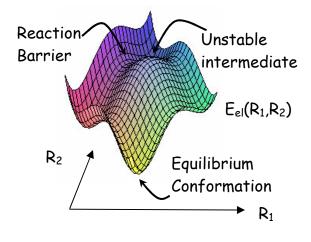# MODERN ELECTRONIC STRUCTURE THEORY

At this point, we have more or less exhausted the list of electronic structure problems we can solve by hand. If we were limited to solving problems manually, there would be a lot of chemistry we wouldn't be able to explain! Fortunately, the advent of fast personal computers allows chemists to routinely use more accurate models of molecular electronic structure. These types of calculations typically play a significant role in interpreting experimental results: calculations can be used to assign spectra, evaluate reaction mechanisms and predict structures of molecules. In this way computation is **complementary** to experiment: when the two agree we have confidence that our interpretation is correct.

The basic idea of electronic structure theory is that, within the Born Oppenheimer approximation, we can fix the M nuclei in our molecule at some positions $\mathbf{R_I}$. Then, we are left with the Hamiltonian for the electrons moving in the effective field created by the nuclei:

$$\hat{H} \equiv -\tfrac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{I=1}^{M}\frac{Z_I}{\left|\hat{\mathbf{r}}_i - \mathbf{R_I}\right|} + \sum_{i<j}^{N}\frac{1}{\left|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j\right|} \qquad \textbf{Eq. 1}$$

Where the first term is the kinetic energy of all N electrons, the second term is the attraction between the electrons and nuclei and the third is the pairwise repulsion between all the electrons. **The central aim of electronic structure 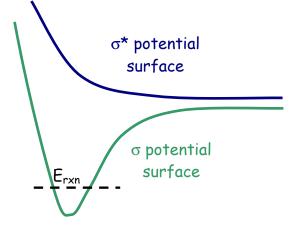theory is to find all the eigenfunctions of this Hamiltonian**. As we have seen, the eigenvalues we get will depend on our choice of the positions of the nuclei – $E_{el}(\mathbf{R_1},\mathbf{R_2},\mathbf{R_3},...\mathbf{R_M})$. As was the case with diatomics, these energies will tell us how stable the molecule is with a given configuration of the nuclei $\{\mathbf{R_I}\}$ – if E_{el} is very low, the molecule will be very stable, while if E_{el} is high, the molecule will be unstable in

that configuration. The energy $E_{el}(\mathbf{R_1},\mathbf{R_2},\mathbf{R_3},...\mathbf{R_M})$ is called the **potential energy surface**, and it contains a wealth of information, as illustrated in the picture at above. We can determine the equilibrium configuration of the

molecule by looking for the minimum energy point on the potential energy surface. We can find metastable intermediate states by looking for local minima – i.e. minima that are not the lowest possible energy states, but which are separated from all other minima by energy barriers. In both of these cases, we are interested in points where $\nabla E_{el} = 0$. Further, the potential surface can tell us about the activation energies between different minima and the pathways that are required to get from the "reactant" state to the "product" state.

Solving the electronic Schrödinger also gives us the electronic wavefunctions $\Psi_{el}(r_1, r_2, r_3, \ldots r_N)$, which allow us to compute all kinds of electronic properties – average positions, momenta, uncertainties, etc – as we have already seen for atoms.

We note that while the Hamiltonian above will have many, many eigenstates, in most cases we will only be interested in the lowest eigenstate – the electronic ground state. The basic reason for this is that in stable molecules, the lowest excited states are usually several eV above the ground state and therefore not very important in chemical reactions where the available energy is usually only tenths of an eV. In cases where multiple electronic states are important, the Hamiltonian above will give us separate potential surfaces E^1_{el}, E^2_{el}, E^3_{el} … and separate wavefunctions Ψ^1_{el}, Ψ^2_{el}, Ψ^3_{el}. The different potential surfaces will tell us about the favored conformations of the molecules in the different electronic states. We have already seen this for H_2^+. When we solved for the electronic states, we got two eigenstates: σ and σ^*. If we put the electron in the σ orbital, the molecule was bound and had a potential surface like the lower surface at right. However, if we put the electron in the $\sigma*$ orbital the molecule was not bound and we got the upper surface.

σ* potential surface

σ potential surface

E_{rxn}

So, at the very least our task is clear cut: solve for the eigenstates of Eq. 1. Unfortunately, this task is also impossible in practice, even on a computer. Thus, over the years chemists have developed a vast array of sophisticated techniques that allow us to at least approximate these solutions to within a tolerable degree of accuracy. Learning all the details of these approximations would require a course unto itself: the derivations of the individual approximations are exceedingly complex, and the sheer number of different approximations that can be made is quite impressive. These detailed derivations teach us a lot about what molecules and properties we should expect our approximations to work for and how we should think about improving our calculations in cases where the theory fails. However, the thing that has really brought computational chemistry into the mainstream is the fact that **one does not have to understand every nuance of a method in order to know how to use it successfully**. It suffices to have a simple, qualitative understanding of how each method works and when it can be applied. Then, coupling that knowledge with a little technical proficiency at using commercial chemistry software packages allows us to run fairly sophisticated calculations on our desktop computer. The next two lectures are intended to prepare us to run these types of calculations.

First, we note that nearly all the popular approximations are still based on MO theory – MO theory on steroids in some cases, but MO theory nonetheless. Thus, there are still 5 steps in the calculation

1) Choose an Atomic Orbital Basis
2) Build the Relevant Matrices
3) Solve the Eigenvalue Problem
4) Occupy the orbitals based on a stick diagram
5) Compute the energy

In a typical calculation, the computer automatically handles steps 2-4 automatically – we don't have to tell it anything at all. It is sometimes helpful to know that the computer is doing these things (e.g. The calculation crashed my computer. What was it doing when it crashed? Oh, it was trying to solve the eigenvalue problem.) but we essentially never have to do them ourselves. This leaves two steps (1 and 5) that require some input from us for the calculation to run properly

<u>Choosing an Atomic Orbital Basis</u>

The first point here is that for real electronic structure calculations, you typically use a basis set that is much larger than your intuition might tell you. For example, for H_2^+ we guessed that we would get a decent result if we wrote:
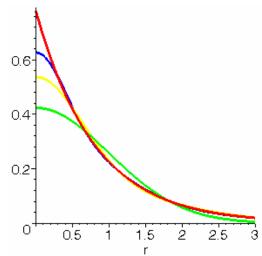
$$\psi = c_1 1sA + c_2 1s_B$$

A basis of this type would be called a minimal basis set, because it is the smallest one that is even close to right. In a more accurate calculation, you might use a basis that looks more like:

$$\psi = c_1 1s_A + c_2 1s_B + c_3 2s_A + c_4 2s_B + c_5 2p_{xA} + c_6 2p_{xB} + c_7 2p_{yA}$$
$$+ c_8 2p_{yA} + c_9 2p_{zA} + c_{10} 2p_{zB} + c_{11} 3s_A + c_{12} 3s_B$$

The reason we use such extended basis sets arises from a point that was discussed earlier concerning MO theory. Because our results are variational, a bigger basis always gets us a lower energy, and therefore a longer AO expansion always gets us closer to the ground state energy. In the worst case, the unimportant basis functions will just have coefficients that are very near zero. While such extended basis sets would be a significant problem if we were working things out by hand, computers have no problem dealing with expansions involving even 10,000 basis functions.

The second important point about the atomic orbitals we use is that they are **not hydrogenic orbitals**. The reason for this is that the two-electron integrals involving hydrogenic orbitals cannot all be worked out analytically, making it very difficult to complete Step 2. Instead of hydrogenic orbitals – which decay like e^{-r} – we will use Gaussian orbitals that decay like $e^{-\alpha r^2}$.

Gaussians do not look very much like hydrogenic orbitals – they don't have a cusp at r=0 and they decay much too fast at large distances. About the only good thing about them is that they have a mximum at r=0 and decay. These differences between Gaussians and hydogenic orbitals are not a problem, though, because we use extended basis sets as emphasized above. Basically, given enough Gaussians, you can expand anything you like – including a hydrogenic orbital, as shown in the picture at right. So

while using Gaussians may mean we have to use a few extra AOs, if we use enough of them we should be able to get the same answer.

So we plan to use relatively large Gaussian basis sets for our calculations. How exactly do we choose those basis sets? Thankfully, a significant amount of trial-and-error research has distilled the choices down to a few key basis set features. Depending on the problem at hand and the accuracy desired we only need to consider three aspects of the AO basis.

Single, Double, Triple, Quadruple Zeta Basis Sets

As we have already discussed for MO theory of diatomics, the smallest basis we can think of for describing bonding would include all the valence orbitals of each atom involved. Thus, for H we had 1 s-function, for C there were 2 s-functions and one set of p's. Similarly, for sulfur we would have needed 3 s-functions and 2 p's …. A basis of this size is called a minimal or single zeta basis. The term "single zeta" refers to the fact that we have only a single set of the valence functions (Note: single valence might seem like a more appropriate name, but history made a different choice). The most important way to expand the basis is to include more than a single set of valence functions. Thus, in a double zeta (DZ) basis set, one would include 2 s-functions for H, 3 s- and 2 p-functions for C and 4 s- and 3 p-functions for S. Qualitatively, we think of these basis functions as coming from increasing the n quantum number: the first s function on each atom is something like 1s, the second something like 2s, the third like 3s …. Of course, since we are using Gaussians, they're not exactly 1s, 2s, 3s … but that's the basic idea. Going one step further, a triple zeta (TZ) basis would have: H=3s, C=4s3p, S=5s4p. For Quadruple zeta (QZ): H=4s, C=5s4p, S=6s5p and so on for 5Z, 6Z, 7Z. Thus, one has:

	H,He	Li-Ne	Na-Ar	Names
Minimal	1s	2s1p	3s2p	STO-3G
DZ	2s	3s2p	4s3p	3-21G,6-31G, D95V
TZ	3s	4s3p	5s4p	6-311G,TZV

Unfortunately, the commonly used names for basis sets follow somewhat uneven conventions. The basic problem is that many different folks develop basis sets and each group has their own naming conventions. At the end of the table above, we've listed a few names of commonly used SZ,DZ and TZ basis sets. There aren't any commonly used QZ basis sets, because once

your basis is that large, it is best to start including polarization functions (see below).

Polarization Basis Functions

Note that no matter how high you go in the DZ, TZ, QZ hierarchy, you will never, for example, get a p-function on hydrogen or a d-function on carbon. These functions tend to be important for describing polarization of the electrons; at a qualitative level, the p-functions aren't as flexible in their angular parts and it's hard to get them to "point" in as many directions as d-functions. Thus, particularly when dealing with directional bonding in molecules, it can be important to include some of these higher angular momentum functions in your AO basis. In this situation the basis set is said to contain some "polarization" functions. The general nomenclature of polarization functions is to add the letter "P" to a basis set with a single set of polarization functions, and "2P" to a basis with two sets. Thus, a DZP basis would have: 2s1p on hydrogen, 3s2p1d on carbon and 4s3p1d on sulfur. A TZP basis set would have 3s1p on hydrogen, 4s3p1d on carbon and 5s4p1d on sulfur.

	H,He	Li-Ne	Na-Ar	Names
DZP	2s1p	3s2p1d	4s3p1d	6-31G(d,p), D95V
TZP	3s1p	4s3p1d	5s4p1d	6-311G(d,p), TZVP

We note that in practice it is possible to mix-and-match different numbers of polarization functions with different levels of zeta basis sets. The nomenclature here is to put (xxx,yyy) after the name of the basis set. "xxx" specifies the number and type of polarization functions to be placed on Hydrogen atoms and "yyy" specifies the number and type of polarization functions to be placed on non-hydrogen atoms. Thus, we would have, for example:

	H,He	Li-Ne	Na-Ar
6-311G(2df,p)	3s1p	4s3p2d1f	5s4p2d1f

Diffuse Functions

Occasionally, and particularly when dealing with anions, the SZ/DZ/TZ/… hierarchy converges very slowly. For anions, this is because the extra electron is only very weakly bound, and therefore spends a lot of time far

from the nucleus. It is therefore best to include a few basis functions that decay very slowly to describe this extra electron.   Functions of this type are called "diffuse" functions. They are still Gaussians ($e^{-\alpha r^2}$), but the value of α is very, very small causing the atomic orbital to decay slowly. Similar to the situation for polarization functions, diffuse functions can be added in a mix-and-match way to standard basis sets. Here, the notation "+" or "aug-" is added to a basis set to show that diffuse functions have been added. Thus, we have basis sets like 3-21++G, 6-31+G(d,p), aug-TZP.

Aside: Transition Metals

Those of you interested in inorganic chemistry will note that no transition metals appear in the tables above. This is not because there aren't basis sets for transition metals – it is just more complicated to compare different transition metal basis sets. First, we note that many of the basis sets above are defined for transition metals. Thus, for example, a 6-31G(d,p) basis on iron is 5s4p2d1f while a TZV basis for iron is 6s5p3d. The reason we didn't include this above is that the idea of "valence" for a transition metal is a subject of debate: is the valence and s- and d- function? An s a p and a d? Hence, depending on who put the basis set together, there will be some variation in the number of functions. However, one still expects the same ordering in terms of quality: TZ will be better than DZ, DZP will be better than a minimal basis, etc. Thus, you can freely use the above basis sets for all the elements between K and Kr without significant modification. Extending the above table for specific basis sets gives:

	K-Ca	Sc-Zn	Ga-Kr
3-21G	5s4p	5s4p2d	5s4p1d
6-31G(d,p)	5s4p1d	5s4p2d1f	N/A
6-311G(d,p)	8s7p2d	N/A	8s7p3d
TZV	6s3p	6s3p2d	6s5p2d

Things also become more complicated when dealing with second row transition metals. Here, relativistic effects become important, because the Schrödinger equation predicts that the 1s electrons in these atoms are actually moving at a fair fraction of the speed of light. Under these circumstances, the Schrödinger equation is not strictly correct and we need

to start considering corrections for relativistic effects. The most efficient way to incorporate the relativity is to use an effective core potential (ECP). An effective core potential removes the core electrons from the problem and replaces them with an effective potential that the valence electrons feel. This potential reflects the combined interaction with the nucleus and the (relativistic) core electrons. Thus, for an ECP we specify both how many core electrons we want to neglect and how many basis functions we want to use to describe the valence electrons. For example, one popular double zeta ECP is the LANL2DZ basis. As an example, for ruthenium LANL2DZ replaces the 28 core electrons ($1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10}$=Argon) with an effective potential and uses a 3s3p2d basis to describe the valence orbitals. Thus, for the second transition series we have (using [CoreSize]/ValenceBasisSize as our shorthand):

	Y-Cd	Hf-Hg
LANL2DZ	[Argon]/3s3p2d	N/A
SDD	[Argon]/8s7p6d	$[Kr4d^{10}4f^{14}]$/8s7p6d

As one progresses further up the periodic table, fewer and fewer basis sets are available, simply because less is known about their chemistry.

This is just a very brief overview of what basis sets are available and how good each one is expected to be. The general idea of using basis sets is to use larger and larger basis sets until the quantity we are computing stops changing. This is based on the idea that we are really using the AO expansion to approximate the exact solution of the Schrödinger equation. If we had an **infinite basis**, then we would get the exact answer, but with a large and finite basis we should be able to get close enough. Note, however, that the calculation will typically take more time in a larger basis than in a smaller one. Thus, we really want to make the basis just big enough to get the answer right, but no larger.

Computing the Energy
For simple MO theory, we used the non-interacting (NI) electron model for the energy:

$$E_{NI} = \sum_{i=1}^{N} E_i = \sum_{i=1}^{N} \int \psi_i(1)\hat{H}\psi_i(1)\,d\tau$$

Where, on the right hand side we have noted that we can write the NI energy as a sum of integrals involving the orbitals. We already know from looking at atoms that this isn't going to be good enough to get us really accurate answers; the electron-electron interaction is just too important. In real calculations, one must choose a method for computing the energy from among several choices.

The Hartree Fock (HF) Approximation

The Hartree-Fock method uses the IPM energy expression we've already encountered:

$$E_{IPM} = \sum_{i=1}^{N} E_i + \sum \tilde{J}_{ij} - \tilde{K}_{ij}$$

$$E_i = \sum_{i=1}^{N} \int \psi_i(1)\hat{H}\psi_i(1)\,d\tau$$

$$\tilde{J}_{ij} \equiv \iint \psi_i^*(1)\psi_j^*(2)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\psi_i(1)\psi_j(2)\,d\mathbf{r}_1 d\mathbf{r}_2 d\sigma_1 d\sigma_2$$

$$\tilde{K}_{ij} \equiv \iint \psi_i^*(1)\psi_j^*(2)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\psi_i(2)\psi_j(1)\,d\mathbf{r}_1 d\mathbf{r}_2 d\sigma_1 d\sigma_2$$

Since the energy contains the average repulsion, we expect our results will be more accurate. However, there is an ambiguity in this expression. The IPM energy above is correct for a determinant constructed out of any set of orbitals $\{\psi_i\}$ and the energy will be different depending on the orbitals we choose. For example, we could have chosen a different set of orbitals, $\{\psi'_i\}$, and gotten a different energy:

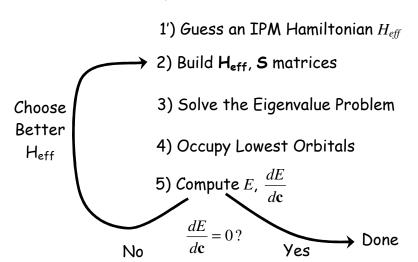$$E'_{NI} = \sum_{i=1}^{N} E'_i + \sum \tilde{J}'_{ij} - \tilde{K}'_{ij}$$

How do we choose the best set of orbitals then? Hartree-Fock uses the variational principle to determine the optimal orbitals. That is, in HF **we find the set of orbitals that minimize the independent particle energy**. These orbitals will be different from the non-interacting orbitals because they will take into account the average electron-electron repulsion terms in the Hamiltonian. Thus, effects like shielding that we have discussed qualitatively will be incorporated into the shapes of the orbitals. This will tend to lead to slightly more spread out orbitals and will also occasionally

change the ordering of different orbitals (e.g. σ might shift below π once interactions are included).

Now, the molecular orbitals (and hence the energy) are determined by their coefficients. Finding the best orbitals is thus equivalent to finding the best coefficients. Mathematically, then, we want to find the orbitals that make the derivative of the IPM energy zero:

$$\frac{\partial E_{IPM}}{\partial c_i^\alpha} = \frac{\partial}{\partial c_i^\alpha} \sum_{i=1}^{N} E_i + \sum \tilde{J}_{ij} - \tilde{K}_{ij} = 0$$

In order to satisfy this condition, one typically resorts to an iterative procedure, where steps 2-5 of our MO procedure are performed repeatedly:

1) Choose an AO Basis

1') Guess an IPM Hamiltonian H_{eff}

2) Build **H**$_{eff}$, **S** matrices

Choose Better H$_{eff}$

3) Solve the Eigenvalue Problem

4) Occupy Lowest Orbitals

5) Compute E, $\dfrac{dE}{d\mathbf{c}}$

$\dfrac{dE}{d\mathbf{c}} = 0$?

No   Yes   Done

Here, HF makes use of the fact that defining an IPM Hamiltonian, H_{eff}, completely determines the molecular orbital coefficients, **c**. Thus, the most convenient way to change the orbitals is actually to change the Hamiltonian that generates the orbitals. The calculation converges when we find the molecular orbitals that give us the lowest possible energy, because then $\dfrac{dE}{d\mathbf{c}} = 0$. These iterations are called self-consistent field (SCF) iterations and the effective Hamiltonian H_{eff} is often called the Fock operator, in honor of one of the developers of the Hartree-Fock approximation.

Generally Hartree Fock is not very accurate, but it is quite fast. On a decent computer, you can run a Hartree Fock calculation on several hundred atoms quite easily, and the results are at least reasonable.

Density Functional Theory (DFT)

Here, we still use a Slater determinant to describe the electrons. Hence, the things we want to optimize are still the MO coefficients c^α. However, we use a different prescription for the energy – one that is entirely based on the electron density. For a single determinant, the electron density, $\rho(\mathbf{r})$ is just the probability of finding an electron at the point $\mathbf{r}$. In terms of the occupied orbitals, the electron density for a Slater Determinant is:

$$\rho(\mathbf{r}) = \sum_{\alpha=1}^{N} \left| \psi^\alpha(\mathbf{r}) \right|^2 \qquad \textbf{Eq. 2}$$

This has a nice interpretation: $\left| \psi_i(\mathbf{r}) \right|^2$ is the probability of finding an electron in orbital i at a point $\mathbf{r}$. So the formula above tells us that for a determinant the probability of finding an electron at a point $\mathbf{r}$ is just the sum of the probabilities of finding it in one of the orbitals at that point.

There is a deep theorem (the Hohenberg-Kohn Theorem) that states:

> There exists a functional $E_v[\rho]$ such that, given the ground state density, ρ_0, $E_v[\rho_0]=E_0$ where E_0 is the **exact ground state energy**. Further, for any density, ρ', that is not the ground state density, $E_v[\rho']>E_0$.

This result is rather remarkable. While solving the Schrödinger Equation required a very complicated 3N dimensional wavefunction $\Psi_{el}(\mathbf{R}_1, \mathbf{R}_2,...\mathbf{R}_N)$, this theorem tells us we only need to know the density - which is a 3D function! – and we can get the exact ground state energy. Further, if we don't know the density, the second part of this theorem gives us a simple way to find it: just look for the density that minimizes the functional E_v.

The unfortunate point is that we don't know the form of the functional E_v. We can prove it exists, but we can't construct it. However, from a pragmatic point of view, we do have very good approximations to E_v, and the basic idea is to choose an approximate (but perhaps very, very good) form for E_v and then minimize the energy as a function of the density. That is, we look for the point where $\dfrac{dE_v}{d\rho} = 0$. Based on Eq. 2 above, we see that ρ just

depends on the MOs and hence on the MO coefficients, so once again we are looking for the set of MO coefficients such that $\frac{dE_v}{d\mathbf{c}} = 0$. Given the similarity between DFT and HF, it is not surprising that DFT is also solved by self consistent field iterations. In fact, in a standard electronic structure code, DFT and HF are performed in exactly the same manner (see flow chart above). The only change is the way one computes the energy and $\frac{dE}{d\mathbf{c}}$.

Now, as alluded to above, there exist good approximation<u>s</u> (note the plural) to E_v. Just as was the case with approximate AO basis sets, these approximate energy expressions have strange abbreviations. We won't go into the fine differences between different DFT energy expressions here. I'll simply note that roughly, the quality of the different functionals is expected to follow:

<div align="center">LSDA < PBE ≈ BLYP < PBE0 ≈ B3LYP</div>

Thus, LSDA is typically the worst DFT approximation and B3LYP is typically among the best. I should mention that this is just a rule of thumb; unlike the case for basis sets where we were approaching a well-defined limit, here we are trying various uncontrolled approximations to an unknown functional. Experience shows us that B3LYP is usually the best, but this need not always be the case.

Finally, we note that the speed of a DFT calculation is about the same as Hartree Fock – both involve self-consistent field iterations to determine the best set of orbitals, and so both take about the same amount of computer time. However, for the same amount of effort, you can get quite accurate results. As a rule of thumb, with B3LYP you can get energies correct to within 3 kcal/mol and distances correct to within .01 Å.
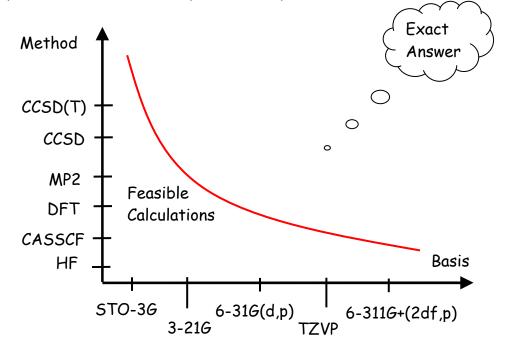
Post-Hartree Fock Calculations

Here, the idea is to employ wavefunctions that are more flexible than a Slater determinant. This can be done by adding up various combinations of Slater determinants, by adding terms that explicitly correlate pairs of electrons (e.g. functions that depend on $\mathbf{r}_1$ and $\mathbf{r}_2$ simulataneously) and a variety of other creative techniques. These approaches are all aimed at

incorporating the **correlation** between electrons – i.e. the fact that electrons tend to spend more time far apart from one another as opposed to close together. This correlation reduces the average repulsion employed in HF and brings us closer to the true ground state energy. In each case, one usually does a Hartree-Fock calculation first and then includes the correlation afterward, leading to the heading of "post-Hartree Fock" methods. Once again, there are a number of acronyms, and we merely assert that the quality of the results goes approximately like:

HF < CASSCF < MP2 < CCSD < MP4 < CCSD(T)

Here, the ordering is rigorous: we have something solid that we are tyring to approximate and going from left to right we are making better and better approximations. On the scale above, DFT typically gives results of about MP2 quality.

As a general rule, post-HF calculations are **much, much more expensive** than HF or DFT and also require bigger basis sets: whereas HF might converge with a DZ or TZ basis, a post-HF calculation might require QZ or even 5Z. Hence, they should only be attempted for relatively small molecules where high accuracy is required

Combining what we have learned, then, the approximations we can make fit nicely into a two-dimensional parameter space:

On the one axis, we have the basis sets we can choose from. On the other, we have the different methods for approximating the energy. The get close to the exact answer, we need to employ a large basis set and an accurate energy method. Unfortunately, both increasing the size of the basis and improving the method tend to slow our calculations down. Given that we don't want to wait years and years to find out the result of a calculation, modern computers therefore limit how accurate our answers can be (as illustrated with the red line above). As we become experts at what is and is not feasible with current computing power, we become better able to get good accuracy for a variety of problems.