TCC Electronic Structure Module
L3: Configuration Interaction Theory

David P. Tew

February 15, 2021

Overview

These lectures introduce electronic structure theory. This is an enormous field that has undergone continuous development for near 100 years and the scope of these lectures is necessarily limited. The focus will be on the central theory as applied to molecules. The aim is to introduce you the basics of molecular electronic structure theory, which will help you understand the level of sophistication of calculations you find in the literature, and provide the foundation for further study in the field.

L1 Exact and inexact wavefunctions

L2 Hartree–Fock theory

L3 Configuration Interaction theory

L4 Møller–Plesset perturbation theory

L5 Coupled-cluster theory

L6 Properties and excited states

L7 Density Functional Theory

L8 Density Functional Approximations

Recommended books

- Good introduction to the topic: Frank Jensen, *Introduction to Computational Chemistry* (Wiley, 2007)
- Excellent concise treatise: Szabo and Ostlund, *Modern Quantum Chemistry* (Dover, 1996)
- Comprehensive reference for quantum chemistry methods (excluding DFT): Helgaker, Jørgensen and Olsen, *Molecular Electronic-Structure Theory* (Wiley, 2002)
- Excellent book for many-body correlation: Bartlett and Shavitt, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory* (CUP, 2009)

1 Introduction

In L2 we introduced Hartree–Fock theory, which provides an important approximation to the ground state of the time-independent non-relativistic electronic Hamiltonian. Hartree–Fock theory is based on the independent electron model, where the n electrons occupy n spin orbitals and each electron moves in the Fermionic mean-field of the other electrons. The correlation among electrons arising from the instantaneous Coulomb repulsion among the electrons is not accounted for in Hartree–Fock theory. This lecture deals with the general topic of electron correlation, arising from both the Fermionic nature of electrons and from their Coulombic repulsion. We will examine the extent to which correlation is and is not present in the Hartree–Fock wavefunction and how configuration interaction theory systematically improves upon Hartree–Fock theory. Two electron systems provide an accessible paradigm within which to illustrate these concepts and we first discuss the helium atom and hydrogen molecule before turning to configuration interaction theory for many-electron systems.

Further reading: this lecture contains an abridged and updated presentation of the material in *Electron correlation: The Many-body Problem a the Heart of Chemistry*, Tew, Klopper and Helgaker, Journal of Computational Chemistry 28, 1307 (2007) [DOI 10.1002/jcc]

2 Lessons from helium

2.1 Fermi correlation

If we consider the two electrons in a helium atom as independent particles we may write the wave function as a Hartree product of two occupied spin-orbitals

$$\Psi_{\mathrm{H}} = \varphi_i(\mathbf{r}_1)\sigma_i(1)\varphi_j(\mathbf{r}_2)\sigma_j(2) \tag{1}$$

The probability of finding the two electrons at two given points in space and with a given pair of spins is proportional to $|\Psi|^2$. For Ψ_{H} this becomes a product of one factor for each electron and in this Hartree description the motion of one electron is completely independent of that of the other electron and their motion is said to be *uncorrelated*.

The Hartree description, however, is fundamentally flawed because it distinguishes between the two electrons, treating them differently. In the Hartree–Fock description the Hartree product is antisymmetrised with respect to all electron permutations so that the Fermionic property $\hat{\pi}_{ij}\Psi = -\Psi$ is properly described, for any pair of electrons ij:

$$\Psi_{\mathrm{HF}} = \frac{1}{\sqrt{2}}\Big[\varphi_i(\mathbf{r}_1)\sigma_i(1)\varphi_j(\mathbf{r}_2)\sigma_j(2) - \varphi_j(\mathbf{r}_1)\sigma_j(1)\varphi_i(\mathbf{r}_2)\sigma_i(2)\Big] \tag{2}$$

This fermionic indistinguishability introduces a large and fundamental correlation between electrons with the same spin. Consider the configurations belonging, respectively, to the lowest singlet and triplet states of helium:

$$\Psi^1_{\mathrm{HF}} = \frac{1}{\sqrt{2}}\varphi_{1s}(\mathbf{r}_1)\varphi_{1s}(\mathbf{r}_2)\Big[\alpha(1)\beta(2) - \beta(1)\alpha(2)\Big] \tag{3}$$

$$\Psi^3_{\mathrm{HF}} = \frac{1}{\sqrt{2}}\Big[\varphi_{1s}(\mathbf{r}_1)\varphi_{2s}(\mathbf{r}_2) - \varphi_{2s}(\mathbf{r}_1)\varphi_{1s}(\mathbf{r}_2)\Big]\alpha(1)\alpha(2) \tag{4}$$

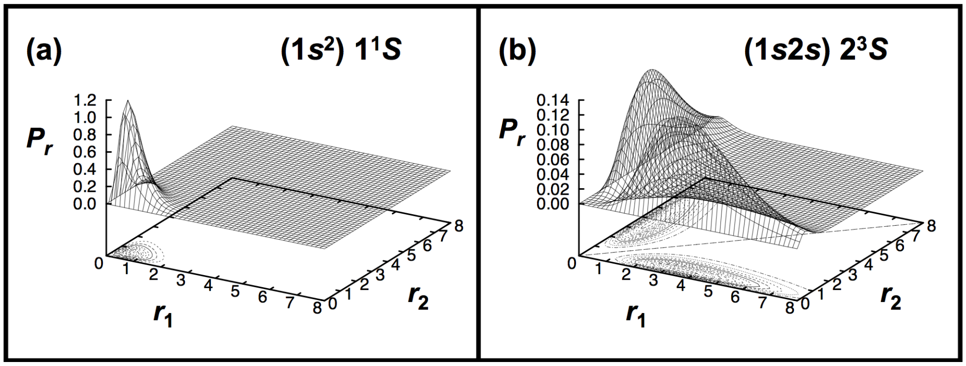

Figure 1: The two-electron radial distribution functions of the helium atom in the orbital approximation.

The pair radial probability distributions are plotted in Figure 1. For the singlet electron pair, the spatial distribution of the electrons is uncorrelated. For the triplet pair, it is much more probable that when one electron is close to the nucleus (in a $1s$ orbital), the other electron is further out (in a $2s$ orbital). There is identically zero probability of the two electrons being found at the same distance from the nucleus. This is a specific example of the phenomenon known as the Fermi hole, an excluded volume where two electrons with the same spin cannot occupy the same point in space. This fundamental Fermi correlation of electrons is correctly described by Hartree–Fock theory.

2.2 Coulomb correlation

The Hartree–Fock description of the ground singlet state of the helium atom is that of an uncorrelated probability distribution of the two electrons. However, since the two electrons are negatively charged, they interact through their Coulombic repulsion resulting in *correlated* motion that reduces the probability of finding the two electrons close together and increases the probability of finding them far apart. Wavefunctions that account for this, over and above any Fermi correlation, are called correlated wavefunctions. The resulting reduction in the electronic energy relative to the Hartree–Fock energy, due to reduced Coulomb repulsion, is called the *correlation energy.*

$$E_{\mathrm{corr}} = E_{\Psi} - E_{\mathrm{HF}} \tag{5}$$

Let us examine Coulomb correlation in more detail through the example of the helium atom. The variational principle alows us to use any wavefunction expansion and vary its parameters to optimise the wavefunction by minimising the energy. As discussed in L1, we can systematically converge towards the exact wavefunction through an expansion of the kind

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \sum_I f_I(\mathbf{x}_1, \mathbf{x}_2) C_I \tag{6}$$

where each of the functions f_I is an antisymmetrised product of orthonormal one-electron functions drawn from a complete set. One such set are the eigenstates of the Fock operator, $\varphi_p(\mathbf{r}_1)\sigma_p(1)$. Using this basis, each function f_I is therefore a Slater determinant for an excited state configuration (often denoted with

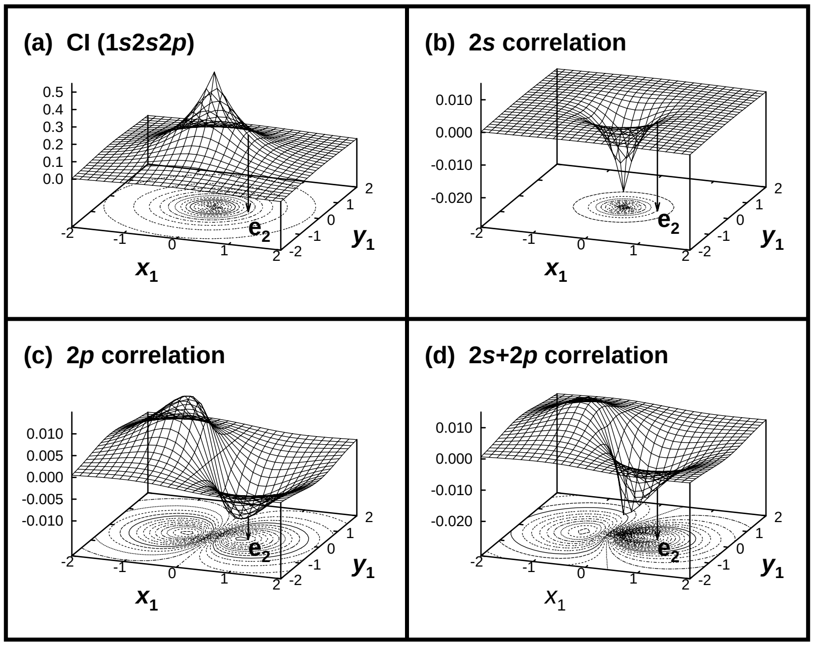

Figure 2: Coulomb correlation in the helium atom. The helium wave function and contributions to it plotted with one electron fixed at 0.65 a_0 from the nucleus, plotted in a plane containing the nucleus and the two electrons (note different scales): (a) the CI wave function calculated from 1s, 2s, and 2p; (b) the contribution to the CI wave function from the 2s orbital; (c) the contribution to the CI wave function from the 2p orbitals; (d) the combined contributions to the CI wave function from 2s and 2p.

short-hand $|pq\rangle$), and the expansion is

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \sum_{p>q} C_{pq} \big[\varphi_p(\mathbf{r}_1)\varphi_q(\mathbf{r}_2)\alpha(1)\beta(2) - \varphi_q(\mathbf{r}_1)\varphi_p(\mathbf{r}_2)\beta(1)\alpha(2)\big] \tag{7}$$

$$|\Psi\rangle = \sum_{p>q} C_{pq}|pq\rangle \tag{8}$$

where the coefficients C_{pq} are found by solving the matrix eigenvalue equation $\mathbf{HC} = E\mathbf{C}$ (c.f. L1 and noting that the configurations are orthonormal). This is called *configuration interaction* theory. The influence of the first few terms in this expansion on the helium wavefunction is depicted in figure 2. Configurations of the tpye $\varphi_p, \varphi_q = 1s1s$, $1s2s$ and $2s2s$ introduce radial flexibility into the wavefunction, where if one electron is close to the nucleus, it is more likely for the other to be farther out. Configurations involving the $2p$ x, y, z orbitals introduce angular flexibility, where if one electron is on one side of the nucleus, it is more likely that the other electron is on the opposite side.

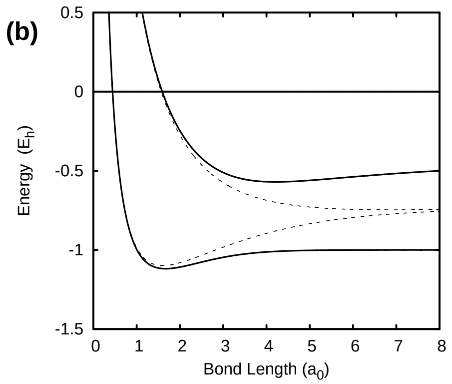

Figure 3: Coulomb correlation in closed-shell states. The dotted lines represent the two single determinants; the full lines represent the variationally optimised linear combinations of these determinants.

3 Lessons from hydrogen

3.1 Static and dynamic correlation

Before we consider configuration interaction for many-electron systems, it is informative to examine another simple system, the hydrogen molecule. The lowest two molecular orbitals for the hydrogen molecule are the sigma bonding and antibonding orbitals, which are plus and minus combinations of $1s$ orbitals on each hydrogen, respectively

$$\sigma_g(\mathbf{r}) = N_g\big[1s_A(\mathbf{r}) + 1s_B(\mathbf{r})\big] \tag{9}$$

$$\sigma_u(\mathbf{r}) = N_u\big[1s_A(\mathbf{r}) - 1s_B(\mathbf{r})\big] \tag{10}$$

In terms of these orbitals the ground state wavefunction has one electron in each of the σ_g spin-orbitals

$$\frac{1}{\sqrt{2}}\sigma_g(\mathbf{r}_1)\sigma_g(\mathbf{r}_2)\big[\alpha(1)\beta(2) - \beta(1)\alpha(2)\big] \tag{11}$$

and the lowest energy wavefunction of the same Σ_g^+ symmetry has one electron in each of the σ_u spin-orbitals

$$\frac{1}{\sqrt{2}}\sigma_u(\mathbf{r}_1)\sigma_u(\mathbf{r}_2)\big[\alpha(1)\beta(2) - \beta(1)\alpha(2)\big] \tag{12}$$

Figure 3 contains a plot of the energy of these two states as a function of the separation of the hydrogen atoms. It also shows the energy from the wavefunctions obtained by mixing these two states, allowing for the effect of Coulomb correlation. At short distances the two single determinant states are well separated in energy, but they converge to the same energy as the molecule dissociates. Close to equilibrium, the single determinant description is good and the effect of Coulomb correlation is to lower the energy slightly. At long distances, the single determinant description is poor and accounting for Coulomb correlation leads to a large mixing of the two determinants and a large correlation energy.

In the regime where a single determinant is qualitatively correct, we describe the Coulomb correlation as dynamic, or weak correlation. In the regime where a single determinant is qualitatively incorrect, we describe the Coulomb correlation as static, or strong correlation. In this regime Hartree–Fock theory breaks down and we require a multi-determinant description to describe the electronic state properly.

4 Many-electron configurations

The simple two-electron systems of hydrogen and helium provide important insights into many-electron systems. They serve as good illustrations for the basic concepts of electron correlation, and in addition, since the Coulomb interactions in the electronic Hamiltonian are pairwise additive, pair-correlations are the largest corrections to Hartree–Fock theory and the characteristics of singlet and triplet pairs carry directly over to the many-electron case.

The n-electron Hartree-Fock wavefunction is an antisymmetrised product of n spin-orbitals corresponding to a particular electron configuration (e.g. the Berylium atom in its ground state has occupied orbitals $1s\alpha 1s\beta 2s\alpha 2s\beta$). The Hartree–Fock wavefunction can be expressed in first- or second-quantised forms

$$|0\rangle = \sqrt{n!}\,\hat{A}\,|\varphi_1\varphi_2\cdots\varphi_n\rangle \tag{13}$$

$$|0\rangle = \hat{a}_n^\dagger\cdots\hat{a}_2^\dagger\hat{a}_1^\dagger|\text{vac}\rangle \tag{14}$$

The spin-orbitals (one-electron states) are the solutions to the Roothan–Hall self-consistent field Hartree–Fock pseudo-eigenvalue equations, where the orbitals are expanded as a linear combination of atomic orbitals

$$\sum_\nu F_{\mu\nu}C_{\nu p} = \sum_\nu S_{\mu\nu}C_{\nu p}\epsilon_p \tag{15}$$

$$\varphi_p = \sum_\mu C_{\mu p}\chi_\mu \tag{16}$$

If there are m atomic spin-orbitals in the expansion, there are m spin-orbitals that satisfy the above equation. To define a n-electron wavefunction we must select n of the possible m spin-orbitals. In the Hartree–Fock wavefunction, the n spin-orbitals with the lowest eigenvalues ϵ_p are selected. These are referred to as the occupied spin-orbitals and the remaining $m-n$ referred to as virtual spin-orbitals. All of the remaining of the m choose n possible configurations are higher in energy and are referred to as excited state configurations. Indeed, the antisymmetrised wavefunctions (Slater determinants) for these configurations can be used to (poorly) approximate excited states of the system.

Occupation number representation:
Any of the m choose n configurations is completely specified by an occupation number vector. In this representation, we place the m spin-orbitals in an ordered list (alpha orbitals first, lowest energy to highest, then beta in the same way). We then use a vector of length m to label the configuration, where each element is zero unless that orbital is occupied in the configuration in which case the element value is one. This specifies the configuration and specifies the antisymmetrised wavefunction.

Notation: We use $i, j, k, \ldots$ to denote the occupied spin-orbitals (those occupied in the HF wavefunction); we use $a, b, c, \ldots$ to denote the virtual spin-orbitals; and we use $p, q, r, \ldots$ when neither is specified.

We may use the Hartree–Fock wavefunction to classify the excited state configurations as being related to the HF state through a replacement of a single orbital, a double replacement, triple, etc.

Single replacements $|{}^{a}_{i}\rangle = a^\dagger_a a_i |0\rangle$

Double replacements $|{}^{ab}_{ij}\rangle = a^\dagger_b a_j a^\dagger_a a_i |0\rangle$

Triple replacements $|{}^{abc}_{ijk}\rangle = a^\dagger_c a_k a^\dagger_b a_j a^\dagger_a a_i |0\rangle$

etc

The process of replacing an occupied orbital with a virtual orbital is called an excitation. An electron in a low-lying one-particle state (occupied spin-orbital) is excited, perhaps through some interaction, into a different one-partilce state with higher energy.

5 Slater–Condon rules

Second quantisation provides a natural framework within which to evaluate the matrix elements of operators over configurations. The operators of particular relevance are the unit operator, for the overlap of two configurations, and the Hamiltonian operator, for the transition matrix element between two configurations.

Let $|P\rangle$ denote a slater determinant of the orbitals $p_1, \ldots, p_n$

$$|P\rangle = \prod_i a^\dagger_{p_r} |\text{vac}\rangle \tag{17}$$

$$\langle P| = \langle \text{vac}| \prod_r a_{p_r} \tag{18}$$

The overlap between two determinants is

$$\langle Q|P\rangle = \langle \text{vac}| \prod_r a_{q_r} \prod_r a^\dagger_{p_r} |\text{vac}\rangle = \delta_{PQ} \tag{19}$$

This results from the fact that the one-electron states (spin-orbtials) are orthonormal and that a non-zero value is only obtained if the states that are created are the same as those that are annihilated.

The full Hamiltonian in second quantisation is

$$\hat{H} = \sum_{pq} h_p^q a_q^\dagger a_p + \tfrac{1}{4} \sum_{pqrs} g_{rs}^{pq} a_p^\dagger a_q^\dagger a_s a_r \tag{20}$$

where the matrix elements $h_p^q = \langle q|\hat{h}|p\rangle$ and $g_{rs}^{pq} = \langle pq|r_{12}^{-1}|rs\rangle - \langle qp|r_{12}^{-1}|rs\rangle$. The expectation value for the core hamiltonian, as an example of one-electron operators, is

$$\langle P| \sum_{rs} h_r^s a_s^\dagger a_r |P\rangle = \sum_{p_r} h_{p_r}^{p_r} \tag{21}$$

$$\langle Q| \sum_{rs} h_r^s a_s^\dagger a_r |P\rangle = h_p^q \tag{22}$$

if $|Q\rangle$ and $|P\rangle$ differ by only the orbitals p and q and is zero if the configurations $|P\rangle$ and $|Q\rangle$ differ by more than one orbital. The expectation value for the Coulomb interaction, as an example of two-electron operators, is

$$\langle P| \sum_{pqrs} g_{rs}^{pq} a_p^\dagger a_q^\dagger a_s a_r |P\rangle = \sum_{p_r q_r} g_{p_r q_r}^{p_r q_r} \tag{23}$$

$$\langle Q| \sum_{pqrs} g_{rs}^{pq} a_p^\dagger a_q^\dagger a_s a_r |P\rangle = \sum_{p_r} g_{p p_r}^{q p_r} \tag{24}$$

$$\langle Q| \sum_{pqrs} g_{rs}^{pq} a_p^\dagger a_q^\dagger a_s a_r |P\rangle = g_{pq}^{rs} \tag{25}$$

where the second equation holds if $|Q\rangle$ and $|P\rangle$ differ by only the orbitals p and q and the third holds if $|Q\rangle$ and $|P\rangle$ differ by two orbitals, pq and rs and zero otherwise.

6 Configuration Interaction Theory

We have already introduced configuration interaction theory for two-electron systems. The extension to n-electron systems is straightforward. Denoting an arbitrary configuration of an n-electron system as $|P\rangle$, the CI wavefunction is

$$|\mathrm{CI}\rangle = \sum_P C_P |P\rangle \tag{26}$$

The Rayleigh–Ritz variational principle is used to determine the coefficients C_P

$$\min_{C_P} \langle E\rangle = \min_{C_P} \frac{\langle \mathrm{CI}|\hat{H}|\mathrm{CI}\rangle}{\langle \mathrm{CI}|\mathrm{CI}\rangle} \tag{27}$$

It is straightforward to show, for example using the technique of Lagrange multipliers, that this is equivalent to solving the generalised matrix eigenvalue problem

$$\sum_Q H_{PQ} C_P = \sum_Q E S_{PQ} C_Q \tag{28}$$

with

$$H_{PQ} = \langle P|\hat{H}|Q\rangle \tag{29}$$

$$S_{PQ} = \langle P|Q\rangle = \delta_{PQ} \tag{30}$$

The eigenvector with the lowest eigenvalue is the best possible description of the ground electronic state achievable using the given atomic orbital basis and the eigenvalue is the energy of that state. Similarly, each of the other eigenvectors/eigenvalues correspond to the wavefunctions/energies of excited states and are also the best possible within the atomic orbital basis.

If all possible m choose n configurations are included in the expansion, the method is referred to as Full CI (FCI). The number of possible configurations grows combinatorially with the size of the system, as does the dimension of the matrix eigenvalue problem. We say that the cost of the method scales exponentially with system size.

To arrive at a method that has polynomial cost with system size and can be applied to larger systems, the number of configurations in the CI expansion can be reduced by making an a priori selection of those that are likely to be the most important. The most common approaches for this are:

- The CIS, CISD, CISDT… hierarchy of methods includes only those configurations that are related to the Hartree–Fock reference by single, double, triple… excitations.
- The CASSCF method selects a small set of “active” orbitals (some occupied and some virtual) within which a manageable FCI calculation is performed. In the CASSCF method the core and active orbitals are additionally optimised self consistently in a manner similar to Hartree–Fock theory.
- The MRCI method where in addition to the CASSCF configuration space, configurations obtained by single and double excitations from any of the CASSCF configurations are also included.