

**Exam: Modern Quantum Chemistry (MQC 2016-2017)**

Date: 25 January 2017

Place: Gorlaeus Lab, lecture room C1

Time/duration: 3 hours (14.00-17.00)

Lecturer: Dr. F. Buda

Second lecturer reviewing the exam: Dr. Mark Somers

This written exam will contribute to 80% of the final grade. The remaining 20% is based on the written reports on the computer exercises.

These theoretical questions are organized in 7 sections and cover most of the material discussed during the course. For each question the associated points are indicated explicitly. The total number of points is 100.

Allowed information sources and tools:

Calculator

Please clearly indicate on each sheet: name and student number

Please write with blue or black ink and as clearly as possible.

**Good Luck!**

## Born-Oppenheimer (BO) approximation and potential energy surface (PES)

1. In the derivation of the Born-Oppenheimer approximation we have used the following ansatz for the total wavefunction:

$$\Psi(\mathbf{R}, \mathbf{r}, t) = \sum_i \psi_i(\mathbf{r}; \mathbf{R}) \chi_i(\mathbf{R}, t)$$

- a. Explain the meaning of each term in this ansatz.
- b. The wavefunctions  $\psi_i(\mathbf{r}; \mathbf{R})$  are required to be eigenfunctions of an operator. Which operator?
- c. What is the physical reason that justifies this factorization of the total wavefunction?

(6 points)

A1. The total wavefunction is factorized in a term that depend on the electronic positions and only parametrically on the nuclear positions (that are considered fixed when solving the electronic Schrödinger equation). The second term is the nuclear wavefunction that depends on the nuclear coordinates and on time. The  $\psi_i(\mathbf{r}; \mathbf{R})$  are eigenfunctions of the electronic Hamiltonian containing the electronic kinetic energy operator, the el-nuc and the el-el potential energy interactions. Usually also the nuc-nuc interaction is included in the electronic part of the Hamiltonian, although this is a constant term given that the nuclear coordinates are fixed. Therefore the expression for the total wavefunction can be seen as a linear expansion on the complete basis set provided by the eigenfunctions of the electronic Hamiltonian. The physical reason that justifies this factorization is the large difference in mass between the electrons and the nuclei, so that it is reasonable to assume that the electrons will adjust instantaneously to the nuclear motion and no exchange of energy between nuclear and electronic degrees of freedom (d.o.f.) occurs (adiabatic approximation).

2. Consider a diatomic molecule, e.g.  $\text{N}_2$ . Describe schematically how can you calculate the ground-state potential energy surface (PES) for this molecule within the Born-Oppenheimer approximation. (3 points)

A2. For a diatomic molecule the PES is described by only one nuclear parameter, i.e. the distance  $R$  between the two atoms. (in general there are  $3N-6$  nuclear d.o.f., but for linear molecules there are  $3N-5$ ). For a fixed value of  $R$  the electronic Hamiltonian is defined and the corresponding time-independent Schrödinger equation (S.E.) has to be solved. The lowest eigenvalue of the S.E. corresponds to the ground-state energy of the molecule for that particular geometry. Then one can consider a different distance and solve again the electronic problem (single point calculation) to get a new energy. This procedure has to be repeated for many values of  $R$  and eventually these data can be fitted (for example with a Morse potential) to generate the whole PES.

3. Imagine now to substitute the nitrogen-14 in this diatomic molecule with the nitrogen-15 isotope ( $^{15}\text{N}$ ). Do you expect that the ground-state PES would be affected by this isotopic substitution? Explain your answer. (3 points)

A3. No, the PES will not change since the mass of the nuclei appears only in the nuclear S.E. and not in the electronic one. Therefore the eigenvalues of the electronic S.E. will not be modified by a different nuclear mass. The isotopic substitution will only affect the nuclear dynamics.

4. In which circumstances the BO approximation breaks down and non-adiabatic couplings have to be taken into account? (3 points)

A4. The BO approximation would break down when the energy difference between different electronic state becomes small and comparable to the typical energy of the nuclear motion ( $\Delta\varepsilon = h\nu_{\text{nuc}}$ ). In this case a change in the nuclear coordinates can induce a change in the electronic state and non-adiabatic coupling terms between different electronic states become important. A noticeable example of such a situation is close to a conical intersection.

### Variational principle and secular equations

5. Consider an expansion of a trial electronic wavefunction  $\Phi$  in a finite set of linearly independent (basis) functions  $\varphi_i$  (Note that in general these basis functions are not orthonormal to each other) :

$$\Phi = \sum_i c_i \varphi_i$$

Derive the expression for the electronic energy based on this trial wavefunction and make use of the definition of overlap integrals and resonance integrals. (4 points)

A5. The electronic energy can be calculated by considering the expectation value of the electronic Hamiltonian with this trial wavefunction:

$$E = \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle} = \frac{\int d\mathbf{r} \left( \sum_i c_i^* \varphi_i^* \right) \hat{H} \left( \sum_j c_j \varphi_j \right)}{\int d\mathbf{r} \left( \sum_k c_k^* \varphi_k^* \right) \left( \sum_l c_l \varphi_l \right)} = \frac{\sum_{i,j} c_i^* c_j \int d\mathbf{r} \varphi_i^* \hat{H} \varphi_j}{\sum_{k,l} c_k^* c_l \int d\mathbf{r} \varphi_k^* \varphi_l} = \frac{\sum_{i,j} c_i^* c_j H_{ij}}{\sum_{k,l} c_k^* c_l S_{kl}}$$

where  $H_{ij}$  are the resonance integrals and  $S_{ij}$  are the overlap integrals.

6. What are the variational parameters in the expression for the energy derived in the previous question? (2 points)

A6. The coefficient  $c_i$  are the variational parameters that can be optimized to minimize the electronic energy.

7. How can we make use of the variational principle to obtain the optimal parameters? (2 points)

A7. According to the variational principle a trial wavefunction will always give an energy higher than the exact ground-state energy. Therefore by looking for a minimum with respect to the coefficient  $c_i$  we can get the best approximation for the ground-state energy and wavefunction. Mathematically this corresponds to impose the condition:

$$\frac{\partial E}{\partial c_m} = 0 \quad \forall m$$

which leads to the secular equations.

8. If we use a basis set comprising M functions, how many states will be derived by solving the secular equations? (2 points)

A8. We obtain as many solutions as the number of basis function  $M$ , since the size of the secular determinant is  $M \times M$ . The lowest energy will correspond to the ground-state energy.

### Hartree and Hartree-Fock approximation

9. Describe the concept of Molecular Orbitals approximation for a many-electron wavefunction  $\Psi(r_1, r_2, \dots, r_n)$   
(3 points)

A9. The MO approximation, also known as the independent particles approximation, corresponds to considering the  $n$  electrons as being independent on each other. Therefore the total wavefunction can be written as a product of single particle terms:

$$\Psi(r_1, r_2, \dots, r_n) = \psi_1(r_1) \cdot \psi_2(r_2) \cdot \dots \cdot \psi_n(r_n)$$

Each of these single particle terms is called a Molecular Orbital.

10. Write the Slater determinant for a molecular system with three electrons. How can you see that this wavefunction satisfies the Pauli principle? (3 points)

A10.

$$\tilde{\Phi}(r_1, r_2, r_3) = \begin{vmatrix} \psi_1(r_1) & \psi_2(r_1) & \psi_3(r_1) \\ \psi_1(r_2) & \psi_2(r_2) & \psi_3(r_2) \\ \psi_1(r_3) & \psi_2(r_3) & \psi_3(r_3) \end{vmatrix}$$

When exchanging two electrons (e.g. 1 and 2) we are exchanging two rows in the Slater determinant. Therefore the sign of the SD will change. This is indeed the condition on the many-electron wavefunction required by the Pauli principle.

11. How is the indistinguishability of quantum particles apparent in the Slater determinant wavefunction? (3 points)

A11. In a Slater determinant each electronic coordinate appears as argument of all the single particle functions. This indicates that the electrons are identical particles and we

cannot distinguish or label them. Therefore we cannot say e.g., that electron 2 is in the orbital  $\psi_2$ , since all other electrons appear as well in that orbital.

12. Write down the expression of a generic Coulomb integral  $J_{ij}$  that appears in the Hartree and Hartree-Fock energy. (2 points)

A12. A Coulomb integral is obtained by taking the integral of the coulomb potential operator (el-el interaction) with the electron density distribution of two molecular orbitals:

$$\underbrace{\left\langle \psi_i(\mathbf{r}_1)\psi_j(\mathbf{r}_2) \left| \frac{1}{r_{12}} \right| \psi_i(\mathbf{r}_1)\psi_j(\mathbf{r}_2) \right\rangle}_{J_{ij}}$$

This integral simply represents the energy of the Coulombic interaction between an electron in orbital  $i$  with an electron in orbital  $j$ .

13. What is the physical meaning of the Hartree-Fock molecular orbital energies  $\epsilon_k$  according to Koopmans theorem? (2 points)

A13. The Koopmans theorem states that within the HF approximation the orbital energy represents the ionization energy, i.e., the energy required to extract an electron from the corresponding molecular orbital.

14. Mention some of the most commonly used basis set functions in molecular quantum mechanical calculations. (Consider also your experience during the computer exercises) (2 points)

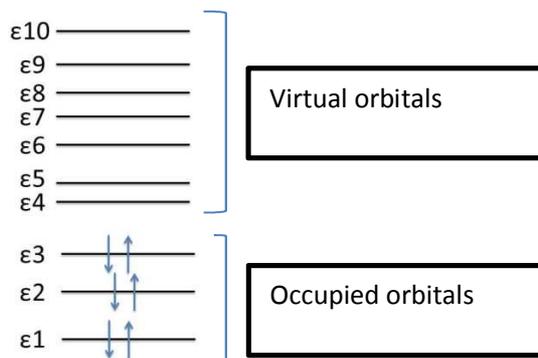
A14. In many QM software packages, the basis set functions are chosen to be atomic-like localized functions. In particular, Slater type orbitals or Gaussian functions are most often used. However, one can also use delocalized basis set, such as plane waves. This choice is more convenient for periodic systems.

### Correlation Energy and Configuration Interaction

15. Let's consider a closed-shell molecular system with 6 electrons and a basis set including  $M=10$  basis functions. Draw a schematic molecular orbital (MO) diagram including all occupied and virtual orbitals obtained solving the Roothaan-Hall

secular equations for the restricted (all electrons spin-paired) Hartree-Fock configuration  $\Phi_{HF}$ . (3 points)

A15.



16. Describe how to build up a Configuration Interaction wavefunction starting from this Hartree-Fock solution. (3 points)

A16. In the full CI method one would write the wavefunction as a linear combination of Slater determinants obtained by moving electrons from the occupied to the virtual orbitals. In this case one can consider to move from 1 up to a maximum of 6 electrons into virtual orbitals to generate all possible Slater determinants.

17. Consider the matrix element of the electronic Hamiltonian  $H_{mm} = \langle \Phi_m | \hat{H} | \Phi_m \rangle$ , where  $\Phi_m$  is a Slater determinant different from the Hartree-Fock determinant. Is  $H_{mm}$  higher or lower than the Hartree-Fock energy? Explain. (2 points)

A17. It will be higher than the Hartree-Fock energy, since the HF energy has been obtained by applying the variational principle using a single Slater determinant wavefunction. Therefore, there is no other Slater determinant which gives a lower energy.

18. Consider now a CISD wavefunction containing a linear combination of three Slater determinants:

$$\Psi_{CISD} = a_0 \Phi_{HF} + a_S \Phi_S + a_D \Phi_D$$

Can the expectation value of the electronic Hamiltonian using  $\Psi_{CISD}$  give a value lower than the Hartree-Fock energy and why? (2 points)

A18. Yes, since there will be matrix elements between different configurations (e.g. between the HF and a doubly exchanged configuration) that can lower the energy. In more physical terms, the interference between these different electronic configurations will contribute to move electrons on average more apart from each other (Coulomb hole) and thus reducing the el-el repulsion energy.

### Density Functional Theory (DFT)

19. What is the central physical quantity in density functional theory and how is it related to the many-electron wavefunction? (2 points)

A19. The central physical quantity in DFT is the electron density  $\rho(\mathbf{r})$ . The electron density can be calculated from the  $n$ -electron wavefunction  $\Psi$  according to the Born interpretation as:

$$\rho(\mathbf{r}_1) = n \int \cdots \int |\Psi(\mathbf{r}_1 s_1 \mathbf{x}_2 \dots \mathbf{x}_n)|^2 ds_1 d\mathbf{x}_2 \cdots d\mathbf{x}_n.$$

The integral runs over  $n-1$  electronic coordinates and  $n$  spins.

20. How is the variational principle formulated within DFT? (2 points)

A20. The second Hohenberg-Kohn theorem is a variational theorem directly derived from the standard variational principle for the wavefunctions. It states that any electron density  $\rho$  different from the exact ground-state density  $\rho_0$  will give an energy higher than the exact ground-state energy  $E_0$ :

$$E[\rho] \geq E[\rho_0] = E_0$$

21. Write down the electron density in terms of the single-particle Kohn-Sham orbitals. (2 points)

A21.

$$\rho(r) = \sum_{i=1}^n |\varphi_i(r)|^2$$

where  $n$  is the number of electrons in the system.

22. In the Kohn-Sham approach the kinetic term in the energy functional is separated into an independent particles contribution ( $T_s$ ) and a kinetic correlation part ( $T_c$ ). Write the expression for the kinetic energy functional  $T_s$  in terms of the set of Kohn-Sham orbitals. (2 points)

A22. The electronic kinetic energy in terms of the Kohn-Sham orbitals  $\varphi_i$  is written as follows:

$$T_s = \sum_i f_i \int dr \varphi_i^* \left( -\frac{\nabla^2}{2} \right) \varphi_i$$

This expression is in atomic units and  $f_i$  is the occupation number of the orbital  $\varphi_i$ .

23. The Kohn-Sham equations can be written as

$$\left( -\frac{1}{2} \nabla^2 + \sum_I \frac{Z_I}{|r - R_I|} + \int \frac{\rho(r')}{|r - r'|} dr' + V_{xc}(\rho(r)) \right) \varphi_i(r) = \varepsilon_i \varphi_i(r),$$

where the index  $i$  runs from 1 to the number of occupied orbitals.

- Discuss shortly the meaning of each term in these equations.
- Explain why we need an iterative self-consistent field (SCF) procedure to solve these equations.

(4 points)

A23. The first term is the kinetic energy operator, the second term is the electron-nucleus interaction potential term, the third is the Coulomb potential and the fourth is the exchange-correlation (XC) potential term. Both the Coulomb and XC potentials depend on the electron density, which in turn is written using the K-S orbitals. Therefore the effective potential in the K-S equation depends on the solutions  $\varphi_i(r)$ . This is why we need to start from an initial guess for the K-S orbitals, build the potential and solve iteratively these equations until a self-consistent solution is found.

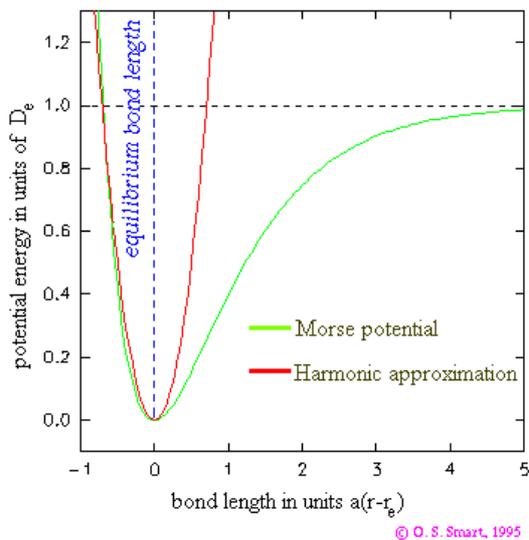
## Force-field and Molecular Dynamics simulations

24. The Morse potential can describe accurately the potential energy of a diatomic molecule. However, in broadly used force-field, such as AMBER, the bond stretching energy term is described by an harmonic potential.

- Discuss in which way these potentials differ from each other (use also a drawing to better illustrate this point)
- Under which circumstances the harmonic potential is still a valid approximation?

(4 points)

A24. The most striking difference between the Morse potential and the harmonic potential (see schematic picture below) is that the Morse potential goes to a finite value for large distances corresponding to the sum of the two atomic energies, while the harmonic potential grows quadratically with increasing distance.



The harmonic approximation is very similar to the Morse potential in the region near the minimum. In normal thermodynamic conditions (room temperature and standard atmospheric pressure) the molecule oscillates around the equilibrium distance with an amplitude that is dictated by the typical  $k_B T$  thermal energy. This energy at room temperature is much smaller than the typical bonding energy of a covalent bond. Therefore during the time evolution, the system will explore

distances close to the equilibrium distance for which the harmonic potential does not differ significantly from the Morse potential.

25. Write down the expression of the Lennard-Jones potential. Explain the physical origin of each term in the potential and the meaning of the parameters in this potential. (4 points)

A25. The Lennard-Jones potential can be written as follows:

$$E_{vdw} = 4\epsilon_{AB} \left[ \left( \frac{\sigma_{AB}}{R_{AB}} \right)^{12} - \left( \frac{\sigma_{AB}}{R_{AB}} \right)^6 \right]$$

The  $R^{-12}$  term, which is the repulsive term, describes Pauli repulsion at short ranges due to overlapping electron orbitals and the  $R^{-6}$  term, which is the attractive long-range term, describes attraction due to van der Waals force, or dispersion force. The parameter  $\epsilon$  is the depth of the potential well;  $\sigma$  is the finite distance at which the inter-particle potential is zero;  $R_{AB}$  is the (variable) distance between the particles.

26. Why the Coulomb interaction is said to be a long-range interaction? You can illustrate your answer by considering for example two atoms characterized in the force-field by effective charges of -0.6 and 0.3, respectively (in atomic units of charge). (3 points)

A26. The Coulomb potential energy between two point charges can be written as follows:

$$E_{electrostatic} = \frac{q^A q^B}{4\pi\epsilon_0 R_{AB}}$$

This potential decays slowly as  $R^{-1}$  with the distance between the two charges and therefore can be experienced also at large distances. For example, in the case of two effective charges -0.6 and 0.3, (in vacuum) the electrostatic energy at a distance of 20

Bohr ( $\sim 1.06$  nm) is equal to 0.009 a.u. ( $\sim 0.24$  eV), which is still about 10 times larger than the thermal energy ( $kT$ ) at room temperature.

27. Derive the Verlet algorithm for the numerical integration in MD simulations.  
(5 points)

A27. We can write two Taylor expansions for the positions about  $r(t)$ , one backward and one forward in time:

$$\begin{aligned} \mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \mathbf{v}(t)\Delta t + (1/2)\mathbf{a}(t)\Delta t^2 + (1/6)\mathbf{b}(t)\Delta t^3 + O(\Delta t^4) \\ \mathbf{r}(t - \Delta t) &= \mathbf{r}(t) - \mathbf{v}(t)\Delta t + (1/2)\mathbf{a}(t)\Delta t^2 - (1/6)\mathbf{b}(t)\Delta t^3 + O(\Delta t^4) \end{aligned}$$

By summing term by term the above equations we obtain:

$$r(t + \Delta t) + r(t - \Delta t) = 2r(t) + a(t)\Delta t^2 + O(\Delta t^4)$$

By substituting now in the second order term the acceleration with the force divided by the mass (according to Newton equation of motion) and by neglecting higher order terms in the expansion we obtain the Verlet algorithm:

$$r(t + \Delta t) = -r(t - \Delta t) + 2r(t) + \frac{F(t)}{M}\Delta t^2$$

28. How can you test the stability and accuracy of the integration algorithm in Molecular Dynamics simulations? (3 points)

A28. One way to check the accuracy and stability of the integration algorithm is to monitor along the simulation the total energy (potential plus kinetic energy) of the system. Since the total energy represents a constant of motion, this value should remain approximately constant and show no systematic drift in time.

29. Consider a molecular system in which the highest normal mode frequency is  $3500 \text{ cm}^{-1}$ . Determine the time step that would be appropriate to perform a

molecular dynamics simulation for this system. (Hint: make use of the relation  $\lambda\nu = c$ , where the speed of light is  $c \cong 3 * 10^{10} \text{ cm/s}$ ). (3 points)

A29. The time step in MD simulations should be taken small enough to describe correctly the fastest oscillations in the system. A rule of thumb is to take the time step  $\Delta t$  an order of magnitude smaller than the shortest period of oscillation. If we consider the mode with wavenumber  $3500 \text{ cm}^{-1}$ , the corresponding period is  $\sim 9.5 \text{ fs}$ . Therefore a reasonable choice for  $\Delta t$  is  $1 \text{ fs}$ .

### **Ab initio Molecular Dynamics (AIMD) and Hybrid QM/MM method**

30. Mention some advantages and disadvantages of AIMD in comparison with empirical force-field MD (3 points)

A30. Advantages: More general applicability and predictive power compared to MD using predefined potentials (force-field). No need to determine force-field parameters. Able to describe chemical bond breaking and formation. Disadvantages: computationally more demanding and therefore limited to relatively small systems (few hundred atoms) and short time scales (few picoseconds).

31. In the Car-Parrinello method for ab initio MD an extended Lagrangian is defined that includes both nuclear and electronic degrees of freedom (d.o.f.) as classical dynamical variables in the system. In particular the Lagrangian contains a classical kinetic energy term for the electronic d.o.f. Does this term correspond to a physically meaningful energy contribution? (2 points)

A31. The fictitious kinetic term associated to the electronic d.o.f. has no direct physical meaning in CPMD. This term has to remain very small compared to the other physical energy terms during the simulation in order to have a meaningful CPMD trajectory. It can be seen as a measure of how much the potential energy deviates from the actual Born-Oppenheimer surface.

32. How is the orthonormalization condition of the Kohn-Sham orbitals included in the Car-Parrinello Lagrangian? (3 points)

A32. The orthonormalization constraint is included using the method of Lagrange multipliers. This corresponds to include the following term in the CPMD Lagrangian:

$$\sum_{i,j} \lambda_{ij} (\langle \phi_i | \phi_j \rangle - \delta_{ij})$$

33. The total Hamiltonian of a system described with the hybrid QM/MM approach can be written as follows:

$$H = H_{QM} + H_{MM} + H_{QM/MM}$$

Explain briefly the meaning of each term and what kind of interactions are included in each part of this Hamiltonian. (4 points)

A33. The first term  $H_{QM}$  is the quantum-mechanical Hamiltonian describing all the energy terms and interactions between the particles included in the QM region. The second term  $H_{MM}$  includes all the interactions between the particles included in the MM region. These interactions are described by an empirical force-field. Finally, the third term  $H_{QM/MM}$  includes all the interactions between particles in the QM region with particles in the MM region. These interactions include primarily: (i) electrostatic interactions between the electronic charge in the QM part and the effective point charges in the MM part; (ii) van der Waals interactions close to the boundary between the two regions.

34. Describe the concept of “link atom” in the context of a QM/MM simulation. In which circumstances do we need to include it? Is the “link atom” interacting with the MM atoms? (4 points)

A34. If the boundary separating the QM and MM region does cut across a chemical bond, then we have to include a so-called link atom to saturate the dangling bond, otherwise the electron density in the QM region will be strongly perturbed by the missing bond. Usually we choose to put the boundary across a

single bond, therefore a H atom can be a reasonable choice as link atom. The link atom is seen by the QM part, but is not interacting with the MM atoms.