

Lecture 37, May 1, 2006

(Some material in this lecture has been adapted from Cramer, C. J. *Essentials of Computational Chemistry*, Wiley, Chichester: 2002; pp. 191-203.)

Solved Homework

Consider first, for an individual vibration

$$U_{\text{vib}} = R \frac{h\omega}{k_{\text{B}}(e^{h\omega/k_{\text{B}}T} - 1)}$$

If we replace the exponential by its power series expansion, truncating after the first power in ω (since ω is going to zero, and thus only the first power term has significant size compared to the others) we have

$$\begin{aligned} U_{\text{vib}} &= R \frac{h\omega}{k_{\text{B}}\left(1 + \frac{h\omega}{k_{\text{B}}T} - 1\right)} \\ &= RT \end{aligned}$$

So, there appears to be no problem with this term blowing up as ω goes to zero.

Now consider

$$S_{\text{vib}} = R \left[\frac{h\omega}{k_{\text{B}}T(e^{h\omega/k_{\text{B}}T} - 1)} - \ln(1 - e^{-h\omega/k_{\text{B}}T}) \right]$$

The first term in brackets is just the one we did above times $1/T$, so we know that it converges to 1 as ω goes to zero. What about the second term? Using the power series expansion, we have

$$\begin{aligned} R \ln(1 - e^{-h\omega/k_{\text{B}}T}) &= R \ln\left(1 - 1 + \frac{h\omega}{k_{\text{B}}T}\right) \\ &= R \ln \omega + R \ln\left(\frac{h}{k_{\text{B}}T}\right) \end{aligned}$$

The second term on the r.h.s. is just a constant, but the first term goes to negative infinity as ω goes to zero. This is actually correct. The entropy of a harmonic oscillator becomes infinite as the vibrational frequency goes to zero (implying that there is really no well anymore; thus, you just have a free particle, and since a free particle has a continuum of states open to it, it has infinite entropy given an infinite volume). However, as the frequency goes to zero (meaning a very “soft” mode in the molecule), the harmonic oscillator is a lousy approximation for the potential energy surface. Typically such modes are torsions (rotations about single bonds), and as such they don’t look like parabolic wells on the PES, they are instead threefold periodic potentials that repeat after one full rotation. The partition function for an internal rotor is different from that for a harmonic oscillator, and does not diverge at zero frequency (a so-called “free rotor”).

Electron Correlation

Multiconfiguration Self-Consistent Field Theory. Hartree-Fock theory makes the fundamental approximation that each electron moves in the static electric field created by all of the other electrons, and then proceeds to optimize orbitals for all of the electrons in a self-consistent fashion subject to a variational constraint. The resulting wave function, when operated upon by the Hamiltonian, delivers as its expectation value the lowest possible energy *for a single-determinantal wave function formed from the chosen basis set.*

So, the question arises of how we might modify the HF wave function to obtain a lower electronic energy when we operate on that modified wave function with the Hamiltonian. By the variational principle, such a construction would be a more accurate wave function. We cannot do better than the HF wave function with a *single* determinant, so one obvious choice is to construct a wave function as a linear combination of multiple determinants, i.e.,

$$\Psi = c_0\Psi_{\text{HF}} + c_1\Psi_1 + c_2\Psi_2 + \dots \quad (37-1)$$

where the coefficients c reflect the weight of each determinant in the expansion and also ensure normalization. For the moment, we will ignore the nature of the determinants, other than the first one, which is the HF determinant. A general expansion does not *have* to include the HF determinant, but since the HF wave function seems to be a reasonable one for many purposes, it is useful to think of it as a leading term in any more complete wave function.

For the majority of the chemical species we’ve discussed thus far, the chief error in the HF approximation derives from ignoring the correlated motion of each electron with every other. This kind of electron correlation is called “dynamical correlation” because it refers to the dynamical character of the electron-electron interactions. Empirically, it is observed that for most systems the HF wave function dominates in the linear combination expressed by eq. 37-1 (i.e., c_0 is much larger than any other

coefficient); even though the correlation energy may be large, it tends to be made up from a large sum of small contributions from other determinants.

However, in some instances, one or more of these other determinants may have coefficients of similar magnitude to that for the HF wave function. It is easiest to illustrate this by consideration of a specific example. Consider the closed-shell singlet wave function for trimethylenemethane (TMM, Figure below). TMM is a so-called non-Kekulé molecule, meaning that the unsaturation in the molecule can't be represented with an integer number of double bonds between unsaturated atoms (in TMM, you can only draw one double bond from the central carbon to one terminal carbon, and after that you're stuck with imagining radicals on the remaining two carbon atoms—of course, there are three equivalent resonance structures for this arrangement, so the MOs must be delocalized). As this is a carbon π system, we could use Hückel theory to analyze the orbitals, and if we do we find that in full D_{3h} symmetry it has two degenerate frontier π orbitals into which only 2 π electrons are available to be placed (after filling the lowest energy orbital with the other two). Following a molecular analog of Hund's rule for atoms, the molecule has a triplet ground state (i.e., the lowest energy state has one spin-aligned electron in each degenerate orbital π_2 and π_3), but here we are concerned with the closed-shell *singlet*.

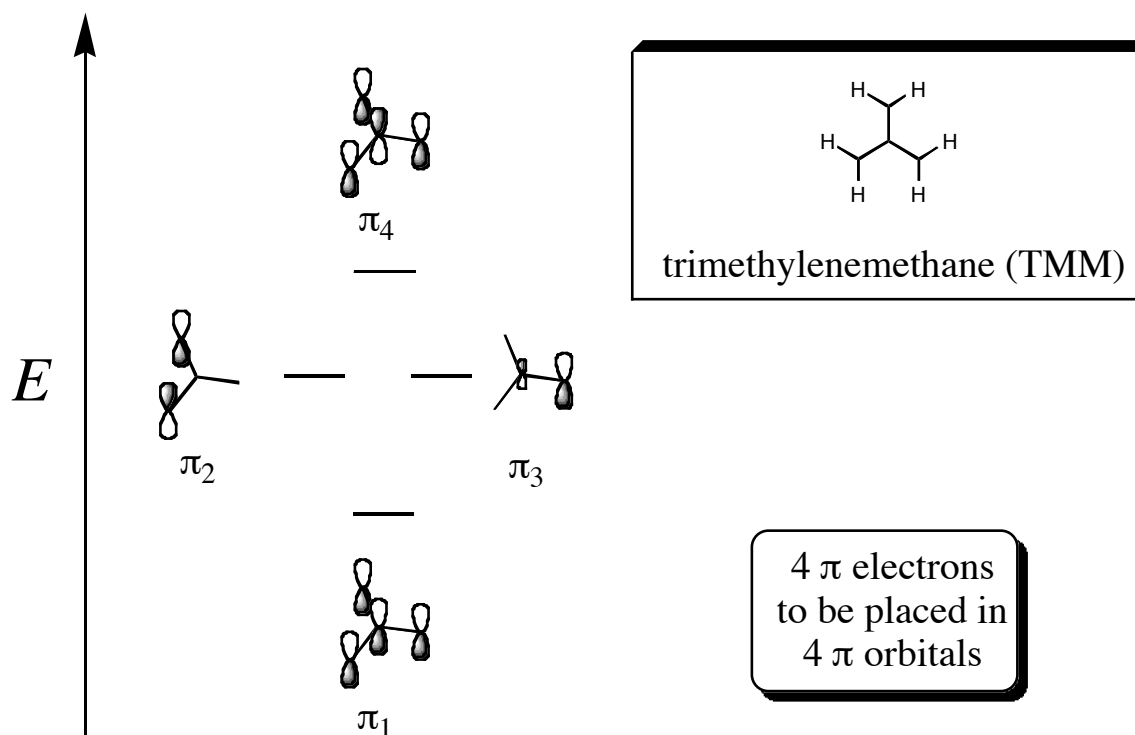


Figure. The π orbital system of TMM. Orbitals π_2 and π_3 are degenerate when TMM adopts D_{3h} symmetry.

If we carry out a restricted HF calculation, one or the other of the degenerate frontier pair must be chosen to be occupied, because the nature of RHF is to doubly occupy every orbital. The calculation will then optimize the shapes of all of the occupied orbitals, and we will end up with a best possible single-Slater-determinantal wave function formed from those MOs. But it should be fairly obvious that an equally good wave function might have been formed if the original guess had chosen to populate the *other* of the two degenerate frontier orbitals. Thus, we might expect each of these two different RHF determinants to contribute with roughly equal weight to an expansion of the kind represented by eq. 37-1. This kind of electron correlation, where different determinants have similar weights because of near (or exact) degeneracy in frontier orbitals, is called “non-dynamical correlation” to distinguish it from dynamical correlation. This emphasizes that the error here is not so much that the HF approximation ignores the correlated motion of the electrons, but rather that the HF process is constructed in a fashion that is intrinsically single-determinantal, which is insufficiently flexible for some systems.

Continuing with our TMM example, let us say that we have carried out an RHF calculation where the frontier orbital that was chosen to be occupied was π_2 . The determinant resulting after optimization will be

$$\Psi_{\text{RHF}} = \left| \cdots \pi_1^2 \pi_2^2 \pi_3^0 \right\rangle \quad (37-2)$$

and orbital π_3 will be empty (i.e., a virtual orbital). We emphasize this by including it in the Slater determinant with an occupation number of zero, although this notation is not standard. We might generate the alternative determinant by keeping the same MOs but simply switching the occupation numbers, i.e.,

$$\Psi_{\pi_2 \rightarrow \pi_3} = \left| \cdots \pi_1^2 \pi_2^0 \pi_3^2 \right\rangle \quad (37-3)$$

An alternative, however, would be to require the RHF calculation to populate π_3 in the initial guess, in which case we would determine

$$\Psi'_{\text{RHF}} = \left| \cdots \pi_1'^2 \pi_2'^0 \pi_3'^2 \right\rangle \quad (37-4)$$

where the prime on the wave function and orbitals emphasizes that, since different orbitals were occupied during the SCF process, the shapes of *all* orbitals will be different comparing one RHF wave function to the other.

If we were to compare the energies of the wave functions from eqs. 37-2, 37-3, and 37-4, we would find the energies of the first and third to be considerably lower than the second. Since the real system has degenerate frontier orbitals, it seems reasonable that the energies of wave functions eq. 37-2 and eq. 37-4 are similar, but why is 37-3 higher? The problem lies in the nature of the SCF process. Only occupied orbitals contribute to the electronic energy—virtual orbitals do not. As such, there is no driving force to

optimize the shapes of virtual orbitals; all that is required is that they be orthogonal to the occupied MOs. Thus, the quality of the shape of orbital π_3 depends on whether it is determined as an occupied or a virtual orbital.

From the nature of the system, however, we would really like π_2 and π_3 to be treated *equivalently* during the orbital optimization process. That is, we would like to find the best orbital shapes for these MOs so as to minimize the energy of the *two*-configuration wave function

$$\Psi_{\text{MCSCF}} = a_1 \left| \cdots \pi_1^2 \pi_2^2 \right\rangle + a_2 \left| \cdots \pi_1^2 \pi_3^2 \right\rangle \quad (37-5)$$

where a_1 and a_2 account for normalization and relative weighting (and we expect them to be equal for D_{3h} TMM). Such a wave function is a so-called “multiconfiguration self-consistent-field” (MCSCF) one, because the orbitals are optimized for a *combination* of configurations (the particular case where the expansion includes only two configurations is sometimes abbreviated TCSCF).

As a technical point, a “configuration” or “configuration state function” (CSF) refers to the molecular spin state and the occupation numbers of the orbitals. For closed-shell singlets, CSFs can always be represented as single determinants with all orbitals doubly occupied. In many open-shell systems, on the other hand, proper CSFs can only be represented by a combination of two or more determinants (recall our old friend the open-shell singlet, which was a linear combination of two single determinants). MCSCF theory is designed to handle *both* multiple configurations and the possible multi-determinantal character of individual configurations.

If we carry out a MCSCF calculation permitting *all* electrons to be redistributed over *all* possible orbitals formed from the basis set, such a calculation is called “full configuration interaction” or “full CI”. Within the choice of basis set, it is the best possible calculation that can be done, because it considers the contribution of every possible CSF.

Indeed, a full CI with an infinite basis set is an “exact” solution of the (non-relativistic, Born-Oppenheimer, time-independent) Schrödinger equation. It is the pinnacle of modern quantum theory (*we got there!*)

Alas, that pinnacle is essentially unreachable for all but the smallest of molecules (perhaps up to 4 atoms), since the number of CSFs in a full CI can be staggeringly large. The trouble is not the number of electrons, which is a constant, but the number of basis functions. Consider methanol (CH_3OH), for example. If we were to use a basis set that has roughly two functions for each atomic orbital of the atoms involved and also puts some d functions (for extra flexibility) on the heavy atoms C and O, the total number of basis functions would be only 38. However, in a full CI that considered every possible way to occupy the 38 MOs with the molecule’s 18 electrons, we would have to optimize

2.4×10^{13} expansion coefficients (!), and this is really a rather small basis set for chemical purposes...

Configuration Interaction with a Single Determinant Reference. In most cases, it is impractical to include all possible excited configurations and do a full CI. However, what if we were to reduce the CI problem by allowing only a limited number of other determinants to be included in the wave function? How many should we include? To proceed in evaluating this question, it is helpful to rewrite eq. 37-1 using a more descriptive notation, i.e.,

$$\Psi = a_0 \Psi_{\text{HF}} + \sum_i^{\text{occ. vir.}} \sum_r a_i^r \Psi_i^r + \sum_{i < j}^{\text{occ. vir.}} \sum_{r < s} a_{ij}^{rs} \Psi_{ij}^{rs} + \dots \quad (37-6)$$

where i and j are occupied MOs in the HF “reference” wave function, r and s are virtual MOs in Ψ_{HF} , and the additional CSFs appearing in the summations are generated by exciting an electron from the occupied orbital(s) indicated by subscripts into the virtual orbital(s) indicated by superscripts. Thus, the first summation on the r.h.s. of eq. 37-6 includes all possible single electronic excitations, the second includes all possible double excitations, etc.

If we assume that we do *not* have any problem with non-dynamical correlation, we may assume that there is little need to reoptimize the HF MOs even if we do not plan to carry out the expansion in eq. 37-6 to its full-CI limit. In that case, the problem is reduced to determining the expansion coefficients for each excited CSF that *is* included.

This situation is rather analogous to the LCAO approach for forming MOs. In LCAO, we need to find the coefficients for basis functions to make an MO. Here, we need to find the coefficients for determinants to make a wave function. So, the usual secular equation approach can be employed. The energies E of N different CI wave functions (i.e., corresponding to different variationally determined sets of coefficients) can be determined from the N roots of the CI secular equation

$$\begin{vmatrix} H_{11} - E & H_{12} & \cdots & H_{1N} \\ H_{21} & H_{22} - E & \cdots & H_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1} & H_{N2} & \cdots & H_{NN} - E \end{vmatrix} = 0 \quad (37-7)$$

where

$$H_{mn} = \langle \Psi_m | H | \Psi_n \rangle \quad (37-8)$$

H is the Hamiltonian operator and the numbering of the CSFs is arbitrary, but for convenience we will take $\Psi_1 = \Psi_{\text{HF}}$ and then all singly-excited determinants, all doubly-excited, etc. Solving the secular equation is equivalent to diagonalizing \mathbf{H} , and permits

determination of the CI coefficients associated with each energy. While this is presented without derivation, the formalism is entirely analogous to that we've used previously in class. Note that different determinants are always orthogonal to one another since the different MOs that are occupied in each one are orthogonal to one another (having been determined from an HF calculation), so $\mathbf{S} = \mathbf{1}$ and energy E only appears in diagonal elements of the CI secular determinant.

To solve eq. 37-7, we need to know how to evaluate matrix elements of the type defined by eq. 37-8. To simplify matters, recall that the Hamiltonian operator is composed only of one- and two-electron operators. Thus, if two CSFs differ in their occupied orbitals by 3 or more orbitals, every possible integral over electronic coordinates hiding in the r.h.s. of eq. 37-8 will include an overlap between at least one pair of different, and hence orthogonal, HF orbitals, and the matrix element will necessarily be zero. For the remaining cases of CSFs differing by two, one, and zero orbitals, one simply has to evaluate the surviving integrals over the one- and two-electron operators in the Hamiltonian.

One simplification is that matrix elements between the HF determinant and singly-excited determinants are always zero. This situation was first proven by Brillouin, so it is called "Brillouin's theorem". To appreciate Brillouin's theorem, let's take a simple example of a two-determinant wave function for a two-electron system with two MOs a and r where the HF determinant is

$$\Psi_{\text{HF}} = |a^2\rangle = a(1)a(2) \quad (37-9)$$

We're considering only the spatial part of the wave function since the Hamiltonian does not depend on spin. The singlet spin function ensures antisymmetry and will integrate to two in all wave function products, even though we are not writing it out explicitly.

The singly excited determinant of interest is

$$\Psi_a^r = |a\bar{r}\rangle = \frac{1}{\sqrt{2}} [a(1)r(2) + a(2)r(1)] \quad (37-10)$$

So, now let us evaluate the matrix element in eq. 37-8

$$\begin{aligned} \langle \Psi_{\text{HF}} | H | \Psi_a^r \rangle &= \frac{1}{\sqrt{2}} \langle a(1)a(2) | H | a(1)r(2) + a(2)r(1) \rangle \\ &= \frac{1}{\sqrt{2}} \left[\langle a(1)a(2) | H | a(1)r(2) \rangle \right. \\ &\quad \left. + \langle a(1)a(2) | H | a(2)r(1) \rangle \right] \\ &= \sqrt{2} \langle a(1)a(2) | H | a(1)r(2) \rangle \end{aligned} \quad (37-11)$$

where the simplification at the end comes from the equivalence of the two different expectation values in the middle line: they differ only by electron labeling.

Evaluation of the final integral (let's assume a single nucleus for simplicity of notation) gives

$$\begin{aligned}
\langle a(1)a(2)|H|a(1)r(2)\rangle &= \\
\left\langle a(1)a(2)\left|-\frac{1}{2}\nabla_1^2-\frac{1}{2}\nabla_2^2-\frac{Z}{r_1}-\frac{Z}{r_2}+\frac{1}{r_{12}}\right|a(1)r(2)\right\rangle &= \\
\left\langle a(1)\left|-\frac{1}{2}\nabla_1^2\right|a(1)\right\rangle\langle a(2)|r(2)\rangle + \left\langle a(2)\left|-\frac{1}{2}\nabla_2^2\right|r(2)\right\rangle\langle a(1)|a(1)\rangle & \\
+ \left\langle a(1)\left|-\frac{Z}{r_1}\right|a(1)\right\rangle\langle a(2)|r(2)\rangle + \left\langle a(2)\left|-\frac{Z}{r_2}\right|r(2)\right\rangle\langle a(1)|a(1)\rangle & \quad (37-12) \\
+ \left\langle a(1)a(2)\left|\frac{1}{r_{12}}\right|a(1)r(2)\right\rangle &= \\
0 + \left\langle a(2)\left|-\frac{1}{2}\nabla_2^2\right|r(2)\right\rangle + 0 + \left\langle a(2)\left|-\frac{Z}{r_2}\right|r(2)\right\rangle + \left\langle a(2)\left|\frac{\langle a(1)|a(1)\rangle}{r_{12}}\right|r(2)\right\rangle &
\end{aligned}$$

The orthonormality of the HF MOs simplifies the first 4 terms. The operator in the last term is the Coulomb operator J_2 , which runs over all other electrons (there's only one in this case) and computes the repulsion between those other electrons and the two basis functions for which it is being evaluated for electron 2.

Note that if we combine the surviving terms into a single Dirac bracket we have

$$\begin{aligned}
\langle a(1)a(2)|H|a(1)r(2)\rangle &= \left\langle a(2)\left|-\frac{1}{2}\nabla_2^2-\frac{Z}{r_2}+J_2\right|r(2)\right\rangle \\
&= \langle a|f|r\rangle \\
&= \langle a|\varepsilon_r r\rangle \\
&= \varepsilon_r \langle a|r\rangle \\
&= 0
\end{aligned} \quad (37-13)$$

where the terms in the first line on the r.h.s. simply define the Fock operator for electron 2 (there is no exchange operator in our example because there are no same-spin electrons to worry about). Since orbital r is an eigenfunction of the fock operator with energy eigenvalue ε_r , the integral simplifies to an overlap integral between a and r , but that is zero.

A formal proof of Brillouin's theorem for the completely general case is notationally more formidable, but still reduces to the matrix element between the HF determinant and the excited determinant being $\langle a | f | r \rangle$, which is always zero.

It is *not* the case that arbitrary matrix elements between *other* determinants differing by only one occupied orbital are equal to zero. Nevertheless, the CI matrix in a broad sense is reasonably sparse, as illustrated in the below figure. With that in mind, next time we will consider the question of which excitations to include in a "non-full" CI.

	Ψ_{HF}	Ψ_i^a	Ψ_{ij}^{ab}	Ψ_{ijk}^{abc}
Ψ_{HF}	E_{HF}	0	dense	0
Ψ_i^a	0	dense	sparse	very sparse
Ψ_{ij}^{ab}	d e n s e	sparse	sparse	extremely sparse
Ψ_{ijk}^{abc}	0	very sparse	extremely sparse	extremely sparse

Figure. Structure of the CI matrix as blocked by classes of determinants. The HF block is the (1,1) position, the matrix elements between the HF and singly excited determinants are zero by Brillouin's theorem, and between the HF and triply and higher excited determinants are zero as well. In a system of reasonable size, remaining regions of the matrix become increasingly sparse, but the number of determinants in each block grows to be extremely large. Thus, the (1,1) eigenvalue is most affected by the doubles, then by the singles, then by the triples, etc.

Homework

To be solved in class:

For molecular hydrogen, H_2 , in a minimal basis, there is a filled σ orbital and a virtual σ^* orbital. A CI calculation includes only state 1, $|\sigma^2\rangle$, and state 2, $|\sigma^{*2}\rangle$. Show that the CI ground-state energy for H_2 is lower than the HF energy (you don't have to reduce anything to actual numbers; this can be done simply from knowledge of the matrix elements in general form).

To be turned in for possible grading May 5: none—look for inspiration for your haiku.