Tutorial Answers

(1) Describe one dichotomy facing computational chemists and identify commonly employed compromises/approximations employed as a result

- one dichotomy is in **accuracy** vs **resources**, a very high level and accurate theory can be used (such as CCSD(T)), but this requires substantial computational resources, often not available.
- The chemist must then evaluate which information is important and what accuracy is required
- Accurate energies (<1 kJ/mol) do require highly accurate methods with very large basis sets and only small systems (<50 atoms) can be studied
- If geometries are desired these can be obtained at a lower level such as using DFT (with less accurate energies) which uses substantially less computational resource and allowing much larger systems to be studied
- In some cases a higher basis-set can be used on a reaction site and lower basis-sets on peripheral components. For example 6-311G(d) at a TM center, and then 6-31G on the ligand substituents, especially if these are simple alkyl groups where the steric constraints are more important. However, this can cause an imbalance in the density, so such approaches must be applied carefully
- Sometimes the whole structure can be simplified, especially if there are large ligands who's purpose is to stabilise a complex eg kinetically labile ligands eg replace tert-Bu with simple Me or Et groups. It is important to ensure the larger ligands do not influence sterics at a reaction center.
- For heavy atoms a pseudo-potential can be used, and for very heavy elements with relativistic effects this is a necessity. PP, like basis-set, come at different levels of accuracy and expense, Stuttgart being more expensive (and accurate) and LANLMB being the most basic.

(2) Explain the composition of the 6-31+G(d) basis set

- this is a Pople basis set
- the 6-31+G(d) basis set is a split valence basis set, meaning the core and valence electrons are treated differently
- where g_i=gaussian, variable coefficients c_i and fixed coefficients d_i
- this basis has six contracted gaussians for each core electron AO, this means a single variable c_1 with 6 fixed coefficients (d_{1-6})

$$AO=c_1(d_1g_1+d_2g_2+d_3g_3+d_4g_4+d_5g_5+d_6g_6)$$

- The core orbitals do not vary much but need to be described well due to the high energy of core orbitals, hence the low variation but larger number of primitive gaussian basis functions
- for the valence region this basis has a set of four gaussians contracted in a pattern of three and then 1 for each valence electron in the following 3-1 pattern, there are two variables c_1 and c_2.

$$AO=c_2(d_7g_7+d_8g_8+d_9g_9) +c_3g_{10}$$

- The valence electrons need to move around and form bonds, it is important there is flexibility in the basis, however a larger number of functions is not required.
- \+ stands for a diffuse function in this case the diffuse functions are only applied to the non-H atoms and this will be (a set of s and p functions with low exponent)
- Diffuse functions allow electron density to spill out of the occupied orbitals further from the atomic center
- (d) stands for a polarization function in this case the polarisation functions are only applied to the non-H. Polarisation functions are a set of functions with higher angular momentum, d functions for the first row
- Polarisation functions allow electrons to effectively polarize, ie shift to one side of the occupied orbital.

(3) Define and then discuss the basis set limit.

- basis sets improve with an increase in number of functions describing each AO
- improvement is identified by a *decrease* in the *total energy* as basis set increases
- reach "basis set limit" where there is minimal decrease in energy for improvement in basis set
- the basis set limit is different for each type of method, eg HF vs MP2 vs CCSD(T), and we require larger basis sets to reach basis set limit for correlated methods
- one problem that can occur is linear dependency within the basis functions, where one basis function can essentially be described by a linear combination of the other basis functions
- the improvement for properties increases much faster than energy as the basis set gets larger and so the basis set limit for a property calculation can be lower, however the increase may not always lead to a systematic improvement
- one of the disadvantages of Pople basis-sets is that the increase basis-set does not lead systematically to the basis set limit, there are other basis sets which can do this

(4) Modern density functionals are defined by the form of their exchange-correlation functional. For the following DFT functionals: SVWN, BLYP, B3LYP, PBE0 **describe** the hierarchy of functional improvements represented by this series. **Explain** one of the improvements made in more modern functionals and give a specific **example** of such a functional.

- functional acronyms are combination of exchange and correlation functionals
    - SVWN, S is exchange VWN is correlation
    - BLYP and B3LYP, B and B3 are exchange and LYP is correlation
    - PBE0, PBE exchange and PBE correlation
- hierarchy
    - SVWN is a local density functional (LDA) dependent only on the density,
    - BLYP is a generalised gradient approximation (GGA) dependent on the density and the gradient of the density,
    - B3LYP and PBE0 are hybrid GGA functionals that contain some contribution of HF exchange
- one of the more modern functional forms, ONE from:
    - Meta GGAs which include higher order derivatives of the density, in this case the second derivative which contributes to the kinetic energy part of the functional. An example is TPSS
    - Range corrected functionals split the exchange term into a short range (traditional DFT) and long range (exact HF) component. An example is ωB97X.
    - Dispersion corrected functionals, include an ad hoc dispersion component, these come in various stages of improvement D, D2, D3 and then D3-BJ. An example is B3LYP-D3
    - Highly empirically parameterised functionals have been developed by fitting to extensive empirical data. An example is the M06 set of functionals, or M06-2X

(5) On the basis of experimental data a reaction mechanism can be proposed. Discuss the information that can be obtained from theoretical calculations to evaluate such a mechanism.

- the computational process is to calculate critical points, stable reactants, products and intermediates as well as transition states. A reaction profile of energy vs reaction coordinate is obtained.
- the geometry of each complex on the reaction path can be determined, structural deformations that occur during a reaction can be identified, key steric interactions can be evaluated, structural data of labile states such as TS/intermediates inaccessible experimentally can be obtained

- relative energy of stable species can be found. The thermodynamics of the overall reaction can be determined. Energy barriers can be determined and overall a specific lowest energy reaction path can be identified. Specific bond breaking/making energies can be determined
- the energy profiles for different mechanisms and structural information, such as if proposed intermediates are stable, can be used to validate one mechanism over another
- experimentally relevant properties such as NMR chemical shifts, IR vibrations can be used to interpret things like through space interactions and bonding within the molecule. TDDFT can give an indication of UV-vis and ECD spectra. Spectral data can be compared with relevant experiments or higher level calculations to establish accuracy.
- properties can also be obtained and used to understand and evaluate the mechanism, electron density for each structure can be examined for information on bonding: relevant information includes partial charges, densities (density derivatives), the electrostatic potential (ESP) at a surface or at given points in the molecule, molecular orbital (MO) interactions, localisation procedures provide localised orbitals and population analysis.

(6) **Identify** the best computational methods to study each of the systems below, **justify** your answer. If quantum chemistry is the best option, **identify** and **justify** good choices for the method and basis-set.

(a) benchmarking acetal and methyl rotational profiles in acetylthiophene

- a *small system and high accuracy* required indicate that high level methods should be used, QCISD(T) or CCSD(T) with the best basis set accessible, going towards the basis-set limit, for stationary points 6-311++G(2df,2pd) is good however for a potential energy surface scan where a large number of calculations are required a slightly lower method/basis may be required, diffuse functions are necessary because of the long-range "through-space" interactions MP2/6-311++G(d,p)

(b) Nitrile hydrogenation by a Mn- PNP amino pincer complex Mn(CO)2HL where $L=HN(CH_2CH_2P(iPr)_2)_2$.

- studying the *reactivity of larger systems*, makes DFT a good choice, due to the TM a functional with a low level of hybrid exchange is best (B3LYP older choice, the meta-GGA M06 newer choice). As we are studying bond breaking/making and TSs a flexible valence basis-set including polarisation and diffuse functions for the movement of electron density in bonds is appropriate, H is not important so we don't need to expand that part of the basis so 6-311G(d) is a good first choice

(c) reorganisation of an enzyme surrounded by water versus a hydrophobic solvent

- to study *1000s of atoms* that are *moving* in an *explicit solvent* molecular dynamics parametrised for biological systems is best

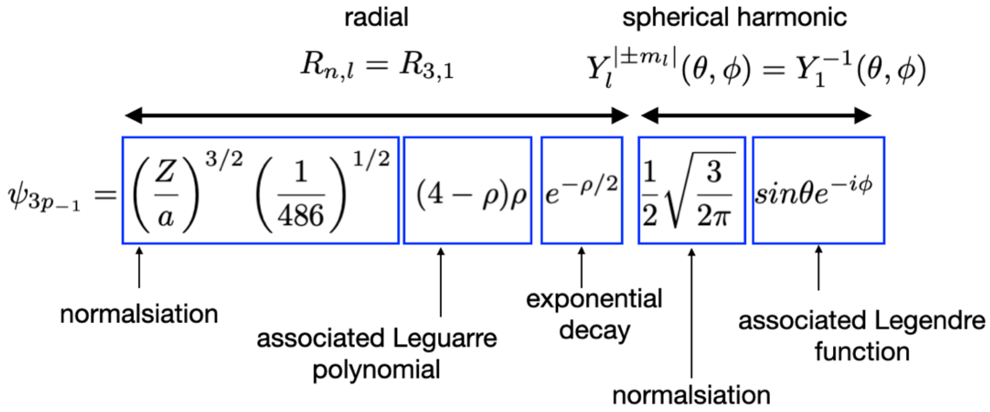
(d) catalytic conversion of N_2 on a close packed (001) Ru catalyst

- *reactivity* requires a method which has explicit electrons, and this is the *solid state*, thus DFT (PBE0 is a good functional) with a plane wave method

(7) For the Hydrogenic wavefunction given below, **define** each of the variables. **Identify** and **explain** the impact of each component/function on the physical form of the wavefunction

$$\psi_{3p_{-1}} = \left(\frac{Z}{a}\right)^{3/2}\left(\frac{1}{486}\right)^{1/2}(4-\rho)\rho e^{-\rho/2}\frac{1}{2}\sqrt{\frac{3}{2\pi}}sin\theta e^{-i\phi}$$

$$\text{where } \rho = \frac{2Z}{na}r \qquad \text{and } a = \frac{4\pi\epsilon_0\hbar^2}{\mu e^2}$$



- define variables
  - r is the radial coordinate, θ, φ are the angular coordinates
  - Z =nuclear charge,
  - ρ is related to the radial component (r)
  - ε_0 is the permittivity of free space
  - μ is the reduced mass
  - π and *ħ* are constants
  - e is the charge on the electron
  - n is the principle quantum number
- individual functional components
  - correct expression of functions with specific quantum numbers
  - 3 is the *principle quantum number* giving the radial shell is the electronic configuration,
  - p is the l=1 *angular quantum number*, p_{-1} has the m_l=-1 *magnetic quantum number*
  - R_{nl} and Y_l^{ml} are the *radial and angular solutions*
  - within R the associated Leguarre function is a *polynomial* and provides the radial nodal structure, and the radial *exponential* provides the envelope for radial decay of the density
  - *normalisation* provides the "magnitude", ensures only 2e occupy each orbital
  - Y is the angular component with the associated Legendre function and exponential ie (θ and φ)giving rise to the pAO shape and phase pattern

Computational Chemistry Equations

Schrodinger Equation (zero potential)

$$-\frac{\hbar^2}{2I}\frac{d^2}{d\phi^2}\Phi(\phi) = E\Phi(\phi)$$

$$-\frac{\hbar^2}{2I}\Lambda^2 Y(\theta,\phi) = EY(\theta,\phi)$$

Hydrogenic Laplacian

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Lambda^2$$

$$\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} = \frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} = \frac{1}{r}\frac{\partial^2}{\partial r^2}r$$

Legendrian

$$\Lambda^2 = \frac{1}{sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{sin\theta}\frac{\partial}{\partial\theta}sin\theta\frac{\partial}{\partial\theta}$$

Angular wavefunctions

$$\Phi(\phi) = Ae^{+im_l\phi} \qquad m_l = 0, \pm1, \pm2\ldots$$

$$\Theta(\theta) = \left\{\left(\frac{2l+1}{2}\right)\frac{(l-|m_l|!)}{(l+|m_l|!)}\right\}^{\frac{1}{2}} P_l^{|m_l|}(cos\theta)$$

Associated Legendre Functions

$$P_0^0(\cos\theta) = 1$$

$$P_1^0(\cos\theta) = \cos\theta$$

$$P_1^1(\cos\theta) = -\sin\theta$$

$$P_2^0(\cos\theta) = \tfrac{1}{2}\left(3\cos^2\theta - 1\right)$$

$$P_2^1(\cos\theta) = -3\sin\theta\cos\theta$$

$$P_2^2(\cos\theta) = 3\sin^2\theta$$

$$P_3^0(\cos\theta) = \tfrac{1}{2}\cos\theta\left(5\cos^2\theta - 3\right)$$

$$P_3^1(\cos\theta) = -\tfrac{3}{2}\left(5\cos^2\theta - 1\right)\sin\theta$$

$$P_3^2(\cos\theta) = 15\cos\theta\sin^2\theta$$

$$P_3^3(\cos\theta) = -15\sin^3\theta.$$

Spherical Harmonics

$$Y_0^0(\theta,\varphi) = \frac{1}{2}\sqrt{\frac{1}{\pi}}$$

$$Y_1^{-1}(\theta,\varphi) = \frac{1}{2}\sqrt{\frac{3}{2\pi}}\sin\theta\, e^{-i\varphi}$$

$$Y_1^0(\theta,\varphi) = \frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta$$

$$Y_1^1(\theta,\varphi) = \frac{-1}{2}\sqrt{\frac{3}{2\pi}}\sin\theta\, e^{i\varphi}$$

$$Y_2^{-2}(\theta,\varphi) = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\theta\, e^{-2i\varphi}$$

$$Y_2^{-1}(\theta,\varphi) = \frac{1}{2}\sqrt{\frac{15}{2\pi}}\sin\theta\cos\theta\, e^{-i\varphi}$$

$$Y_2^0(\theta,\varphi) = \frac{1}{4}\sqrt{\frac{5}{\pi}}(3\cos^2\theta - 1)$$

$$Y_2^1(\theta,\varphi) = \frac{-1}{2}\sqrt{\frac{15}{2\pi}}\sin\theta\cos\theta\, e^{i\varphi}$$

$$Y_2^2(\theta,\varphi) = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\theta\, e^{2i\varphi}$$

Hydrogenic Radial Equations, Generalised Leguarre Functions

$$R_{10} = \left(\frac{Z}{a}\right)^{3/2} 2e^{-\rho/2}$$

$$R_{20} = \left(\frac{Z}{a}\right)^{3/2}\left(\frac{1}{8}\right)^{1/2}(2-\rho)e^{-\rho/2}$$

$$R_{21} = \left(\frac{Z}{a}\right)^{3/2}\left(\frac{1}{24}\right)^{1/2}\rho e^{-\rho/2}$$

$$R_{30} = \left(\frac{Z}{a}\right)^{3/2}\left(\frac{1}{243}\right)^{1/2}(6-6\rho+\rho^2)e^{-\rho/2}$$

$$R_{31} = \left(\frac{Z}{a}\right)^{3/2}\left(\frac{1}{486}\right)^{1/2}(4-\rho)\rho e^{-\rho/2}$$

$$R_{32} = \left(\frac{Z}{a}\right)^{3/2}\left(\frac{1}{2430}\right)^{1/2}\rho^2 e^{-\rho/2}$$

Real and Imaginary orbitals

$$p_0 = p_z = \left(\frac{3}{4\pi}\right)^{1/2} R_{n,1}(r)cos\theta$$

$$p_{+1} = -\left(\frac{3}{8\pi}\right)^{1/2} R_{n,1}(r)sin\theta e^{+i\phi}$$

$$p_{-1} = \left(\frac{3}{8\pi}\right)^{1/2} R_{n,1}(r)sin\theta e^{-i\phi}$$

$$p_x = \frac{1}{\sqrt{2}}(p_{-1} - p_{+1}) = \left(\frac{3}{4\pi}\right)^{1/2} R_{n,1}(r)sin\theta cos\phi$$

$$p_y = \frac{i}{\sqrt{2}}(p_{-1} - p_{+1}) = \left(\frac{3}{4\pi}\right)^{1/2} R_{n,1}(r)sin\theta sin\phi$$