

T3. Ab Initio Hartree-Fock Calculations

The point of the empirical parameters in semiempirical calculations was to cut down on the number of electron-electron repulsion integrals that needed to be computed. In ab initio calculations one simply calculates them all. This inevitably means that ab initio calculations take much longer than semiempirical ones. Some modification of basis sets is generally employed in order to make the calculations tractable.

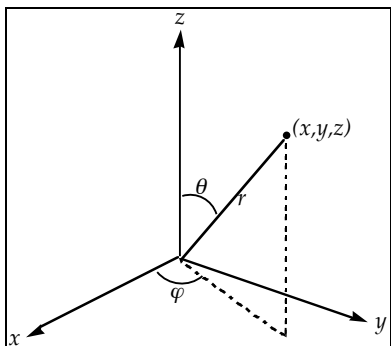
We have specified before that our molecular orbitals are commonly expressed as linear combinations of atomic orbitals (the LCAO approximation):

$$\psi_i = \sum_{\mu} c_{\mu i} \phi_{\mu}$$

but we have not said what the analytical forms of the atomic orbitals are. For the hydrogen atom or for monatomic ions with only one electron, the atomic orbitals can be found exactly.

$$\phi_{n,l,m_l} = R_{n,l}(r)Y_{l,m_l}(\theta,\varphi)$$

The wave functions are conveniently expressed in spherical polar coordinates:



The indices n , l , and m_l are the principal, azimuthal, and magnetic quantum numbers, respectively. The first few examples of the functions $R_{n,l}$ and $Y_{l,m}$ are tabulated below and on the next page. The $R_{n,l}$ functions define the radial part of each wave function, while the $Y_{l,m}$ functions (which are called spherical harmonics) define the angular part. The true spherical harmonics are complex. They can be used in linear combinations that are real, but these linear combinations do not have the proper m_l

quantum number.

n	l	$R_{n,l}$
1	0	$2\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-\frac{Zr}{a_0}}$
2	0	$\left(\frac{Z}{2a_0}\right)^{\frac{3}{2}} \left(2 - \frac{Zr}{a_0}\right) e^{-\frac{Zr}{2a_0}}$
2	1	$\left(\frac{Z}{2a_0}\right)^{\frac{3}{2}} \frac{1}{\sqrt{3}} \frac{Zr}{a_0} e^{-\frac{Zr}{2a_0}}$
3	0	$\left(\frac{Z}{3a_0}\right)^{\frac{3}{2}} \frac{2}{27} \left(27 - \frac{18Zr}{a_0} + \frac{2Z^2r^2}{a_0^2}\right) e^{-\frac{Zr}{3a_0}}$
3	1	$\left(\frac{Z}{3a_0}\right)^{\frac{3}{2}} \frac{4}{27\sqrt{2}} \left(\frac{6Zr}{a_0} - \frac{Z^2r^2}{a_0^2}\right) e^{-\frac{Zr}{3a_0}}$
3	2	$\left(\frac{Z}{3a_0}\right)^{\frac{3}{2}} \frac{4}{27\sqrt{10}} \frac{Z^2r^2}{a_0^2} e^{-\frac{Zr}{3a_0}}$

Y_{l,m_l} (complex)	Y_l (real)
$Y_{0,0} = \frac{1}{\sqrt{4\pi}}$	$s = \frac{1}{\sqrt{4\pi}}$
$Y_{1,0} = \sqrt{\frac{3}{4\pi}} \cos \theta$	$p_z = \sqrt{\frac{3}{4\pi}} \cos \theta$
$Y_{1,1} = \sqrt{\frac{3}{8\pi}} \sin \theta e^{i\varphi}$	$p_y = \sqrt{\frac{3}{4\pi}} \sin \theta \sin \varphi$
$Y_{1,-1} = \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\varphi}$	$p_x = \sqrt{\frac{3}{4\pi}} \sin \theta \cos \varphi$
$Y_{2,0} = \sqrt{\frac{5}{16\pi}} (1 - 3 \cos^2 \theta)$	$d_{z^2} = \sqrt{\frac{5}{16\pi}} (1 - 3 \cos^2 \theta)$
$Y_{2,\pm 1} = \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\varphi}$	$d_{xz} = \sqrt{\frac{15}{4\pi}} \sin \theta \cos \theta \cos \varphi$
$Y_{2,\pm 2} = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\varphi}$	$d_{yz} = \sqrt{\frac{15}{4\pi}} \sin \theta \cos \theta \sin \varphi$
	$d_{x^2-y^2} = \sqrt{\frac{15}{16\pi}} \sin^2 \theta \cos 2\varphi$
	$d_{xy} = \sqrt{\frac{15}{16\pi}} \sin^2 \theta \sin 2\varphi$

In practice, the radial nodes make little difference to the description of bonding, but considerably complicate the calculation of electron-electron repulsion integrals. Slater proposed a series of radial functions lacking nodes:

$$R_{n,l} = (2\xi)^{n+\frac{1}{2}} [(2n)!]^{-\frac{1}{2}} r^{n-1} e^{-\xi r}$$

$$\xi = \frac{Z-s}{n^*}$$

Z = atomic number

s = shielding parameter

n^* = effective principal quantum number

n	n^*
1	1.0
2	2.0
3	3.0
4	3.7
5	4.0
6	4.2

The shielding parameter s is given an empirical value based on n and l , and on the number of other electrons of various types (s, p, d etc.) in the atom.

In combination with the angular functions, these radial functions generate a set of atomic-orbital-like basis functions that have subsequently become known as Slater-type-orbitals (STOs). However, for large molecules even the STOs do not provide a computationally tractable basis set. They are consequently commonly further approximated by Gaussian-type orbitals (GTOs), which have a $\exp[-\zeta r^2]$ distance dependence, instead of $\exp[-\zeta r]$. The reason is that the electron-electron repulsion integrals can then be solved analytically, whereas with STOs they can only be solved numerically. The down side is that one GTO

is a very bad approximation to a STO – the GTO falls off too quickly with increasing r , and lacks the proper “cusp” at the nucleus. This problem is handled by using several GTOs to represent one STO. Even though this greatly increases the total number of basis functions, it still turns out to be computationally quicker than trying to use the STO functions themselves.

When we come to consider molecules rather than atoms, the wave functions for the electrons do not look much like atomic orbitals. Even though we use the LCAO approximation, there is no reason to restrict ourselves to the orbitals that we consider “natural” for isolated atoms. Thus, while a 1s orbital is an exact (aside from relativity) and sufficient description of the electron wave function of a ground-state hydrogen atom, a good description of the electron wave function for a hydrogen atom in a molecule may require that we used p or d orbitals, and will generally be improved if we have several orbitals of each type in our basis set. However, for every orbital that we add to the basis set, the calculation gets bigger. There are groups of more-or-less standard basis sets from which one can select on the basis of the size of the molecule and the availability of computational resources.

A General Representation of Gaussian Basis Sets

A common, if not completely specific, way of specifying the basis set used in a calculation is as illustrated in the following example: (14s,9p,5d)/[8s,4p,2d]. This is sometimes called a contraction scheme – it shows how the primitive Gaussians are contracted together to simulate STOs. The particular example illustrated here means that for the atom in question there were 8 s -type functions, 4 sets of p -type functions and 2 sets of d -type functions. The 8 s -type functions were approximated by linear combinations of 14 Gaussians, the 4 sets of p -type functions by 9 sets of Gaussians, and the 2 sets of d -type functions by 5 sets of Gaussians. The reason this is not completely specific is that it doesn't tell us how the Gaussians were divided up to make the basis functions. A more specific specification would be one such as this: [62111111/5211/32]. The s , p , and d functions are separated from each other by slashes. Within each group the number of digits tells us how many basis functions of that type there are, and the value of the digit tells us how many primitive Gaussians were used to represent each basis function.

The Pople Basis Sets

J.A. Pople has been one of the principal contributors to ab initio electronic structure theory. He and his coworkers have developed a series of basis sets, with names that one commonly sees in the chemistry literature:

STO-3G: This is the smallest and least accurate of the Pople basis sets. It uses three Gaussian functions to represent each Slater-type orbital, and uses only the atomic valence basis functions. Thus for hydrogen it would be (3s)/[1s] and for carbon and other elements in that row in the periodic table it would be (6s,3p)/[2s,1p], or [33/3] in the more specific representation.

3-21G: This is the smallest of the so-called *split-valence* basis sets. It recognizes that the bonding in molecules is not well described by the use of just atomic valence basis functions, and so provides the valence shell with two sets of each type of orbital, having different ζ values – i.e. different distance dependencies. One often hears such a basis set described as being of “valence double-zeta” quality. For hydrogen, the contraction scheme is (3s)/[2s] or [21]. For elements

in the carbon row the contraction scheme is (6s,3p)/[3s,2p] or [321/21]. The total number of primitive Gaussians required for each carbon atom is 15: that's 3 for the 1s orbital, 2 for 2s and 1 for 2s', 2 each for 2p_x, 2p_y, and 2p_z, and 1 each for 2p_x', 2p_y', and 2p_z'.

6-31G* or 6-31G(d): This is one of the most commonly used basis sets for medium-sized organic molecules. It is the first in the Pople series to use *polarization functions*, i.e. basis functions of higher angular momentum than would be necessary for the wave function of the hydrogenic atom. In this case, the polarization functions are *d* functions on atoms in the carbon row of the periodic table. For hydrogen the contraction scheme is (4s)/[2s] or [31] and for carbon it is (10s,4p,1d)/[3s,2p,1d] or [631/31/1]. By default in ab initio programs such as Gaussian 03 and GAMESS, the *d* functions are the six Cartesian *d* functions $d_{x^2}, d_{y^2}, d_{z^2}, d_{xz}, d_{yz},$ and d_{xy} rather than the more familiar spherical-harmonic *d* orbitals. Thus, each carbon atom is represented by 28 primitive Gaussians.

6-31G or 6-31G(d,p):** This is like 6-31G(d) except that there are now *p*-type polarization functions added to the hydrogens.

6-311G* or 6-311G(d): This is a basis set of valence triple-zeta quality. Its contraction scheme is (11s,5p,1d)/[4s,3p,1d] or [6311/311/1]. For some reason, the default in Gaussian 03 uses 5 spherical-harmonic *d* functions instead of the 6 Cartesian *d* functions used for the 6-31G(d) basis set. With the default setting, there are then 31 primitive Gaussians for each carbon atom.

6-31+G* or 6-31+G(d): The "+" basis sets add so-called *diffuse functions* to the previously described mix. The diffuse functions represent very loosely held electrons, and so the orbitals are big (small ζ). They are also typically *s*-type functions. In 6-31+G(d) they are added only to the non-hydrogen atoms. In 6-31++G(d) they are added to hydrogens as well. They are defined for larger basis sets too, so 6-311++G(d,p) is a recognized combination. Diffuse-function basis sets are primarily of value for describing anions and electronic excited states, although there is evidence that their inclusion is useful in most density-functional calculations, too (*vide infra*).

The Dunning Correlation-Consistent Basis Sets

The Pople basis sets were created at time when Hartree-Fock calculations were the best most people could do. The exponents (ζ) and the contraction coefficients (linear-combination coefficients for the primitive Gaussians) were optimized for HF-type calculations. Subsequently, more sophisticated calculations, designed to capture at least some electron correlation (see T4 and T5), have become computationally feasible. Dunning has therefore created a completely new group of basis sets whose adjustable parameters have been optimized for post-Hartree-Fock calculations. They are the cc-pVDZ, cc-pVTZ, etc. basis sets. The "cc" stands for "correlation consistent," the "p" stands for "polarization," indicating that polarization functions are included in all of the basis sets, and the "VDZ" or "VTZ" stand for "valence-double-zeta" or "valence-triple-zeta," respectively. Here are the sets of basis functions in each:

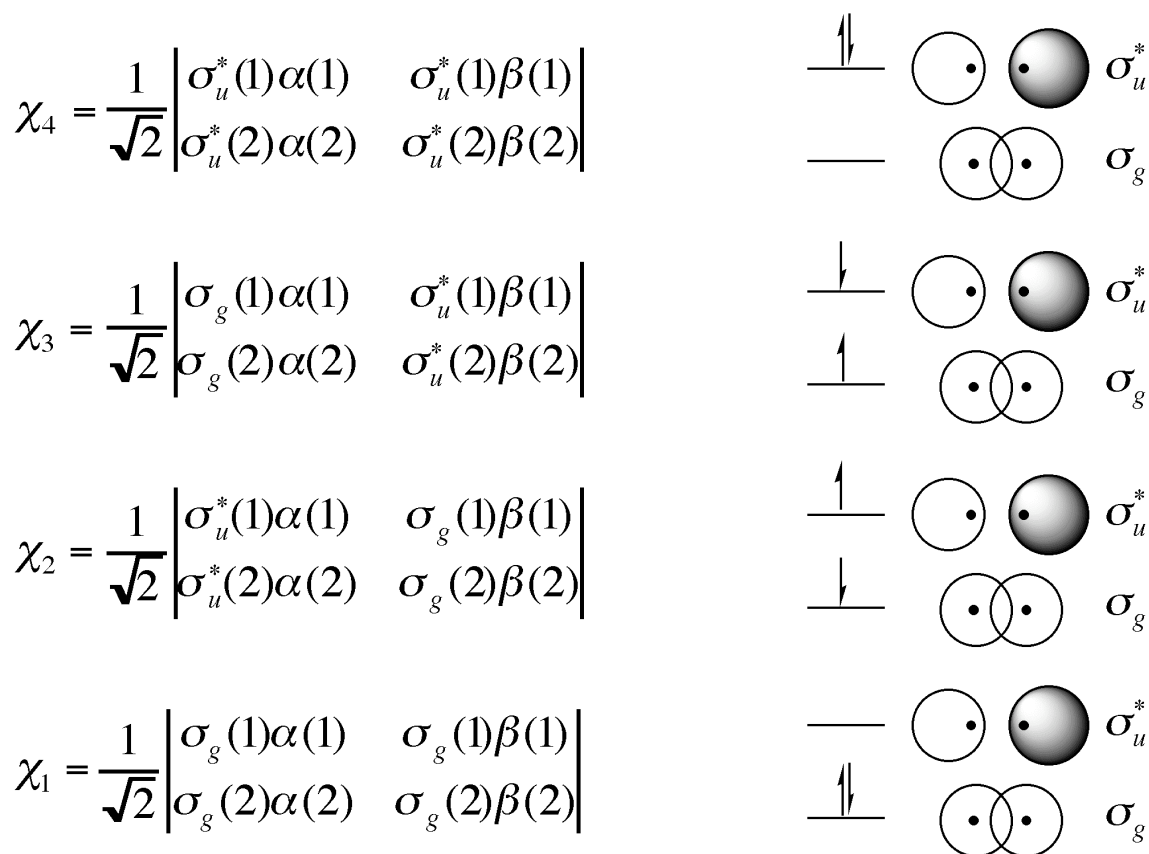
Elements	cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z
H, He	2s,1p	3s,2p,1d	4s,3p,2d,1f	5s,4p,3d,2f,1g
Li – Ne	3s,2p,1d	4s,3p,2d,1f	5s,4p,3d,2f,1g	6s,5p,4d,3f,2g,1h
Na – Ar.	4s,3p,1d	5s,4p,2d,1f	6s,5p,3d,2f,1g	7s,6p,4d,3f,2g,1h

As usual, these basis functions are contracted from Gaussians. Thus, for example, a carbon atom in the cc-pVTZ basis set has the following contraction scheme: (22s,7p,2d,1f)/[4s,3p,2d,1f].

T4. Post-Hartree-Fock Calculations

T4.1 Introduction to the Electron Correlation Problem

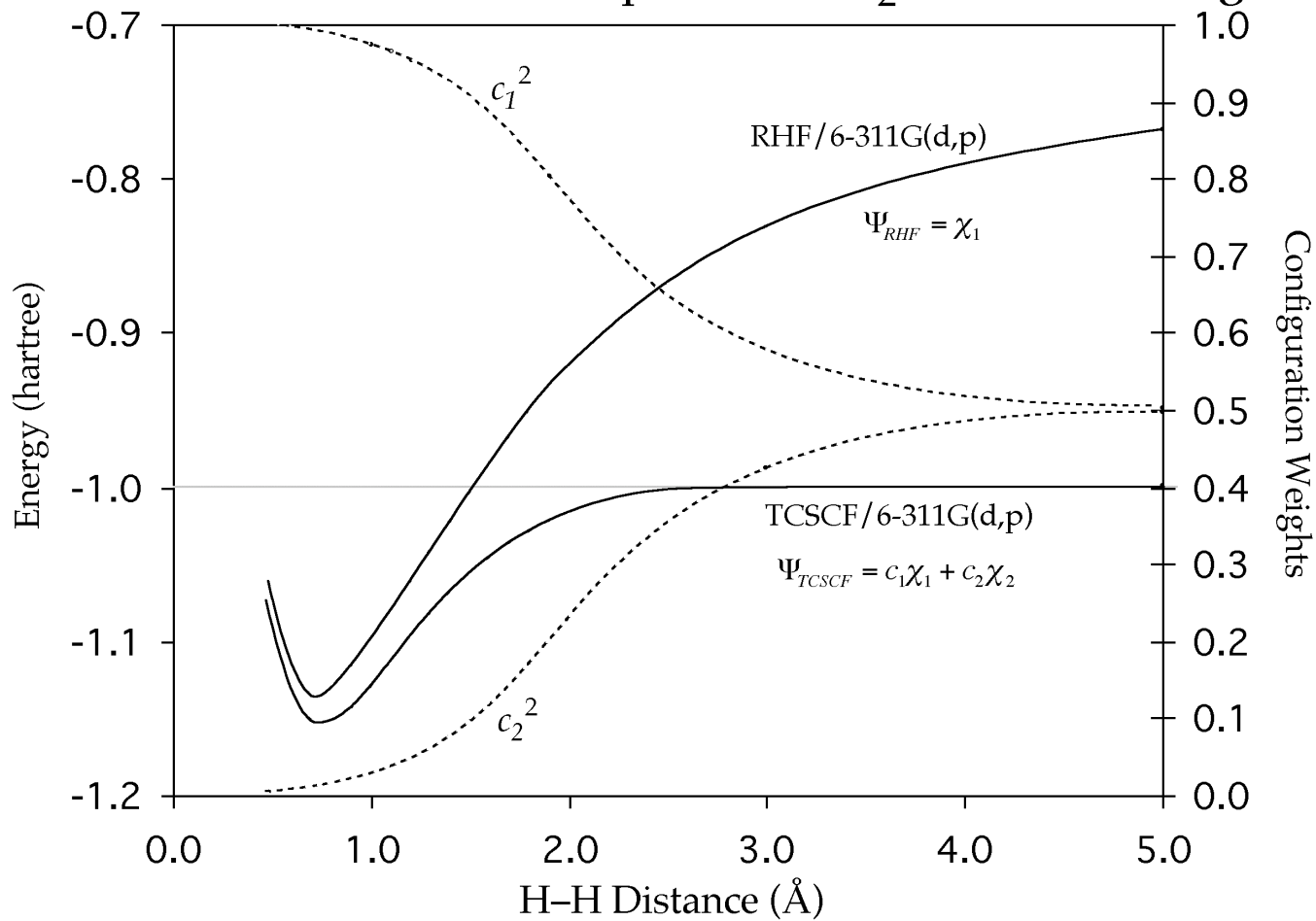
Let's review the RHF description of the H_2 molecule. This is a simple case to analyze in detail because we have only two electrons and two molecular orbitals amongst which to distribute them. There are four ways to do this with singlet coupling (i.e. with one α -spin and one β -spin electron):



Each of the configurations $\chi_2 - \chi_4$ can be viewed as being derived from the configuration χ_1 by one or more substitutions of σ_u^* for σ_g in the Slater determinant. Configurations χ_2 and χ_3 are single-substitution configurations (a substitution replaces σ_g by σ_u^* in an entire column of the determinant), while χ_4 is a double-substitution configuration.

The RHF description of the bonding in H_2 uses just the configuration χ_1 . One could imagine a better description coming from some linear combination of χ_1 and χ_4 (the other two configurations have different symmetry and could not mix), but how much of an improvement would this be? The answer is shown on the next page. (Note: there is a subtle but important point about how the TCSCF calculations shown in the figure are done. See the MCSCF section T4.5.)

RHF vs TCSCF Descriptions of H₂ Bond Breaking



$$\chi_1 = \frac{1}{\sqrt{2}} \begin{vmatrix} \sigma_g(1)\alpha(1) & \sigma_g(1)\beta(1) \\ \sigma_g(2)\alpha(2) & \sigma_g(2)\beta(2) \end{vmatrix}$$

$$\chi_2 = \frac{1}{\sqrt{2}} \begin{vmatrix} \sigma_u^*(1)\alpha(1) & \sigma_u^*(1)\beta(1) \\ \sigma_u^*(2)\alpha(2) & \sigma_u^*(2)\beta(2) \end{vmatrix}$$

One sees that for the H_2 molecule at its equilibrium geometry, the RHF description agrees pretty well with the TCSCF description. Consistent with this, the contribution from the double-substitution configuration (i.e. the magnitude of c_2^2) is quite small at this geometry. However, the RHF description is completely inadequate for describing the breaking of the H–H bond. By definition, the potential energy of an electron in an isolated ground-state hydrogen atom is exactly -0.5 hartree, and so the potential energy of an H_2 molecule should asymptotically approach -1 hartree as we stretch the H–H bond. The TCSCF wave function does that, but the RHF one does not. Indeed there is no real sign of asymptotic leveling off of the energy even out to 5 \AA ! The reason is that the RHF wave function describes a pair of electrons that spend half their time in the vicinity of one H nucleus or the other *no matter how far apart the nuclei are!* In other words, the dissociation limit for the RHF wave function is $1/2(H \bullet \bullet H)$ and $1/2(H^+ -H)$. The latter contribution shouldn't be there in the completely dissociated state, but it should be there in H_2 molecule. The TCSCF wave function allows us to add in a second configuration that subtracts out the ionic contribution in a smooth way as the H–H distance increases.

The incorrect description of bond breaking by RHF wave functions is due to their neglect of *electron correlation*. The particular variety that is needed to describe a pair of dissociated H atoms is called *nondynamic electron correlation*. In brief, any time one has a number of electrons that are incompletely filling a set of nearly-degenerate energy levels, the wave function cannot be properly described by a single Slater determinant. There will be several determinants (two in the case we just discussed) that have linear-combination coefficients with nearly equal absolute magnitude. The kinds of species that require this sort of description are singlet-state biradicals (triplet states can usually be reasonably described by a single UHF determinant), singlet radical pairs, and antiaromatic annulenes such as cyclobutadiene, and, most importantly, many transition states.

The need for multideterminant wave functions for description of nondynamic electron correlation provides a real problem for the Dewar-type semiempirical MO models. These were supposed to have their electron correlation built in by parameterization. The problem is that there exists no parameterization that can describe nondynamic electron correlation properly. If one uses a multideterminant wave function (which can be done in MOPAC) then one double counts some of the electron correlation, making the resulting species too stable. If one uses single-determinant wave functions then the entire description of the species is wrong. Notice, however, that the Pople-type semiempirical methods (CNDO and INDO) do not have this problem because they were never parameterized to include electron correlation in the first place.

Even when there is not a near-degeneracy issue, the neglect of electron correlation in RHF wave functions can still have negative consequences. For example, important aspects of non-bonded interactions, such as van der Waals forces, can only be described with wave functions that include electron correlation. If one chose to address such problems with multideterminant wave functions, one would find that the RHF configuration would have by far the largest weight in the linear combination, but although the weights of the other configurations may be small, there would be a lot of them. Together these small individual contributions from a large number of configurations can have an important overall effect. This is generally known as *dynamic electron correlation*.

The various methods that have been devised for handling the electron-correlation problem differ in how well they deal with the dynamic and nondynamic components, as described in the following sections.

T4.2 Configuration Interaction (CI)

We can imagine writing our multideterminant wave function as follows:

$$\Psi = C_0\chi_0 + \sum_i \sum_a C_i^a \chi_i^a + \sum_{i < j} \sum_a \sum_b C_{ij}^{ab} \chi_{ij}^{ab} + \sum_{i < j < k} \sum_a \sum_b \sum_c C_{ijk}^{abc} \chi_{ijk}^{abc} + \dots$$

The C s are just linear-combination coefficients, and the χ s are Slater determinants. The first one, χ_0 , is just the HF determinant. It is called the reference configuration, because all of the other determinants are defined with respect to it, using the molecular orbitals that were optimized for the HF solution. The second term contains all of the single-substitution determinants, where the subscript i and superscript a mean that molecular orbital i of the reference configuration is replaced by virtual orbital a . (The virtual orbitals are the empty molecular orbitals in the HF calculation.) The third term contains all of the double-substitution determinants, and so on.

If we could evaluate the energy of such a wave function, we would get the *exact* solution to the Schrödinger equation for the particular basis set that we have selected. In other words, we would have completely recovered from having made the orbital approximation in the first place. However, this so-called Full Configuration Interaction (FCI) is not practical for most molecules, as we will see.

In order to find the energy of our wave function, we need to evaluate:

$$\frac{\int \Psi^* H \Psi d\tau}{\int \Psi^* \Psi d\tau}$$

as usual. We can substitute the expression for Ψ in terms of the Slater determinants, and then use the Variation Principle to optimize the CI coefficients. The result will look superficially like the one that we got for simple Hückel theory:

$$\sum_{j=1}^n C_j (H_{kj} - ES_{kj}) = 0$$

Or, in matrix formulation :

$$\mathbf{E} = \mathbf{C}^{-1} \mathbf{H} \mathbf{C}$$

However, the coefficients that we are trying to optimize are not the LCAO coefficients that define how the molecular orbitals are made up of the atomic orbitals; rather they are CI coefficients that define how the CI wave function is made up of Slater determinants. Let's look at the CI matrix, \mathbf{H} , that we are trying to diagonalize:

	0	1	2	3	4	5
0						
1						
2						
3						
4						
5						

All of the unshaded blocks of the CI matrix are zero. The 0,1 and 1,0 blocks corresponding to integrals of the type $\int \chi_0 H \chi_i^a d\tau$, are zero because of Brillouin's theorem, which says that single-substitution configurations, by themselves, make no contribution to improving the HF wave function. (Overall they make a contribution because the 1,2 and 2,1 blocks are nonzero.) The remaining zero blocks of the CI matrix have that value because H is a function containing only one- and two-electron terms. If the two Slater determinants in a block H_{ij} differ in more than two molecular orbitals, the matrix element vanishes.

If we could diagonalize the CI matrix, we would get the energies and wave functions of all of the electronic states of a particular spin symmetry (singlets, triplets, etc.), and the results would be the best possible (nonrelativistic) solutions we could get for whatever basis set we have used. The reason that is not practical is the vast number of configurations that one generates by permuting electrons among molecular orbitals. For a water molecule described with a 6-31G(d) basis set, there are approximately 10^{10} singlet Slater determinants one can write. That means our CI matrix would be of dimension $10^{10} \times 10^{10}$. Not even modern computers can diagonalize matrices of that size. So, what is to be done?

T4.2.1 Limited Configuration Interaction

One could imagine truncating the full CI expression by ignoring all determinants that involve more than a certain number of substitutions from the reference determinant. These limited versions of configuration are given abbreviations, CIS, CID, CISD, CISDT, CISDTQ, etc. that indicate which sets of substitutions have been retained (S=single, D=double, T=triple, Q=quadruple). Each CI calculation has a number of roots, corresponding to different electronic states. Because of Brillouin's theorem, the smallest truncation that can improve the ground-state wave function is CID. If you request a CIS calculation in the Gaussian03 package, it automatically assumes you are asking for higher roots of the CI determinant, since the lowest energy one would be identical with the HF solution.

Limited CI is variational, meaning that an improvement of the level of theory is guaranteed to give a lower energy (eg. $E_{\text{CISDT}} < E_{\text{CISD}}$). However, limited CI has the major flaw that it is not *size consistent*. Size consistency means that the following intuitively obvious criterion should be satisfied: the total energy of two atoms or molecules, A and B, separated by a near infinite distance, must be equal to the sums of the energies of A and B calculated separately. The reason this criterion is not satisfied for limited CI is easy to see. Suppose one did a CID calculation on A and then one on B. Those determinants involving *both* double substitutions on A and double substitutions on B would be included in the separate calculations, but would be excluded in a CID calculation on A and B

together (no matter the distance between them) because they would then be counted as quadruple substitutions.

A device for getting around this problem has been discovered. It is called Quadratically-convergent Configuration Interaction or QCI. The technical details of how this is done will not be covered here, but it should be recognized that the improvement comes at the cost of some computational overhead.

Limited CI and QCI calculations provide a good way of capturing dynamic electron correlation. They do not do such a good job on nondynamic electron correlation, for reasons explained in section T4.5.

T4.3 Møller-Plesset Theory

The Møller-Plesset approach is to treat the electron correlation problem as an exercise in perturbation theory. In perturbation theory one treats the Hamiltonian of a system of interest (\mathbf{H}) as being derivable from a simpler Hamiltonian ($\mathbf{H}^{(0)}$) for which the solution to the Schrödinger equation is known.

$$\mathbf{H} = \mathbf{H}^{(0)} + \lambda\mathbf{V}$$

Here \mathbf{V} is the perturbing potential, and λ is some number $\ll 1$.

The wave function, Ψ , that would be the eigenfunction of \mathbf{H} is written in terms of the (known) wave function for $\mathbf{H}^{(0)}$, plus a series of correction terms:

$$\Psi = \Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \dots$$

Similarly,

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots$$

Perturbation theory permits $\Psi^{(1)}$ and $E^{(1)}$ to be calculated from \mathbf{V} and $\Psi^{(0)}$. Then $\Psi^{(2)}$ and $E^{(2)}$ can be calculated from \mathbf{V} and $\Psi^{(1)}$, and so on. Each correction term gets harder to calculate, but also smaller in magnitude, and so at some point the series can be truncated.

A common use of this approach is to determine the interaction between molecules by treating the wave function of the interacting system as a perturbed version of the wave functions of the separate reactants.

In Møller-Plesset theory the electron-electron repulsion is treated as a perturbation on the one-electron (i.e. extended-Hückel) Hamiltonian. The first-order correction provides the Hartree-Fock solution. Higher-order corrections incorporate increasing amounts of electron correlation. They are commonly specified by the letters MP and a number, indicating where the perturbation-theory expansion was terminated.

MP2 theory is roughly equivalent to CISD in terms of the amount of electron correlation that it captures. However it is much quicker to evaluate than the full CISD expression. It also has the advantage over CISD of being size consistent. However, it has one big flaw: CISD was accomplished by using the exact Hamiltonian on a truncated-expansion wave function. It was therefore variational. MP theory uses a truncated Hamiltonian, and is therefore not variational. That means that one cannot use the calculated energy as a measure of the quality of the calculation. Indeed, there is no objective measure of quality

– one just has to hope that increasing numbers of correction terms lead to an increasingly accurate result.

As implemented in Gaussian 03, MP4 theory comes in two forms: MP4(SDQ) and MP4(SDTQ). As the letters in parentheses imply, the former excludes but the latter includes triple-substitution determinants. Inclusion of triples greatly increases the size of the computation, but can also significantly improve the quality of the results. The highest level MP theory available in Gaussian 03 is MP5, but it is impractical for molecules of a size interesting to organic chemists.

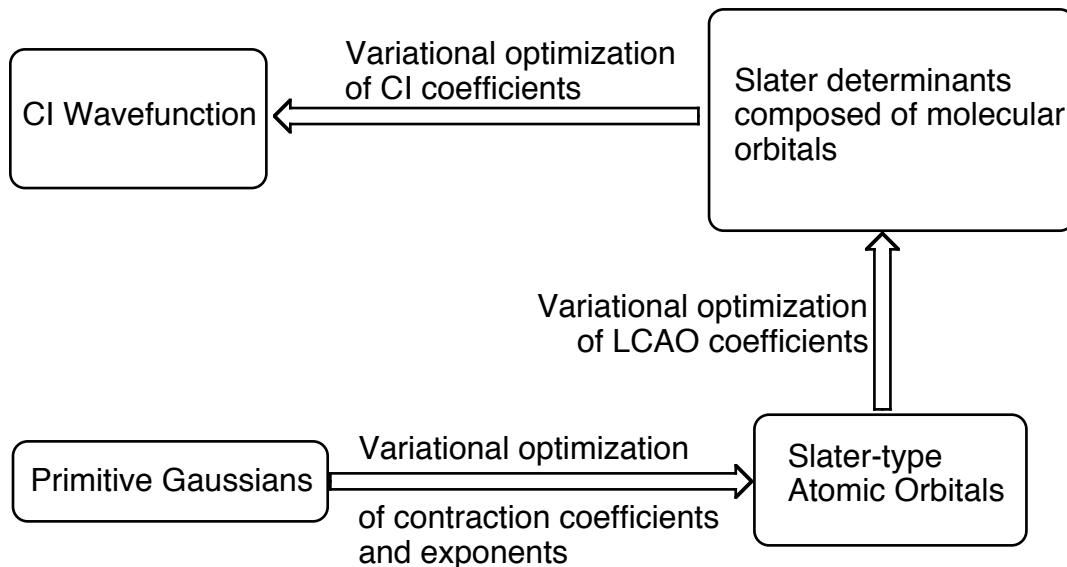
It is important to note that Møller-Plesset theory, being predicated on the idea of a small perturbation, is valid only for systems in which the reference configuration is dominant. It cannot be used reliably on systems in which nondynamic electron correlation is important. It is also only valid for the lowest-energy state of each spin multiplicity.

T4.4 Coupled-Cluster Theory

Through a clever reformulation of the CI expansion in terms of exponential operators, coupled cluster theory manages to solve the size consistency problem of limited CI. That is CCSD (couple clusters with single and double substitutions) actually includes the missing quadruple-substitution terms (and some higher terms too) of CISD. The price for this reformulation is that the method is no longer variational. Nevertheless, CCSD(T) (the parentheses around the T means that the triple-substitution clusters are only included approximately) is currently considered the best practical method for calculating dynamic electron correlation of moderate-sized molecules. Like the various limited-CI methods it does not do a very good job with nondynamic electron correlation. The reason is explained in the next section.

T4.5 MCSCF, CASSCF, and GVB Calculations

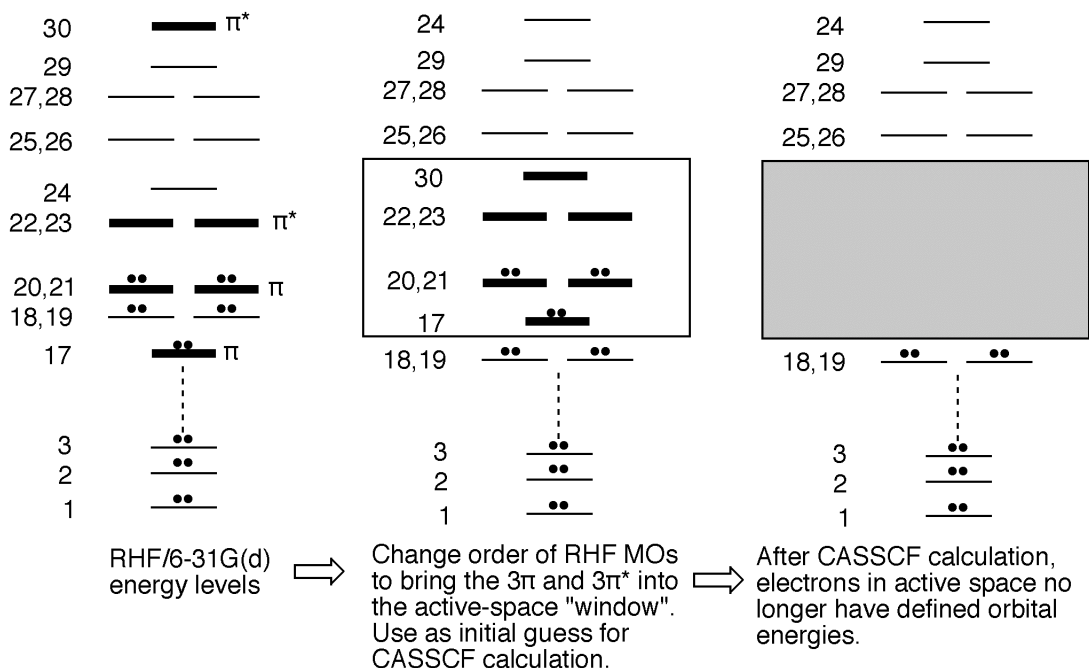
It is useful at this point to remind ourselves of the many linear-combination coefficients that go into a CI calculation. They are summarized in the schematic diagram below:



In a typical calculation, the Gaussian contraction coefficients and exponents are optimized just once, and define the basis set for the calculation. Note however, that the Dunning correlation-consistent basis sets were constructed with the recognition that the earlier Pople-type basis sets had been variationally optimized only for HF calculations, and that a different set of coefficients and exponents might be more suited to post-HF calculations.

In any of the post-HF methods that we have discussed so far, the variational optimization of the LCAO coefficients is undertaken only for the reference (HF) configuration. The other determinants use the same LCAO coefficients, even though those would certainly not be optimum if one were to variationally optimize them for each configuration. This does not present a serious error when the reference configuration makes the dominant contribution, but it can be a problem when there are two or more configurations that make contributions of similar magnitude (i.e. when nondynamic electron correlation is important). The most notorious species exhibiting such behavior are singlet-state biradicals. For such situations one should use Multi-Configuration Self-Consistent Field (MCSCF) methods, in which the LCAO coefficients are individually optimized for each determinant, and the CI coefficients are optimized simultaneously. Not surprisingly, this is much more work than just optimizing the CI coefficients, and really not computationally tractable if many configurations are included in the CI.

A common solution is to restrict the MCSCF computation to an "active space" of orbitals that encompasses the key sites of bond making and breaking. Such restricted MCSCF calculations are commonly called CASSCF, for Complete Active Space Self Consistent Field. The CASSCF terminology generally includes two numbers in parentheses, separated by a comma, such as CASSCF(4,4) or CASSCF(12,10). The first specifies the number of electrons included in the MCSCF calculation, and the second specifies the number of orbitals. The procedure for carrying out a CASSCF(6,6)/6-31G(d) calculation on benzene is indicated in the following diagram:



GVB theory does something like CASSCF. In its most common form, GVB-PP (Generalized Valence Bond – Perfect Pairing), it retains the electron-pair bond concept of VB theory. Electrons are permuted among the σ and σ^* orbitals (or π and π^* orbitals) of an electron-pair bond, but not between the orbitals of one atom pair and another. The simplest possible CASSCF calculation CASSCF(2,2) is thus identical in its result to GVB-PP(1). These methods are also known as TCSCF for Two-Configuration Self Consistent Field. This is the minimum level of theory that can capture the nondynamic electron correlation of a simple (unconjugated) singlet-state biradical.

The principal problem with CASSCF is that it provides no correlation for electrons outside of the active space. Thus the dynamic component of electron correlation is poorly treated by such methods.

T4.6 Multireference CI and Multireference Perturbation Theory

In order to do a good job with *both* dynamic and nondynamic electron correlation, one needs to do CI or perturbation theory from multiple reference configurations. One does a CASSCF calculation first and uses the resulting multideterminant wave function to provide reference configurations for the calculation of dynamic electron correlation. In their calculation on cyclopropane stereomutation, Getty, *et al.*) did TCSCF-CISD – i.e. CI with single and double substitutions from a two-configuration reference wave function.

Multireference CI is very computationally intensive, and far from user friendly, and so in recent years a number of efforts to formulate multireference second-order perturbation theory have been reported. There are four readily available versions of such methods CASMP2 (in Gaussian 03), MCQDPT (in GAMESS), CASPT2 (in MOLCAS), and RS2C (in MOLPRO). Most people think that CASPT2 does the best job. A number of people (BKC included) think that the CASMP2 in Gaussian 03 is just plain wrong! None of these methods is easily able to do geometry optimization or vibrational-frequency calculation, and so the standard approach is to run CASPT2(m,n) // CASSCF(m,n).

T4.7 Composite Methods

Gaussian 03 includes a number of methods designed to approximate the result of high-level electron correlation calculations with large basis sets. All of these methods involve additivity schemes, with some empirical parameters included.

The Gaussian models G1, G2, and G3:

G2, for example, tries to estimate the result of a QCISD(T)/6-311+G(3df,2p) calculation by combining results of QCISD(T) with smaller basis sets and lower-level correlation approximations with the big basis set.

The “Complete Basis Set” or CBS models CBS-4, CBS-QB3, and CBS-APNO:

The CBS methods use extrapolation schemes to try to estimate what one would get with an infinitely large basis set and a high-level correlation method. CBS-QB3 is practical for molecules with up to about 10 heavy (i.e. non-hydrogen) atoms, and generally gives properties such as bond energies within about ± 1 kcal/mol.

Note: None of the composite methods can handle nondynamic electron correlation very well, and so they are unreliable for things such as singlet biradicals, and may be questionable for transition states.

T5. Density Functional Theory

For a good introductory reference see: Kohn, W.; Becke, A.D.; Parr, R.G. *J. Phys. Chem.* **1996**, *100*, 12974.

Density functional theory (DFT) takes a completely different approach to calculation of electronic structure from either MO or VB theory. Whereas the latter two consider the multielectron wave function, $\Psi(1,2\dots n)$ to be the fundamental function from which all else can be derived, DFT, as its name implies, assigns that role to $\rho(r)$, the electron density. The two are related by:

$$\rho(r) = n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \Psi^*(1,2\dots n) \Psi(1,2\dots n) dr_1 dr_2 \dots dr_{n-1}$$

i.e. $\rho(r)$ is the probability of finding one electron out of n in a volume element between r and $r + dr$ (for some reason, the τ of MO theory traditionally gets replaced by a confusing r in DFT). However, this relationship does not have to serve as a *definition* of $\rho(r)$. It could equally well serve as a (not very convenient) definition of Ψ .

More importantly, the first Hohenberg-Kohn Theorem proves that all physical properties of a system can be uniquely determined from $\rho(r)$, without need to calculate Ψ . So how are we going to find $\rho(r)$? An answer comes from writing the expression for the energy of a system of electrons as a functional of ρ :

$$E(\rho) = F(\rho) + \int \rho(r)v(r) dr$$
$$F(\rho) = T(\rho) + U_{ee}(\rho)$$

Here $E(\rho)$ is the energy, $F(\rho)$ is called the Hohenberg-Kohn functional, $T(\rho)$ is the kinetic-energy functional (which can be calculated exactly) and $U_{ee}(\rho)$ is the electron-electron repulsion functional (which cannot, except for the case of the uniform "electron gas" in which $v(r)=0$). The function $v(r)$ is the "external" potential experienced by the electrons (usually due just to the nuclei). "External" means the potential beyond that exerted by the electrons on each other. The second Hohenberg-Kohn Theorem proves that there exists an equivalent to the Variation Principle in MO theory, i.e. any trial density $\rho(r)$ will give an energy $E(\rho)$ that is greater than E_0 , the true ground-state energy of the system with density $\rho_0(r)$. Thus in searching for the best density, one seeks to minimize the energy of the system, just as one does in searching for the best Ψ in variational MO theory.

The best practical approach for optimizing the density thus far has come from the Kohn-Sham approach. They postulate a set of orbitals, $\varphi_j(r)$, describing a set of *noninteracting* electrons (in the Hartree-Fock sense – i.e. the electrons feel each other's average field but do not respond explicitly to one another) whose density is equal to the true density for the molecule:

$$\rho(r) = \sum_j |\varphi_j(r)|^2$$

Now the expression for $F(\rho)$ becomes:

$$F(\rho) = T_s(\rho) + \frac{1}{2} \int \frac{\rho(r)\rho(r')}{|r-r'|} dr dr' + E_{xc}(\rho)$$

This turns out to be just a reformulation of the Hartree-Fock expression except for the $E_{xc}(\rho)$ term, which is known as the *exchange correlation energy*. If an expression for this were known, then its addition would turn the Hartree-Fock approximation into an exact theory, aside from relativity. Needless to say, an analytical expression for $E_{xc}(\rho)$ is not known.

Most of the recent research in DFT has been involved in coming up with various approximate expressions for $E_{xc}(\rho)$. The need to approximate $E_{xc}(\rho)$ means that the various DFT methods are not variational. Furthermore, many of the expressions for $E_{xc}(\rho)$ are empirical, and so one winds up with sets of different DFT methods and no objective way of choosing one over another.

That said, the method known as B3LYP has been extremely popular in recent years, and seems to have given very good results for a variety of molecules. (The B,L,Y, and P stand for Becke, Lee, Yang, and Parr, whose functionals are used in this method. The exchange correlation functional is divided into exchange and correlation parts. The correlation part comes from L, Y, and P, and the exchange part from B. The 3 comes from the fact that this one of Becke's exchange functionals has three parameters in it.) B3LYP is sometimes called a hybrid method, because the expression for Becke's exchange functional requires that a HF calculation be carried out.

The great appeal of DFT methods is that they generally require little more effort than a HF calculation, but can include dynamic electron correlation quite well. If an exact expression for $E_{xc}(\rho)$ were known, nondynamic electron correlation would also be included. In reality, it has been much debated whether the approximate expressions for $E_{xc}(\rho)$ currently available do include nondynamic electron correlation. A number of authors (notably Houk at UCLA) have used UB3LYP (i.e. B3LYP with unrestricted Kohn-Sham orbitals) to carry out calculations on singlet biradicals. These empirically seem to give good results in a number of cases, but the $\langle S^2 \rangle$ value that one calculates is invariably (and necessarily) near 1.0 instead of 0. Several theorists (notably Davidson at Indiana University) think that UB3LYP is not reliable for systems needing good descriptions of nondynamic electron correlation. In summary, the debate is between those who say "It works" and those who say "But it shouldn't."

*Note: The designation "B3LYP" is unfortunately not unique. Different programs for doing density functional calculations interpret it in different ways. The problem is described in: Hertwig, R. H.; Koch, W. Chem. Phys. Lett. **1997**, 268, 345. It stems from two different suggestions for the form of the exchange functional in the same paper; some programs followed one suggestion and some the other. For example, Gaussian 03 and GAMESS have different definitions. The differences are not huge, but if you are trying to compare your results with something in the literature, you do need to make certain that you used the same implementation!*

Summary of Methods for Treating Electron Correlation

We have seen a large number of ways of trying to deal with the electron-correlation problem. They all have their pros and cons. In the end, the questions of cost, size consistency, variational character and ability to handle dynamic and/or nondynamic electron correlation have to be weighed against each other. Here is a table that can help when trying to pick a method for a particular problem.

Method	Dynamic Correlation	Nondynamic Correlation	Size Consistent	Variational	Scaling ^a
HF	Poor – None	None	Yes	Yes	n^2
B3LYP	Very Good	Controversial	Yes	No	$\sim n^3$
MP2	Fair	None	Yes	No	nN^4
CISD	Good	Fair	No	Yes	n^2N^4
CCSD	Very Good	Fair	Yes	No	n^2N^4
MP4	Excellent	None	Yes	No	n^3N^4
CCSD(T)	Excellent	Good	Yes	No	n^3N^4
CASSCF	None	Excellent	No	Yes	b
CASPT2	Good	Excellent	No	No	b

^a n is the number of occupied orbitals and N the number of virtual orbitals. The scaling expressions are approximate and based on the assumption that $N \gg n$.

^b The number of configurations contributing to a CASSCF wave function in C_1 symmetry is given by:

$$\frac{(2S+1)[(\mu+1)!]^2}{(\mu+1)(\mu - \frac{\nu}{2} - S)!(1 + \frac{\nu}{2} + S)!(\frac{\nu}{2} - S)!(1 + \mu - \frac{\nu}{2} + S)!}$$

where S is the total spin (i.e. $1/2$ of the number of unpaired electrons), ν is the number of active electrons, and μ is the number of active orbitals. For higher symmetry species, the number of configurations will be smaller. For CASPT2 the number of configurations is the same, but then the calculation will take roughly nN^4 times longer because of the MP2-like correction for dynamic electron correlation.

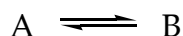
E 5. Isotope Effects

The theory of *kinetic* isotope effects, in which isotopic substitution alters the *rate* of a reaction is best developed by beginning with *equilibrium* isotope effects, such as those shown below:



Why is there an effect from isotopic substitution at all, and why does its magnitude depend on the nature of the isotope?

From statistical thermodynamics:



$$K = \frac{Q_B}{Q_A} e^{-[\varepsilon_0(B) - \varepsilon_0(A)]/k_B T}$$

$$Q = \sum_i g_i e^{-\varepsilon_i/k_B T}$$

Q = partition function

ε_0 = lowest energy level

k_B = Boltzmann's constant

g_i = degeneracy of energy level i

According to the Born-Oppenheimer approximation:

$$Q_{TOT} = Q_{TRANS} Q_{ROT} Q_{VIB} Q_{ELEC} Q_{NUC}$$

For Q_{ELEC} with a non-degenerate ground state, define $\varepsilon_0 = 0$:

$$Q_{ELEC} = e^{-\varepsilon_0/k_B T} + g_1 e^{-\varepsilon_1/k_B T} + g_2 e^{-\varepsilon_2/k_B T} + \dots$$

Usually $\varepsilon_1 \gg k_B T$, so all terms beyond the first are ~ 0 . Hence $Q_{ELEC} \approx 1$. Similarly, $Q_{NUC} \approx 1$.

$$Q_{TRANS} = \frac{(2\pi M k_B T)^{3/2} V}{h^3}$$

M = mass of molecule

h = Planck's constant

V = volume of container

From the rigid rotor solutions to the Schrödinger equation:

$$Q_{ROT} = \frac{\left\{ \pi (8\pi^2 k_B T)^3 I_A I_B I_C \right\}^{1/2}}{\sigma h^3}$$

I_A = moment of inertia around principal axis A

σ = external symmetry number (= number of equivalent representations of the molecule attainable purely by rigid rotation).

From the harmonic oscillator solutions to the Schrödinger equation:

$$Q_{VIB} = \prod_i^{3N-6} \frac{e^{-h\nu_i/k_B T}}{(1 - e^{-h\nu_i/k_B T})}$$

ν_i = frequency of mode i

N = number of atoms in molecule

These approximations give quite a good account of equilibrium isotope effects. For the two examples given, one calculates values of 6.35 and 1.030, respectively. The isotope effects arise because the different masses of the isotopes affect Q_{TRANS} directly (through M), Q_{ROT} through the moments of inertia, and Q_{VIB} through the mass dependence of the vibrational frequencies.

Kinetic isotope effects are generally explained within the context of Transition State Theory:

$$k_1 = \kappa_1 \frac{k_B T}{h} \frac{Q_1^\ddagger}{Q_1} e^{(\epsilon_0 - \epsilon_0^\ddagger)/k_B T}$$

$$k_2 = \kappa_2 \frac{k_B T}{h} \frac{Q_2^\ddagger}{Q_2} e^{(\epsilon_0 - \epsilon_0^\ddagger)/k_B T}$$

k_1 = rate constant for reaction of lighter isotope

k_2 = rate constant for reaction of heavier isotope

$\kappa_{1,2}$ = transmission coefficients Q^\ddagger = partition function for transition state

Substituting as before for the partition functions gives:

$$\frac{k_1}{k_2} = \frac{\kappa_1}{\kappa_2} \frac{\sigma_1 \sigma_1^\ddagger}{\sigma_2 \sigma_2^\ddagger} \cdot MMI \cdot EXC \cdot ZPE$$

MMI = Mass, moment - of - inertia term:

$$MMI = \frac{(M_1^\ddagger M_2)_{3/2} (I_{A1}^\ddagger I_{B1}^\ddagger I_{C1}^\ddagger I_{A2} I_{B2} I_{C2})^{1/2}}{(M_2^\ddagger M_1)_{3/2} (I_{A2}^\ddagger I_{B2}^\ddagger I_{C2}^\ddagger I_{A1} I_{B1} I_{C1})^{1/2}}$$

EXC = Excitation term (factored out of vibrational partition function):

$$EXC = \frac{\prod_i^{3N-6} \{(1 - e^{-u_{i1}})/(1 - e^{-u_{i2}})\}}{\prod_i^{3N-7} \{(1 - e^{-u_{i1}^\ddagger})/(1 - e^{-u_{i2}^\ddagger})\}} \quad u_i = \frac{h\nu_i}{k_B T}$$

ZPE = Zero - point - energy term (also from vibrational partition function):

$$ZPE = \frac{\exp\left\{\sum_i^{3N-6} (u_{i1} - u_{i2})/2\right\}}{\exp\left\{\sum_i^{3N-7} (u_{i1}^\ddagger - u_{i2}^\ddagger)/2\right\}}$$

If we assume that $\kappa_1 \approx \kappa_2$; $M_2^\ddagger/M_1^\ddagger \approx M_2/M_1$; $I_2^\ddagger/I_1^\ddagger \approx I_2/I_1$, then $MMI \approx 1$.

For most organic molecules, $\tilde{\nu}_i \geq 500 \text{ cm}^{-1}$, meaning $e^{-u_i} < 0.08$ at 298 K. Hence $EXC \approx 1$. (This is generally true for organic reactions, but see: Slaughter, L. M.; Wolczanski, P. T.; Klinckman, T. R.; Cundari, T. R. *J. Am. Chem. Soc.* **2000**, *122*, 7953 for evidence that it may not be reliably true in organometallic reactions.)

Kinetic isotope effects are therefore traceable primarily to differences in ZPE between transition states and reactants:

$$\frac{k_1}{k_2} \approx \frac{\sigma_1 \sigma_2^\ddagger}{\sigma_2 \sigma_1^\ddagger} \frac{\exp\left\{\sum_i^{3N-6} (u_{i1} - u_{i2})/2\right\}}{\exp\left\{\sum_i^{3N-7} (u_{i1}^\ddagger - u_{i2}^\ddagger)/2\right\}}$$

1. The isotopically heavier molecule always has the lower ZPE.
2. As the force constant of a particular mode increases, the difference in ZPE for the two isotopes increases.
3. As a consequence of 1 and 2, the heavier isotope always prefers to be in the more strongly bound position.
4. The difference in ZPE for two isotopes decreases as the atomic number of the element increases.

Calculation of isotope effects with Gaussian 03

Many of the papers reporting the use of *ab initio* or density functional theory to compute isotope effects cite the application of specific programs (such as "QUIVER") for the purpose. However, unless tunneling corrections need to be made, this isn't necessary. All one has to do is to calculate activation free energies for the isotopic isomers. The isotope effect is then given by:

$$\frac{k_1}{k_2} = e^{(\Delta G_2^\ddagger - \Delta G_1^\ddagger)/RT}$$

Within the WebMO interface, this can be accomplished in the following way:

1. Calculate the free energy of the reactant(s) for the unlabeled molecule(s) by geometry optimizing and then carrying out a vibrational frequency computation. When you calculate the vibrational frequencies, click the "Save Checkpoint File" box in the "Advanced" panel.
2. Do the same thing for the transition state.
3. Since isotopic substitution does not affect the Hessian within the limits of the Born-Oppenheimer approximation, it is unnecessary to repeat the lengthy computation of the Hessian. It's that information that is being saved in the checkpoint files for steps 1 & 2. To access the information select "New Job Using This Geometry" from the vibrational frequency calculations on one of your unlabeled species. In the job options window, select "Use Checkpoint File" in the "Advanced" panel, and choose the

correct checkpoint file from the dropdown menu. Next, select the “Preview” panel and hit the “Generate” button. You will now need to edit the Gaussian input file to request that the force-constant matrix (i.e. Hessian) be read in, and to specify where and what your isotopes are. You do this by adding immediately after the “FREQ” keyword: `=(READFC,READISO)`. Then, after the last two numbers in the input file (which specify the charge and spin multiplicity of your molecule, and so would be 0 1 for a neutral singlet) *leave one blank line* followed by the temperature in K and pressure in atmospheres, followed by a list of atomic masses in amu, in the order in which your atoms were numbered in the unlabeled molecule. For example, if I had done a B3LYP/6-31G(d) frequency calculation on ethylene, with the two carbons numbered 1 and 2, and I wanted to repeat the calculation for a single ¹³C-labeled isotopomer at 298.15 K and 1 atmosphere, the input file would look like this:

```
%CHK=output.chk
%NOSAVE
#N B3LYP/6-31G(d) FREQ=(READFC,READISO) GEOM=CHECKPOINT
<Blank line>
C2H4
<Blank line>
0 1
<Blank line>
298.15 1.0
13
12
1
1
1
1
<Blank line>
```

The integer masses for the atoms are automatically replaced by exact masses for the isotopes in the program. All of the blank lines are important, which is why they are shown explicitly in this example. When you click on the “Submit Job” instruction, you should find that the new calculation is complete in a second or two, no matter how long the original calculation might have taken. Furthermore, the free energy that you see reported should be slightly different from the one for the unlabeled molecule.

- Repeat this exercise for each of your reactant(s) and for the transition state. Thereby calculate an activation free energy for the unlabeled reaction and one for the reaction with your chosen isotope(s). These two quantities are respectively the ΔG_1^\ddagger and ΔG_2^\ddagger in the equation for the isotope effect (although make sure you have converted from hartrees to the units of R – typically kcal/mol). If your reaction is bimolecular, you don’t need to worry about the standard state, because the unlabeled and isotopically substituted reactions will be referred to the same standard state (which defaults to 1 atmosphere in Gaussian 03), and so any correction for a change of standard state would cancel out in the comparison.