CHEM6085: Density Functional Theory

Lecture 8

Gaussian basis sets

C.-K. Skylaris

Solving the Kohn-Sham equations on a computer

- The SCF procedure involves solving the Kohn-Sham single-electron equations for the molecular orbitals

$$\left(-\frac{1}{2}\nabla^2 + V_{KS}(\mathbf{r})\right)\psi_j(\mathbf{r}) = \varepsilon_j\psi_j(\mathbf{r})$$

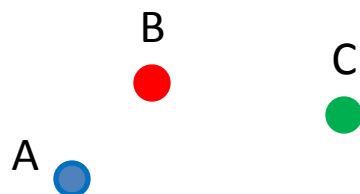- Where the Kohn-Sham potential of the non-interacting electrons is given by

$$V_{\mathrm{KS}}(\mathbf{r}) = V_{\mathrm{ext}}(\mathbf{r}) + V_{\mathrm{Coul}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r})$$

$$= V_{\mathrm{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}' + \frac{\delta E_{\mathrm{xc}}}{\delta n}$$

- We all have some experience in solving equations on paper but how we do this with a computer?

UNIVERSITY OF
Southampton
School of Chemistry

Linear Combination of Atomic Orbitals (LCAO)

- We will **express the MOs** as a linear combination of atomic orbitals (LCAO)

- The strength of the LCAO approach is its general applicability: it can work on any molecule with any number of atoms
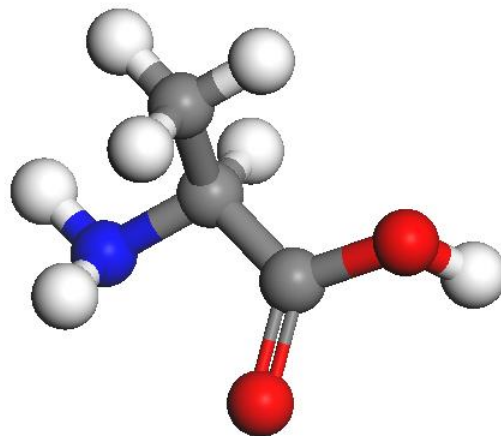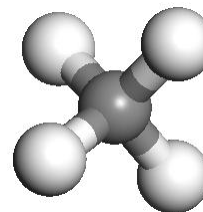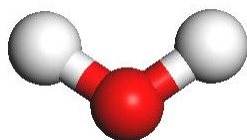
Example:

B

C

A

$$\psi(\mathbf{r}) = c_1\chi_1(\mathbf{r}) + c_2\chi_2(\mathbf{r}) + c_3\chi_3(\mathbf{r}) + c_4\chi_4(\mathbf{r}) + c_5\chi_5(\mathbf{r}) + c_6\chi_6(\mathbf{r}) + c_7\chi_7(\mathbf{r}) + \dots$$

AOs on atom A   AOs on atom B   AOs on atom C

Example: find the AOs from which the MOs of the following molecules will be built

Basis functions

- We can take the LCAO concept one step further:

  - Use a larger number of AOs (e.g. a hydrogen atom can have more than one s AO, and some p and d AOs, etc.). This will achieve a more flexible representation of our MOs and therefore more accurate calculated properties according to the variation principle

  - Use AOs of a particular mathematical form that simplifies the computations (but are not necessarily equal to the exact AOs of the isolated atoms)

- We call such sets of more general AOs **basis functions**

- Instead of having to calculate the mathematical form of the MOs (impossible on a computer) the problem is reduced to determining the MO expansion coefficients in terms of the basis functions

molecular (spatial) orbital $\longrightarrow$ $$\psi_i(\mathbf{r}) = \sum_{\alpha=1}^{N_{BF}} G_\alpha(\mathbf{r})\, C_{\alpha i}$$ $\longleftarrow$ coefficient (a number, to be determined by the SCF calculation)
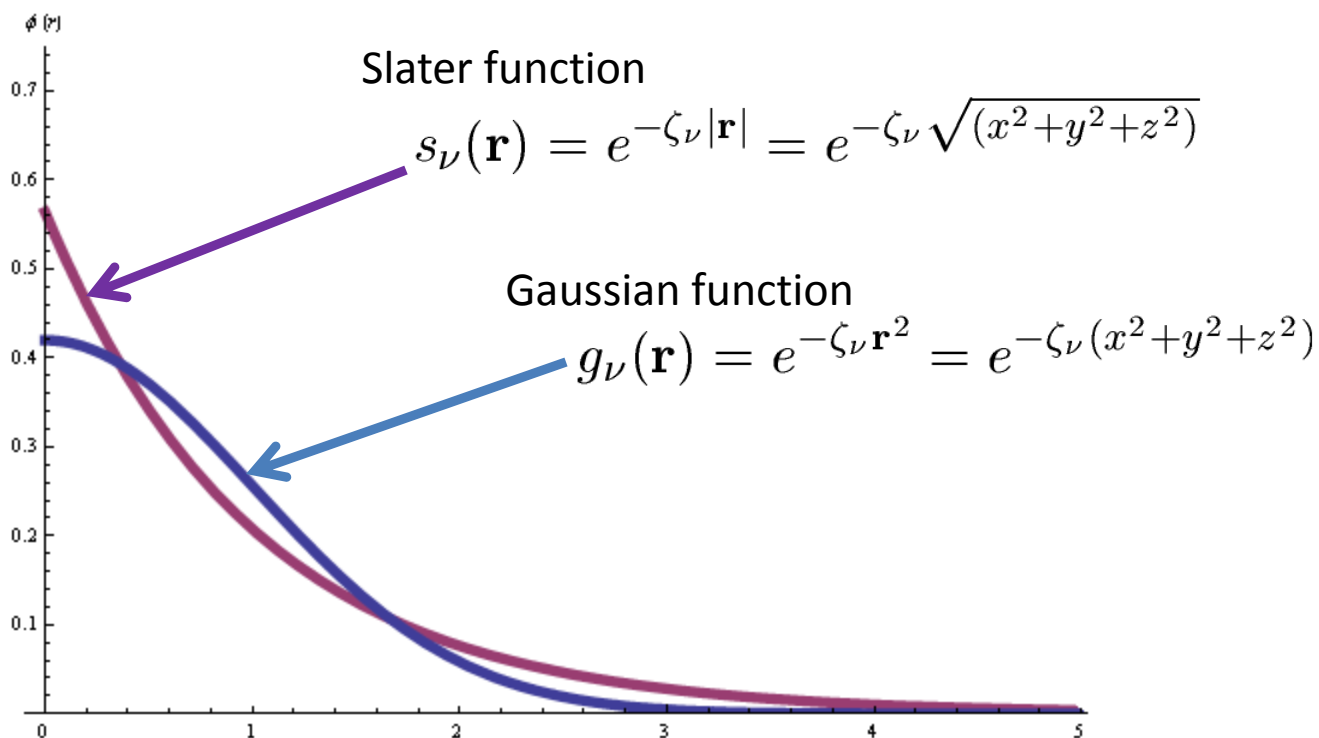
basis function (fixed)

UNIVERSITY OF
Southampton
School of Chemistry

Gaussian basis functions

Functions that resemble hydrogen AOs (Slater functions) are very suitable for expanding MOs because they have the correct shape:

- Near the nucleus (cusp)
- Far from the nucleus (decay like e^{-ar})

However, Gaussian functions are preferred in practice because they allow for efficient computation of molecular integrals (simpler formulas)

Slater function

$$s_\nu(\mathbf{r}) = e^{-\zeta_\nu |\mathbf{r}|} = e^{-\zeta_\nu \sqrt{(x^2+y^2+z^2)}}$$

Gaussian function

$$g_\nu(\mathbf{r}) = e^{-\zeta_\nu \mathbf{r}^2} = e^{-\zeta_\nu (x^2+y^2+z^2)}$$

School of Chemistry

Primitive Gaussian functions

- In the jargon of quantum chemistry a single Gaussian function is called a **primitive Gaussian function**, or **primitive GTO** (Gaussian Type Orbital)

- Some programs use Cartesian primitive GTOs:

$$g_\nu(\mathbf{r}) = x^k y^m z^n e^{-\zeta_\nu \mathbf{r}^2} = x^k y^m z^n e^{-\zeta_\nu (x^2 + y^2 + z^2)}$$

- Other programs use Spherical primitive GTOs

$$g_\nu(\mathbf{r}) = Y_{l,m}(\theta, \phi) \, r^l \, e^{-\zeta_\nu r^2} = Y_{l,m}(\theta, \phi) \, r^l \, e^{-\zeta_\nu (x^2 + y^2 + z^2)}$$

- Spherical and Cartesian functions are the same for up to l=1 (p functions) but differ slightly for l=2 or higher

UNIVERSITY OF
Southampton
School of Chemistry

Cartesian primitive Gaussian functions

Similar to atomic orbitals, we define **Cartesian Gaussian** atomic functions by an angular momentum "quantum number" ℓ and the numbers $\boldsymbol{k}$, $\boldsymbol{m}$ and $\boldsymbol{n}$ *where* $\boldsymbol{l=k+m+n}$

s functions

$$l = 0 \quad \Big\{ \quad e^{-\zeta \mathbf{r}^2}$$

p functions

$$l = 1 \quad \left\{ \begin{array}{l} xe^{-\zeta \mathbf{r}^2} \\ ye^{-\zeta \mathbf{r}^2} \\ ze^{-\zeta \mathbf{r}^2} \end{array} \right.$$

d functions

$$l = 2 \quad \left\{ \begin{array}{l} x^2 e^{-\zeta \mathbf{r}^2} \\ y^2 e^{-\zeta \mathbf{r}^2} \\ z^2 e^{-\zeta \mathbf{r}^2} \\ xy e^{-\zeta \mathbf{r}^2} \\ xz e^{-\zeta \mathbf{r}^2} \\ zy e^{-\zeta \mathbf{r}^2} \end{array} \right.$$

etc.

Spherical primitive Gaussian functions

Similar to atomic orbitals, we define **Spherical Gaussian** atomic functions by an angular momentum "quantum number" ℓ and its components $m=-l,...,l$

s functions

$$l = 0 \left\{ e^{-\zeta \mathbf{r}^2} \right.$$

d functions

p functions

$$l = 1 \left\{ \begin{array}{l} Y_{1,-1}(\theta,\phi)\, r\, e^{-\zeta r^2} \\ Y_{1,0}(\theta,\phi)\, r\, e^{-\zeta r^2} \\ Y_{1,1}(\theta,\phi)\, r\, e^{-\zeta r^2} \end{array} \right.$$

$$l = 2 \left\{ \begin{array}{l} Y_{2,-2}(\theta,\phi)\, r^2\, e^{-\zeta r^2} \\ Y_{2,-1}(\theta,\phi)\, r^2\, e^{-\zeta r^2} \\ Y_{2,0}(\theta,\phi)\, r^2\, e^{-\zeta r^2} \\ Y_{2,1}(\theta,\phi)\, r^2\, e^{-\zeta r^2} \\ Y_{2,2}(\theta,\phi)\, r^2\, e^{-\zeta r^2} \end{array} \right.$$

etc.

UNIVERSITY OF
Southampton
School of Chemistry
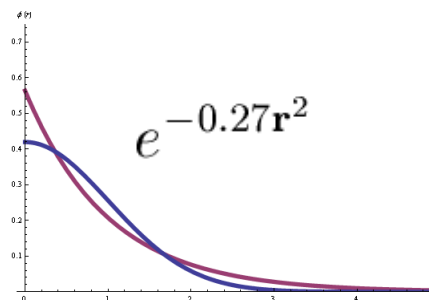
Contracted Gaussian functions

- In practice, fixed linear combinations of "primitive" Gaussian functions are used

- These are called "Contracted Gaussians" (**CGs**):

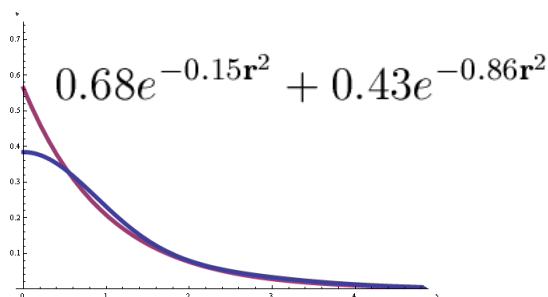$$G_\alpha(\mathbf{r}) = \sum_{\nu=1}^{N_\alpha} c_\nu \, g_\nu(\mathbf{r})$$

- The simplest kind of CGs are the **STO-nG** basis sets

- These basis sets attempt to approximate Slater-type orbitals (STOs) by **n** primitive Gaussians

E.g. STO-nG functions for the 1s orbital of a hydrogen atom

STO-1G
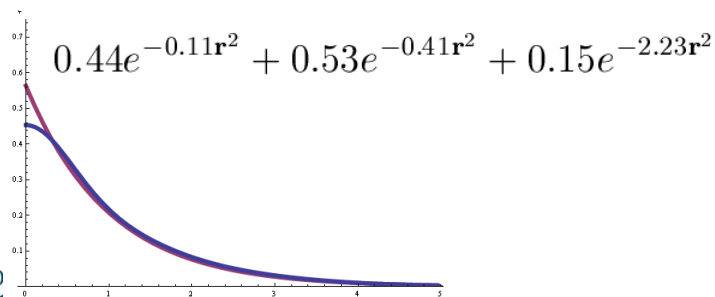
$$e^{-0.27\mathbf{r}^2}$$

STO-2G

$$0.68e^{-0.15\mathbf{r}^2} + 0.43e^{-0.86\mathbf{r}^2}$$

STO-3G

$$0.44e^{-0.11\mathbf{r}^2} + 0.53e^{-0.41\mathbf{r}^2} + 0.15e^{-2.23\mathbf{r}^2}$$

UNIVERSITY O
Southampton
School of Chemistry

Gaussian basis sets

The STO-nG basis sets are rather unsatisfactory as they include only one contracted Gausssian (CG) per atomic orbital. Improved basis sets are obtained by including:

- More than one CG per atomic orbital, e.g.: **DZ** ("double zeta"), **TZ** ("triple zeta"), **QZ** ("quadruple zeta")

- One CG per "core" atomic orbital and more than one for the valence atomic orbitals, e.g.: **SV**, **3-21G**, **4-31G**, **6-31G**, **6-311G**

Write down how many CGs some of the above basis sets will include for the following atoms: H, C, S

And for the following molecules: H_2O, CH_4

Polarisation and diffuse functions

Increasing the number of CGs per atomic orbital will never result in a good quality basis set. Other types of CGs need to be included, such as:

- CGs of angular momentum higher than in the valence orbitals of each atom. These "**polarisation functions**" enhance the "flexibility" of atoms to form chemical bonds in any direction and hence improve calculated molecular structures. Examples: **3-21G*, 6-31G*, 6-31G**, DVP, TZP, cc-pVDZ, cc-pVTZ**

- CGs which extend further from the nucleus than the atomic orbitals. Such "**diffuse functions**" improve the predicted properties of species with extended electronic densities such as anions or molecules forming hydrogen bonds. Examples: **4-31+G, 6-31+G**

- Basis sets are considered "balanced" when they include both polarisation and diffuse functions. Examples: **6-31+G*, 6-311++G**, aug-cc-pVDZ**
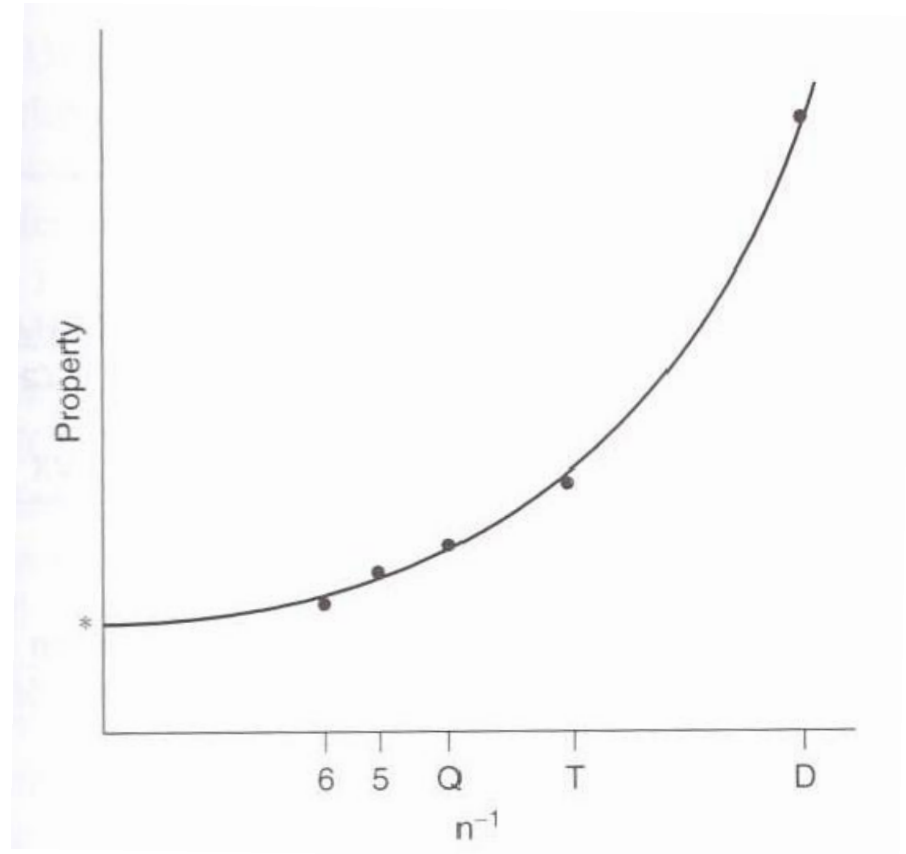
Write down how many CGs some of the above basis sets will include for the following atoms: H, F, S and molecules: H_2O, CH_4
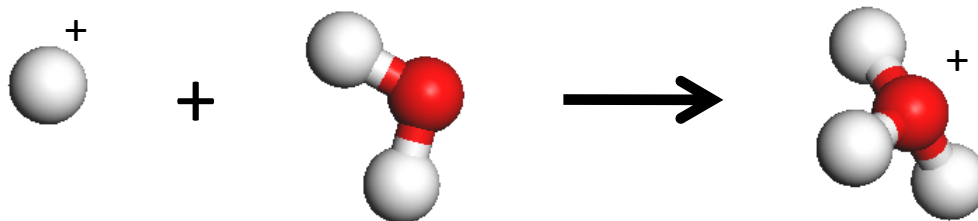
The complete basis set limit

- Basis sets are an approximation introduced in order to solve the KS equations for the MOs on a computer

- The MOs obtained are solutions of the Kohn-Sham equations only within the "function space" of the basis set used (so solutions within the STO-3G set of functions, or the 6-31G set, etc.)

- Improving the quality of the basis set requires increasing the number of CGs

- A **complete** basis set can represent exactly **any molecular orbital**

- Unfortunately, complete basis sets tend to have an infinite number of functions and are therefore not practical for calculations

Extrapolation to the complete basis set limit

- We can estimate the complete basis set result by systematically increasing the number of basis functions and extrapolating to an infinite-size basis set

- The cc-pVDZ, cc-pVTZ, cc-pVQZ, etc, basis sets are an example of a systematic series of basis sets that can be extrapolated to the complete basis set limit

Example calculations: Protonation energy of a water molecule

Basis set	H_2O energy (E_h)	H_3O^+ energy (E_h)	Protonation energy (E_h)	Protonation energy (kcal/mol)
STO-3G	-75.3133	-75.6817	-0.3684	-231.2
STO-6G	-76.0366	-76.4015	-0.3649	-229.0
6-31G	-76.3852	-76.6721	-0.2869	-180.1
6-31++G	-76.4000	-76.6753	-0.2753	-172.7
6-31G**	-76.4197	-76.7056	-0.2859	-179.4
6-31++G**	-76.4341	-76.7078	-0.2738	-171.8

Example: the available choices of basis set in a quantum chemistry program

```
3-21g                          aug-cc-pvqz                    cc-pvtzfi_sf_fw            lanl2dzdp_ecp_polarization
3-21++g                        aug-cc-pvqz_diffuse            cc-pvtzfi_sf_lc            lanl2dz_ecp
3-21gs                         aug-cc-pvt+dz                  cc-pvtzfi_sf_sc            mclean_chandler_vtz
3-21++gs                       aug-cc-pvt+dz_diffuse          cc-pvtz-fit2-1             midi!
3-21gsp                        aug-cc-pvtz                    cc-pvtzpt_sf_fw            midi_huzinaga
3-21gs_polarization            aug-cc-pvtz_diffuse            cc-pvtzpt_sf_lc            mini_huzinaga
4-22gsp                        bauschlicher_ano               cc-pvtzpt_sf_sc            mini_scaled
4-31g                          binning-curtiss_1d_polarization cc-pvtzseg-opt            nasa_ames_ano
6-311g                         binning-curtiss_df_polarization chipman_dzp_+_diffuse     nasa_ames_cc-pcv5z
6-311++g2d_2p                  binning_curtiss_sv             core_val._functions_cc-pcv5z   nasa_ames_cc-pcvqz
6-311g2df_2pd                  binning_curtiss_svp            core_val._functions_cc-pcv6z   nasa_ames_cc-pcvtz
6-311++g3df_3pd                binning_curtiss_vtz            core_val._functions_cc-pcvdz   nasa_ames_cc-pv5z
6-311gs                        binning_curtiss_vtzp           core_val._functions_cc-pcvqz   nasa_ames_cc-pvqz
6-311+gs                       blaudeau_polarization          core_val._functions_cc-pcvtz   nasa_ames_cc-pvtz
6-311gs_polarization           cc-pcv5z                       crenbl_ecp                 partridge_uncontr._1
6-311gss                       cc-pcv6z                       crenbs_ecp                 partridge_uncontr._2
6-311++gss                     cc-pcvdz                       d-aug-cc-pv5z              partridge_uncontr._3
6-311gss_polarization          cc-pcvqz                       d-aug-cc-pv5z_diffuse      pople_2d_2p_polarization
6-31g                          cc-pcvtz                       d-aug-cc-pv6z              pople_2df_2pd_polarization
6-31++g                        cc-pv5+dz                      d-aug-cc-pv6z_diffuse      pople_3df_3pd_polarization
6-31g3df_3pd                   cc-pv5z                        d-aug-cc-pvdz              pople-style_diffuse
6-31g-blaudeau                 cc-pv5z_dk                     d-aug-cc-pvdz_diffuse      pv6z
6-31gs                         cc-pv5zfi_sf_fw                d-aug-cc-pvqz              qmmm_zhang_3-21g_ecp
6-31+gs                        cc-pv5zfi_sf_lc                d-aug-cc-pvqz_diffuse      qmmm_zhang_6-31gs_ecp
6-31++gs                       cc-pv5zfi_sf_sc                d-aug-cc-pvtz              sadlej_pvtz
6-31gs-blaudeau                cc-pv5zpt_sf_fw                d-aug-cc-pvtz_diffuse      sbkjc_vdz_ecp
6-31gs_polarization            cc-pv5zpt_sf_lc                demon_coulomb_fitting      sdb-aug-cc-pvqz
6-31gss                        cc-pv5zpt_sf_sc                dgauss_a1_dft_coulomb_fitting   sdb-aug-cc-pvqz_diffuse
6-31++gss                      cc-pv6+dz                      dgauss_a1_dft_exchange_fitting  sdb-aug-cc-pvtz
6-31gss_polarization           cc-pv6z                        dgauss_a2_dft_coulomb_fitting   sdb-aug-cc-pvtz_diffuse
ahlrichs_coulomb_fitting       cc-pvd+dz                      dgauss_a2_dft_exchange_fitting  sdb-cc-pvqz
ahlrichs_polarization          cc-pvdz                        dhms_polarization          sdb-cc-pvtz
ahlrichs_pvdz                  cc-pvdz_dk                     dunning-hay_diffuse        sto-2g
ahlrichs_tzv                   cc-pvdzfi_sf_fw                dunning-hay_double_rydberg  sto-3g
ahlrichs_vdz                   cc-pvdzfi_sf_lc                dunning-hay_rydberg        sto-3gs
ahlrichs_vtz                   cc-pvdzfi_sf_sc                dz_+_double_rydberg_dunning-hay  sto-3gs_polarization
aug-cc-pcv5z                   cc-pvdz-fit2-1                 dz_dunning                 sto-6g
aug-cc-pcvdz                   cc-pvdzpt_sf_fw                dzp_+_diffuse_dunning      stuttgart_rlc_ecp
aug-cc-pcvqz                   cc-pvdzpt_sf_lc                dzp_dunning                stuttgart_rsc_1997_ecp
aug-cc-pcvtz                   cc-pvdzpt_sf_sc                dzp_+_rydberg_dunning      stuttgart_rsc_ano_ecp
aug-cc-pv5+dz                  cc-pvdzseg-opt                 dz_+_rydberg_dunning       stuttgart_rsc_segmented_ecp
aug-cc-pv5+dz_diffuse          cc-pvq+dz                      dzvp2_dft_orbital          sv_+_double_rydberg_dunning-hay
aug-cc-pv5z                    cc-pvqz                        dzvp_dft_orbital           sv_dunning-hay
aug-cc-pv5z_diffuse            cc-pvqz_dk                     feller_misc._cvdz          svp_+_diffuse_dunning-hay
aug-cc-pv6+dz                  cc-pvqzfi_sf_fw                feller_misc._cvqz          svp_+_diffuse_+_rydberg
aug-cc-pv6+dz_diffuse          cc-pvqzfi_sf_lc                feller_misc._cvtz          svp_dunning-hay
aug-cc-pv6z                    cc-pvqzfi_sf_sc                gamess_pvtz                svp_+_rydberg_dunning-hay
aug-cc-pv6z_diffuse            cc-pvqzpt_sf_fw                gamess_vtz                 sv_+_rydberg_dunning-hay
aug-cc-pvd+dz                  cc-pvqzpt_sf_lc                glendening_polarization    tz_dunning
aug-cc-pvd+dz_diffuse          cc-pvqzpt_sf_sc                hay-wadt_mb_n+1_ecp        tzvp_dft_orbital
aug-cc-pvdz                    cc-pvqzseg-opt                 hay-wadt_vdz_n+1_ecp       wachters+f
aug-cc-pvdz_diffuse            cc-pvt+dz                      hondo7_polarization        wtbs
aug-cc-pvq+dz                  cc-pvtz                        huzinaga_polarization
aug-cc-pvq+dz_diffuse          cc-pvtz_dk                     lanl2dzdp_ecp
```

Basis sets on the Web

- Many kinds of basis sets have been developed over the years
- Most are available for download from websites, such as
  - https://bse.pnl.gov/bse/portal
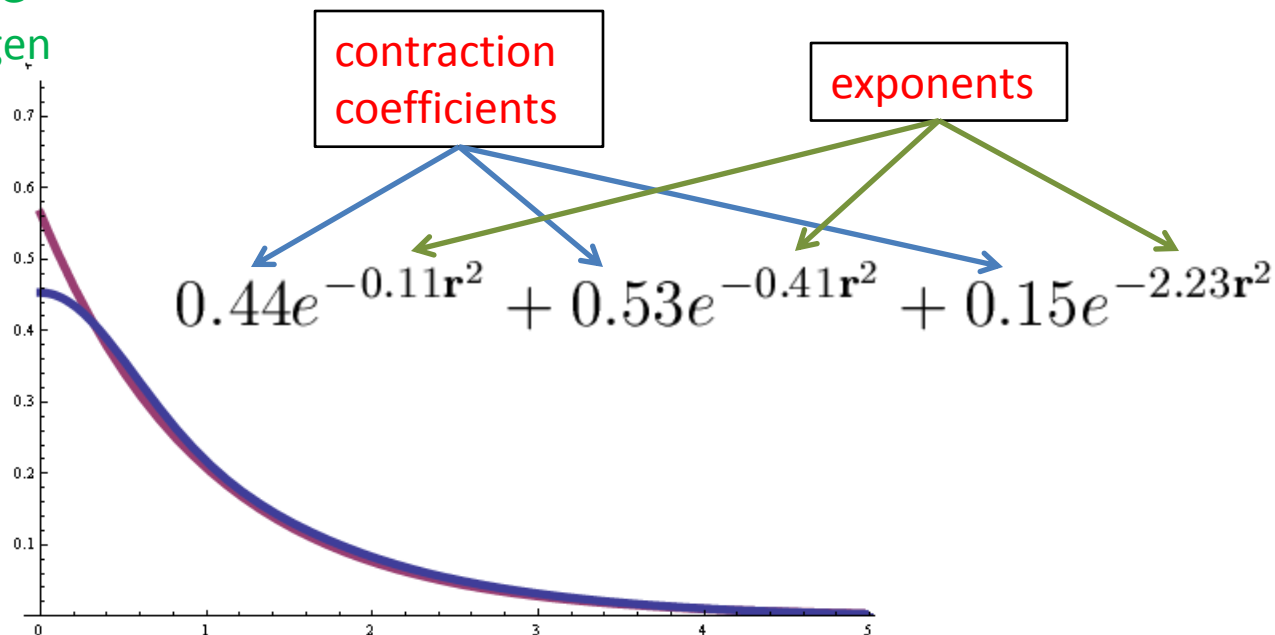 - http://www.emsl.pnl.gov/forms/basisform.html

Exponents and contraction coefficients

Basis sets are essential input data for calculations. Gaussian bases are represented by two kinds of numbers:
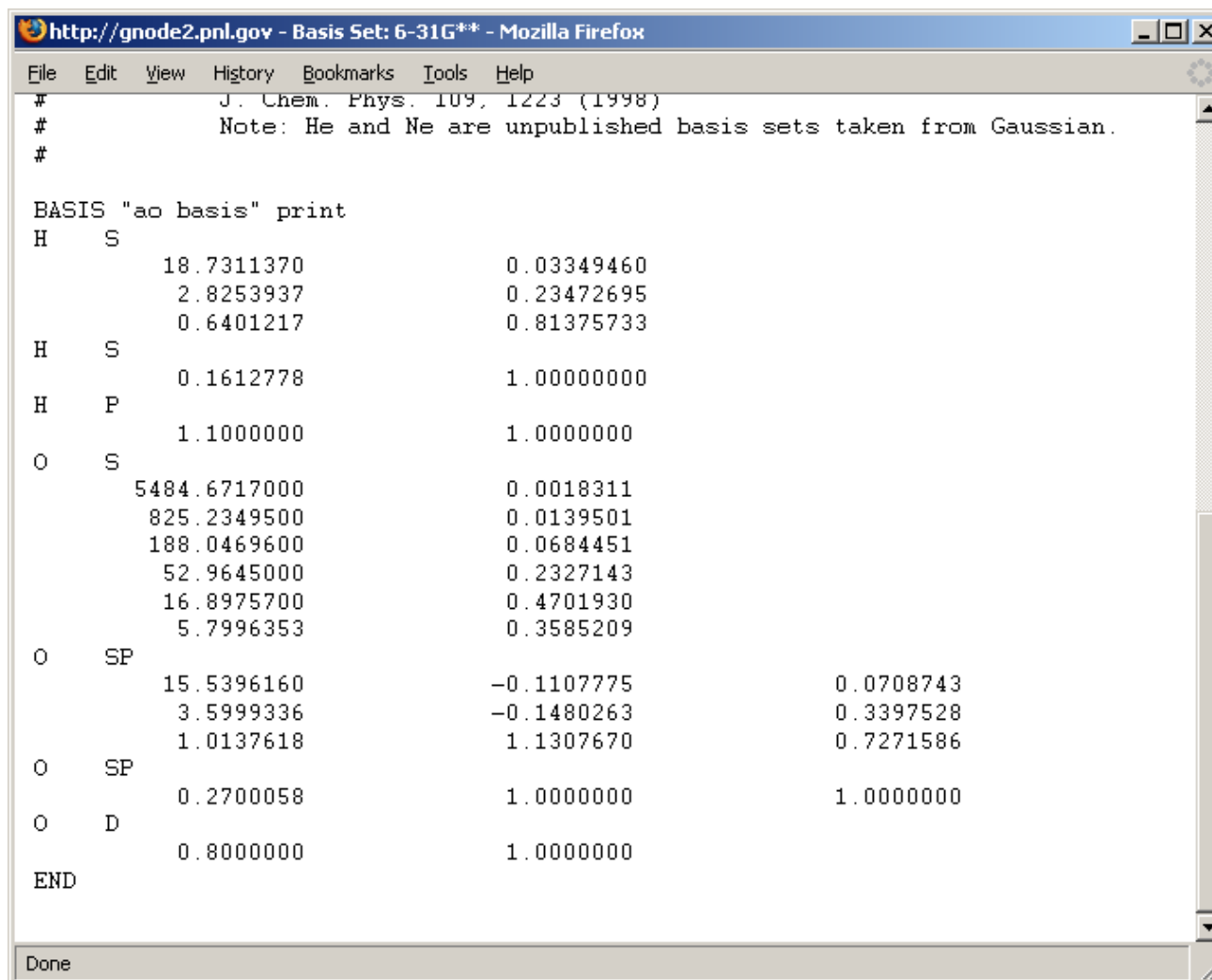
1) Exponents
2) Contraction coefficients

$$G_\alpha(\mathbf{r}) = \sum_{\nu=1}^{N_\alpha} c_\nu \, e^{-\zeta_\nu \mathbf{r}}$$

Example: STO-3G basis for hydrogen

contraction coefficients

exponents

$$0.44 e^{-0.11\mathbf{r}^2} + 0.53 e^{-0.41\mathbf{r}^2} + 0.15 e^{-2.23\mathbf{r}^2}$$

Downloading basis sets

Example: Data needed for calculation on water with 6-31G** basis

Homework

1) Describe how many and what type (s, p, d) primitive and contracted Gaussians you will have in the STO-2G, DZP and 3-21G basis sets for the O atom.

2) The 3-21G* basis set for a carbon atom can be input into a quantum chemistry program using the following parameters (exponents and contraction coefficients):

```
C     S
     172.256000000         0.06176690
      25.910900000         0.35879400
       5.533350000         0.70071300
C     SP
       3.664980000        -0.39589700        0.23646000
       0.770545000         1.21584000        0.86061900
C     SP
       0.195857000         1.00000000        1.00000000
```

Which of the above parameters describe functions for the core electrons, valence electrons and for polarisation? Describe how you can "uncontract" the basis set and what effect this would have on your calculations.

3) Substitute the expression for the basis set expansion of a molecular orbital into the Schrödinger equation for the Kohn-Sham orbitals and derive a matrix representation of the Schrödinger equation, involving the "matrix elements" of the Kohn-Sham Hamiltonian $H_{\alpha\beta} = \int G_\alpha^*(\mathbf{r})\hat{h}_{KS}G_\beta(\mathbf{r})d\mathbf{r}$ and the overlap matrix of the basis functions (which are not orthogonal). This matrix equation can be solved on a computer to obtain the orbital expansion coefficients (diagonalisation of the Hamiltonian matrix) and is part of the traditional SCF procedure.

5-minute quiz

Name :

Date :

Surname:

1) Why are Slater functions more suitable than Gaussians as basis functions?

2) What do we mean when we say that a basis set contains "polarisation functions" and what do we mean by "diffuse functions"?

3) What is a "double-zeta basis set"? Would you expect a triple-zeta basis set to give better results than STO-3G and why?

UNIVERSITY OF
Southampton
School of Chemistry