TCC Electronic Structure Module
L6: Properties and excited states

David P. Tew

February 21, 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 order to relate results from quantum chemical calculations to experiment, we need to compute quantities that are directly available from measurements that characterise a molecules behaviour or spectroscopic signature. In lectures 1–5 we have focused on methods for computing the energy of the ground state of the electrons. This lecture concerns how to use the same methodologies to compute molecular properties.

2 Properties as energy derivatives

Many molecular properties are the response of a molecule to an external stimulus and can be computed as derivatives of the energy with respect to the perturbation. For example, consider a molecule in a static electric field $\mathcal{E}$. The Hamiltonian is

$$\hat{H}(\mathcal{E}) = \hat{H}(0) - \mu \cdot \mathcal{E} \tag{1}$$

If we treat the field as a weak perturbation, we can Taylor expand the energy around the field free case

$$E(\mathcal{E}) = E(0) + \frac{dE}{d\mathcal{E}}\bigg|_0 \mathcal{E} + \frac{1}{2}\frac{d^2E}{d\mathcal{E}^2}\bigg|_0 \mathcal{E}\mathcal{E} + \dots \tag{2}$$

By definition the first derivative is the electronic dipole moment, the second derivative is the electronic polarisibility tensor and so on

$$\mu = \frac{dE}{d\mathcal{E}}\bigg|_0 \tag{3}$$

$$\alpha = \frac{d^2E}{d\mathcal{E}^2}\bigg|_0 \tag{4}$$

Similarly, other molecular properties are derivatives of the energy with respect to other perturbations:

$\frac{dE}{d\mathbf{R}_I}$	Force on nucleus I (for MD or locating stationary points)
$\frac{d^2E}{d\mathbf{R}_I d\mathbf{R}_J}$	Harmonic force constants (for IR spectra)
$\frac{d^2E}{d\mathbf{R}_I d\mathcal{E}}$	Dipole derivative (for IR intensities)
$\frac{d^2E}{d\mathbf{B} d\mathbf{B}}$	Magnetisability
etc	

Moreover, the transition energy between the ground electronic state and an excited state can be identified by examining the frequency dependence of the polarisibility when applying an oscillating electric field. When the frequency matches the transition energy, the molecule can transition between the states and becomes infinitely polarisable.

Let us turn to the task at hand, which is to compute a derivative of the energy obtained from an optimised wavefunction. We begin by some general considerations. In a quantum chemical calculation, the energy is a function of external parameters x and wavefunction parameters λ: $\varepsilon(x, \lambda)$. x refers to the molecular geometry and any other perturbational parameter such as electric field and λ refers to the MO coefficients and CI coefficients or excitation amplitudes. If the wavefunction parameters have been determined variationally, then

$$\frac{\partial \varepsilon(x,\lambda)}{\partial \lambda} = 0 \qquad \text{for all x} \tag{5}$$

Examples of fully variational methods are Hartree–Fock theory, full configuration interaction theory or MCSCF theory. Truncated CI methods such as CISD or MRCI are variational with respect to the CI coefficients, but not with respect to the MO coefficients. Methods where the energy is obtained by projective formula such as perturbation theory or coupled cluster theory are neither variational with respect to the MO coefficients nor the amplitudes. Let us deal first with fully variational methods.

3 Derivatives of variational methods

The energy is obtained by inserting the optimal λ^* into the energy function.

$$E(x) = \varepsilon(x, \lambda^*) \tag{6}$$

The first derivative at $\lambda = \lambda^*$ is simply

$$\frac{dE(x)}{dx} = \frac{\partial \varepsilon(x,\lambda)}{\partial x} + \frac{\partial \varepsilon(x,\lambda)}{\partial \lambda}\frac{\partial \lambda}{\partial x} \tag{7}$$

$$= \frac{\partial \varepsilon(x,\lambda)}{\partial x} \tag{8}$$

For variational wavefunctions the derivative of the energy with respect to the wavefunction parameters is zero and we do not need to compute the first-order response $\partial\lambda/\partial x$ of the wavefunction parameters to the change in x. The second derivative only slightly more complicated

$$\frac{d^2E(x)}{dx^2} = \frac{\partial^2 \varepsilon(x,\lambda)}{\partial x^2} + 2\frac{\partial^2 \varepsilon(x,\lambda)}{\partial x \partial \lambda}\frac{\partial \lambda}{\partial x} + \frac{\partial^2 \varepsilon(x,\lambda)}{\partial \lambda^2}\left(\frac{\partial \lambda}{\partial x}\right)^2 \tag{9}$$

Note that the term $\partial\varepsilon(x,\lambda)/\partial\lambda(\partial^2\lambda/\partial x^2)$ is missing because it vanishes. We need the first-order response $\partial\lambda/\partial x$ of the wavefunction to calculate the second-derivative. In general, the derivatives of the wavefunction to order n determine the derivatives of the energy to order $2n+1$.

The first-order response $\partial\lambda/\partial x$ is the change in the wavefunction parameters due to the perturbation. It is obtained by differentiating the equations that determine the electronic wavefunction parameters, which for a variational method is $\partial\varepsilon(x,\lambda)/\partial\lambda = 0$. This results in a set of linear equations, called the response equations, that determine $\partial\lambda/\partial x$.

$$\frac{d}{dx}\left(\frac{\partial \varepsilon(x,\lambda)}{\partial \lambda}\right) = \frac{\partial^2 \varepsilon(x,\lambda)}{\partial x \partial \lambda} + \frac{\partial^2 \varepsilon(x,\lambda)}{\partial \lambda^2}\left(\frac{\partial \lambda}{\partial x}\right) = 0 \tag{10}$$

3.1 The Hellman–Feynman Theorem

The result in Eq. (8) is a particular case of a more general result known as the Hellman–Feynman theorem. According to the Hellman–Feynman theorem, the derivative of a variational energy of a normalised wavefunction is

$$\frac{d}{dx}\langle\Psi|\hat{H}|\Psi\rangle = \langle\Psi|\frac{\partial\hat{H}}{\partial x}|\Psi\rangle \tag{11}$$

This is always true if Ψ is an eigenfunction of $\hat{H}$ (Exercise 3). It is also true for approximate wavefunctions when the parameters of Ψ have been determined variationally.

As an example of the application of the Hellman–Feynman theorem, let us consider again the Hamiltonian of a molecule in a static electric field $\mathcal{E}$

$$\hat{H}(\mathcal{E}) = \hat{H}(0) - \hat{\mu}\cdot\mathcal{E} \tag{12}$$

The dipole moment in the z direction is

$$\frac{d}{d\mathcal{E}_z}\langle\Psi(\mathcal{E})|\hat{H}(\mathcal{E})|\Psi(\mathcal{E})\rangle\Big|_0 = -\langle\Psi(0)|\hat{\mu}_z|\Psi(0)\rangle \tag{13}$$

For Hartree–Fock and truncated CI wavefunctions where the wavefunction parameters have been determined variationally, the dipole moment can be evaluated as an expectation value of the dipole moment operator using the wavefunctions optimised for the molecule in the absence of an electric field.

However, for calculations in a finite basis, the Hellman–Feynman theorem does not apply for a geometric distortion. This is because the parameters of the atomic orbital basis are non-variational parameters and additional terms appear in the derivative due to the dependence of the basis functions on the location of the nuclei. These additional corrections are sometimes referred to as Pulay terms.

4 Derivatives of non-variational methods

For non-variational methods such as perturbation theory and coupled-cluster theory, straightforward differentiation of the energy function leads to equations that require the first-order response $\partial\lambda/\partial x$ of the wavefunction parameters to the perturbation x. A substantially more elegant formulation is obtained, where the dependence on $\partial\lambda/\partial x$ is eliminated in the same way as for variational theories, by employing the technique of Lagrange multipliers.

The Lagrangian function depends on the wavefunction parameters λ, and their Lagrange multipliers $\bar{\lambda}$

$$L(x,\lambda,\bar{\lambda}) = E(x,\lambda) + \bar{\lambda}\cdot\Omega(x,\lambda) \tag{14}$$

where $\Omega(x,\lambda) = 0$ are the equations used to determine the non-variational parameters λ. The Lagrangian can be made stationary with respect to both λ and $\bar{\lambda}$ and at stationarity, the Lagrangian is equal to the energy.

$$\frac{\partial L(x,\lambda,\bar{\lambda})}{\partial\lambda} = 0 \tag{15}$$

$$\frac{\partial L(x,\lambda,\bar{\lambda})}{\partial\bar{\lambda}} = 0 \tag{16}$$

when $\lambda = \lambda^*$ and $\bar{\lambda} = \bar{\lambda}^*$. Thus, although we have doubled the number of wavefunction parameters, all of the parameters are variational in the Lagrangian formulation. The stationarity condition for λ is the equation that determines the Lagrange multipliers.

As before, the energy is obtained by inserting the optimal λ^* and $\bar{\lambda}^*$ into the Lagrangian function. The total derivative of the energy can be evaluated as the total derivative of the Lagrangian evaluated at λ^* and $\bar{\lambda}^*$.

$$\frac{dE(x)}{dx} = \frac{\partial L(x, \lambda, \bar{\lambda})}{\partial x} + \frac{\partial L(x, \lambda, \bar{\lambda})}{\partial \lambda} \frac{\partial \lambda}{\partial x} \tag{17}$$

$$= \frac{\partial L(x, \lambda, \bar{\lambda})}{\partial x} \tag{18}$$

Let us consider a simplified coupled-cluster Lagrangian, where we have ignored the fact that the HF orbital coefficients are not variational parameters for the coupled-cluster wavefunction. This is called the orbital-unrelaxed approach.

$$L(x, t, \bar{t}) = \langle 0|e^{-\hat{T}} \hat{H}(x) e^{\hat{T}}|0\rangle + \bar{t}_\mu \langle \mu|e^{-\hat{T}} \hat{H}(x) e^{\hat{T}}|0\rangle \tag{19}$$

We can define a bra state $\langle \Lambda|$ and a ket state $|\mathrm{CC}\rangle$

$$\langle \Lambda| = \langle 0|(1 + \sum_\mu \bar{t}_\mu \hat{\tau}^\dagger_\mu) e^{-\hat{T}} \tag{20}$$

$$|\mathrm{CC}\rangle = e^{\hat{T}}|0\rangle \tag{21}$$

and re-write the Lagrangian as

$$L(x, t, \bar{t}) = \langle \Lambda|\hat{H}(x)|\mathrm{CC}\rangle \tag{22}$$

Because all the amplitudes t_μ and their multipliers $\bar{t}_\mu$ are variational parameters, this is a variational coupled-cluster energy and obeys the Hellman–Feynman theorem so that

$$L(x, t, \bar{t}) = \langle \Lambda|\hat{H}(x)|\mathrm{CC}\rangle \tag{23}$$

$$\frac{dE}{dx} = \langle \Lambda|\frac{\partial \hat{H}}{\partial x}|\mathrm{CC}\rangle \tag{24}$$

We can express the variational coupled-cluster energy in terms of *variational* or *relaxed* density matrices.

$$E = \sum_{pq} \bar{D}^\Lambda_{pq} h_{pq} + \frac{1}{4} \sum_{pqrs} \bar{d}^\Lambda_{pqrs} g^{rs}_{pq} \tag{25}$$

$$\bar{D}^\Lambda_{pq} = \langle \Lambda|\hat{a}^\dagger_p \hat{a}_q|\mathrm{CC}\rangle \tag{26}$$

$$\bar{d}^\Lambda_{pqrs} = \langle \Lambda|\hat{a}^\dagger_p \hat{a}^\dagger_q a_s a_r|\mathrm{CC}\rangle \tag{27}$$

The same density matrices can be used with the Hellman–Feynman formula to compute first-order properties.

5 Excited States

The most straightforward way to compute excited state energies and therefore electronic transitions is to use configuration interaction theory and solve for as many roots as required

$$\mathbf{H}\mathbf{C}_k = E_k \mathbf{C}_k \tag{28}$$

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

The excited state wavefunctions are eigenfunctions of the CI Hamiltonian and are therefore orthogonal to each other as required. Truncated CI theory works well either for small systems or for transition energies where the size extensivity errors are similar in the ground and excited states and cancel. If the main correlation effects can be captured by a small set of active virtual orbitals, the CASSCF approach provides a practical way forwards since it is FCI in the active space and is therefore size extensive. Moreover, perturbative corrections can be applied to account for correlation outside the active space (CASPT2).

Hartree–Fock theory and coupled-cluster theory both involve writing down a trial wavefunction expression and optimising the parameters. Repeating this procedure using a different orbital occupancy to target excited states is possible provided that the excited states are orthogonal to the ground state for reasons of symmetry or spin. In other words, the targeted excited state is the ground state of that symmetry or spin quantum number. For general excited states this approach is not successful since by optimising the wavefunction parameters the wavefunction either optimises to the ground state or to a local minimum that is not guaranteed to be a good approximation to the state of interest.

One approach to obtain excited state energies at the level of coupled-cluster theory is to apply time dependent perturbation theory and search for poles in the frequency dependent polarisibility. This is called linear response coupled-cluster theory. There is an alternative approach where one obtains excited states as eigenvalues of the similarity transformed Hamiltonian using the cluster amplitudes of the ground state.

$$\mathbf{H}\mathbf{c}_k = E_k\mathbf{c}_k \tag{30}$$

$$H_{\mu\nu} = \langle\mu|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\nu\rangle \tag{31}$$

Since $\mathbf{H}$ is not Hermitian, we must solve for either left or right eigenvectors (in this case right eigenvectors). This is called equation of motion coupled-cluster theory. It turns out that the transition energies are equal to those of linear response.

6 Further Exercises

1. Expand the CCSD wavefunction $e^{\hat{T}_1+\hat{T}_2}|0\rangle$ in terms of configurations up to quadruple excitations to express the CI coefficients of the determinants $|{}^{a}_{i}\rangle$, $|{}^{ab}_{ij}\rangle$, $|{}^{abc}_{ijk}\rangle$ and $|{}^{abcd}_{ijkl}\rangle$ in terms of t^i_a and t^{ij}_{ab}.

2. Consider a system of n non-interacting atoms, where the Hamiltonian is $\sum_I^n H^I$. The Hartree–Fock reference determinant is $|0^1 0^2 \dots 0^n\rangle$ where $|0^I\rangle$ is the determinant for the I^{th} atom. Show that if the coupled cluster amplitudes for each atom $\mathbf{t}^I$ satisfy the coupled-cluster equations

   $$0 = \langle\mu^I|e^{-\hat{T}^I}\hat{H}^I e^{\hat{T}^I}|0^I\rangle$$
   $$E^I = \langle 0^I|e^{-\hat{T}^I}\hat{H}^I e^{\hat{T}^I}|0^I\rangle$$

   then the amplitude equation for the whole system is also satisfied

   $$\langle\mu|e^{-\hat{T}}\hat{H}e^{\hat{T}}|0\rangle = 0$$

   and hence show that the coupled cluster energy is size extensive.

3. Prove the Hellman–Feynman theorem for exact wavefunctions. Hint: you should use the fact that $\hat{H}\Psi = E\Psi$ and $d/dx\langle\Psi|\Psi\rangle = 0$ for all values of x.