Potential Energy Surfaces in Hybrid Quantum Mechanical/Molecular Mechanical Methods

A. M. TOKMACHEV, A. L. TCHOUGRÉEFF

Karpov Institute of Physical Chemistry, Vorontsovo pole 10, Moscow 103064, Russia

ABSTRACT: The problem of conjunction between quantum and classical parts in hybrid quantum mechanical/molecular mechanical methods is considered. The form of the junction is deduced with use of the perturbation expansion on the assumption that the wave function underlying molecular mechanics is of the antisymmetrized product of strictly localized geminals 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. © 2001 John Wiley & Sons, Inc. Int J Quantum Chem 84: 39–47, 2001

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

Introduction

Hybrid quantum mechanical (QM)/molecular mechanical (MM) methods are very popular in modern research on large molecular systems (particularly biological systems) [1]. 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 the reaction center (RC). For that reason it seems practical to use approximations of different levels to describe different parts of the system when its potential energy surface (PES) is considered. The smaller part of the molecular system with an actually transforming electronic structure (RC) must be described by a thorough (preferably, 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 of reproducing only general features of the PES. The physical reason for 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 self-consistent field (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

Correspondence to: A. L. Tchougréeff; e-mail: andrei@cc.nifhi.ac.ru.

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 when applied to 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 cannot be neglected. It leads to various forms of the junction between quantum and classical subsystems accepted in many hybrid schemes implemented in the literature [2–17]. 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, the authors of Ref. [15] represent the junction between subsystems as a sum of electrostatic and van der Waals interactions with adjusted parameters, in Ref. [16] "junction dummy atoms" are introduced, and in Ref. [17] the intersystem Coulomb and exchange integrals are taken as linear combinations of exponential functions with subsequent parameterization.

From our 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 on the sequential separation of electronic variables related to respective subsystems. This approach to the junction was originally proposed in Ref. [18]. It is based on the perturbation expansion over specially constructed wave functions of the group function method and uses an 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. [18] supposed that the parameters of the MM method must be renormalized due to interaction with the quantum subsystem. Another construction of hybrid schemes was proposed in Ref. [19]: the MM method remains constant and interactions modify only the Hamiltonian of the quantum subsystem. The effective Hamiltonian of the quantum subsystem was explicitly constructed in Ref. [20]. The main idea of all these works was 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 [21]) the effective Hamiltonian for the reactive subsystem which takes into account the interactions between subsystems perturbatively. To accomplish this, 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 in turn can be chosen as a group function product [22] with the groups describing local two electron bonds, lone pairs, conjugated π systems, etc. Particular formulae are written on an assumption that the wave function for the inert subsystem is of the form of an antisymmetrized product of strictly localized geminals (APSLG) [23–26]. 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 seems to be a natural starting point for a transition from quantum mechanics to molecular mechanics. The effective Hamiltonian of the reactive subsystem in the presence of the "classical" subsystem described by the APSLG method has been explicitly constructed in Ref. [20]. The aim of the present work is 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.

PES of Composite Quantum-Classical System

Let us consider the application of the effective electron Hamiltonian method to the separation of the electronic variables, the aim of which is to construct 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:

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

The distinction between the R and M subsystems is formalized by ascribing each of the one-electron basis states of the entire system either to the R or 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 to 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:

$$\Psi_k^{\text{appr}} = \Phi_k^R \wedge \Phi_{00}^M. \quad (2)$$

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 [23–26]. Therefore, it has the form

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

where each geminal (two-electron wave function) $g_m^+ |0\rangle$ 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 (nonhydrogen) atom:

$$t_{m\sigma} = \sum_{i\in\{s,x,y,z\}} h_{mi}^A a_{i\sigma}, \quad (4)$$

where $|t_{m\sigma}\rangle$ is a spin orbital with the spin projection σ, assigned to the mth bond and located on the right or 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 atomic orbital (AO) basis orbitals s, p_x, p_y, and p_z of the atom A to the hybrid orbitals (HOs), and $a_{i\sigma}$ are op-

erators 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 $|l_{m\sigma}\rangle$

$$g_m^+ = u_m r_{m\alpha}^+ r_{m\beta}^+ + v_m l_{m\alpha}^+ l_{m\beta}^+ + w_m \left(r_{m\alpha}^+ l_{m\beta}^+ + l_{m\alpha}^+ r_{m\beta}^+\right) \quad (5)$$

with the normalization condition $u_m^2 + v_m^2 + 2w_m^2 = 1$. These functions are introduced in the theory of Weinbaum [27].

The approximate wave function [Eq. (2)] can be obtained from the exact wave function of the whole system in the "simplified group function" representation [28–30],

$$\Psi_k^{\text{exact}} = \sum_{n_M n_R} \sum_{i_M i_R} C_{i_M i_R}^k (n_M n_R) \Phi_{i_M 0}^M (n_M) \wedge \Phi_{i_R}^R (n_R),$$

$$(n_M + n_R = N_e), \quad (6)$$

by two subsequent projections. The first one projects the space where the exact wave function [Eq. (6)] resides on the space of functions with a fixed number of electrons in each of the subsystems (provided that $n_M + n_R = N_e$), thus eliminating 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. This projection leads to the wave function of the group function method. 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 $1_R \otimes |\Phi_{00}^M\rangle\langle\Phi_{00}^M|$. 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 effective Hamiltonian is constructed to fulfill the requirement that its eigenvalues acting in the restricted subspace coincide with those of the original Hamiltonian acting in the full space. The first projection yields the effective Hamiltonian of the form

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

where

$$V^{rr}(q, E) = V^r(q)QR(q, E)QV^r(q) \quad (8)$$

with the resolvent operator

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

The second projection gives the operator acting in the space of functions of the type in Eq. (2):

$$\begin{aligned}
\mathcal{H}_{\text{eff}}(q, E, \omega) = e^2 \sum_{\substack{A \neq B \\ A \in R, B \in M}} Z_A^R Z_B^M R_{AB}^{-1} \\
+ \mathcal{PP}\big[H_0^R(q) + H_0^M(q) + \mathcal{W}(q, E) \\
+ \mathcal{V}^M(q) + \mathcal{W}(q, E)P\mathcal{QR}(q, \omega) \\
\times \mathcal{QPW}(q, E)\big]P\mathcal{P},
\end{aligned}$$

(10)

where

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

and the resolvent operator

$$\mathcal{R}(q, \omega) = (\omega - QHQ)^{-1}. \quad (12)$$

The energy of the kth state of the combined system can be obtained by averaging the effective Hamiltonian over the approximate wave function [Eq. (2)]:

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

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

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

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:

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

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

$$V^c(q) = \sum_{\substack{pp' \in R \\ mm' \in M}} \sum_{\sigma\tau} (p_\sigma p_\sigma' \| m_\tau m_\tau') p_\sigma^+ m_\tau^+ m_\tau' p_\sigma', \quad (16)$$

where

$$(p_\sigma p_\sigma' \| m_\tau m_\tau') = (pp' | mm') - \delta_{\sigma\tau}(pm' | mp') \quad (17)$$

and the indices 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 HOs r_k and l_k in the M subsystem. The zero differential overlap (ZDO) approximation results in 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:

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

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

$$P_k^{\sigma pp'} = \langle \Phi_k^R | p_\sigma^+ p_\sigma' | \Phi_k^R \rangle. \quad (19)$$

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 [25, 26]. Therefore the elements of the one-electron density matrix for the M subsystem are

$$\begin{aligned}
P_m^{tt'} = \langle \Phi_{00}^M | t_{m\sigma}^+ t_{m\sigma}' | \Phi_{00}^M \rangle = \langle 0 | g_m t_{m\sigma}^+ t_{m\sigma}' g_m^+ | 0 \rangle, \\
P_m^{rr} = u_m^2 + w_m^2; \qquad P_m^{ll} = v_m^2 + w_m^2; \\
P_m^{rl} = P_m^{lr} = (u_m + v_m)w_m,
\end{aligned}$$

(20)

where amplitudes u_m, v_m, and w_m were introduced by Eq. (5). It is also convenient to introduce reduced Coulomb integrals:

$$Y_{pp'mm'} = \sum_\tau (p_\sigma p_\sigma' \| m_\tau m_\tau') = 2(pp' | mm') - (pm' | mp'). \quad (21)$$

The Coulomb electron interaction between the subsystems contributes to the energy

$$\langle \Psi_k^{\text{appr}} | V^c(q) | \Psi_k^{\text{appr}} \rangle = \sum_{pp' \in R} \sum_{t_m t_m' \in M} Y_{pp'i_m j_m} P_m^{tt'} \sum_\sigma P_k^{\sigma pp'}. \quad (22)$$

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

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

For frontier atoms the integral γ_{AB} must be replaced by $Y_{ppt_mt_m}$, which cannot 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

$$\mathcal{V}^M(q) = -e^2 \sum_B \frac{Z_{B'}^M}{|r - R_{B'}|}$$

$$= -\sum_{pp' \in R} \sum_{B \in M} V_B^{pp'} Z_B^M \sum_\sigma p_\sigma^+ p_\sigma', \quad (24)$$

$$V_B^{pp'} = -e^2 \int d^3r \, \frac{\psi_p^*(r)\psi_{p'}(r)}{|r - R_{B'}|}.$$

The averaging of this expression over the approximate wave function gives

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

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

$$\langle \Psi_k^{\text{appr}} | \mathcal{V}^R(q) | \Psi_k^{\text{appr}} \rangle = -2 \sum_{t_m t_m' \in M} \sum_{A \in R} V_A^{t_m t_m'} Z_A^R P_m^{tt'}. \quad (26)$$

It can be noted that the contributions of Eqs. (23), (25), and (26) 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:

$$\Delta E_k^{\text{el}-\text{st}} = e^2 \sum_{\substack{A \neq B \\ A \in R, \, B \in M}} Z_A^R Z_B^M R_{AB}^{-1}$$

$$+ \langle \Psi_k^{\text{appr}} | V^c(q) + \mathcal{V}^M(q) + \mathcal{V}^R(q) | \Psi_k^{\text{appr}} \rangle. \quad (27)$$

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

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

where the effective atomic charges are defined by

$$Q_A^{Rk} = \sum_{p \in R} \sum_\sigma P_k^{\sigma pp} - Z_A^R,$$

$$Q_B^M = 2 \sum_{t_m \in B} P_m^{tt} - Z_B^M. \quad (29)$$

The contribution to the PES from the operator $V^{rr}(q, E)$ corresponds to the second order of the per-

turbation theory with respect to the intersystem one-electron transfers. To estimate it we present the corresponding resolvent operator in the approximate form

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

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. (30)] are expressed through the ionization potentials (IPs) and the electron affinities (EAs) of the respective subsystems:

$$E_i - E = \begin{cases} I_\mu - A_\rho - g_{\mu\rho} \\ I_\rho - A_\mu - g_{\rho\mu}. \end{cases} \quad (31)$$

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 [31]

$$\begin{aligned} b_{m\sigma} &= x_m l_{m\sigma} + y_m r_{m\sigma} \\ a_{m\sigma} &= -y_m l_{m\sigma} + x_m r_{m\sigma} \end{aligned} \quad (x_m^2 + y_m^2 = 1), \quad (32)$$

where x_m and y_m are chosen to represent the mth geminal in the form

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

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

$$\begin{aligned} u_m &= U_m y_m^2 + V_m x_m^2 \\ v_m &= U_m x_m^2 + V_m y_m^2 \\ w_m &= (U_m - V_m) x_m y_m \end{aligned} \quad (34)$$

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

$$\begin{aligned} & b_{m\sigma}^+ \prod_{l \neq m} g_l^+ |0\rangle, \qquad a_{m\sigma}^+ \prod_{l \neq m} g_l^+ |0\rangle, \\ & b_{m\sigma}^+ r_{m-\sigma}^+ l_{m-\sigma}^+ \prod_{l \neq m} g_l^+ |0\rangle, \\ & a_{m\sigma}^+ r_{m-\sigma}^+ l_{m-\sigma}^+ \prod_{l \neq m} g_l^+ |0\rangle. \end{aligned} \quad (35)$$

The corresponding IPs and EAs entering Eq. (31) are estimated in Ref. [20]. It is convenient to introduce

some vacuum averages for the M subsystem:

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

These averages can be readily evaluated:

$$\theta_{rb}^m = y_m w_m + x_m v_m, \qquad \theta_{ra}^m = x_m w_m - y_m v_m,$$
$$\theta_{lb}^m = -x_m w_m - y_m u_m, \qquad \theta_{la}^m = y_m w_m - x_m u_m,$$
$$\eta_{rb}^m = y_m u_m + x_m w_m, \qquad \eta_{ra}^m = x_m u_m - y_m w_m,$$
$$\eta_{lb}^m = y_m w_m + x_m v_m, \qquad \eta_{la}^m = x_m w_m - y_m v_m. \tag{37}$$

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

$$\Delta E_k^{rr}$$
$$= \langle \Psi_k^{\mathrm{appr}} | V^{rr}(q, E) | \Psi_k^{\mathrm{appr}} \rangle$$
$$= -\sum_{pp' \in R} \sum_{m \in M} \sum_{ij \in \{r,l\}} \sum_{f \in \{a,b\}} v_{pi_m} v_{p'j_m}$$
$$\times \sum_{\sigma} \Bigg(\sum_{\rho \in ImO_R(N_R-1)} \frac{\langle \Phi_k^R | p_\sigma^+ | \rho \rangle \langle \rho | p_\sigma' | \Phi_k^R \rangle \theta_{if}^m \theta_{jf}^m}{I_{k\rho} - A_{mf} - g_{k\rho f_m}}$$
$$+ \sum_{\rho \in ImO_R(N_R+1)} \frac{\langle \Phi_k^R | p_\sigma | \rho \rangle \langle \rho | p_\sigma'^+ | \Phi_k^R \rangle \eta_{if}^m \eta_{jf}^m}{I_{mf} - A_{k\rho} - g_{f_m k\rho}} \Bigg). \tag{38}$$

This expression can be simplified if the IPs and EAs are replaced by their average values:

$$\Delta E_k^{rr} \approx -\sum_{pp' \in R} \sum_{m \in M} \sum_{ij \in \{r,l\}} \sum_{f \in \{a,b\}} v_{pi_m} v_{p'j_m}$$
$$\times \sum_{\sigma} \left(\frac{P_k^{\sigma pp'} \theta_{if}^m \theta_{jf}^m}{I_k - A_{mf} - g_{kf_m}} + \frac{(1 - P_k^{\sigma p'p}) \eta_{if}^m \eta_{jf}^m}{I_{mf} - A_k - g_{f_m k}} \right). \tag{39}$$

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

$$\langle \Psi_k^{\mathrm{appr}} | \mathcal{W}(q, E) P Q \mathcal{R}(q, \omega) Q P \mathcal{W}(q, E) | \Psi_k^{\mathrm{appr}} \rangle$$
$$= \langle \Psi_k^{\mathrm{appr}} | V^c(q) P Q R(q, \omega) Q P V^c(q) | \Psi_k^{\mathrm{appr}} \rangle$$
$$+ \langle \Psi_k^{\mathrm{appr}} | \mathcal{V}^R(q) P Q R(q, \omega) Q P \mathcal{V}^R(q) | \Psi_k^{\mathrm{appr}} \rangle$$
$$+ 2 \langle \Psi_k^{\mathrm{appr}} | V^c(q) P Q R(q, \omega) Q P \mathcal{V}^R(q) | \Psi_k^{\mathrm{appr}} \rangle$$
$$+ 2 \langle \Psi_k^{\mathrm{appr}} | V^c(q) P Q R(q, \omega) Q P V^{rr}(q, E) | \Psi_k^{\mathrm{appr}} \rangle$$
$$+ 2 \langle \Psi_k^{\mathrm{appr}} | V^{rr}(q, E) P Q R(q, \omega) Q P \mathcal{V}^R(q) | \Psi_k^{\mathrm{appr}} \rangle$$
$$+ \langle \Psi_k^{\mathrm{appr}} | V^{rr}(q, E) P Q R(q, \omega) Q P V^{rr}(q, E) | \Psi_k^{\mathrm{appr}} \rangle. \tag{40}$$

The first three contributions of Eq. (40) 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 that are all higher than the first excitation energy in the M (classical) subsystem, which in 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 Eq. (40) is the second-order correction due to the Coulomb repulsion operator. To write down this contribution it is convenient to introduce dynamic polarization propagators [22] for the R- and M-subsystems,

$$\Pi_{pp'qq'}^{Rk}(\omega) = \sum_{\rho \neq k} \frac{(\epsilon_\rho - \epsilon_k) \langle \Phi_k^R | p^+ p' | \rho \rangle \langle \rho | q^+ q' | \Phi_k^R \rangle}{(\epsilon_\rho - \epsilon_k)^2 - \omega^2},$$
$$\Pi_{mm'nn'}^{M}(\omega) = \sum_{\mu \neq 0} \frac{\epsilon_\mu \langle \Phi_{00}^M | m^+ m' | \mu \rangle \langle \mu | n^+ n' | \Phi_{00}^M \rangle}{\epsilon_\mu^2 - \omega^2}, \tag{42}$$

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. [20]. Using those notations we obtain

$$\langle \Psi_k^{\mathrm{appr}} | V^c(q) P Q \mathcal{R}(q, 0) Q P V^c(q) | \Psi_k^{\mathrm{appr}} \rangle$$
$$= \sum_{pp'qq' \in R} \sum_{i_m i_n' j_n j_m' \in M}$$
$$\times \left\{ \sum_{\substack{\sigma\tau \\ \sigma'\tau'}} (p_\sigma p_\sigma' \| i_{m\sigma'} i_{n\sigma'}') (q_\tau q_\tau' \| j_{n\tau'} j_{m\tau'}') \right\}$$
$$\times \Bigg[P_k^{\sigma pp'} P_k^{\tau qq'} \Pi_{i_{m\sigma'} i_{n\sigma'}' j_{n\tau'} j_{m\tau'}'}^{M}(0)$$
$$+ \frac{2}{\pi} \int \Pi_{p_\sigma p_\sigma' q_\tau q_\tau'}^{Rk}(iu) \Pi_{i_{m\sigma'} i_{n\sigma'}' j_{n\tau'} j_{m\tau'}'}^{M}(iu) \, du \Bigg]. \tag{43}$$

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 compared with

those of the R subsystem:

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

i.e., it is a sum of products of the static polarization propagator of the M subsystem and transition electron densities of the R subsystem. The contribution in Eq. (44) 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 states in the R subsystem. The operator $\mathcal{V}^R(q)$ can be written in the form

$$\mathcal{V}^R = -\sum_{mm' \in M} \sum_\sigma m^+_\sigma m'_\sigma \sum_{A \in R} V^{mm'}_A Z^R_A$$
$$\approx -\sum_{B \in M} \sum_{m \in B \cap M} \sum_\sigma m^+_\sigma m_\sigma \sum_{\substack{A \in R \\ A \neq B}} \gamma_{AB} Z^R_A. \tag{45}$$

The contribution to the energy is

$$\Delta E^{RR}_k = \langle \Psi^{\text{appr}}_k | \mathcal{V}^R(q) P Q \mathcal{R}(q,0) Q P \mathcal{V}^R(q) | \Psi^{\text{appr}}_k \rangle$$
$$= \sum_{i_m i'_n j_n j'_m \in M} \sum_{AA' \in R} V^{i_m i'_n}_A V^{j_n j'_m}_{A'} Z^R_A Z^R_{A'}$$
$$\times \sum_{\sigma\tau} \Pi^M_{i_{m\sigma}i'_{n\sigma}j_{n\tau}j'_{m\tau}}(0). \tag{46}$$

Let us denote the atom at the right end of the mth bond as R_m and that at the left end as L_m. If $AA' \notin \{R_m, L_m\}$ the contribution of Eq. (46) can be approximately rewritten in the physically transparent form in terms of the bond polarizability tensors $\hat{\alpha}^{(m)}(\omega)$ [20]:

$$-\sum_{m \in M} \sum_{\substack{AA' \in R \\ AA' \notin \{R_m, L_m\}}} Z^R_A Z^R_{A'} \big(\nabla V(R^{(m)}_{A'}) | \hat{\alpha}^{(m)}(0) | \nabla V(R^{(m)}_A)\big), \tag{47}$$

where $V(R^{(m)}_A)$ is the potential induced by a unit charge placed on the atom A at the center of the mth bond. It is worth mentioning that the bond polarizabilities $\hat{\alpha}^{(k)}(0)$ are tabulated, for example, in Ref. [32].

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

$$\Delta E^{cR}_k = \langle \Psi^{\text{appr}}_k | V^c(q) P Q \mathcal{R}(q,0) Q P \mathcal{V}^R(Q) | \Psi^{\text{appr}}_k \rangle$$
$$= -\sum_{\sigma\sigma'\tau} \sum_{pp' \in R} P^{\sigma pp'}_k \sum_{A \in R} Z^R_A$$

$$\times \sum_{i_m i'_n j_n j'_m \in M} (p_\sigma p'_\sigma \| i_{m\sigma'} i'_{n\sigma'}) V^{j_n j'_m}_A$$
$$\times \Pi^M_{i_{m\sigma'}i'_{n\sigma'}j_{n\tau}j'_{m\tau}}(0). \tag{48}$$

The contributions to the energy ΔE^{cR}_k and ΔE^{RR}_k and the first contribution of Eq. (43) can be grouped in the total correction ΔE^{coul}_k, 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

$$\Delta E^{\text{coul}}_k = \sum_{m \in M} \sum_{ii' \in \{r,l\}} \sum_{AA' \in R} \gamma_{AI_m} \gamma_{A'I'_m} Q^R_A Q^R_{A'}$$
$$\times \sum_{\sigma\tau} \Pi^M_{i_{m\sigma}i'_{m\sigma}i'_{m\tau}i_{m\tau}}(0), \tag{49}$$

where I_m denotes the atom of the mth bond corresponding to the ith HO of the bond (R_m or L_m).

The next two contributions of Eq. (40), 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 electron transfers between the subsystems are replaced by average ones, depending only on the number of bonds taking part in the electron transfer. The cross term between the $V^c(q)$ and $V^{rr}(q,E)$ is

$$\Delta E^{crr}_k = \langle \Psi^{\text{appr}}_k | V^c(q) P Q \mathcal{R}(q,0) Q P V^{rr}(q,E) | \Psi^{\text{appr}}_k \rangle$$
$$\approx \sum_{pqq' \in R} \sum_{i_m j_m j'_m \in M} v_{qj_m} v_{q'j'_m} \sum_{\sigma\sigma'\tau} (p_\sigma p_\sigma \| i_{m\sigma'} i_{m\sigma'})$$
$$\times \Bigg\{ (I_m - A_k - g_{mk})^{-1}$$
$$\times \bigg[P^{\sigma pp}_k (1 - P^{\tau q'q}_k) \Pi^M_{i_{m\sigma'}i_{m\sigma'}j_{m\tau}j'_{m\tau}}(0)$$
$$- \frac{2}{\pi} \int \Pi^{Rk}_{p_\sigma p_\sigma q'_\tau q_\tau}(iu)$$
$$\times \Pi^M_{i_{m\sigma'}i_{m\sigma'}j_{m\tau}j'_{m\tau}}(iu)\, du \bigg]$$
$$- (I_k - A_m - g_{km})^{-1}$$
$$\times \bigg[P^{\sigma pp}_k P^{\tau qq'}_k \Pi^M_{i_{m\sigma'}i_{m\sigma'}\tilde{j}_{m\tau}\tilde{j}'_{m\tau}}(0)$$
$$+ \frac{2}{\pi} \int \Pi^{Rk}_{p_\sigma p_\sigma q_\tau q'_\tau}(iu)$$
$$\times \Pi^M_{i_{m\sigma'}i_{m\sigma'}\tilde{j}_{m\tau}\tilde{j}'_{m\tau}}(iu)\, du \bigg] \Bigg\}, \tag{50}$$

where $\tilde{j} = r$ for $j = l$ and $\tilde{j} = l$ for $j = r$. The integral expressions can be replaced by sums using Eq. (44).

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

$$\Delta E_k^{Rrr} = \left\langle \Psi_k^{\mathrm{appr}} \middle| \mathcal{V}^R(q) P \mathcal{Q} \mathcal{R}(q, 0) \mathcal{Q} P V^{rr}(q, E) \middle| \Psi_k^{\mathrm{appr}} \right\rangle$$

$$\approx - \sum_{\sigma\tau} \sum_{pp' \in R} \sum_{A \in R} Z_A^R \sum_{j_m j_m' \in M} v_{pj_m} v_{p'j_m'}$$

$$\times \left\{ (I_m - A_k - g_{mk})^{-1} \left(1 - P_k^{\tau p' p}\right) \right.$$

$$\times \left(\gamma_{AR_m} \Pi_{r_{m\sigma} r_{m\sigma} j_{m\tau} j_{m\tau}'}^M (0) \right.$$

$$\left. + \gamma_{AL_m} \Pi_{l_{m\sigma} l_{m\sigma} j_{m\tau} j_{m\tau}'}^M (0) \right) \quad (51)$$

$$- (-1)^{\delta_{jj'}} (I_k - A_m - g_{km})^{-1} P_k^{\tau p p'}$$

$$\times \left(\gamma_{AR_m} \Pi_{r_{m\sigma} r_{m\sigma} \tilde{j}_{m\tau} \tilde{j}_{m\tau}'}^M (0) \right.$$

$$\left. \left. + \gamma_{AL_m} \Pi_{l_{m\sigma} l_{m\sigma} \tilde{j}_{m\tau} \tilde{j}_{m\tau}'}^M (0) \right) \right\}. \quad (52)$$

The last contribution to the energy of Eq. (40) is of the fourth order with respect to the perturbation operator (counting four small operators of electron transfer) and may be totally neglected.

Discussion

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

We try to justify the form of the junction between subsystems. This problem is obviously close to that of intermolecular forces. At the same time there are some important differences: the hybrid schemes require consideration of boundary atoms, the special form of the wave function of inert subsystem is required, and the resonance terms cannot be neglected. The complete neglect of the junction, i.e., the use of the bare (ab initio) Hamiltonian for the R subsystem and a standard molecular mechanical procedure without renormalization leads to the energy $E_{k0}^R + E_0^M$ 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:

$$\Delta E_k \approx \Delta E_k^{\mathrm{el-st}} + \Delta E_k^{rr} + \Delta E_k^{\mathrm{disp}}$$
$$+ \Delta E_k^{\mathrm{coul}} + \Delta E_k^{crr} + \Delta E_k^{Rrr}, \quad (53)$$

where the main contribution $\Delta E_k^{\mathrm{el-st}}$ is of the first order with respect to the perturbation operator and originates from the electrostatic interaction between the subsystems; the next three 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 the interaction of the Coulomb electron–electron interaction with the interaction of the electrons of the M subsystem and cores of the R subsystem; and the second term can be regarded as the dispersion interaction between subsystems. Moreover, the contributions $\Delta E_k^{\mathrm{disp}}$ and $\Delta E_k^{\mathrm{coul}}$ 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. (53) 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 the Coulomb interaction between electrons of the M subsystem with electrons and cores of the R subsystem.

The APSLG approach used for description of the classical part of the whole system is not valid for description of highly delocalized groups. The APSLG form of the M subsystem wave function affects only the precise form of the resolvents entering the formulae presented. The expressions for the corrections to the energy in the R subsystem written with use of these resolvents remain valid irrespective of the specific form of the group function used for their calculation.

References

1. CECAM-NSF. Int J Quantum Chem 1996, 60(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 Am Chem Soc 1992, 114, 1606.

18. Tchougréeff, A. L. Khim Fiz 1997, 16, 62 (in Russian); Chem Phys Rep 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. McWeeny, R. Methods of Molecular Quantum Mechanics, 2nd ed.; Academic Press: London, 1992.

23. Surján, P. R. In Theoretical Models of Chemical Bonding, Part 2; Maksić, Z. B., Ed.; Springer: Heidelberg, 1989.

24. Wu, W.; McWeeny, R. J Chem Phys 1994, 101, 4826.

25. Tokmachev, A. M.; Tchougréeff, A. L. Zh Fiz Khim 1999, 73, 347 (in Russian); Russ J Phys Chem 1999, 73, 320 (in English).

26. Tokmachev, A. M.; Tchougréeff, A. L. J Comput Chem, to appear.

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

28. Constanciel, R. In Localization and Delocalization in Quantum Chemistry; Chalvet, O.; Daudel, R.; Diner, S; Malrieu, J.-P., Eds.; D. Reidel: Dordrecht, The Netherlands, 1975.

29. Schipper, P. Aust J Chem 1986, 40, 635.

30. Tchougréeff, A. L.; Misurkin, I. A. DAN SSSR 1986, 291, 1977.

31. Tokmachev, A. M.; Tchougréeff, A. L.; Misurkin, I. A. Int J Quantum Chem 2001, 000, 000.

32. Le Fevre, R. J. W. In Advances in Physical Organic Chemistry; Gold, V., Ed.; Academic Press: New York, 1965; Vol. 3.