Potential Energy Surfaces in Hybrid QM/MM Methods

A.M. Tokmachev and A.L. Tchougréeff

Karpov Institute of Physical Chemistry

Vorontsovo pole 10, Moscow 103064 RUSSIA

October 5, 2000

Abstract

The problem of conjunction between quantum and classical parts in the hybrid QM/MM methods is considered. The form of the junction is deduced with use of the perturbation expansion in the assumption that the wave function underlying molecular mechanics is of the APSLG type. The renormalization of the potential energy surfaces of the combined system due to interaction between quantum and classical subsystems is represented as a sum of contributions with transparent physical sense.

Key words: hybrid QM/MM methods, perturbation expansion, effective Hamiltonian, potential energy surfaces, strictly localized geminals.

1 Introduction.

The hybrid quantum mechanical/molecular mechanical (QM/MM) methods became very popular in modern research on large molecular systems (particularly, biological ones) [?]. These approaches are based on the well justified experimental observation: chemical transformations are usually local, *i.e.* restricted to a small region of the entire molecular system called reaction center (RC). For that reason it seems to be practical to use approximations of different level in order to describe different parts of the system when its potential energy surface (PES) is considered. The smaller part of the molecular system with actually transforming electronic structure (RC) must be described by a thorough (preferrably, highly correlated) quantum chemical approach while the rest of the molecular system can be described either by molecular mechanics or by a faster quantum chemical method capable to reproduce only general features of the PES. The physical reason of that difference is the relative importance of electron correlations for correct description of the processes of the bond formation and bond cleavage - those where the uncorrelated SCF approach gives wrong asymptotic behavior of the electronic wave function of the system.

Two important questions immediately arise for such hybrid schemes. The first one is about the margin between the QM and MM subsystems. The important restriction must be set on the subsystems. Elaborated quantum chemical approaches (as well as the molecular mechanical schemes) work properly only being applied to the objects with an integer number of electrons. Therefore, distributing electrons between the subsystems is restricted by the condition of small fluctuations of numbers of electrons in the subsystems.

The second question arises from the fact that the classical part of the molecule modifies the PES of the reactive (quantum) subsystem. The interaction between subsystems can not be neglected. It leads to various forms of the junction between quantum and classical subsystems accepted in many hybrid schemes implemented by now in the literature [?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?]. At the same time it is noteworthy that the form of the junction in these approaches is usually taken *ad hoc*, without proper substantiation. For example, authors of Ref. [?] represent the junction between subsystems as a sum of electrostatic and van der Waals interactions with adjusted parameters; in Ref. [?] "junction dummy atoms" are introduced; in Ref. [?] the intersystem Coulomb and exchange integrals are taken as linear combinations of exponential functions with subsequent parameterization.

From out point of view the correct form of the junction between the quantum and classical parts of the molecular system must be constructed on the basis of the quantum chemical description of the whole system and of the sequential separation of electronic variables related to respective subsystems. This approach was originally proposed in Ref. [?]. It is based on the perturbation expansion and uses effective electron Hamiltonian technique. The margin between quantum and classical subsystems is chosen so that the chemical bonds are not broken and can be assigned only to one subsystem. The construction of hybrid schemes in Ref. [?] supposed that the parameters of the MM method must be renormalized due to interaction with quantum subsystem. Another construction of hybrid schemes was proposed in Ref. [?]: MM method remains constant and interactions modify only Hamiltonian of quantum subsystem. The effective Hamiltonian of quantum subsystem was explicitly constructed in Ref. [?]. The main idea of all these works is to single out the subset of electronic variables responsible for a chemical transformation to be described with use of a QM approach and to construct (using the Löwdin partition technique [?]) the effective Hamiltonian for the reactive subsystem which takes into account the interactions between subsystems perturbatively. To perform this program *i.e.* to construct the correction to the bare Hamiltonian for the RC it is necessary to assume some form of the electronic wave function of the inert environment underlying its MM description. This function has been chosen to be of the form of antisymmetrized product of strictly localized geminals (APSLG) [?, ?, ?, ?]. The APSLG based approach operates by local quantities such as chemical bonds and electron lone pairs, describes the ground states of the organic molecules with closed electron shells and high excitation energies, and provides the natural starting point for a transition from the quantum mechanics to the molecular mechanics (to be described elsewhere [?]). The effective Hamiltonian of the reactive subsystem in the presence of "classical" subsystem described by the APSLG method has been explicitly constructed in Ref. [?]. The present work is aimed to construct the PES of the molecular system described by the hybrid method with the APSLG wave function underlying the MM description of the classical subsystem.

2 PES of composite quantum-classical system.

Let us consider the application of the effective electron Hamiltonian method to the separation of the electronic variables which is aimed to constructing the PES for molecular systems in the hybrid QM/MM methods. We denote the reactive subsystem (described by quantum mechanics) by subscript R and the inert subsystem (described by MM) by subscript M. The Hamiltonian for the whole molecular system can be divided into parts corresponding to separate subsystems and to the interaction between them:

$$H = H^R(q) + H^M(q) + V^c(q) + V^r(q) + \frac{e^2}{2} \sum_{\substack{A \neq B \\ A \in R \\ B \in M}} Z_A^R Z_B^M R_{AB}^{-1}. \qquad (1)$$

The distinction between the R- and M-subsystems is formalized by ascribing each one-electron basis state of the entire system either to the R- or to the M-subsystem. The operators $H^R(q)$ and $H^M(q)$ are then polynomials with respect to the fermion operators creating or annihilating electrons in the one-electron states ascribed to the respective subsystems. On the other hand the operators containing the products of the fermion operators ascribed to the different subsystems are the interaction operators. We restrict ourselves by only the Coulomb V^c and the resonance V^r (electron transfer) interactions. The core charges are also subdivided between two subsystems. The contributions to the Hamiltonian can be further specified. The Hamiltonian for the R-subsystem is a sum of the Hamiltonian for the free R-subsystem (without environment) H_0^R and of the operator $\mathcal{V}^M$ of attraction of electrons of the R-subsystem to the cores of the M-subsystem. Analogously, the Hamiltonian for the M-subsystem H^M can be presented as a sum of the Hamiltonian for free M-subsystem H_0^M and of the attraction of electrons of the M-subsystem to the cores of the R-subsystem $\mathcal{V}^R$.

The wave function for the whole system is then approximated by the antisymmetrized product of the wave function of the reactants (which can be taken as a linear combination of configurations) Φ_k^R (where k is the number of the electronic state of the R-subsystem) and the ground state wave function of the inert environment to be calculated without reactants Φ_{00}^M:

$$\Psi_k^{appr} = \Phi_k^R \wedge \Phi_{00}^M. \tag{2}$$

The wave function Φ_{00}^M is the approximation to the lowest eigenstate of the Hamiltonian H_0^M. It is assumed to be obtained in the framework of the APSLG method [**?**, **?**, **?**, **?**]. Therefore, it has the form:

$$\left|\Phi_{00}^M\right\rangle = \prod_m g_m^+ \left|0\right\rangle, \tag{3}$$

where each geminal (two-electron wave function) $g_m^+ \left|0\right\rangle$ is assigned either to a chemical bond or to an electron lone pair.

To construct these geminals a unitary transformation must be performed from the initial atomic basis set to the hybrid one for each heavy (non hydrogen) atom:

$$t_{m\sigma} = \sum_{i\in\{s,x,y,z\}} h^A_{mi} a_{i\sigma}, \tag{4}$$

where $|t_{m\sigma}\rangle$ is a spin-orbital with the spin projection σ, assigned to the m-th bond and located on the right or the left end of the bond ($|r_{m\sigma}\rangle$ or $|l_{m\sigma}\rangle$); $h^A \in SO(4)$ is a matrix of transforming the AO basis orbitals s, p_x, p_y, and p_z of the atom A to the hybrid orbitals(HO's), and $a_{i\sigma}$ are operators annihilating an electron on a spin-orbital of the AO basis (the HO basis set is thus orthonormal).

Each geminal is a linear combination with varying amplitudes of three possible singlet states (two ionic and one covalent — Heitler-London type) constructed from the hybrid spin-orbitals $|r_{m\sigma}\rangle$ and $|l_{m\sigma}\rangle$:

$$g^+_m = u_m r^+_{m\alpha} r^+_{m\beta} + v_m l^+_{m\alpha} l^+_{m\beta} + w_m (r^+_{m\alpha} l^+_{m\beta} + l^+_{m\alpha} r^+_{m\beta}) \tag{5}$$

with the normalization condition $u^2_m + v^2_m + 2w^2_m = 1$. These functions are introduced in the theory by S. Weinbaum [**?**].

The approximate wave function Eq. (**??**) can be obtained from the exact wave function of the whole system

$$\Psi^{exact}_k = \sum_{n_M n_R} \sum_{i_M i_R} C^k_{i_M i_R}(n_M n_R) \Phi^M_{i_M 0}(n_M) \wedge \Phi^R_{i_R}(n_R), \ (n_M + n_R = N_e) \tag{6}$$

by two subsequent projections: first one projects the space where the exact wave function Eq. (**??**) resides onto the space of functions with fixed number of electrons in each of the subsystems (provided that $n_M + n_R = N_e$)

eliminating by this the fluctuations of numbers of electrons in the subsystems. The corresponding projection operator is denoted as P and $Q = 1 - P$ is the corresponding complementary projection operator. The second projection takes into account the fact that the chemically inert environment is described by its ground state wave function (it is also true when the wave function underlying the MM is used). Therefore, the second projection is onto the space of states with the ground state of the M-subsystem taken as a multiplier. The corresponding projection operator $\mathcal{P}$ is $1_R \otimes \left|\Phi_{00}^M\right\rangle \left\langle\Phi_{00}^M\right|$. Analogously we introduce the complementary projection operator $\mathcal{Q} = 1 - \mathcal{P}$.

The projections imply the transition from the exact Hamiltonian for the entire molecular system to the relevant effective Hamiltonians. The construction of the effective Hamiltonian is performed to fulfill the requirement that the eigenvalues of the effective Hamiltonian acting in the restricted subspace coincide with those of the original Hamiltonian acting in the full space. First projection yields the effective Hamiltonian of the form:

$$\begin{aligned} H_{eff}(q,E) = PH^R(q)P + PH^M(q)P + PV^c(q)P + \\ + PV^{rr}(q,E)P + \tfrac{e^2}{2} \sum_{\substack{A \neq B \\ A \in R, B \in M}} Z_A^R Z_B^M R_{AB}^{-1}, \end{aligned} \tag{7}$$

where

$$V^{rr}(q,E) = V^r(q) Q R(q,E) Q V^r(q) \tag{8}$$

with the resolvent operator

$$R(q,E) = (E - QHQ)^{-1}. \tag{9}$$

The second projection gives the operator acting in the space of functions of the type Eq. (**??**):

$$\mathcal{H}_{eff}(q,E,\omega) = \frac{e^2}{2} \sum_{\substack{A\neq B \\ A\in R, B\in M}} Z_A^R Z_B^M R_{AB}^{-1} + \mathcal{P}P[H_0^R(q) + H_0^M(q) + \\ + \mathcal{W}(q,E) + \mathcal{V}^M(q) + \mathcal{W}(q,E)P\mathcal{Q}\mathcal{R}(q,\omega)\mathcal{Q}P\mathcal{W}(q,E)]P\mathcal{P}, \tag{10}$$

where

$$\mathcal{W}(q,E) = V^c(q) + V^{rr}(q,E) + \mathcal{V}^R(q) \tag{11}$$

and the resolvent operator

$$\mathcal{R}(q,\omega) = (\omega - \mathcal{Q}H\mathcal{Q})^{-1}. \tag{12}$$

The energy of the k-th state of the combined system can be obtain by averaging the effective Hamiltonian over the approximate wave function Eq. (**??**):

$$E_k = \langle \Psi_k^{appr} | \, \mathcal{H}_{eff}(q,E,\omega) \, | \Psi_k^{appr} \rangle \,. \tag{13}$$

This energy obviously differs from that obtained by using *ab initio* Hamiltonian for R-subsystem and unrenormalized MM description for M-subsystem:

$$E_k = (E_{k0}^R + E_0^M) + \Delta E_k, \tag{14}$$

where the quantities E_{k0}^R and E_0^M are the averages of the free subsystem Hamiltonians over the functions Φ_k^R and Φ_{00}^M, respectively. The correction to the *ab initio* energy (junction between quantum and classical parts) arises from interaction between the subsystems:

$$\Delta E_k = \frac{e^2}{2} \sum_{\substack{A \neq B \\ A \in R, B \in M}} Z_A^R Z_B^M R_{AB}^{-1} +$$
$$+ \langle \Psi_k^{appr} | \, \mathcal{W}(q,E) + \mathcal{V}^M(q) + \mathcal{W}(q,E) P \mathcal{Q} \mathcal{R}(q,\omega) \mathcal{Q} P \mathcal{W}(q,E) \, | \Psi_k^{appr} \rangle \, . \tag{15}$$

To perform the averaging the explicit form of the operators of the Coulomb and resonance interactions is needed. The operator of the Coulomb interaction is taken to conserve the number of electrons in the subsystems and thus can be written as:

$$V^c(q) = \sum_{\substack{pp' \in R \\ mm' \in M}} \sum_{\sigma\tau} (p_\sigma p'_\sigma \parallel m_\tau m'_\tau) p_\sigma^+ m_\tau^+ m'_\tau p'_\sigma, \tag{16}$$

where

$$(p_\sigma p'_\sigma \parallel m_\tau m'_\tau) = (pp' \mid mm') - \delta_{\sigma\tau} (pm' \mid mp') \tag{17}$$

and the indici pp' and mm' refer to the spatial parts of one-electron states in the R- and M-subsystems, respectively. In the latter case the one-electron states can be taken as the HO's r_k and l_k in the M-subsystem. The ZDO approximation results in its turn in further simplification due to the restrictions: $p = p'$ and $m = m'$.

The operator of resonance interaction describes the one-electron transfers from the M-subsystem to the R-subsystem or *vice versa*:

$$V^r(q) = \sum_{\substack{p \in R \\ m \in M}} v_{pm}(q) \sum_\sigma (p_\sigma^+ m_\sigma + m_\sigma^+ p_\sigma). \tag{18}$$

It is convenient to introduce one-electron density matrices for each subsystem. The elements of this matrix for the R-subsystem are:

$$P_k^{\sigma pp'} = \left\langle \Phi_k^R \right| p_\sigma^+ p'_\sigma \left| \Phi_k^R \right\rangle . \tag{19}$$

The APSLG structure of the M-subsystem assumes that the one-electron densities differ from zero only for orbitals belonging to the same geminal. Also, the one-electron density matrix of the M-subsystem is spin-independent [**?**, **?**]. Therefore the elements of the one-electron density matrix for the M-subsystem are:

$$\begin{gathered} P_m^{tt'} = \left\langle \Phi_{00}^M \right| t_{m\sigma}^+ t'_{m\sigma} \left| \Phi_{00}^M \right\rangle = \langle 0 | \, g_m t_{m\sigma}^+ t'_{m\sigma} g_m^+ \, | 0 \rangle \, ; \\ P_m^{rr} = u_m^2 + w_m^2 ; P_m^{ll} = v_m^2 + w_m^2 ; P_m^{rl} = P_m^{lr} = (u_m + v_m) w_m . \end{gathered} \tag{20}$$

It is also convenient to introduce reduced Coulomb integrals:

$$Y_{pp'mm'} = \sum_\tau (p_\sigma p'_\sigma \parallel m_\tau m'_\tau) = 2(pp' \mid mm') - (pm' \mid mp'). \tag{21}$$

The Coulomb electron interaction between the subsystems contributes to the energy:

$$\langle \Psi_k^{appr} | \, V^c(q) \, | \Psi_k^{appr} \rangle = \sum_{pp' \in R} \sum_{t_m t'_m \in M} Y_{pp' i_m j_m} P_m^{tt'} \sum_\sigma P_k^{\sigma pp'} . \tag{22}$$

If the ZDO approximation is used and orbitals are attached to different atoms ($A \neq B$), this contribution to the energy can be rewritten in the form:

$$2 \sum_{\substack{A \in R \\ B \in M}} \gamma_{AB} \left(\sum_{t_m \in B} P_m^{tt} \right) \left(\sum_{p \in A} \sum_\sigma P_k^{\sigma pp} \right) . \tag{23}$$

In the case of frontier atoms the integral γ_{AB} must be replaced by $Y_{ppt_mt_m}$ which can not be factorized from the sums over t_m and p, respectively. The operator of attraction of electrons in the R-subsystem to the cores of the M-subsystem $\mathcal{V}^M$ can be written as

$$\begin{array}{c} \mathcal{V}^M(q) = -e^2 \sum\limits_{B'} \frac{Z^M_{B'}}{|r-R_{B'}|} = - \sum\limits_{pp'\in R} \sum\limits_{B\in M} V_B^{pp'} Z_B^M \sum\limits_{\sigma} p^+_\sigma p'_\sigma, \\ V_B^{pp'} = -e^2 \int d^3r \frac{\psi^*_p(r)\psi_{p'}(r)}{|r-R_{B'}|}. \end{array} \tag{24}$$

The averaging of this expression over the approximate wave function gives:

$$\langle \Psi_k^{appr} | \, \mathcal{V}^M(q) \, | \Psi_k^{appr} \rangle = - \sum_{pp'\in R} \sum_{B\in M} V_B^{pp'} Z_B^M \sum_{\sigma} P_k^{\sigma pp'}. \tag{25}$$

In the case of the ZDO approximation the $p = p'$ condition applies. Analogously, the contribution to the energy from the attraction of electrons in the M-subsystem to the cores of the R-subsystem is:

$$\langle \Psi_k^{appr} | \, \mathcal{V}^R(q) \, | \Psi_k^{appr} \rangle = -2 \sum_{t_m t'_m \in M} \sum_{A\in R} V_A^{t_m t'_m} Z_A^R P_m^{tt'}. \tag{26}$$

It can be noted that the contributions Eqs. (**??**), (**??**), and (**??**) can be combined with the repulsion of the cores of the R- and M-subsystems yielding the total electrostatic contribution to the renormalization of the PES. Each of these four contributions is large but the total electrostatic correction is relatively small:

$$\begin{array}{c} \Delta E_k^{el-st} = \frac{e^2}{2} \sum\limits_{\substack{A\neq B \\ A\in R, B\in M}} Z_A^R Z_B^M R_{AB}^{-1} + \\ + \langle \Psi_k^{appr} | \, V^c(q) + \mathcal{V}^M(q) + \mathcal{V}^R(q) \, | \Psi_k^{appr} \rangle \, . \end{array} \tag{27}$$

If the ZDO approximation is used the electrostatic contribution from different atoms of the R- and M-subsystems acquires the familiar form

$$\sum_{A\in R}\sum_{B\in M}\gamma_{AB}Q_A^{Rk}Q_B^M, \tag{28}$$

where the effective atomic charges are defined by:

$$\begin{aligned} Q_A^{Rk} &= \sum_{p\in R}\sum_{\sigma}P_k^{\sigma pp} - Z_A^R, \\ Q_B^M &= 2\sum_{t_m\in B}P_m^{tt} - Z_B^M. \end{aligned} \tag{29}$$

The contribution to the PES from the operator $V^{rr}(q,E)$ corresponds to the second order of the perturbation theory with respect to the intersystem one-electron transfers. To estimate it we present the corresponding resolvent operator in the approximate form:

$$R(E) = \sum_{i\in \mathrm{Im}Q}\frac{|i\rangle\langle i|}{E-E_i}, \tag{30}$$

where $|i\rangle$ are the states with one electron transferred from the M-subsystem to the R-subsystem or *vice versa.* Each state $|i\rangle$ is an antisymmetrized product of the ionized states $|\mu\rangle$ and $|\rho\rangle$ of the M- and R-subsystems respectively. The energy denominators in the resolvent expression Eq. (**??**) are expressed through the ionization potentials (IP) and the electron affinities (EA) of the respective subsystems:

$$E_i - E = \left\{ \begin{array}{l} I_\mu - A_\rho - g_{\mu\rho} \\ I_\rho - A_\mu - g_{\rho\mu} \end{array} \right. . \tag{31}$$

The approximate form of ionized states of the M-subsystem as it appears from the APSLG approximation is the following. The ionized states $|\mu\rangle$ are assumed to be obtained by removing or adding an electron from or to bonding and antibonding orbitals of the geminals of the M-subsystem. These orbitals can be written as [**?**]:

$$\left\{ \begin{array}{l} b_{m\sigma} = x_m l_{m\sigma} + y_m r_{m\sigma} \\ \\ a_{m\sigma} = -y_m l_{m\sigma} + x_m r_{m\sigma} \end{array} \right. ; \ (x_m^2 + y_m^2 = 1), \tag{32}$$

where x_m and y_m are chosen to represent the m-th geminal in the form:

$$g_m^+ = U_m b_{m\alpha}^+ b_{m\beta}^+ + V_m a_{m\alpha}^+ a_{m\beta}^+; \ (U_m^2 + V_m^2 = 1). \tag{33}$$

The coefficient sets $(U_m, V_m; x_m, y_m)$ and (u_m, w_m, v_m) are uniquely related:

$$\left\{ \begin{array}{l} u_m = U_m y_m^2 + V_m x_m^2 \\ v_m = U_m x_m^2 + V_m y_m^2 \\ w_m = (U_m - V_m)\, x_m y_m \end{array} \right. \tag{34}$$

and the APSLG-structure of the wave function can be expressed both in the basis of hybrid orbitals and in the basis of bond orbitals. The approximate wave functions of the ionized states of the M-subsystem have the form:

$$\begin{array}{c} b_{m\sigma}^+ \prod\limits_{l \neq m} g_l^+ \mid 0\rangle, \ a_{m\sigma}^+ \prod\limits_{l \neq m} g_l^+ \mid 0\rangle, \\ b_{m\sigma}^+ r_{m-\sigma}^+ l_{m-\sigma}^+ \prod\limits_{l \neq m} g_l^+ \mid 0\rangle, \ a_{m\sigma}^+ r_{m-\sigma}^+ l_{m-\sigma}^+ \prod\limits_{l \neq m} g_l^+ \mid 0\rangle. \end{array} \tag{35}$$

The corresponding IP's and the EA's entering the Eq. (**??**) are estimated in Ref. [**?**]. It is convenient to introduce some vacuum averages for the M-subsystem:

$$\begin{array}{c} \theta_{if}^m = \langle 0|\, g_m i_{m\alpha} f_{m\beta}^+ r_{m\alpha}^+ l_{m\alpha}^+ \,|0\rangle\,, \\ \eta_{if}^m = \langle 0|\, g_m i_{m\alpha}^+ f_{m\beta}^+ \,|0\rangle\,. \end{array} \tag{36}$$

These averages can be readily evaluated:

$$\begin{aligned}
\theta^m_{rb} &= y_m w_m + x_m v_m, \ \ \theta^m_{ra} = x_m w_m - y_m v_m; \\
\theta^m_{lb} &= -x_m w_m - y_m u_m, \ \ \theta^m_{la} = y_m w_m - x_m u_m; \\
\eta^m_{rb} &= y_m u_m + x_m w_m, \ \ \eta^m_{ra} = x_m u_m - y_m w_m; \\
\eta^m_{lb} &= y_m w_m + x_m v_m, \ \ \eta^m_{la} = x_m w_m - y_m v_m.
\end{aligned} \tag{37}$$

Using these notations we obtain the contribution to the energy from the operator $V^{rr}(q,E)$:

$$\begin{aligned}
\Delta E^{rr}_k = \langle \Psi^{appr}_k | \, V^{rr}(q,E) \, | \Psi^{appr}_k \rangle = - \sum_{pp' \in R} \sum_{m \in M} \sum_{ij \in \{r,l\}} \sum_{f \in \{a,b\}} v_{pi_m} v_{p'j_m} \times \\
\times \sum_\sigma \Bigl(\sum_{\rho \in ImO_R(N_R-1)} \frac{\left\langle \Phi^R_k \middle| p^+_\sigma \middle| \rho \right\rangle \left\langle \rho | p'_\sigma | \Phi^R_k \right\rangle \theta^m_{if} \theta^m_{jf}}{I_{k\rho} - A_{mf} - g_{k\rho f m}} + \\
+ \sum_{\rho \in ImO_R(N_R+1)} \frac{\left\langle \Phi^R_k | p_\sigma | \rho \right\rangle \left\langle \rho \middle| p'^+_\sigma \middle| \Phi^R_k \right\rangle \eta^m_{if} \eta^m_{jf}}{I_{mf} - A_{k\rho} - g_{f_m k\rho}} \Bigr).
\end{aligned} \tag{38}$$

This expression can be simplified if the IP's and EA's are replaced by their average values:

$$\begin{aligned}
\Delta E^{rr}_k \approx - \sum_{pp' \in R} \sum_{m \in M} \sum_{ij \in \{r,l\}} \sum_{f \in \{a,b\}} v_{pi_m} v_{p'j_m} \times \\
\times \sum_\sigma \left(\frac{P^{\sigma pp'}_k \theta^m_{if} \theta^m_{jf}}{I_k - A_{mf} - g_{kfm}} + \frac{\left(1 - P^{\sigma p' p}_k\right) \eta^m_{if} \eta^m_{jf}}{I_{mf} - A_k - g_{f_m k}} \right).
\end{aligned} \tag{39}$$

The contribution from the second projection can be presented as a sum:

$$
\begin{aligned}
&\langle\Psi_k^{appr} \mid \mathcal{W}(q,E)P\mathcal{Q}\mathcal{R}(q,\omega)\mathcal{Q}P\mathcal{W}(q,E) \mid \Psi_k^{appr}\rangle = \\
&= \langle\Psi_k^{appr} \mid V^c(q)P\mathcal{Q}\mathcal{R}(q,\omega)\mathcal{Q}PV^c(q) \mid \Psi_k^{appr}\rangle + \\
&+\langle\Psi_k^{appr} \mid \mathcal{V}^R(q)P\mathcal{Q}\mathcal{R}(q,\omega)\mathcal{Q}P\mathcal{V}^R(q) \mid \Psi_k^{appr}\rangle + \\
&+2\langle\Psi_k^{appr} \mid V^c(q)P\mathcal{Q}\mathcal{R}(q,\omega)\mathcal{Q}P\mathcal{V}^R(q) \mid \Psi_k^{appr}\rangle + \\
&+2\langle\Psi_k^{appr} \mid V^c(q)P\mathcal{Q}\mathcal{R}(q,\omega)\mathcal{Q}PV^{rr}(q,E) \mid \Psi_k^{appr}\rangle + \\
&+2\langle\Psi_k^{appr} \mid V^{rr}(q,E)P\mathcal{Q}\mathcal{R}(q,\omega)\mathcal{Q}P\mathcal{V}^R(q) \mid \Psi_k^{appr}\rangle + \\
&+\langle\Psi_k^{appr} \mid V^{rr}(q,E)P\mathcal{Q}\mathcal{R}(q,\omega)\mathcal{Q}PV^{rr}(q,E) \mid \Psi_k^{appr}\rangle .
\end{aligned}
\tag{40}
$$

The first three contributions of Eq. (**??**) are large enough but the perturbation expansion is valid because these contributions are grouped into the relatively small correction. At the same time, for convenience we consider these contributions separately. The resolvent can be approximated by the sum over the states:

$$
\mathcal{R}(q,\omega) = \sum_{\substack{\rho \\ \mu\neq 0}} \frac{|\rho,\mu\rangle\,\langle\rho,\mu|}{\omega-\omega_{\rho\mu}}. \tag{41}
$$

We assume that the dependence of the resolvent on ω is weak and the ω values of interest are much smaller than the resolvent poles which are all higher than the first excitation energy in the M- (classical) subsystem which in its turn is higher than the excitation energies of interest in the R-subsystem. Therefore we can pass to the limit $\omega \to 0$ in the above resolvent. The first contribution to the expression Eq. (**??**) is the second order correction due to the Coulomb repulsion operator. To write down this contribution it is convenient to introduce dynamic polarization propagators [**?**] for the R- and

M-subsystems:

$$\begin{array}{l}\Pi^{Rk}_{pp'qq'}(\omega)=\sum\limits_{\rho\neq k}\frac{(\epsilon_\rho-\epsilon_k)\left\langle\Phi^R_k\middle|p^+p'\middle|\rho\right\rangle\left\langle\rho\middle|q^+q'\middle|\Phi^R_k\right\rangle}{(\epsilon_\rho-\epsilon_k)^2-\omega^2},\\ \Pi^M_{mm'nn'}(\omega)=\sum\limits_{\mu\neq 0}\frac{\epsilon_\mu\left\langle\Phi^M_{00}\middle|m^+m'\middle|\mu\right\rangle\left\langle\mu\middle|n^+n'\middle|\Phi^M_{00}\right\rangle}{\epsilon^2_\mu-\omega^2},\end{array}\tag{42}$$

where ϵ_ρ and ϵ_μ are the energies of excitations in the R- and M-subsystems, respectively. The excitation energies and matrix elements necessary for calculating the polarization propagator of the M-subsystem are given in Ref. [**?**]. Using those notations we obtain:

$$\begin{array}{c}\langle\Psi^{appr}_k\mid V^c(q)P\mathcal{QR}(q,0)\mathcal{Q}PV^c(q)\mid\Psi^{appr}_k\rangle=\\ =\sum\limits_{pp'qq'\in R}\sum\limits_{i_m i'_n j_n j'_m\in M}\{\sum\limits_{\substack{\sigma\tau\\ \sigma'\tau'}}(p_\sigma p'_\sigma\parallel i_{m\sigma'}i'_{n\sigma'})(q_\tau q'_\tau\parallel j_{n\tau'}j'_{m\tau'})\times\\ \times\left[P^{\sigma pp'}_k P^{\tau qq'}_k\Pi^M_{i_{m\sigma'}i'_{n\sigma'}j_{n\tau'}j'_{m\tau'}}(0)+\frac{2}{\pi}\int\Pi^{Rk}_{p_\sigma p'_\sigma q_\tau q'_\tau}(iu)\Pi^M_{i_{m\sigma'}i'_{n\sigma'}j_{n\tau'}j'_{m\tau'}}(iu)du\right].\end{array}\tag{43}$$

The expression using integrals is not useful in practice. At the same time we can approximate it taking into account the essentially higher energies of excitations in the M-subsystem in comparison with those of the R-subsystem:

$$\begin{array}{c}\frac{2}{\pi}\int\Pi^{Rk}_{p_\sigma p'_\sigma q_\tau q'_\tau}(iu)\Pi^M_{i_{m\sigma'}i'_{n\sigma'}j_{n\tau'}j'_{m\tau'}}(iu)du\approx\\ \approx\Pi^M_{i_{m\sigma'}i'_{n\sigma'}j_{n\tau'}j'_{m\tau'}}(0)\sum\limits_{k'\neq k}\left\langle\Phi^R_k\left|p^+_\sigma p'_\sigma\right|\Phi^R_{k'}\right\rangle\left\langle\Phi^R_{k'}\left|q^+_\tau q'_\tau\right|\Phi^R_k\right\rangle,\end{array}\tag{44}$$

i.e. it is a sum of products of the static polarization propagator of M-subsystem and transition electron densities of R-subsystem. The contribution Eq. (**??**) mainly corresponds to the dispersion interaction between the subsystems ΔE^{disp}_k.

The contribution to the energy of the second order in operators $\mathcal{V}^R(q)$ does not change the relative energies of the states in the R-subsystem and, therefore, is k-independent, where k is the number of electronic state in the R-subsystem. The operator $\mathcal{V}^R(q)$ can be written in the form:

$$\begin{aligned}\mathcal{V}^R &= -\sum_{mm'\in M}\sum_{\sigma} m^+_\sigma m'_\sigma \sum_{A\in R} V_A^{mm'} Z_A^R \\ &\approx -\sum_{B\in M}\sum_{m\in B\cap M}\sum_{\sigma} m^+_\sigma m_\sigma \sum_{\substack{A\in R\\ A\neq B}} \gamma_{AB} Z_A^R.\end{aligned} \tag{45}$$

The contribution to the energy is:

$$\begin{aligned}\Delta E_k^{RR} &= \langle \Psi_k^{appr} \mid \mathcal{V}^R(q) P \mathcal{Q}\mathcal{R}(q,0)\mathcal{Q}P\mathcal{V}^R(q) \mid \Psi_k^{appr}\rangle = \\ &= \sum_{i_m i'_n j_n j'_m \in M}\sum_{AA'\in R} V_A^{i_m i'_n} V_{A'}^{j_n j'_m} Z_A^R Z_{A'}^R \sum_{\sigma\tau} \Pi^M_{i_{m\sigma} i'_{n\sigma} j_{n\tau} j'_{m\tau}}(0).\end{aligned} \tag{46}$$

Let us denote the atom at the right end of the m-th bond as R_m and that at the left end as L_m. In the case if $AA' \notin \{R_m, L_m\}$ the contribution Eq. (**??**) can be approximately rewritten in the physically transparent form in terms of the bond polarizability tensors $\widehat{\alpha}^{(m)}(\omega)$ [**?**]:

$$-\sum_{m\in M}\sum_{\substack{AA'\in R\\ AA'\notin\{R_m,L_m\}}} Z_A^R Z_{A'}^R (\nabla V(R_{A'}^{(m)}) \left|\hat{\alpha}^{(m)}(0)\right| \nabla V(R_A^{(m)})), \tag{47}$$

where $V(\mathbf{R}_A^{(m)})$ is the potential induced by a unit charge placed on the atom A at the center of the m-th bond. It is worth to mention that the bond polarizabilities $\hat{\alpha}^{(k)}(0)$ are tabulated, for example, in Ref. [**?**].

The cross term between the $V^c(q)$ and $\mathcal{V}^R(q)$ equals to:

$$\begin{aligned}\Delta E_k^{cR} &= \langle \Psi_k^{appr} \mid V^c(q) P\mathcal{Q}\mathcal{R}(q,0)\mathcal{Q}P\mathcal{V}^R(Q) \mid \Psi_k^{appr}\rangle = \\ &= -\sum_{\sigma\sigma'\tau}\sum_{pp'\in R} P_k^{\sigma pp'} \sum_{A\in R} Z_A^R \sum_{i_m i'_n j_n j'_m\in M} (p_\sigma p'_\sigma \parallel i_{m\sigma'} i'_{n\sigma'}) V_A^{j_n j'_m} \Pi^M_{i_{m\sigma'} i'_{n\sigma'} j_{n\tau} j'_{m\tau}}(0).\end{aligned} \tag{48}$$

The contributions to the energy ΔE_k^{cR}, ΔE_k^{RR} and the first contribution of Eq. (**??**) can be grouped in the total correction ΔE_k^{coul}, which corresponds to the second order interaction between charges in the R-subsystem due to polarization in the M-subsystem. If the ZDO approximation is used this contribution can be written as:

$$\Delta E_k^{coul} = \sum_{m \in M} \sum_{ii' \in \{r,l\}} \sum_{AA' \in R} \gamma_{AI_m} \gamma_{A'I'_m} Q_A^R Q_{A'}^R \sum_{\sigma\tau} \Pi^M_{i_{m\sigma} i'_{m\sigma} i'_{m\tau} i_{m\tau}}(0), \qquad (49)$$

where I_m denotes the atom of the m-th bond corresponding to the i-th HO of the bond (R_m or L_m).

The next two contributions of Eq. (**??**), including the operator $V^{rr}(q,E)$, correspond to the third order in the perturbation expansion. Therefore, they may be only roughly estimated. The main assumptions are the ZDO approximation and that the energies of an electron transfers between the subsystems are replaced by average ones depending only on the number of bond taking part in the electron transfer. The cross term between the $V^c(q)$ and $V^{rr}(q,E)$ is:

$$\begin{gathered}
\Delta E_k^{crr} = \langle \Psi_k^{appr} \mid V^c(q) P \mathcal{Q} \mathcal{R}(q,0) \mathcal{Q} P V^{rr}(q,E) \mid \Psi_k^{appr} \rangle \approx \\
\approx \sum_{pqq' \in R} \sum_{i_m j_m j'_m \in M} v_{qj_m} v_{q'j'_m} \sum_{\sigma\sigma'\tau} (p_\sigma p_\sigma \parallel i_{m\sigma'} i_{m\sigma'}) \times \\
\times \{ (I_m - A_k - g_{mk})^{-1} [P_k^{\sigma pp} (1 - P_k^{\tau q' q}) \Pi^M_{i_{m\sigma'} i_{m\sigma'} j_{m\tau} j'_{m\tau}}(0) - \\
- \tfrac{2}{\pi} \int \Pi^{Rk}_{p_\sigma p_\sigma q'_\tau q_\tau}(iu) \Pi^M_{i_{m\sigma'} i_{m\sigma'} j_{m\tau} j'_{m\tau}}(iu) du] - \\
- (I_k - A_m - g_{km})^{-1} [P_k^{\sigma pp} P_k^{\tau q q'} \Pi^M_{i_{m\sigma'} i_{m\sigma'} \widetilde{j}_{m\tau} \widetilde{j}'_{m\tau}}(0) + \\
+ \tfrac{2}{\pi} \int \Pi^{Rk}_{p_\sigma p_\sigma q_\tau q'_\tau}(iu) \Pi^M_{i_{m\sigma'} i_{m\sigma'} \widetilde{j}_{m\tau} \widetilde{j}'_{m\tau}}(iu) du] \},
\end{gathered} \qquad (50)$$

where $\widetilde{j} = r$ for $j = l$ and $\widetilde{j} = l$ for $j = r$. The integral expressions can be replaced by sums using Eq. (??).

Analogously, the cross term between $\mathcal{V}^R(q)$ and $V^{rr}(q,E)$ equals to:

$$\begin{aligned}
\Delta E_k^{Rrr} &= \langle \Psi_k^{appr} \mid \mathcal{V}^R(q) P \mathcal{Q} \mathcal{R}(q,0) \mathcal{Q} P V^{rr}(q,E) \mid \Psi_k^{appr} \rangle \approx \\
&\approx -\sum_{\sigma\tau} \sum_{pp' \in R} \sum_{A \in R} Z_A^R \sum_{j_m j'_m \in M} v_{pj_m} v_{p'j'_m} \times \\
&\times \{ (I_m - A_k - g_{mk})^{-1} (1 - P_k^{\tau p' p}) \times \\
&\times \left(\gamma_{AR_m} \Pi^M_{r_{m\sigma} r_{m\sigma} j_{m\tau} j'_{m\tau}}(0) + \gamma_{AL_m} \Pi^M_{l_{m\sigma} l_{m\sigma} j_{m\tau} j'_{m\tau}}(0) \right) - \\
&- (-1)^{\delta_{jj'}} (I_k - A_m - g_{km})^{-1} P_k^{\tau p p'} \times \\
&\times \left(\gamma_{AR_m} \Pi^M_{r_{m\sigma} r_{m\sigma} \widetilde{j}_{m\tau} \widetilde{j}'_{m\tau}}(0) + \gamma_{AL_m} \Pi^M_{l_{m\sigma} l_{m\sigma} \widetilde{j}_{m\tau} \widetilde{j}'_{m\tau}}(0) \right) \}.
\end{aligned} \tag{51}$$

The last contribution to the energy of Eq. (??) is of the fourth order with respect to the perturbation operator (counting four small operators of electron transfer) and may be totally neglected.

3 Discussion

Large molecular systems are very important from the practical point of view. Polymers and many biological molecules contain hundreds and thousands atoms. Modern *ab initio* methods do not allow to calculate the electronic structure of such systems because the computational time for these methods grows as $N^4 \div N^7$, where N is number of one-electron basis functions. Even in the case of the standard semiempirical methods based on the Hartree-Fock approximation this growth is proportional to N^3. It explains the especial role of the hybrid QM/MM schemes in the computational chemistry. There is a large diversity of the hybrid QM/MM schemes proposed in the literature due to the form of junction between quantum and classical subsystems. It leads to divergence of results obtained with different junctions. The results of the calculations seems to be ambiguous due to uncontrolled influence of the junction on the calculated PES of the molecular system.

We try to justify the form of the junction between subsystems. The complete neglect of the junction, *i.e.* the use of the bare ("*ab initio*") Hamiltonian for the R-subsystem and standard molecular mechanical procedure without renormalization leads to the energy $E^R_{k0} + E^M_0$ instead of the correct value E_k. Taking into account the interaction between the system (junction) results in a correction to the PES, which can be represented as a sum of contributions:

$$\Delta E_k \approx \Delta E_k^{el-st} + \Delta E_k^{rr} + \Delta E_k^{disp} + \Delta E_k^{coul} + \Delta E_k^{crr} + \Delta E_k^{Rrr}, \qquad (52)$$

where the main contribution ΔE_k^{el-st} is of the first order with respect to the perturbation operator and originates from the electrostatic interaction between the subsystems; the next four contributions are of the second order and follow from the coupling of one-electron transfers between subsystems, of Coulomb electron-electron interactions between them, of interactions between electrons of the M-subsystem and cores of the R-subsystem; and of interaction of the Coulomb electron-electron interaction with the interaction of the electrons of the M-subsystem and cores of the R-subsystem; the second term can be regarded as dispersion interaction between subsystems. Moreover, the contributions ΔE_k^{disp} and ΔE_k^{coul} can be considered as the second order interaction between the electronic polarization in the M-

subsystem and the polarized R-subsystem. The last two contributions to Eq. (**??**) correspond to the third order in the interaction between the subsystems originating from the coupling two projection procedures. Physically it corresponds to interaction between two one-electron transfers and Coulomb interaction between electrons of the M-subsystem with electrons and cores of the R-subsystem.

One of us (A.M.T.) acknowledges financial support from the Haldor Topsøe A/S.

References

[1] CECAM-NSF Meeting on QC/MM methods, Int J Quantum Chem 1996, 60, No 6, Special Issue.

[2] Aqvist J.; Warshel A. Chem Rev 1993, 93, 2523.

[3] Bakowics D.; Thiel W.; J Comput Chem 1996, 17, 87.

[4] Carmer C.S.; Weiner B.; Frenklach M. J Chem Phys 1993, 99, 1356.

[5] Field M.J.; Bach P.A.; Karplus M. J Comput Chem 1990, 11, 300.

[6] Field M.J. In Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications, van Gunsteren W.F.; Weiner P.K.; Wilkinson A.J., eds. ESCOM: Leiden, 1993.

[7] Gao J.L. In Reviews in Computational Chemistry Lipkowitz K.B.; Boyd D.B., eds. VCH Publishers: New York, 1996.

[8] Gao J.L. J Phys Chem 1992, 96, 537.

[9] Gao J.L.; Xia X.F. Science 1992, 258, 631.

[10] Luzhkov V.; Warshel A. J Comput Chem 1992, 13, 199.

[11] Maseras F.; Morokuma K. J Comput Chem 1995, 16, 1170.

[12] Stanton R.V.; Hartsough D.S.; Merz K.M., Jr. J Comput Chem 1995, 16, 113.

[13] Thompson M.A. J Phys Chem 1995, 99, 4794.

[14] Tuñon I.; Martins-Costa T.C.; Millot C.; Ruiz-Lopez M.F.; Rivail J.-L. J Comput Chem 1996, 17, 19.

[15] Truong T.N.; Stefanovich E.V. Chem Phys Lett 1996, 256, 348.

[16] Morokuma K.; Froese R.D.J.; Humbel S.; Svensson M.; Matsubara T.; Sieber S. J Phys Chem 1996, 100, 2573.

[17] Bernardi F.; Olivucci M.; Robb M.A. J Amer Chem Soc 1992, 114, 1606.

[18] Tchougréeff A.L. Khim Fiz 1997, 16, 62 (in Russian); Chem Phys Reps 1997, 16, 1035 (in English).

[19] Tchougréeff A.L. Phys Chem Chem Phys 1999, 1, 1051.

[20] Tokmachev A.M.; Tchougréeff A.L.; Misurkin I.A. J Mol Struct (Theochem) 2000, 506, 17.

[21] Löwdin P.-O. J Math Phys 1962, 3, 969.

[22] Surján P.R. The Concept of the Chemical Bond, In Theoretical Models of Chemical Bonding; Part 2 Z.B. Maksić, ed. Springer: Heidelberg, 1989.

[23] Wu W.; McWeeny R. J Chem Phys 1994, 101, 4826.

[24] Tokmachev A.M., Tchougréeff A.L. Zhurn Fiz Khim 1999, 73, 347 (in Russian); Russ J Phys Chem 1999, 73, 320 (in English).

[25] Tokmachev A.M.; Tchougréeff A.L. J Comput Chem (Accepted).

[26] Tokmachev A.M.; Tchougréeff A.L. (In preparation).

[27] Weinbaum S. J Chem Phys 1933, 3, 547.

[28] Tokmachev A.M.; Tchougréeff A.L.; Misurkin I.A. Int J Quantum Chem (Submitted).

[29] McWeeny R. Methods of Molecular Quantum Mechanics; 2nd Edition; Academic Press: London, 1992.

[30] Le Fevre R.J.W. Molecular Refractivity and Polarizability, In Advances in Physical Organic Chemistry, Vol. 3, Gold V. ed. Academic Press: London and New York, 1965.