Name:

Q1) What is self-interaction error and what role does it play in density functional theory? (25 points)

DFT computes the electron-electron interaction energy by *starting* from the classical expression for the interaction of a continuous charge density with itself, i.e.,

$$\frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2$$

and then correcting for errors associated with that approximation. One such error is manifest immediately if we consider a one-electron system. In that case, in the physical system, there is *zero* electron-electron repulsion, but the integral show above still delivers a positive value for the interaction of the one electron's worth of density with itself. That is self-interaction error, and it is a very large error. The exchange functional in DFT must eliminate this error, in addition to other associated with the classical density approximation.

Hartree-Fock theory, by construction, has exchange integrals that exactly cancel the self-interaction, so including some HF character in the exchange functional reduces self-interaction error, but it is only eliminated if 100% HF exchange is used. However, functionals including 100% exact exchange are not particularly useful (based on extensive testing), and most hybrid functionals have somewhere between 20 and 56% exact exchange.

Q2) What advantages were there associated with using a pseudopotential on gold in the second problem set as opposed to an all-electron basis set? (25 points)

Pseudopotentials provide 2 key features that can be useful. First, they reduce the total number of electrons for which molecular orbitals must be found. As the orbitals that are eliminated are usually rather dull core orbitals, this contributes to the speed of the calculation. Second, in the case of heavy atoms, the pseudopotentials can be fit so that they mimic core electrons that are behaving in a proper relativistic sense. That is, the core electrons in heavy elements are influenced by (scalar) relativistic effects associated with the very high velocity of the innermost s electrons, and this influences the radial extent of all of the orbitals, and thereby the potential that the nucleus plus the sum of the core electrons exerts on the outermost electrons. If such relativistic effects are not captured in the pseudopotential, then an all-electron method must put them in in some other way in order to be physically realistic.

Q3) What is "dispersion"? How do different modeling techniques—both classical and quantum mechanical—include (or fail to include) dispersion? What kinds of "mistakes" would you expect to make in molecular modeling if your model fails accurately to account for dispersion? (25 points)

Dispersion is the induced-dipole-induced-dipole interaction associated with correlated electronic motion (i.e., dynamic correlation). It is the largest contributing component of medium to long-range (van der Waals) molecular interactions that fall off as r^{-6} (the others being thermally averaged dipole-dipole interactions and dipole-induced-dipole interactions). Dispersion interactions may also be called "London interactions" or "London forces".

Classical force fields tend to include a specific non-electrostatic, non-bonded term, one component of which is designed to include dispersion (e.g., the "6" part of a 6-12 potential). Semiempirical NDDO models that include "corrections" to nuclear repulsion (like AM1 and PM3) effectively do the same, albeit with gaussian functions at molecular contact distances.

Hartree-Fock cannot predict dispersion because it is intrinsically associated with dynamical electron correlation. So, most HF interactions between neutral molecules are unrealistically repulsive. MP2 and other correlated post-HF models, on the other hand, are fine for dispersion, as they include the necessary correlation.

DFT functionals generally fail for dispersion because an exponentially decaying density (which is realistic) cannot be used in a functional in a fashion that gives r^{-6} behavior for attraction. So, like HF, most (the Mo6 series of functionals is an exception) modern functionals are too repulsive for intermolecular interactions. To address this, post hoc dispersion terms are often added, e.g., to make B3LYP-D from B3LYP a non-bonded term looking much like that in a force field is included.

The error one would make with a poor treatment of dispersion is to underestimate non-bonded attractive forces. Thus, for instance, the binding constant of a guest to a host would be predicted to be too small without including dispersion accurately.

Q4) Discuss how one might go about computing the 298 K heat of formation ($\Delta H_{f,298}^{\circ}$) of gaseous 2-methylmorpholine, focusing, obviously, on approaches covered in class (or the reading/videos) so far. Assign a level of confidence to the various protocols that you suggest, if not necessarily in quantitative terms, at least in a "best to worst" characterization. Note how computational constraints might play a role in limiting your range of choices, if at all. (25 points)

An ideal answer would begin with MM, where strain energies may be added to heat-of-formation equivalents for strain-free atom types (determined, for example, from Benson's equivalents or analogs), provided such atomic values are available. At the NDDO level of theory, the MNDO, AM1, PM3, etc. models equate electronic energies with enthalpies, and sum an assembly energy (the opposite of an atomization energy) with experimental atomic heats of formation to compute molecular heats of formation (thereby ignoring zero-point vibrational energy, for example, except to the extent that it is absorbed into parameterization). Other levels of theory, e.g., ab initio HF or post-HF levels, proceed by assembling the enthalpy from the calculations necessary to derive an ideal-gas, rigid-rotator, harmonicoscillator partition function for the molecule, and again summing the assembly enthalpy with the atomic heats of formation to derive a molecular value. Post-HF levels may be used only for the electronic energy portion of the enthalpy (i.e., they may use an HF geometry and thus HF moments of inertia and vibrational frequencies for the rotational and vibrational partition functions, respectively).

A very good answer would then turn to accuracy, and note that force fields are good for what they are good for (and sometimes they are very good), while semiempirical models have substantial noise (accuracy of perhaps \pm 7 kcal/mol might be about right for moderately sized organicky molecules), HF would be disastrous (as the assembly energy is very wrong owing to failure to account for dynamical electron correlation), and post-HF levels can achieve sub-kcal/mol accuracy but quickly become very expensive, thereby limiting their use to rather small molecules. By using isodesmic equations, substantial cancellation of errors in the underlying theory can be achieved. Thus, for instance, one might *compute* the enthalpy change for methylcyclohexane plus morpholine \rightarrow cyclohexane plus 2-methylmorpholine and use the *experimental* gas-phase heats of formation for methylcyclohexane, morpholine, and cyclohexane to derive a best predicted value for 2-methylmorpholine.