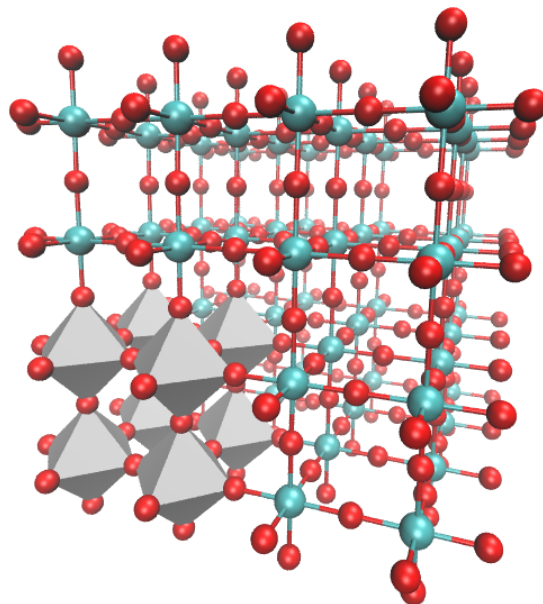


Basis Sets for Molecular Orbital Calculations



Prof. Lee M. Thompson

Introduction

Revision

- Roothan-Hall-Hartree-Fock determines optimum form of molecular orbitals by determining contribution of known functions

$$\hat{F}_i \sum_{\mu}^{N_{\text{basis}}} C_{\mu i} |\phi_{\mu}\rangle = \epsilon_i \sum_{\mu}^{N_{\text{basis}}} C_{\mu i} |\phi_{\mu}\rangle$$

- The molecular orbitals are described as a linear combination of atomic orbitals (LCAO)

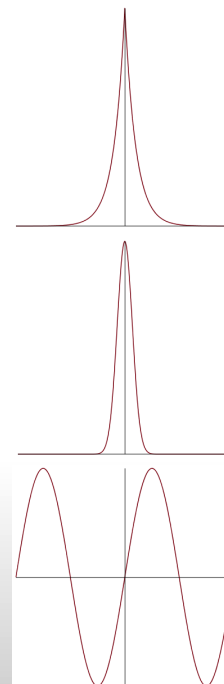
$$|\psi_i\rangle = \sum_{\mu}^{N_{\text{basis}}} C_{\mu i} |\phi_{\mu}\rangle$$

- We haven't yet said anything about the form of these known functions (atomic orbitals) - what functions should we use?

Basis Functions

Types of functions

- In theory we can use any kind of function that we want (so long as they are a complete set)
- In practice we want to use a basis that describes molecular orbitals with the fewest functions
 - Slater-type orbitals
 - Gaussian-type orbitals
 - Plane-waves

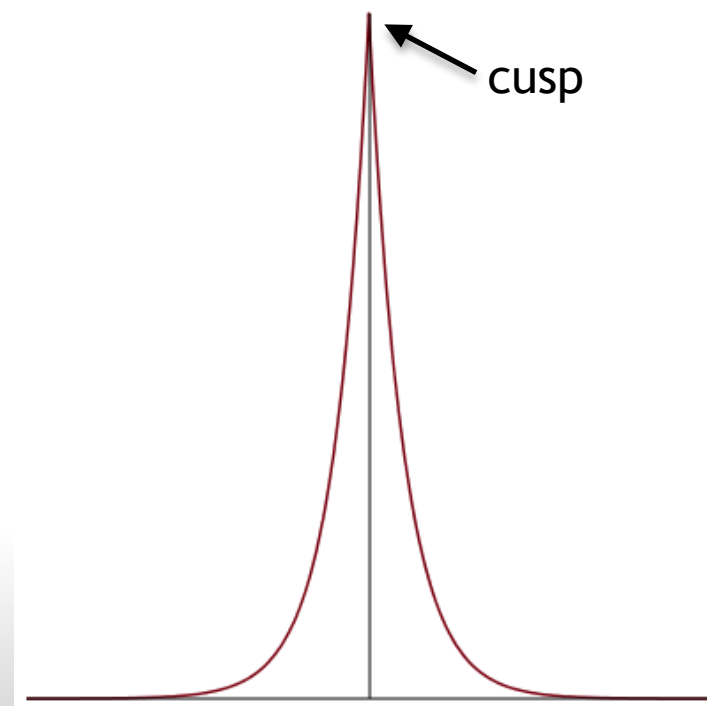


Slater-Type Functions

- A good guess for the form of the molecular orbitals is that they can be described as a (linear) combination of atomic orbitals
- Atomic orbitals are the wavefunctions obtained from solving the Schrödinger equation for an atom
- We can find the exact solutions in the case of the hydrogen atom - use solutions that resemble the radial part

$$|\phi_\mu\rangle = A_\mu e^{-\alpha_\mu |\mathbf{r}|}$$

$$|\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}$$

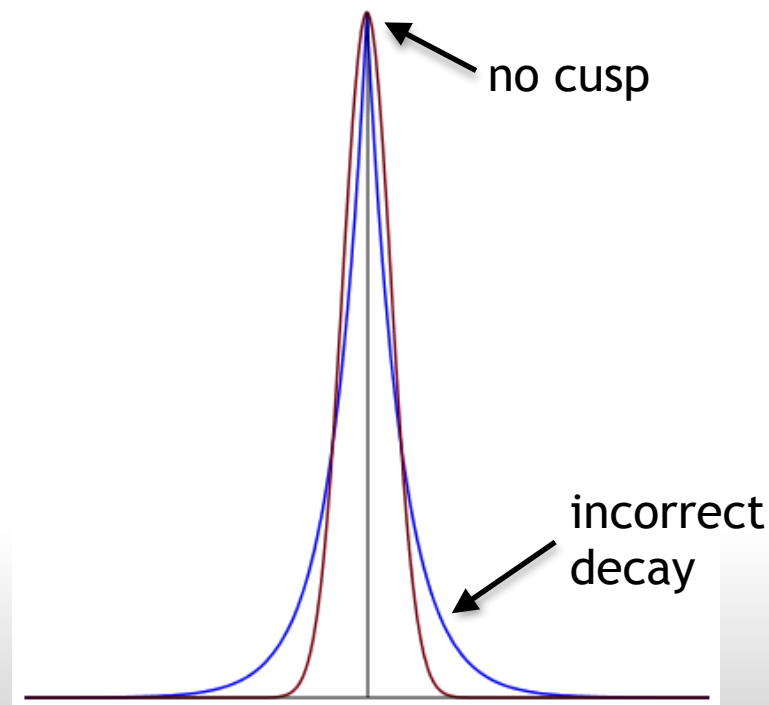


Gaussian-Type Functions

- Gaussian functions are an alternative basis
- We require more of these to approximate a molecular orbital than a Slater function (~3x)
- This is because the shape is not as good
- BUT Gaussian orbitals have some very useful properties that mean they are much faster computationally (~5x)

$$|\phi_\mu\rangle = A_\mu e^{-\alpha_\mu \mathbf{r}^2}$$

$$\mathbf{r} = (\mathbf{x}^2 + \mathbf{y}^2 + \mathbf{z}^2)$$





Gaussian Product Theorem

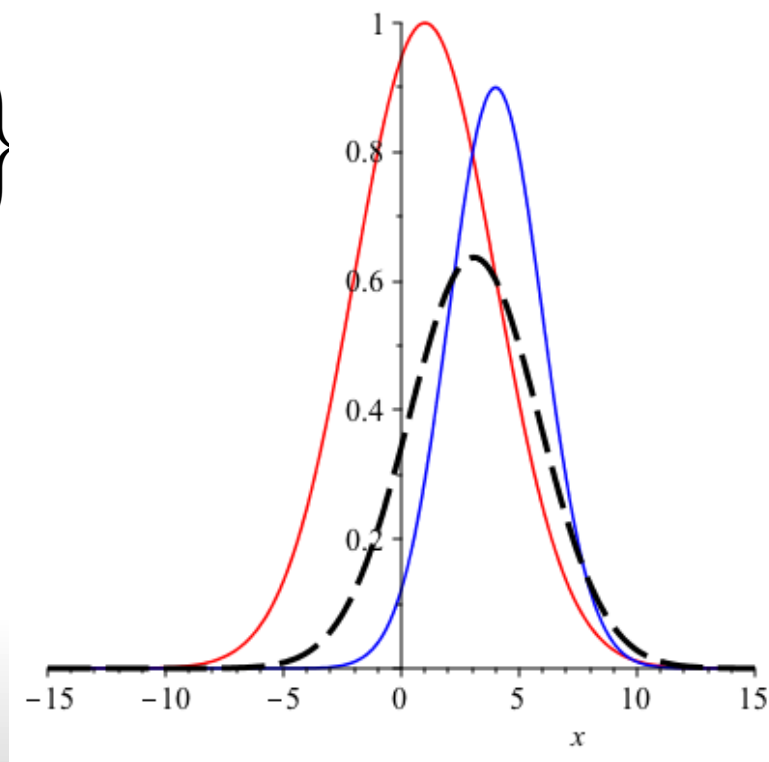
Product of two Gaussians is also a Gaussian

$$|\phi_\mu\rangle = A_\mu e^{-\alpha_\mu \mathbf{r}^2} \Rightarrow A_\mu \exp\left\{-\frac{1}{2\sigma_\mu^2}(\mathbf{r} - \mathbf{r}_\mu)^2\right\}$$

$$|\phi_\mu\phi_\nu\rangle = A_\mu A_\nu \exp\left\{-\frac{(\mathbf{r}_\mu - \mathbf{r}_\nu)^2}{2(\sigma_\mu^2 + \sigma_\nu^2)}\right\} \exp\left\{-\frac{\mathbf{r}_{\mu\nu}^2}{2\sigma_{\mu\nu}^2}\right\}$$

$$\alpha_{\mu\nu} = \frac{\sigma_\mu^2 \sigma_\nu^2}{\sigma_\mu^2 + \sigma_\nu^2}$$

$$\mathbf{r}_{\mu\nu} = \frac{\mathbf{r}_\mu \sigma_\nu^2 + \mathbf{r}_\nu \sigma_\mu^2}{\sigma_\mu^2 + \sigma_\nu^2}$$





Cartesian vs. Spherical

- The solutions of the hydrogen atom Schrödinger equation have both radial part and angular part

$$|\psi_{nlm_l}(\mathbf{r})\rangle = Y_{lm_l}(\theta, \phi)R_{nl}(r)$$

↙ spherical harmonics

- Unlike hydrogenic orbitals, neither Slater or Gaussian orbitals have radial nodes - can obtain with linear combinations
- Gaussian angular component can be included as spherical functions or cartesian functions

$$|\phi_\mu\rangle = A_\mu Y_{lm_l}(\theta, \phi) r^{2n-2-l} e^{-\alpha_\mu \mathbf{r}^2}$$

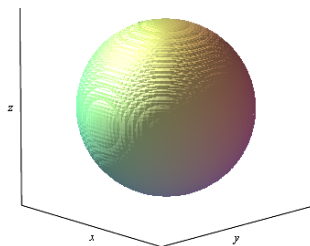
$$|\phi_\mu\rangle = A_\mu x^{l_x} y^{l_y} z^{l_z} e^{-\alpha_\mu \mathbf{r}^2}$$

- There are fewer Gaussians in spherical coordinates than cartesian (cartesian Gaussians are linearly dependent)

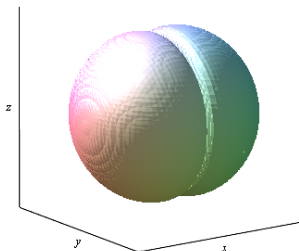


Cartesian vs. Spherical

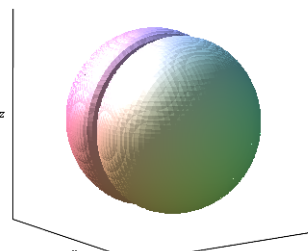
$$l = l_x + l_y + l_z$$



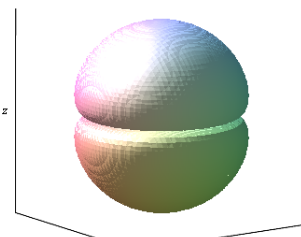
$$|\phi_\mu\rangle = A_\mu x^0 y^0 z^0 e^{-\alpha_\mu r^2}$$



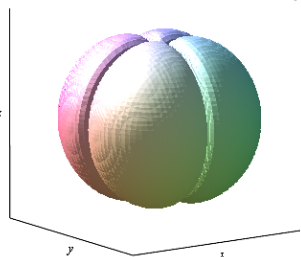
$$|\phi_\mu\rangle = A_\mu x^1 y^0 z^0 e^{-\alpha_\mu r^2}$$



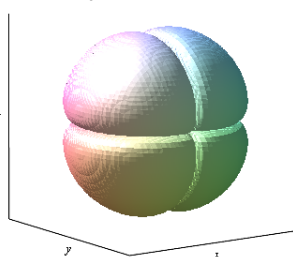
$$|\phi_\mu\rangle = A_\mu x^0 y^1 z^0 e^{-\alpha_\mu r^2}$$



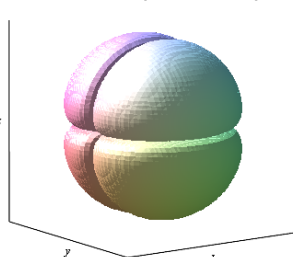
$$|\phi_\mu\rangle = A_\mu x^0 y^0 z^1 e^{-\alpha_\mu r^2}$$



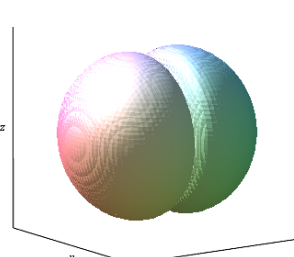
$$|\phi_\mu\rangle = A_\mu x^1 y^1 z^0 e^{-\alpha_\mu r^2}$$



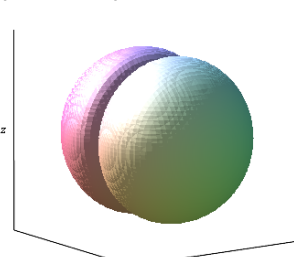
$$|\phi_\mu\rangle = A_\mu x^1 y^0 z^1 e^{-\alpha_\mu r^2}$$



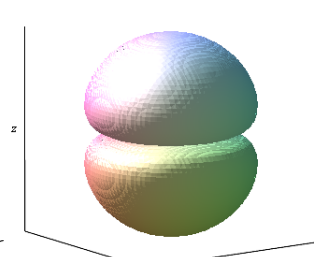
$$|\phi_\mu\rangle = A_\mu x^0 y^1 z^1 e^{-\alpha_\mu r^2}$$



$$|\phi_\mu\rangle = A_\mu x^2 y^0 z^0 e^{-\alpha_\mu r^2}$$



$$|\phi_\mu\rangle = A_\mu x^0 y^2 z^0 e^{-\alpha_\mu r^2}$$



$$|\phi_\mu\rangle = A_\mu x^0 y^0 z^2 e^{-\alpha_\mu r^2}$$

Contracted Gaussians

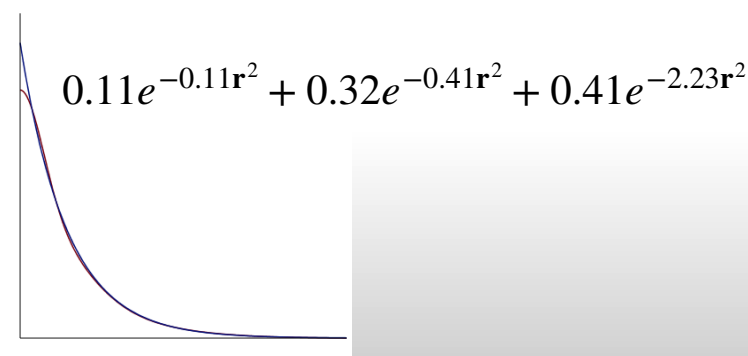
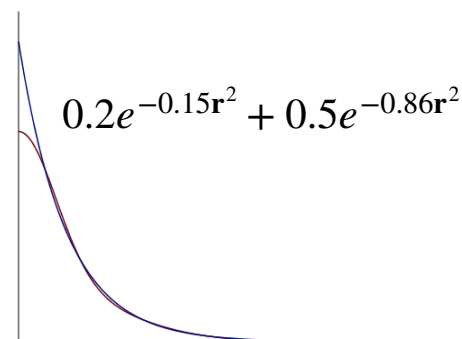
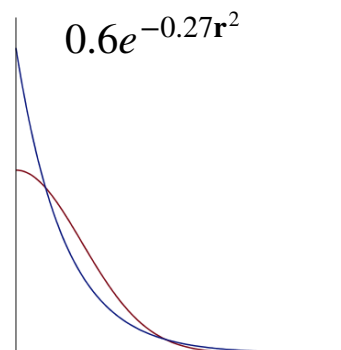
- Individual Gaussian functions referred to as **primitive** Gaussians
- To obtain required accuracy, use fixed linear combinations that approximate Slater function
- Fixed linear combinations are known as **contracted** Gaussians

$$|\psi_i\rangle = \sum_{\mu}^{N_{\text{basis}}} C_{\mu i} |\phi_{\mu}\rangle = \sum_{\mu}^{N_{\text{basis}}} C_{\mu i} G_{\mu}(\mathbf{r})$$

molecular orbital \swarrow SCF optimized coefficients \swarrow atomic orbital \swarrow contracted Gaussian

$$G_{\mu}(\mathbf{r}) = \sum_{\xi}^{N_{\text{prim}}} A_{\xi\mu} g_{\xi}(\mathbf{r})$$

fixed coefficients \swarrow primitive Gaussian





Gaussian Basis Sets

Minimal basis sets

- Atomic orbitals modeled as a single contracted Gaussian are called **minimal** basis sets
- A famous series of minimal basis set are labelled STO-nG, where n is the number of primitive Gaussians in the contraction - in particular STO-3G
- Minimal basis sets used for developing algorithms and initialization of calculations but do not have enough flexibility to describe MOs accurately

How many contracted and primitive Gaussian functions does the STO-3G basis set have for the following atoms and molecules

- Hydrogen (H), Fluorine (F), Sulfur (S)
- Water (H₂O), Methane (CH₄)

Gaussian Basis Sets

Decontracted basis sets

- To add more variational flexibility we can **decontract** the contracted Gaussian

$$|\psi_i\rangle = \sum_{\mu}^{N_{\text{basis}}} C_{\mu i} \sum_{\xi}^{N_{\text{prim}}} A_{\xi\mu} g_{\xi}(\mathbf{r}) \Rightarrow \sum_{\mu}^{N_{\text{basis}}} \sum_{\eta}^{N_{\text{cont}}} C_{\mu i}^{\eta} \sum_{\xi}^{N_{\text{prim}}^{\eta}} A_{\xi\mu} g_{\xi}(\mathbf{r})$$

- Number of integral evaluations remains the same (depends on the number of primitives)
- Number of atomic orbitals has increased, so MO coefficient matrix and Fock matrix are larger (depends on number of contracted Gaussians)
- We now have two basis functions describing the same atomic orbital
- E.g. decontract STO-3G into two basis functions - 2G and 1G

Gaussian Basis Sets

Split-valence

- Nomenclature describing decontraction of basis sets is “(single/double/triple/quadruple etc.) zeta”
- Core molecular orbitals look very much like atomic orbitals and are not affected by bonding
- Valence molecular orbitals look very different to atomic orbitals and require flexibility to describe this change
- Split-valence basis sets use a minimal basis for the core orbitals and a deconstructed basis set for the valence orbitals
- A famous series are the Pople split-valence basis sets
 - 3-21G, 4-31G, 6-31G, 6-311G

How many contracted and primitive Gaussian functions does the 6-31G basis set have for the following atoms and molecules

- Hydrogen (H), Fluorine (F), Sulfur (S)
- Water (H₂O), Methane (CH₄)

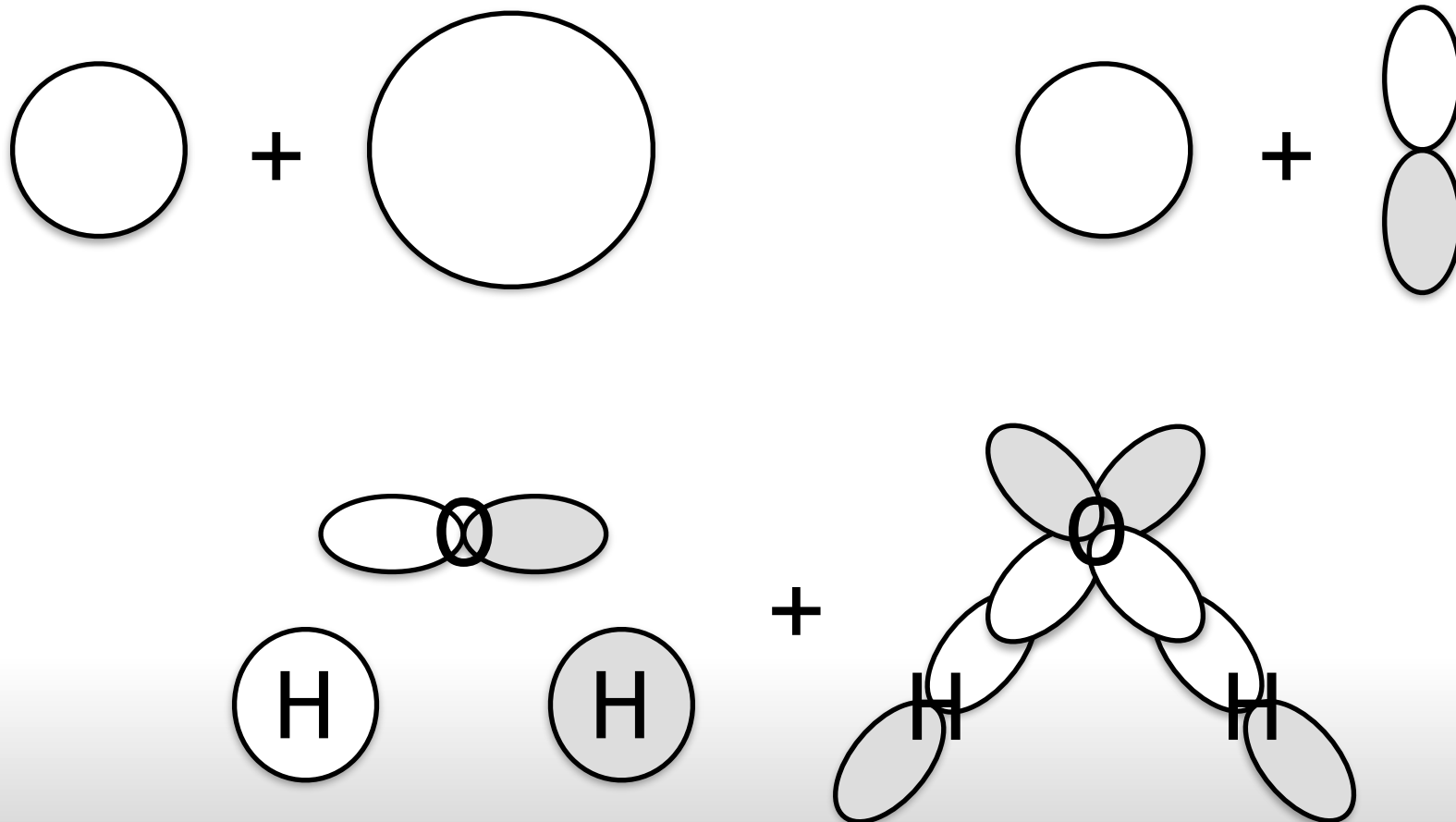
Gaussian Basis Sets

Polarization and diffuse functions

- Even split-valence basis functions do not generally have sufficient flexibility for accurate results. Other types of contracted Gaussians must be included.
- **Polarized** basis sets include Gaussian functions of higher angular momentum than the valence orbitals of each atom. For example they include d orbitals on oxygen
- Have more flexibility to describe bonding and so give better molecular structures
- Examples of polarized basis sets include 3-21G*, 6-31G*, 6-311G**, cc-pVDZ, cc-pVTZ
- **Diffuse** basis sets include Gaussian functions which extend further from the nucleus than the valence orbitals. For example they include 3p orbitals on oxygen
- Required for describing extended electron density such as in anions or hydrogen bonding
- Examples of diffuse basis sets include 4-31+G, 6-31+G
- Balanced basis sets include both polarization and diffuse functions, e.g. 6-31+G*, 6-311++G**, aug-cc-pVDZ

Gaussian Basis Sets

Polarization and diffuse functions





Complete Basis Sets

Complete basis set limit

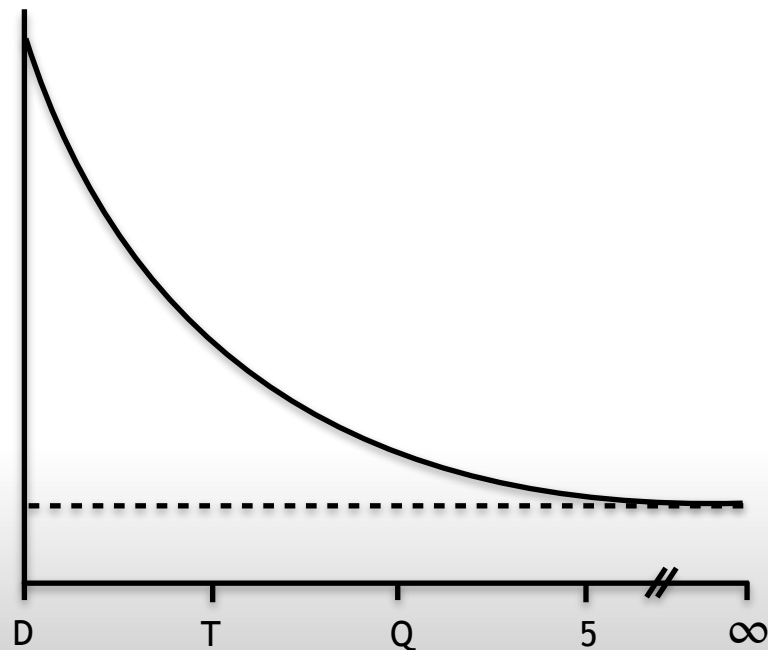
- Basis sets are an approximation required for numerical computation
- Basis sets are complete in the infinite limit so we obtain the exact result (for the given approximations in the model) only when we use infinite number of basis functions
- Molecular orbitals obtained in a finite basis set are the exact solutions for that function space
- The more basis functions we have, the closer we approximate the infinite limit and the better the result we obtain
- Improving the quality of the basis set requires more contracted Gaussian functions so increases the cost of the calculation
- Not possible to reach infinite limit in practice so always perform calculations in a finite basis



Complete Basis Sets

Extrapolation to the infinite basis result

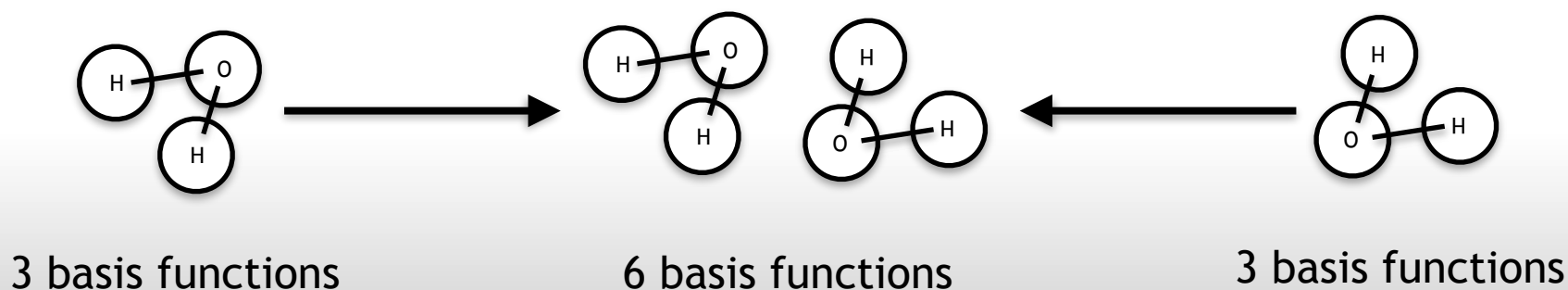
- Fortunately due to the variational theorem we know that the larger the basis, the closer we are to the correct result
- We can extrapolate the result at the infinite basis set limit by extrapolating from calculations with finite basis sets of different sizes
- cc-pVDZ, cc-pVTZ, cc-pVQZ, cc-pV5Z, cc-pV6Z etc. is a systematic basis set series that can be extrapolated to the complete basis set limit



Basis Set Superposition Error

Dissociating fragments

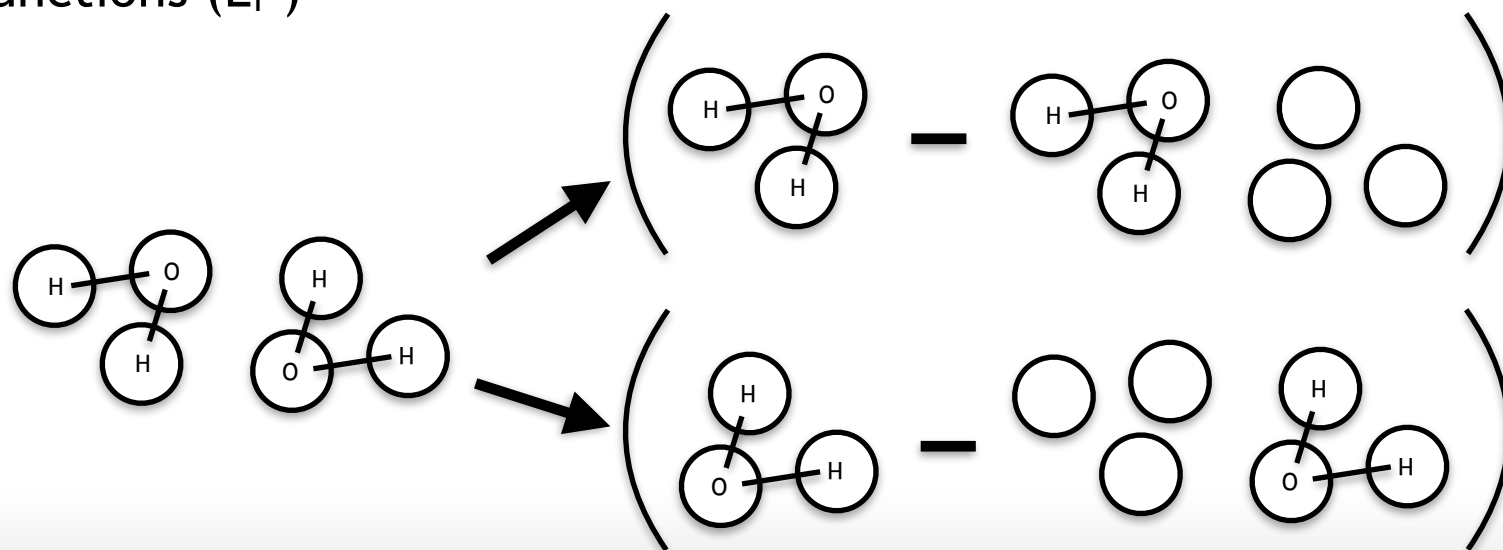
- Additional basis functions increase variational flexibility and obtain lower energies due to variational theorem
- Two isolated fragments will (should) have the twice the energy as one fragment
- But if two fragments are brought close enough together for the basis set to interact, each fragment can “borrow” basis functions from the other monomer
- BSSE artificially lowers the energy of the interacting system and so results in overbinding



Basis Set Superposition Error

Counterpoise correction

- Errors due to BSSE can be estimated through counterpoise correction
- Subtract energy of one fragment with only basis functions centered on the fragment atoms (E_i) from energy of fragment using all basis functions (E_i^*)



$$\Delta E_{cc} = \sum_i^{N_{\text{frag}}} (E_i^* - E_i)$$

$$\Delta E_{\text{int}} = E_{\text{total}} - \Delta E_{cc} - \sum_i^{N_{\text{frag}}} E_i$$



Summary

- Choice of “atomic orbitals” to be made
- Arbitrary functions so want most computationally efficient choice
- Slater functions, gaussian functions, plane waves etc.
- Gaussians do not display cusp or correct decay
- Gaussian product rule enable rapid evaluation of integrals
- Angular and radial part - cartesian or spherical coordinates
- Contracted gaussians approximate Slater orbitals
- Minimal basis sets, decontracted basis sets, split valence basis
- Polarization and diffuse basis functions
- Complete basis set limit and extrapolation
- Basis set superposition error and counterpoise correction