core region.
3. Personal experience, literature, and convention. Experience is a wonderful guide, and the literature is a great resource to take advantage of other's experience. Many tabulations (like the CCCBDB) to help guide. And certainly approaches, like HF/6-31G**, B3LYP/6-311G**, ..., are venerable and well calibrated. In general, though, careful use means the user tests the sensitivity of computed results to basis set (and other approximations).
4. Balance of approximations. Little sense to converge your results with respect to basis set beyond the reliability of the model (e.g. Hartree-Fock) that you are using.
5. Consider types of atoms involved. Not all standard basis sets cover all atoms.
6. Computational resources ultimately dictate what you are able to do.

In very accurate calculations, often apply *extrapolations* to complete basis set (CBS) limit

Basis set resources

- <https://bse.pnl.gov/bse/portal>
- Davidson and Feller, "Basis Set Selection for Molecular Calculations," *Chem. Rev.* **1986**, 86, 681-696.
- *Gaussian* manual

11. Performance details of SCF methods

Basis is often *orthonormalized* to eliminate overlap from H-F-R equation; allows equations to be solved by matrix diagonalization.

Initial guesses for **P** are obtained by solving an approximate Hamiltonian (like extended Hückel). Always beware! Initial guess can influence final converged state.

Because the number of 2-electron integrals grows as N^4 , they are sometimes calculated as needed "on-the-fly", called *direct SCF*.

The SCF procedure is an optimization problem: find set of coefficients that minimizes the total energy. As discussed above, success depends on a reasonable initial guess for **P** and judicious updating. Various strategies can be used to speed and stabilize convergence, like damped mixing of previous cycles.

Second-order SCF is a convergence acceleration method that requires calculation or estimation of the first- and second-derivatives of the energy with respect to the orbital coefficients. See e.g. Chaban et al., *Theor. Chem. Accts.* **1997**, 97, 88-95.

Pulay's "*direct inversion in the iterative subspace*", or "*DIIS*," is a popular and powerful acceleration procedure that extrapolates from several previous Fock matrices to predict optimal next Fock to diagonalize.

12. Open-shell systems

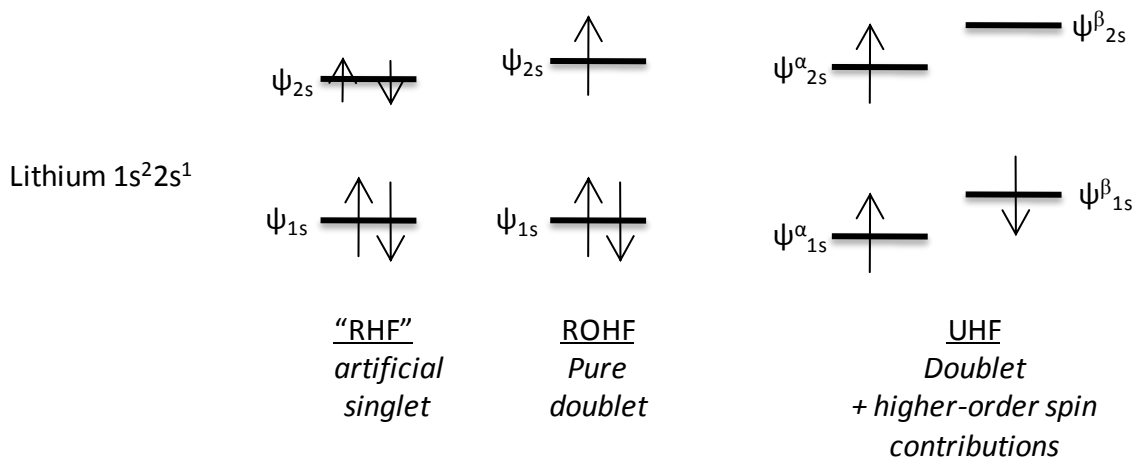
All this development has assumed a system with all electrons paired. Much of interesting chemistry, like transition metal systems or breaking chemical bonds, involve unpaired electrons. Theory has to be elaborated to treat these.

Unrestricted Hartree-Fock

“UHF” creates separate “spin-orbital” for each electron. All electrons interact through Coulomb interaction, but for like electrons, diminished through exchange interaction. UHF advantage is that it is easy and convenient; disadvantages are that you have twice as many orbitals to keep track of, and the resultant wavefunction is not an eigenfunction of the spin operator, so they are not “pure” spin states. Described as “spin-contaminated.”

Restricted Open-Shell Hartree-Fock

“ROHF” retains the same spatial orbitals for spin-up and spin-down electrons but only singly occupies some orbitals. Advantage is that solutions are pure spin functions. Disadvantage is that some spin states can only be described by linear combinations of Slater determinants (recall open-shell examples above), calculations become more cumbersome, and poorer solutions than UHF for given amount of computational effort.



Note that minority spin UHF orbitals are higher in energy because their Coulomb interactions are diminished by fewer exchange interactions.

What goes into an Electronic Structure Calculation?

1. Choose structure (Born-Oppenheimer approximation)
2. Choose electronic structure method (H-F, H-F-S, ...)
3. Choose electronic state
 - a. Charge/# electrons
 - b. Spin-multiplicity
4. Choose basis set and possible core potentials
5. Choose initial guess of orbital coefficients (c_j)
 - a. Determines initial charge density
 - b. Determines initial wavefunctions
 - c. Determines initial hamiltonian
6. Choose convergence criteria
 - a. Change in coefficients < tolerance
 - b. Change in energy < tolerance
7. Choose iteration scheme
 - a. Direct coefficient update...not!
 - b. Damped update
 - c. Optimization-based
 - i. First-order (steepest descents, conjugate gradient)
 - ii. Second-order (quadratically convergent)
 - iii. Extrapolation
 1. DIIS-direct inversion in the iterative subspace
8. Run
9. CHECK RESULTS!!!!!!!