

CHEMISTRY 4021/8021
MIDTERM EXAM KEY — SPRING 2010

1. Molecular Mechanics (50 points)

You are constructing a new molecular mechanics force field for molecules used in the preparation of the Group 13/Group 15 semiconductor indium phosphide; such precursor molecules are composed of H, C, In, and P. Discuss what parameters you will need to include in your force field definition and how you might go about determining optimal values. You need not define a full potential energy function (although you are welcome to do so if you want to) but you should be reasonably specific about the nature of the various parameters.

A generic force field is likely to have bond stretch, angle bend, torsion, electrostatic, and non-bonded interaction terms. Thus, there will be 10 potential kinds of bonds (X–Y, where X,Y are H, C, In, or P), roughly 28 kinds of angles (a complete list is not necessary), some 80+ or so torsions, and 10 each non-bonded and electrostatic pairwise interactions. For bond stretches and angle bends, if the force field is harmonic at least a force constant and an equilibrium value will be needed for each combination (each dependent on the specific atoms involved in the geometric term). For torsions, amplitudes and phase angles will be needed for each Fourier term (four-atom dependence). Non-bonded and electrostatic interactions will require either pairwise terms or individual terms if combining rules are used (thus, e.g., σ and ϵ for a Lennard-Jones potential and a partial charge q if electrostatics derive from q - q interactions). The dielectric constant in the electrostatic term may also be a varied parameter. (The number of parameters will increase if different types of C, In, and/or P atoms are envisioned in the force field definition.)

Optimal parameters will derive either from fitting to experiment using a necessarily arbitrary penalty function or from fitting to high-level (i.e., trustworthy) theoretical data if experimental data are not available. The data in question are likely to be structural, and possibly energetic (e.g., conformational energy differences where known, or heats of formation if atom-equivalents are available for the force field atom types). Optimizing parameters for a subset of atoms first (e.g., C and H) and then relaxing the prior parameters at the same time as the new ones are optimized may deliver a more robust model but one might choose to keep the early parameters fixed for practical reasons (or because one trusts the CH data more than the P and/or In data, for instance).

Other answers are certainly possible if one envisions force field functions for particular terms different from those that I have chosen above.

2. Two-electron Integrals in Semiempirical and ab Initio Hartree-Fock Theory (50 points)

- a. How would the generic two-electron integral $(\mu\nu | \lambda\sigma)$ be written in standard calculus notation? Define all symbols and variables.

$$\iint \varphi_{\mu}(\mathbf{r}_1)\varphi_{\nu}(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \varphi_{\lambda}(\mathbf{r}_2)\varphi_{\sigma}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2$$

where $\mu, \nu, \lambda,$ and σ index atomic orbital basis functions, on each of up to four particular atomic centers, the amplitudes of which are evaluated at positions \mathbf{r}_1 and \mathbf{r}_2 as integration proceeds over all space for these two variables. The product of the various functions is divided by the distance between the two positions for every pair of points evaluated. (Cartesian expansion of this integral would also be fine.)

- b. What is the physical meaning (in “plain” English) of the two-electron integral $(\mu\mu | \lambda\lambda)$?

This integral corresponds to the repulsion between one electron in orbital μ and a second electron in orbital λ (since the square of the atomic orbital basis functions is equivalent to an electron density).

- c. In the complete neglect of differential overlap (CNDO) approximation, one takes $(\mu\nu | \lambda\sigma) = \delta_{\mu\nu}\delta_{\lambda\sigma}(\mu\mu | \lambda\lambda)$. Explain the effect of this approximation in a practical sense. How is the value of $(\mu\mu | \lambda\lambda)$ itself computed? (A general explanation as opposed to a specific formula or formulas is fine for the last part of the question.)

The CNDO approximation annihilates all integrals that are not of the form $(\mu\mu | \lambda\lambda)$. Thus, the number of two-electron integrals is reduced from N^4 to N^2 , where N is the number of atomic orbital basis functions. This integral is not directly computed, but instead approximated by a simple function of the distance between atoms A and B where basis function μ is on atom A and basis function λ is on atom B (this function is equal to r^{-1} at large distance r). For the single center case (both functions on A), the integral is determined from atomic ionization potential and electron affinity data.

(use back of page if required)

- d. Various semiempirical levels of theory adopt less severe approximations than CNDO for the two-electron integrals. Discuss any one such difference.

In the INDO approximation and its variants, the same center integrals ($ss | ss$), ($ss | pp$), ($pp | pp$), ($pp | p'p'$), and ($sp | sp$) are treated each as individual parameters, instead of the first four being represented by a single term and the final one being set equal to zero (as in CNDO).

In the NDDO approximation, one goes beyond this to set

$$(\mu\nu | \lambda\sigma) = \delta_{AB}\delta_{CD}(\mu\nu | \lambda\sigma)$$

where μ is on atom A, ν is on atom B, λ is on atom C, and σ is on atom D (and the Kronecker delta notation implies that the integral is set to zero if either atom A and atom B are not identical or atom C and atom D are not identical). This permits 100 individual integrals to survive if A/B and C/D are heavy atoms carrying one s and three p functions. These 100 non-zero integrals are computed by treating products of basis functions on the atomic center like multipole moments centered on the atomic position and then computing the classical multipole interactions.

- e. What is the motivation for semiempirical levels of theory to adopt approximations for the computation of two-electron integrals?

There are two key motivations. Probably the most important is the opportunity to reduce the enormous number of integrals (N^4 , where N is the number of basis functions) by setting so many equal to zero, thereby enhancing computational efficiency. The second is that ab initio HF theory is known to have large errors associated with ignoring electron correlation. By adopting parameteric forms for the integrals in combination with other approximations involving empirically fitted parameters (for the other terms in the Fock operator), it is possible (in principle) for semiempirical theory to be more accurate than ab initio HF theory.

- f. Ignoring symmetry, how many two electron integrals must be evaluated in a calculation of the energy for a fixed geometry of CH₃OH at the HF/3-21G level? (Showing intermediate steps in arriving at a final answer may avoid loss of credit for any simple arithmetic error.)

The 3-21G basis set has two 1s functions on each H atom (total of $4 \times 2 = 8$ such functions) and 9 functions (a 1s, two 2s, two 2p_x, two 2p_y, and two 2p_z) on each heavy atom (total of $2 \times 9 = 18$ basis functions), so the total number of basis functions N is 26. Thus, the total number of two-electron integrals, ignoring symmetry, is 26^4 or 456,976.

(use back of page if required)

3. Ab Initio Wave Function Theory and Molecular Properties

Select by letter from the list at the bottom of the next page the proper answer to the following questions. There is only one best answer to each question. (5 points each)

- (1) The post-HF model with the most favorable scaling, namely, N^5 , where N is the number of basis functions.

G

- (2) The single-reference post-HF “gold standard” that scales as N^7 , where N is the number of basis functions.

O

- (3) A model is said to have this property if the energy computed for five infinitely separated molecules is equal to five times the energy computed for a single molecule.

M

- (4) To compute the statistical mechanical partition function, these are required.

E

- (5) Something to which correlated methods are more sensitive than the HF model.

J

(2 points partial credit awarded for R)

- (6) Computation of the mixed second derivative of the energy with respect to an external magnetic field **B** and a nuclear spin **I** predicts this.

H

(2 points partial credit awarded for I)

- (7) A useful chemical idea that does *not* have a uniquely defined quantum mechanical operator, so that many different models exist for its prediction.

D

(use back of page if required)

- (8) A model that permits no spin polarization of doubly-occupied orbitals but guarantees wave functions that are eigenfunctions of S^2 .

C

- (9) The proper value of S^2 for a singlet state.

F

- (10) The number of imaginary frequencies associated with a transition-state structure.

A

A: 1	B: UHF	C: ROHF
D: Partial atomic charge	E: An optimized geometry with associated vibrational frequencies	F: 0
G: MP2	H: Nuclear chemical shift	I: Hyperfine coupling constant
J: Basis-set incompleteness	K: 2	L: CISD
M: Size extensivity	N: MNDO	O: CCSD(T)
P: $\pi/2$	Q: Electric dipole moment	R: Spin contamination
S: Acid reflux	T: Slater-type orbitals	U: MP4SDQ

(use back of page if required)