Post-Hartree-Fock Methods

Alston J. Misquitta

Centre for Condensed Matter and Materials Physics
Queen Mary, University of London

March 5, 2014

Many electron basis I

Hartree–Fock is an approximation that is exact only for a one-electron system.

Q: Show that for a one-electron system the Fock operator is exactly the same as the exact one-electron Hamiltonian. Hence, the HF solutions are exact. Is this something you would have expected?

We know that the Hartree–Fock ground state energy E_0 will be an *upper bound* to the exact ground-state energy $\mathcal{E}_0$. The difference

$$E_{\rm corr} = \mathcal{E}_0 - E_0$$

is called the *correlation energy*. This energy is often large enough that we cannot neglect it. The question now is how do we calculate $E_{\rm corr}$?

Many electron basis II

To calculate E_{corr} we need to use what are known as post-Hartree–Fock methods. Amongst the main-stream approaches are

- **Configuration Interaction (CI)**: Create a linear expansion of determinants created from the HF ground-state determinant by exciting electrons into the virtual space.
- **Perturbation Theory (MP*n*)**: Develop a perturbative expansion starting with the Hartree–Fock ground state as the zeroth order solution.
- **Coupled-cluster methods (CC)**: Like CI, but with infinite-order summations.
- **Quantum Montecarlo (QMC)**: Variational Montecarlo (VMC), Diffusion Montecarlo (DMC), full configuration interaction quantum montecarlo (FCIQMC).

Many electron basis III

- **Density-functional theory (DFT)**: Reformulate the problem in terms of the electron density.
- **Density matrix theory**: Variant of the above that recognises that uses the two-electron density.
- **Greens function methods**.....

Here we will look at CI, CC, MP2 and (finally) DFT. The other methods require another course altogether!

Configuration Interaction I

CI: Increase the flexibility in the wavefunction by including in addition the the HF ground state, excited states.

Q: What are excited states and how to we form them?

Configuration Interaction II

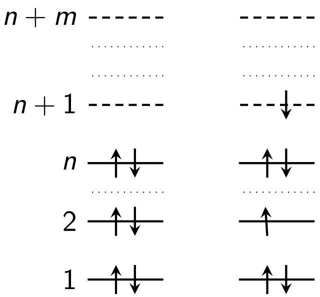


Figure : Left: HF ground state configuration. Right: An example of an excited state configuration. If there are n occupied levels (2 electrons each, so $N = 2n$) and m virtual (un-occupied) levels, in how many ways can we form excited states? Each of these states will correspond to a Slater determinant.

Configuration Interaction III

The Full CI (FCI) wavefunction:

$$
\begin{aligned}
|\Psi\rangle &= c_0|\Psi_0\rangle + \sum_{ia} c_i^a|\Psi_i^a\rangle + \sum_{i<j,a<b} c_{ij}^{ab}|\Psi_{ij}^{ab}\rangle + \cdots \\
&= c_0|\Psi_0\rangle + c_S|S\rangle + c_D|D\rangle + \cdots
\end{aligned}
$$

where electrons are excited from the *occupied* orbitals $i, j, k, \cdots$ to the *virtual* orbitals $a, b, c, \cdots$.

Points to note about CI:

- Because of Brillouin's theorem, there is no (direct) coupling between the HF ground state $|\Psi_0\rangle$ and the single excitations.
- There is no coupling between $|\Psi_0\rangle$ and triples, quadruples, etc. Similarly, singles do not mix with quadruples, quintuples etc.

Q: Why?

Configuration Interaction IV

- The double excitations have the largest effect on the correlation energy because they mix directly with the HF g.s.
- Next in importance are the quadruples.
- Full CI scales exponentially with system size, so it is typically limited to small systems for reference calculations only.

Q: To see why the scaling is so extreme calculate the number of single determinants in the CI expansion.

Configuration Interaction V

CI expansion in *intermediate normalized form* ($c_0 = 1$):

$$|\Psi\rangle = |\Psi_0\rangle + \sum_{ia} c_i^a |\Psi_i^a\rangle + \sum_{i<j,a<b} c_{ij}^{ab} |\Psi_{ij}^{ab}\rangle + \cdots$$

As defined $|\Psi\rangle$ is not normalized. (To see this calculate $\langle\Psi|\Psi\rangle$.) Notice that this form satisfies the intermediate normalization condition

$$\langle\Psi_0|\Psi\rangle = 1$$

Now we have

$$H|\Psi\rangle = \mathcal{E}_0|\Psi\rangle$$

Configuration Interaction VI

rewrite this as

$$(H - E_0)|\Psi\rangle = (\mathcal{E}_0 - E_0)|\Psi\rangle = E_{\mathrm{corr}}|\Psi\rangle$$

Taking the inner product with $\langle\Psi_0|$ we get

$$\langle\Psi_0|(H - E_0)|\Psi\rangle = E_{\mathrm{corr}}\langle\Psi_0|\Psi\rangle = E_{\mathrm{corr}}$$

Now stick in the expansion for the CI wavefunction $|\Psi\rangle$ into this to get:

$$E_{\mathrm{corr}} = \langle\Psi_0|(H - E_0)|\Psi\rangle = \sum_{i<j,a<b} c_{ij}^{ab}\langle\Psi_0|H|\Psi_{ij}^{ab}\rangle$$

Q: Derive this expression for the correlation energy.

Configuration Interaction VII

So we have an expression for the correlation energy in CI and it seems to depend on the coefficients of the doubly excited states only.

Q: How can this be so? Does this mean that we do not need the other terms in the CI expansion? Read Szabo & Ostlund Sec. 4.1.1 to find an answer to this puzzle or, better yet, work it our yourself.

Configuration Interaction VIII

FCI:

$$
\begin{aligned}
|\Psi\rangle &= c_0|\Psi_0\rangle + \sum_{ia} c_i^a|\Psi_i^a\rangle + \sum_{i<j,a<b} c_{ij}^{ab}|\Psi_{ij}^{ab}\rangle + \cdots \\
&= c_0|\Psi_0\rangle + c_S|S\rangle + c_D|D\rangle + \cdots
\end{aligned}
$$

- GOOD This expansion will lead to the *exact* energy *within the basis set used*.
- BAD There are too many determinants!

Configuration Interaction IX

If there are M spin orbitals and N electrons ($N < M$) then how many n-tuple excited determinants will we have?

We have N occupied SOs and $M - N$ virtual SOs. We can select n electrons from the occupied set in $N!/(n!(N-n)!)$ ways. These n electrons can be placed in the virtual space in $(M-N)!/(n!(M-N-n)!)$ ways. So the total number of n-tuple excited determinants is

$$\frac{N!(M-N)!}{(n!)^2(N-n)!(M-N-n)!}$$

which gets large very quickly. For large M,N and n you can use Stirling's approximation to show that the number of exctited determinants scales exponentially.

Configuration Interaction X

One solution to the problem is to use only some of the many determinants. For example we could use only double excitations. This leads to the CID method.

$$
\begin{aligned}
|\Psi^{\mathrm{CID}}\rangle &= |\Psi_0\rangle + \sum_{ij,ab} c_{ij}^{ab} |\Psi_{ij}^{ab}\rangle \\
&= |\Psi_0\rangle + c_D |D\rangle
\end{aligned}
$$

BAD This theory, like all truncated CI methods, is not *size extensive*.

Size-extensivity: If $E(N)$ is the energy of N *non-interacting* identical systems then a method is size-extensive if $E(N) = N \times E(1)$.

Configuration Interaction XI

Q: Is CID size-extensive?

If $\hat{T}_2$ is an operator that creates all double excitations, then we can write the CID wavefunction as

$$\begin{aligned}
|\Psi^{\mathrm{CID}}\rangle &= |\Psi_0\rangle + \sum_{ij,ab} c_{ij}^{ab} |\Psi_{ij}^{ab}\rangle \\
&= |\Psi_0\rangle + c_D |D\rangle \\
&= (1 + \hat{T}_2)|\Psi_0\rangle
\end{aligned}$$

Configuration Interaction XII

The CID wavefunction for each of two identical non-interacting systems will be of that form, so the combined wavefunction will be

$$
\begin{aligned}
|\Psi_A^{\mathrm{CID}}\rangle|\Psi_B^{\mathrm{CID}}\rangle &= (1+\hat{T}_2(A))|\Psi_0^A\rangle(1+\hat{T}_2(B))|\Psi_0^B\rangle \\
&= (1+\hat{T}_2(A)+\hat{T}_2(B)+\hat{T}_2(A)\hat{T}_2(B))|\Psi_0^A\rangle|\Psi_0^B\rangle
\end{aligned}
$$

The last excitation term is a quadruple excitation so it will not be present in the CID wavefunction for the combined A and B systems. Therefore

$$E^{\mathrm{CID}}(AB) \neq E^{\mathrm{CID}}(A) + E^{\mathrm{CID}}(B).$$

CISD: Examples I

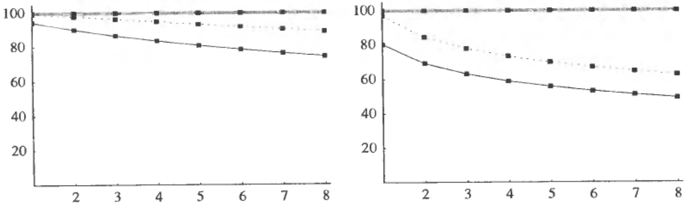


Fig. 11.1. The cc-pVDZ correlation energy per water monomer (as a percentage of the FCI correlation energy) as a function of the number of noninteracting water molecules. The grey line represents the FCI energy, the full black line the CISD energy, and the dotted line the CISD energy with the Davidson correction applied. The calculations have been carried out for the C_{2v} water molecule at the OH separations $R = R_{\mathrm{ref}}$ (left) and $R = 2R_{\mathrm{ref}}$ (right); see Table 11.2.

All figures from: *Molecular Electronic Structure Theory* by Helgaker *et al.*.

CISD: Examples II

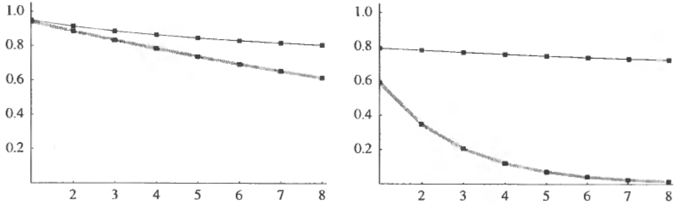


Fig. 11.2. The weight of the Hartree–Fock determinant in the FCI wave function (grey line) and in the CISD wave function (black line) as a function of the number of noninteracting water monomers in the cc-pVDZ basis. The plot on the left corresponds to the molecular equilibrium geometry; the plot on the right represents a situation where the OH bonds have been stretched to twice the equilibrium bond distance. For details on the calculations, see Table 11.2.

CISD: Examples III

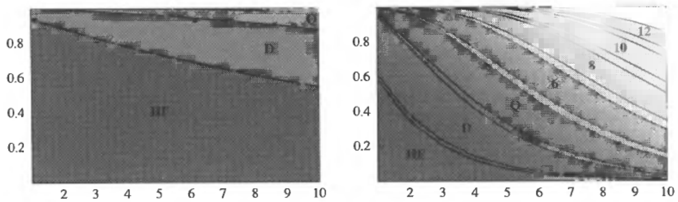


Fig. 11.3. The weights of the excitation levels in the FCI wave function for a system containing one to ten noninteracting water monomers. The plot on the left corresponds to the molecular equilibrium geometry in the cc-pVDZ basis; the plot on the right represents a situation where the OH bonds have been stretched to twice the equilibrium bond distance. For details on the calculations, see Table 11.2.

Coupled-cluster Theory I

The problem of truncated CI methods is severe enough that using them is very problematic. A resolution to the problem is the class of coupled-cluster theories. In these the wavefunction is defined as:

$$|\Psi^{\mathrm{CC}}\rangle = \exp(\hat{T})|\Psi_0\rangle$$

where $\hat{T}$ is an appropriate excitation operator.

Coupled-cluster Theory II

For example, in CCD theory we use $\hat{T}_2$ as the excitation operator. This gives:

$$
\begin{aligned}
|\Psi^{\mathrm{CCD}}\rangle &= \exp(\hat{T}_2)|\Psi_0\rangle \\
&= (1 + \hat{T}_2 + \frac{1}{2!}\hat{T}_2\hat{T}_2 + \cdots)|\Psi_0\rangle
\end{aligned}
$$

The first two terms give us CID theory. The rest are needed to make CCD size-extensive:

$$
\begin{aligned}
\exp(\hat{T}_2(A))|\Psi_0^A\rangle \times \exp(\hat{T}_2(B))|\Psi_0^B\rangle &= \exp(\hat{T}_2(A) + \hat{T}_2(B))|\Psi_0^A\rangle|\Psi_0^B\rangle \\
&\equiv \exp(\hat{T}_2(AB))|\Psi_0^A\rangle|\Psi_0^B\rangle
\end{aligned}
$$

Coupled-cluster Theory III

- GOOD CC theories can be systematically improved.
- GOOD CCSD(T) is a very accurate theory. Here single and double excitations are included as described above and triple excitations are included through a perturbative treatment.
- GOOD Size-extensive.
- BAD Computationally very expensive: CCSD(T) scales as $\mathcal{O}(N^7)$. So double the system size and the calculation costs 128 times more.
- BAD (kind of!) These are single-determinant theories as described. If the system is multi-configurational (more than one state contributing dominantly) the standard CC methods are not appropriate.

Integrals needed for MBPT I

Doubly excited states:
$\Psi_{ij}^{ab} = (N!)^{-1/2}\mathcal{A}\{\chi_1 \cdots \chi_{i-1}\chi_a\chi_{i+1}\cdots\chi_{j-1}\chi_b\chi_{j+1}\cdots\}$

$$\boxed{\langle\Psi_{ij}^{ab}|\sum_{kl}\frac{1}{r_{kl}}|\Psi\rangle = (ai|bj) - (aj|bi)}$$

As there is contribution from the one-electron part we have

$$\boxed{\langle\Psi_{ij}^{ab}|H|\Psi\rangle = (ai|bj) - (aj|bi)}$$

Møller–Plesset Perturbation Theory I

The HF single-determinant wavefunction is *not* an eigenstate of the Hamiltonian:

$$H|\Psi_0\rangle \neq E_0|\Psi_0\rangle.$$

But it is an exact eigenstate of the *Hartree–Fock Hamiltonian*

$$H_0 = \sum_{i}^{N} f(i)$$

Q: Show that it is. And show that the eigenvalue is $E_0^{(0)} = \sum_i \epsilon_i$.

Møller–Plesset Perturbation Theory II

Now we can write $\mathcal{H} = H_0 + \mathcal{V}$, where $\mathcal{V}$ is the *fluctuation operator* defined as

$$\begin{aligned} \mathcal{V} &= \mathcal{H} - H_0 \\ &= \sum_{i,j>i} \frac{1}{r_{ij}} - \sum_i v^{\mathrm{HF}}(i). \end{aligned}$$

That is, $\mathcal{V}$ is the difference in the exact two-electron interaction operator and the approximate HF operator.
We can treate $\mathcal{V}$ as a perturbation to H_0 and expand the energy using Raleigh–Schrödinger perturbation theory:

Møller–Plesset Perturbation Theory III

- Split the Hamiltonian into two parts:

  $$\mathcal{H} = H_0 + \lambda \mathcal{V}$$

  where H_0 is a Hamiltonian which we know how to solve and $\mathcal{V}$ contains that troublesome parts. We expect $\mathcal{V}$ to be a perturbation so it must be small in some sense.
  λ is a complex number that will be 1 for the physical solution.
- Let the solutions of H_0 be:

  $$H_0 \Psi_i^{(0)} = E_i^{(0)} \Psi_i^{(0)}$$

  Here the '0' indicates that these eigenvalues and eigenfunctions are of *zeroth*-order in the perturbation $\mathcal{V}$. We will use the short-form:

  $$|\Psi_i^{(0)}\rangle \equiv |i\rangle$$

Møller–Plesset Perturbation Theory IV

- Express the solutions of $\mathcal{H}$ in a power-series:

$$\Psi_i = \Psi_i^{(0)} + \lambda\Psi_i^{(1)} + \lambda^2\Psi_i^{(2)} + \cdots = \sum_n \lambda^n \Psi_i^{(n)}$$

$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \cdots = \sum_n \lambda^n E_i^{(n)}$$

- Energies can be calculated by collecting terms at various orders:

$$\begin{aligned} E_i^{(0)} &= \langle\Psi_i^{(0)}|H_0|\Psi_i^{(0)}\rangle = \langle i|H_0|i\rangle \\ E_i^{(1)} &= \langle\Psi_i^{(0)}|\mathcal{V}|\Psi_i^{(0)}\rangle = \langle i|\mathcal{V}|i\rangle \\ E_i^{(2)} &= \langle\Psi_i^{(0)}|\mathcal{V}|\Psi_i^{(1)}\rangle \end{aligned}$$

Møller–Plesset Perturbation Theory V

etc. for higher order terms.

- The first-order correction to the wavefunction is given by

$$|\Psi_i^{(1)}\rangle = \sum_{n\neq 0} \frac{|n\rangle\langle n|\mathcal{V}|i\rangle}{E_i^{(0)} - E_n^{(0)}}$$

- So we get

$$E_i^{(2)} = \sum_{n\neq 0} \frac{|\langle n|\mathcal{V}|i\rangle|^2}{E_i^{(0)} - E_n^{(0)}}$$

- This is all we need to derive the experession for MP2.
- Note: We still have not decided what the excited states are.

Møller–Plesset Perturbation Theory VI

Many-body perturbation theory (MBPT) starts from Hartree–Fock theory:

$$H_0 = \sum_{i=1}^{N} f(i) = \sum_{i=1}^{n} \left(h(i) + v^{\mathrm{HF}}(i) \right) \tag{1}$$

where $h(i) = -\frac{1}{2}\nabla_i^2 - \sum_\alpha \frac{Z_\alpha}{r_{i\alpha}}$ and

$$\mathcal{V} = \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} - \sum_{i=1}^{N} v^{\mathrm{HF}}(i) \tag{2}$$

Unlike v^{HF}, the perturbation $\mathcal{V}$ is a 2-electron operator.

Møller–Plesset Perturbation Theory VII

MBPT energy at λ^0:

$$E_0^{(0)} = \langle 0|\mathcal{F}|0\rangle = \sum_{a\in\mathrm{occ}} \epsilon_a$$

At first-order we get:

Q: Show that $E_0^{(1)} = \langle\Psi_0|V|\Psi_0\rangle = -\frac{1}{2}\sum_{ij}\langle ij||ij\rangle$

The sum of $E_0^{(0)}$ and $E_0^{(1)}$ is just the Hartree–Fock ground state energy (something we saw in the last lecture):

$$E^{\mathrm{HF}} = E_0^{(0)} + E_0^{(1)} \tag{3}$$

This means that we need to get to at least second-order in perturbation theory to go beyond the Hartree–Fock description.

Møller–Plesset Perturbation Theory VIII

At second order, we have

$$E_0^{(2)} = \sum_{n \neq 0} \frac{|\langle n|\mathcal{V}|0\rangle|^2}{E_0^{(0)} - E_n^{(0)}}$$

We will take state $|0\rangle$ to be the HF g.s., i.e., $|0\rangle = |\Psi_0\rangle$. But what about the excited states $|n\rangle$?

- Can they be single excitations: $|\Psi_i^a\rangle$? To see why not evaluate the matrix element $\langle\Psi_i^a|\mathcal{V}|\Psi_0\rangle$.
- They cannot be triple or higher excitations as matrix elements of these with the HF g.s. are zero. Why?
- So they have to be double excitations: $|\Psi_{ij}^{ab}\rangle$.

Møller–Plesset Perturbation Theory IX

Q: Show the following results:

$$H_0|\Psi_{ij}^{ab}\rangle = (E_0^{(0)} - (\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b))|\Psi_{ij}^{ab}\rangle$$

$$\langle\Psi_{ij}^{ab}|\mathcal{V}|\Psi_0\rangle = \langle ij||ab\rangle$$

We therefore get the second-order MBPT energy expression:

$$E_0^{(2)} = \sum_{i<j}^{\text{occ}} \sum_{a>b}^{\text{vir}} \frac{[\langle ij||ab\rangle]^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{4}$$

This expression is termed as MBPT2 or MP2. The latter name comes from the other name for this kind of perturbation theory: Møller–Plesset perturbation theory.

Møller–Plesset Perturbation Theory X

- BAD A problem with Møller–Plesset perturbation theory: it diverges! See Olsen *et al.* J. Chem. Phys. **112**, 9736 (2000) for details. We now rarely go beyond MP2 in practical calculations.
- GOOD MP2 contains correlation.
- BAD But not enough correlation. Problems with systems with small HOMO-LUMO gaps (band gap — HOMO is highest occupied MO and LUMO is lowest unoccupied MO).
- GOOD (kind of!) It has a computational cost of $\mathcal{O}(N^5)$. I.e., double the system in size and it will cost 32 times more computational power.
- GOOD MBPT is *size-consistent*

Apparent and intrinsic errors I

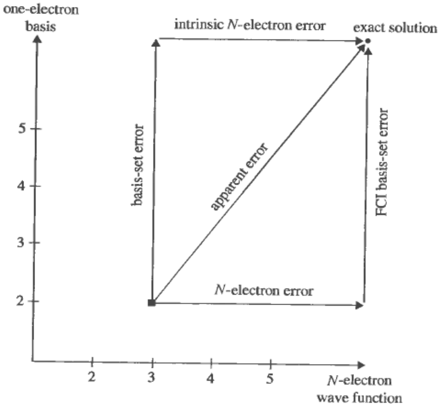


Fig. 15.1. The errors in quantum-chemical calculations.

Apparent and intrinsic errors II

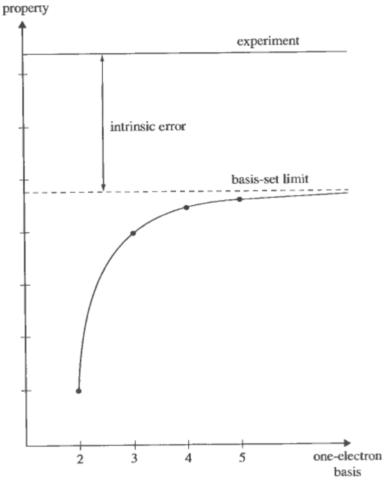


Fig. 15.2. The dependence of the calculated property on the quality of the basis.

Apparent and intrinsic errors III

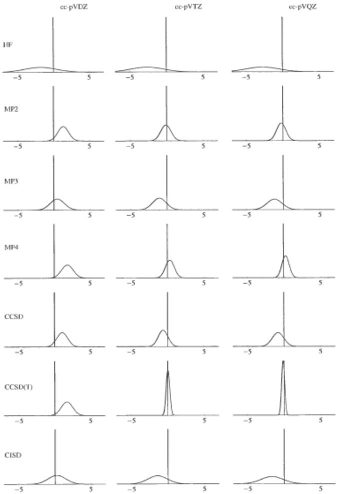


Fig. 15.4. Normal distributions of the errors in the calculated bond distances (pm). For ease of comparison, all distributions have been normalized to one and plotted against the same horizontal and vertical scales.

Apparent and intrinsic errors IV

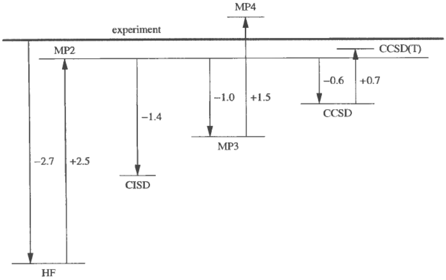


Fig. 15.5. The relationships between the calculated equilibrium bond distances for the standard models (in pm).

Apparent and intrinsic errors V

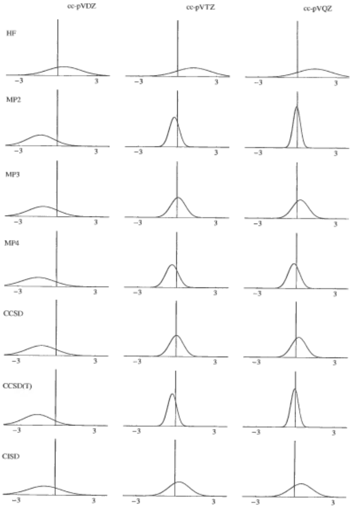


Fig. 15.7. Normal distributions of the errors in the calculated bond angles (in degrees) relative to the experimental values in Table 15.8.

Which method and basis? I

The Gold Standard:

CCSD(T) / aug-cc-pVTZ (or larger)

If not, use MP2, but with caution! Or else, use DFT (next lecture).

Many electron basis I

Practical considerations: CI, CC and MP2 are expensive!

Method	Cost	Description
HF	N^4	Starting point for correlated methods.
MP2	N^5	OK. Useful for optimizations.
CISD	N^6	Not size-consistent.
CCSD(T)	N^7	Very accurate.

All correlated methods require large basis sets with high angular functions.

The electron–electron cusp I

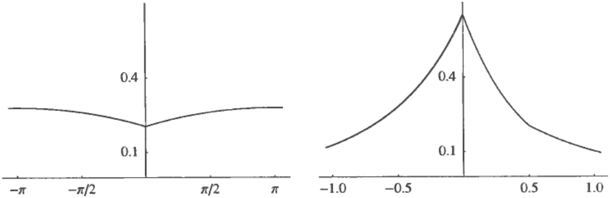


Fig. 7.5. The electronic and nuclear cusps of the ground-state helium atom with one electron fixed at a position $0.5a_0$ from the nucleus (atomic units). On the left, the wave function is plotted on a circle of radius $0.5a_0$ centred at the nucleus; on the right, the wave function is plotted along the axis through the nucleus and the fixed electron.

Correlation is the description of this kink.

All figures from "Molecular Electronic Structure Theory" by Helgaker, Jorgensen and Olsen

The electron–electron cusp II

Principle and partial wave expansions.

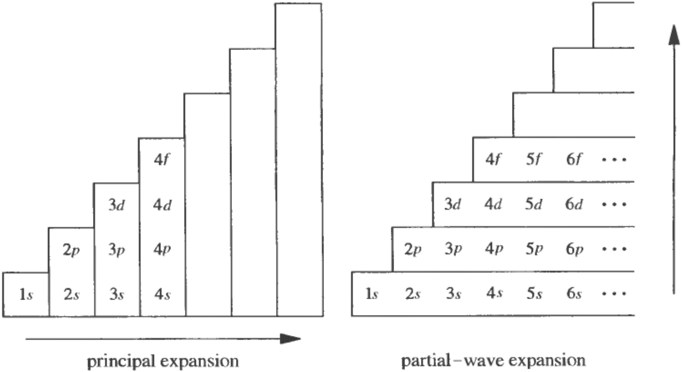
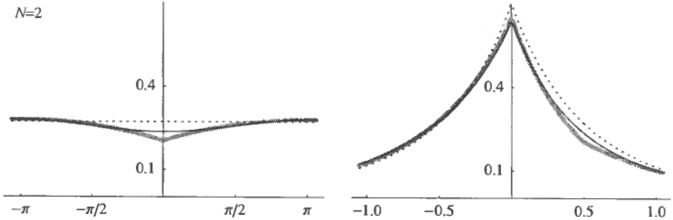


Fig. 7.9. The principal and partial-wave expansions.

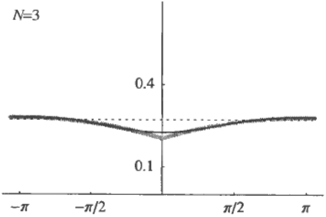
The electron–electron cusp III

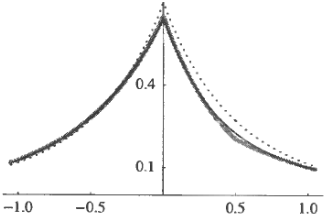
He_2: Principle wave expansion with one electron fixed at 0.5Å:
$n = 2$



The electron–electron cusp IV

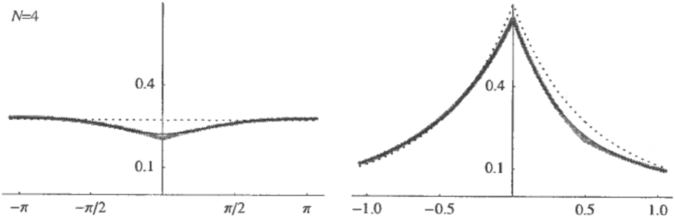
He$_2$: Principle wave expansion with one electron fixed at 0.5Å:
$n = 3$





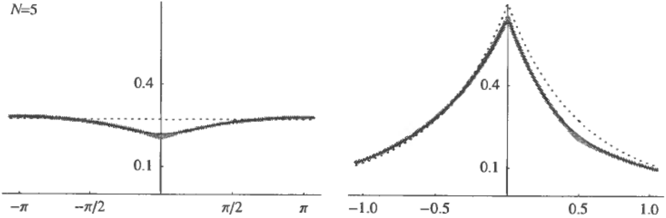
The electron–electron cusp V

He_2: Principle wave expansion with one electron fixed at 0.5Å:
$n = 4$



The electron–electron cusp VI

He_2: Principle wave expansion with one electron fixed at 0.5Å:
$n = 5$



These cusps make correlation hard. Not only do the methods scale poorly with size, but we need rather large basis sets to get sensible results.