

Approximate electrostatic interaction operator for QM/MM calculations

Nicolas Ferré¹, János G. Ángyán^{*}

Équipes de Chimie et Biochimie Théoriques, Unité Mixte de Recherche CNRS/UHP 7565, Institut Nancéen de Chimie Moléculaire, Université Henri Poincaré, B.P. 239, F-54506 Vandœuvre-lès-Nancy, France

Received 11 January 2002; in final form 19 February 2002

Abstract

The electrostatic interaction energy between a quantum chemical charge distribution and an arbitrary external potential is approximated using the expectation value of distributed multipole operators fitted to the electrostatic potential. This operator, by construction, yields the best possible representation of the interaction energy in the framework of a limited multi-centered multipole expansion of the interacting charge distributions. The use of this operator may be particularly advantageous in quantum/classical mixed calculations, because it ensures a smooth transition between quantum chemical and molecular mechanical modeling of the electrostatic interaction between subsystems. © 2002 Published by Elsevier Science B.V.

1. Introduction

Electrostatic embedding is a widespread method to model molecular subsystems in interaction with a polar environment. Provided that the environment is represented by a set of point charges, the interaction energy calculation can be easily realized by using the electrostatic potential integrals [1,2]. In principle, such a procedure could be generalized to dipoles, quadrupoles, etc. with the help of electrostatic field, field-gradient, etc. integrals. In spite of its relative simplicity and easy

implementation, electrostatic embedding via a set of point charges has at least two major drawbacks. The first is that the interchangeability of the classical and quantum subsystems is not satisfied. Suppose that the quantum chemical subsystem is represented by a continuous electron density, while the classical part is approximated by a set of point multipoles. If the role of the subsystems is interchanged, i.e. we replace the charge density function by a point multipole representation and vice versa, the calculated electrostatic interaction energies may be appreciably different, by about 10% of the total electrostatic interaction energy (see below). The second difficulty is that the representation of the external potential in terms of a finite set of point multipoles is not always directly accessible. The most common example is the Madelung potential, which can be determined by

^{*} Corresponding author.

E-mail address: angyan@lctn.uhp-nancy.fr (J.G. Ángyán).

¹ Present address: Dipartimento di Chimica, Università di Siena, via Aldo Moro, 53100 Siena, Italy.

some Ewald-type method [3–5]. One can either consider explicitly the integrals between Gaussian orbitals and the analytic expression of the Ewald sum, as derived e.g. by Saunders et al. [6] or replace the analytic calculation of the Ewald sum by an appropriately selected set of point charges/multipoles [1,3,7], that are fitted to reproduce the Ewald sums in a least squares sense. Still, both of these procedures suffer from the interchangeability problem. Furthermore, the fitted point charge set that represents the Madelung-potential has a non-straightforward relationship to the geometrical parameters of the global system, rendering the analytical calculation of the forces non-obvious.

The lack of interchangeability of quantum and classical subsystems can be traced back to the lack of symmetry in the electrostatic interaction operator itself. In the Longuet–Higgins form [8] the complete electrostatic interaction operator is

$$\hat{V}_{AB} = \int \int d\mathbf{r} d\mathbf{r}' \hat{\rho}_A(\mathbf{r}) T(\mathbf{r}, \mathbf{r}') \hat{\rho}_B(\mathbf{r}'), \quad (1)$$

where $\hat{\rho}_A(\mathbf{r})$ and $\hat{\rho}_B(\mathbf{r}')$ have the same analytical expression

$$\hat{\rho}_X(\mathbf{r}) = \sum_{a \in X} Z_a - \sum_{i \in X} \delta(\mathbf{r} - \mathbf{r}_i), \quad (2)$$

therefore the exchange $A \rightleftharpoons B$ leaves the interaction operator invariant. The ‘conventional’ procedure followed in QM/MM calculations [9–11] consists of replacing the above charge density operator, Eq. (2) of one of the subsystems, usually that of the environment (B), by an approximate one, $\hat{\rho}_B^{\text{appr}}$, which means that if we exchange the roles of A and B

$$\begin{aligned} & \int \int d\mathbf{r} d\mathbf{r}' \hat{\rho}_A(\mathbf{r}) T(\mathbf{r}, \mathbf{r}') \hat{\rho}_B^{\text{appr}}(\mathbf{r}') \\ & \neq \int \int d\mathbf{r} d\mathbf{r}' \hat{\rho}_A^{\text{appr}}(\mathbf{r}) T(\mathbf{r}, \mathbf{r}') \hat{\rho}_B(\mathbf{r}'), \end{aligned} \quad (3)$$

i.e. the invariance of the interaction operator is not preserved any more.

One way to re-establish the symmetry of the interaction operator is to use the same *approximate charge density operator* for both subsystems. It will be shown that if we require that this approximate charge density operator be of multipolar nature and reproduce at the best interaction

energy, the optimal approximate charge density operator has the sum of potential-fitted distributed multipole moments as its expectation value.

Several algorithms exist which allow one to assign point multipoles to atomic centers in the molecule. The most popular techniques are based either on a partitioning of the real molecular space [12] or of the atomic basis function space, [13–15] or on some least squares fitting techniques [16]. Although the partitioning of the molecular space according to the theory of atoms in molecules [12] has the advantage of being transferable, it may lead to quite a lengthy multipolar expansion depending on the shape and the distribution of the included charge in each individual atomic domain. Among the partitioning techniques based on the specific properties of the Gaussian basis function products, the method proposed by Stone is the most efficient to keep the order of expansion low on each of the centers [13]. However, even this method requires relatively high-order multipoles for a faithful reproduction of the electrostatic potential [17]. The charge (multipole) models fitted to the potential provide a reasonable compromise between compactness and precision [16]. **As has been shown earlier, the potential fitted models are able to mimic the effect of intrinsically higher order atomic multipoles, essentially by spreading out tiny fractional charges and lower order multipoles on the surrounding atoms** [17,18]. Since an important requirement in the design of force fields for large-scale molecular dynamics simulations is to keep the order of the multipole expansion at a low level, the popularity of the potential-fitted charge (multipole) models can be easily understood.

In spite of the numerous examples for the incorporation of atomic charge (multipole) models in the electrostatic embedding or solvent effect problems, the question of an atomic multipole *interaction operator* for the description of the deformation of the molecular charge distribution in an external field has not been studied to our knowledge. A slightly reminiscent method has been suggested by Blöchl, [19,20] based on the fit of charge density by a set of spherical Gaussian distributions in the context of plane wave calculations. In the present paper we propose an approximation scheme which minimizes the error of

a low-order multipole model. Explicit formulae are given for the interaction operator matrix elements, the interaction energy itself, as well as for the energy gradients, i.e. the forces.

2. Theory

Let us consider the molecular charge distribution as the expectation value of the operator defined in Eq. (1), i.e.

$$\varrho(\mathbf{r}) = \sum_a Z_a \delta(\mathbf{r} - \mathbf{r}_a) - \sum_{\mu\nu} P_{\mu\nu} \chi_\mu^*(\mathbf{r}) \chi_\nu(\mathbf{r}), \quad (4)$$

which is a sum of the contributions from the nuclei, represented by the charges Z_a , and from the electrons, expressed by the first-order density matrix in terms of the basis functions, $\chi_\mu(\mathbf{r})$. The molecular charge distribution is in interaction with an external perturbing potential, $V(\mathbf{r})$. The question is whether it is possible to replace the charge distribution $\varrho(\mathbf{r})$, by an approximate one, $\tilde{\varrho}(\mathbf{r})$, constructed from atom-centered point multipoles.

The general form of such a point multipole distribution can be written following the equivalent charge distribution concept of Bertaut [21] as

$$\tilde{\varrho}(\mathbf{r}) = \sum_a \delta(\mathbf{r} - \mathbf{r}_a) \sum_{\ell\kappa} \left[\frac{1}{(2\ell-1)!!} \right] Q_{\ell\kappa}^a R_{\ell\kappa}(\mathbf{V}), \quad (5)$$

where the summation is over the multipole centers, a , $Q_{\ell\kappa}^a$ are the real spherical harmonics distributed multipole components [22], $R_{\ell\kappa}(\mathbf{V})$ is a regular solid harmonic function [23] where the \mathbf{V} operator replaces the usual argument, \mathbf{r} , the double factorial means the product of odd numbers up to $(2\ell-1)$, i.e. $(2\ell-1)!! = 1 \cdot 3 \cdot 5 \cdots (2\ell-1)$ and $\delta(\mathbf{r} - \mathbf{r}_a)$ is the Dirac delta function, which reflects the point-like distribution of $\tilde{\varrho}(\mathbf{r})$. The free parameters, $Q_{\ell\kappa}^a$, can be determined by minimizing the difference of the approximate and exact interaction energies with an external potential, $V(\mathbf{r})$.

2.1. Determination of the fitted multipoles

The expression of the interaction energy between $\varrho(\mathbf{r})$, the exact charge distribution, and a set of point multipoles $\{q_{\ell\kappa}^k\}$ localized at positions \mathbf{r}_k is

$$E_{\text{int}} = \sum_k \sum_{\ell\kappa} q_{\ell\kappa}^k V_{\ell\kappa}^k, \quad (6)$$

where $V_{\ell\kappa}^k$ are the Taylor expansion coefficients of the electrostatic potential of the exact charge distribution $\varrho(\mathbf{r})$ at the point \mathbf{r}_k

$$V_{\ell\kappa}^k = \int d\mathbf{r}' \varrho(\mathbf{r}') T_{\ell\kappa,00}(\mathbf{r}_k, \mathbf{r}'). \quad (7)$$

If we substitute the approximate ‘equivalent’ charge distribution, $\tilde{\varrho}(\mathbf{r})$, we obtain the expression of the approximate electrostatic potential, $\tilde{V}_{\ell\kappa}^k$

$$\tilde{V}_{\ell\kappa}^k = \sum_a \sum_{\ell'\kappa'} Q_{\ell'\kappa'}^a T_{\ell\kappa,\ell'\kappa'}(\mathbf{r}_k, \mathbf{r}_a). \quad (8)$$

The optimal distributed multipole parameters can be obtained by minimizing the difference between the approximate and exact interaction energies, based on the following least squares function, weighted by the environment multipole moments $q_{\ell\kappa}^k$, as

$$\sum_k \sum_{\ell\kappa} q_{\ell\kappa}^k \left(\tilde{V}_{\ell\kappa}^k - V_{\ell\kappa}^k \right)^2. \quad (9)$$

The corresponding standard normal equations are in their matrix form

$$\mathbf{T}^\dagger \mathbf{q} \mathbf{T} \mathbf{Q} = \mathbf{T}^\dagger \mathbf{q} \mathbf{V}. \quad (10)$$

The matrix \mathbf{T}^{KA} , and the vectors \mathbf{q}^K , \mathbf{Q}^A , \mathbf{V}^K depend on the collective indices, $K = \{\ell\kappa\}^k$ and $A = \{\ell'\kappa'\}^a$. The potential-fitted point multipoles obtained by this procedure are specific to the set of external point charges through the weighting factors and also by the fact that the fit is restricted to those points where the external potential sources are located.

The matrix elements of the atomic multipole component operators can be determined by considering the following relationship, obtained from the formal solution of Eq. (10) and substituting the electrostatic potential from Eqs. (4) and (7)

$$\mathbf{Q}^A = \mathbf{Z}^A - \sum_B \sum_K \mathbf{G}^{AB} \mathbf{T}^{KB} \mathbf{q}^K \sum_{\mu\nu} P_{\mu\nu} V_{\mu\nu}^K, \quad (11)$$

where $\mathbf{G} = (\mathbf{T}^\dagger \mathbf{q} \mathbf{T})^{-1}$. The electrostatic potential matrix elements are defined as $V_{\mu\nu}^K = \langle \mu | T_{\ell\kappa,00}(\mathbf{r}_k, \mathbf{r}) | \nu \rangle$ and \mathbf{Z}^A is the nuclear charge or higher multipole moment component in the case of

united atoms. As any expectation value, the fitted atomic multipole moment components can also be written in the general form

$$\mathcal{Q}^A = Z^A - \sum_{\mu\nu} P_{\mu\nu} \mathcal{Q}_{\mu\nu}^A, \quad (12)$$

and the matrix elements, $\mathcal{Q}_{\mu\nu}^A$, are obtained by identification as

$$\mathcal{Q}_{\mu\nu}^A = \sum_B \sum_K G^{AB} T^{KB} q_{\mu\nu}^K. \quad (13)$$

Regardless of the weighting, we could have obtained the same result by fitting individually the matrix elements of the electrostatic potential, $V_{\mu\nu}^K$ by the matrix elements of $\hat{\mathcal{Q}}^A \equiv \hat{\mathcal{Q}}_{\ell\kappa}^A$ the operator of the atomic multipole components. This charge density operator, constructed with the help of $\hat{\mathcal{Q}}_{\ell\kappa}^A$ will be called the *electrostatic potential fitted* (ESPF) operator.

In the following **we drop the weighting**, i.e. we replace \mathbf{G} by $(\mathbf{T}^\dagger \mathbf{T})^{-1}$. Using the shorthand matrix notation, by introducing the vector of the N_{grid} electrostatic potential matrix elements, $\mathbf{V}_{\mu\nu}$, the vector of the N_{comp} distributed atomic multipole moments, $\mathbf{Q}_{\mu\nu}$, as well as the $N_{\text{grid}} \times N_{\text{comp}}$ matrix of the electrostatic interaction kernels, \mathbf{T} , the fitted atomic moment matrix elements can be expressed as

$$\mathbf{Q}_{\mu\nu} = (\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \mathbf{V}_{\mu\nu}. \quad (14)$$

2.2. Interaction with external potential

The matrix elements of the interaction operator with the external potential

$$\Delta H_{\mu\nu} = \sum_A \mathcal{Q}_{\mu\nu}^A V^A \quad (15)$$

is the sum of the $\mathcal{Q}_{\mu\nu}^A$ atomic multipole matrix elements with the corresponding electrostatic potential and field components, $V^A = V_{\ell\kappa}(\mathbf{r}_a)$ taken at the positions where the multipole moments are located

$$V_{\ell\kappa}(\mathbf{r}_a) = \sum_k \sum_{\ell'\kappa'} T_{\ell\kappa, \ell'\kappa'}(\mathbf{r}_a, \mathbf{r}_k) q_{\ell'\kappa'}^k. \quad (16)$$

The contraction of the above interaction energy operator matrix, $\Delta H_{\mu\nu}$, with the density matrix $P_{\mu\nu}$,

and taking into account Eq. (12) leads to the usual expression of the classical electrostatic interaction energy

$$\Delta E = \sum_a \sum_k \sum_{\ell\kappa} \sum_{\ell'\kappa'} \mathcal{Q}_{\ell\kappa}^a T_{\ell\kappa, \ell'\kappa'}(\mathbf{r}_a, \mathbf{r}_k) q_{\ell'\kappa'}^k. \quad (17)$$

We see that the continuity between the classical and quantum chemical electrostatic interaction potentials is ensured by the present method.

The quantum chemical energy expression should be corrected by the interaction energy term

$$\Delta E = \sum_{\mu\nu} P_{\mu\nu} \Delta H_{\mu\nu} + \sum_A Z^A V^A. \quad (18)$$

Of course, the total change of the energy for a wave function optimized in the presence of the external perturbation includes the polarization effects related to the change of the density matrix.

2.3. Forces

In the following we present the equations needed to calculate the analytic derivatives of the electrostatic perturbation of the Hamiltonian.

The derivative of the total energy E_0 is corrected by the derivative of ΔE , defined in Eq. (18)

$$\begin{aligned} \frac{\partial E}{\partial x} &= \frac{\partial E_0}{\partial x} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x} \Delta H_{\mu\nu} + \sum_{\mu\nu} P_{\mu\nu} \frac{\partial \Delta H_{\mu\nu}}{\partial x} \\ &\quad + \sum_A Z^A \frac{\partial V^A}{\partial x}. \end{aligned} \quad (19)$$

The derivative of the density matrix $P_{\mu\nu}$ can be treated in the usual way. In the case of the Hartree–Fock method, for example, the explicit calculation of the density matrix derivative can be avoided owing to the self-consistency conditions. The derivative of the electrostatic potential and its derivatives at the nuclear sites are the classical electrostatic field and higher field gradient components. Finally, the derivatives of the ESPF interaction Hamiltonian matrix elements can be calculated as

$$\partial \Delta H_{\mu\nu} = \sum_a \sum_{\ell\kappa} \left(\partial \mathcal{Q}_{\mu\nu}^{a, \ell\kappa} V_{\ell\kappa}(\mathbf{r}_a) + \mathcal{Q}_{\mu\nu}^{a, \ell\kappa} \partial V_{\ell\kappa}(\mathbf{r}_a) \right). \quad (20)$$

The term involving the derivative of the external potential can be combined with the nuclear

contribution leading to the physically plausible expression comprising the net atomic multipole moment multiplied by the electric field or its derivatives at the nuclear position. The derivative of the atomic multipole *matrix element* contains on the one hand the derivative of the electrostatic potential matrix elements, and the derivative of the least square solution matrix, on the other

$$\partial Q_{\mu\nu} = \partial(\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \mathbf{V}_{\mu\nu} + (\mathbf{T}^\dagger \mathbf{T})^{-1} \partial \mathbf{T}^\dagger \mathbf{V}_{\mu\nu} + (\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \partial \mathbf{V}_{\mu\nu}. \quad (21)$$

Using the relationship

$$\partial(\mathbf{T}^\dagger \mathbf{T})^{-1} = -(\mathbf{T}^\dagger \mathbf{T})^{-1} (\partial \mathbf{T}^\dagger \mathbf{T} + \mathbf{T}^\dagger \partial \mathbf{T}) (\mathbf{T}^\dagger \mathbf{T})^{-1} \quad (22)$$

the derivatives of the inverse matrix $\partial(\mathbf{T}^\dagger \mathbf{T})^{-1}$ can be easily found from the derivatives $\partial \mathbf{T}$ and $\partial \mathbf{T}^\dagger$ that involve only higher electrostatic tensor components.

2.4. Computational aspects

The implementation of the ESPF operator, provided that one has the electrostatic potential integrals and their higher derivatives on a grid of points around the quantum chemical subsystem, is quite straightforward. Developing the interaction operator, Eq. (15) leads to

$$\Delta H_{\mu\nu} = \sum_A \sum_K \left((\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \right)^{AK} V_{\mu\nu}^K V^A. \quad (23)$$

Summing first on $A = \{ \ell'_{K'}^a \}$ multipole centers and components, this expression is simplified to a sum of electrostatic potential integrals weighted by the quantity B^K , which is the product of the quantity in parentheses and the electrostatic potential (and its derivatives), V^A , of the environment at the expansion centers

$$\Delta H_{\mu\nu} = \sum_K B^K V_{\mu\nu}^K. \quad (24)$$

The vector B^K involves only geometry-dependent quantities, and it should be calculated only once for a given molecular geometry and environment. Its derivative, that enters in the expression of the forces, is obtained in a straightforward manner using the results outlined above.

In the present work, the ESPF interaction operator has been incorporated into the GAUSSIAN 98 package, [24] through the creation of a new link and the modification of another one. The electronic charge distribution can be fitted by multipoles of order 0 or 1. The evaluation of the electrostatic potential is performed on grids of various types: Merz–Kollman–Singh, [25,26] CHelp, [27], CHelpG [28] and PNT grids [29]. The first three are standard options in GAUSSIAN, the last one comes from the GRID program [29].

3. Results

3.1. Interchangeability of subsystems

As mentioned before, the interchangeability characterizes the continuity between the descriptions used for the quantum system and for the external potential. Thus it depends on the length of the multipolar expansion for both subsystems, i.e. the one that generates the external potential and the one that fits the electrostatic potential. Obviously, **these expansions should be of the same length, otherwise they would violate trivially the interchangeability.** This happens if one incorporates the surrounding point charges through the calculation of $\langle \mu | \frac{q_k}{r_k} | v \rangle$ integrals: the perturbing potential is **calculated through a truncated atomic multipole development**, while the interaction operator consists of the non-expanded charge density. This case will be designated as non-expanded electrostatic (NEES) operator, which obviously does not satisfy the interchangeability condition.

The behavior of the electrostatic potential fitted (ESPF) operator with respect to the interchangeability property has been checked on two molecular complexes, illustrated on Fig. 1. The first one is a formamide–water complex (denoted FW in the following), while the second one is a water dimer (denoted WW in the following). In both cases, one of the molecules is treated quantum mechanically while the other one is treated classically, using point charges, and also charges and dipoles in the case of water. These point charges (and dipoles) were fitted to the electrostatic potential of the

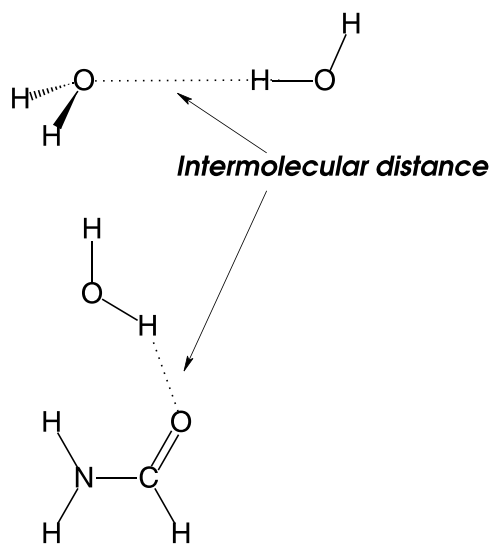


Fig. 1. Definition of intermolecular distance in the water dimer and in the formamide–water complex.

isolated molecule using the GRID program [29]. The interaction between the quantum chemically treated and the classical molecule is taken into account by either the ESPF operator presented in this paper, or by the ‘conventional’ NEES operator. Then the roles are inverted: the first molecule is treated classically while the second one is treated quantum mechanically.

In order to ensure independence of our results with respect to the grid shape, the same grids were used to fit the classical multipole model and the ESPF operator parameters. The initial guess of the density matrix was always the optimized wave function of the isolated molecule and the total energy of the first iteration corresponded to the first order electrostatic interaction energy (Fig. 2). The polarization energy is obtained from the fully converged one by subtraction of the energy calculated in the first iteration.

Several intermolecular distances, as defined in Fig. 1, have been considered, while the internal geometry of each molecule has been kept frozen. Hartree–Fock calculations were performed using the 6-31G** basis set.

The interchangeability property of the interaction operators can be the best illustrated by plotting the difference of the first order electrostatic

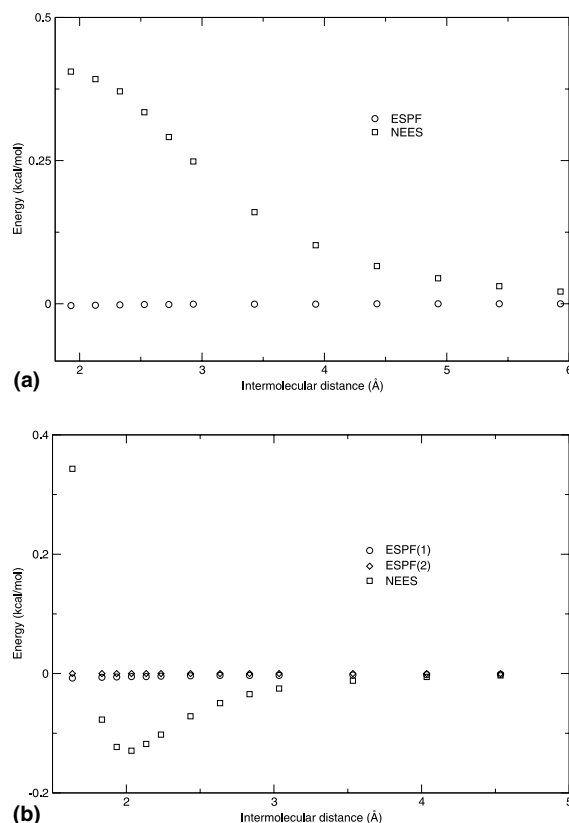


Fig. 2. First order (electrostatic) energy differences for the (a) formamide–water and (b) water–water complexes calculated by the ESPF and NEES operators. ESPF(1) designates the atomic charge operator, ESPF(2) designates the atomic charges and dipoles operator.

energies, obtained from the two cases with inverted roles of the subsystems. Using the ESPF operator, the difference is practically zero. The tiny deviations are due to the limited numerical precision of the fitting procedures. The results obtained by the NEES operator deviate considerably from this interchangeability. The fact of treating the first or second subsystem as ‘active’ one, results in a difference of 0.2–0.4 kcal/mol of the electrostatic interaction energy.

The polarization energy, obtained as the difference of the converged and of the first order electrostatic energies, is considerably smaller than the electrostatic energies themselves: in these two complexes it represents less than 15% of the electrostatic energy. However, one should not forget

that the estimation of the polarization of the wave function and the corresponding stabilization constitutes the major goal of the QM/MM type calculations. The notion of interchangeability is less obvious for this case, since $E_{\text{pol}}(\text{A} \leftarrow \text{B})$ and $E_{\text{pol}}(\text{B} \leftarrow \text{A})$ can be inherently different on physical grounds and therefore their difference does not characterize the interchangeability in the same sense. The polarization energy components calculated for the FW and WW complexes are plotted in Fig. 3.

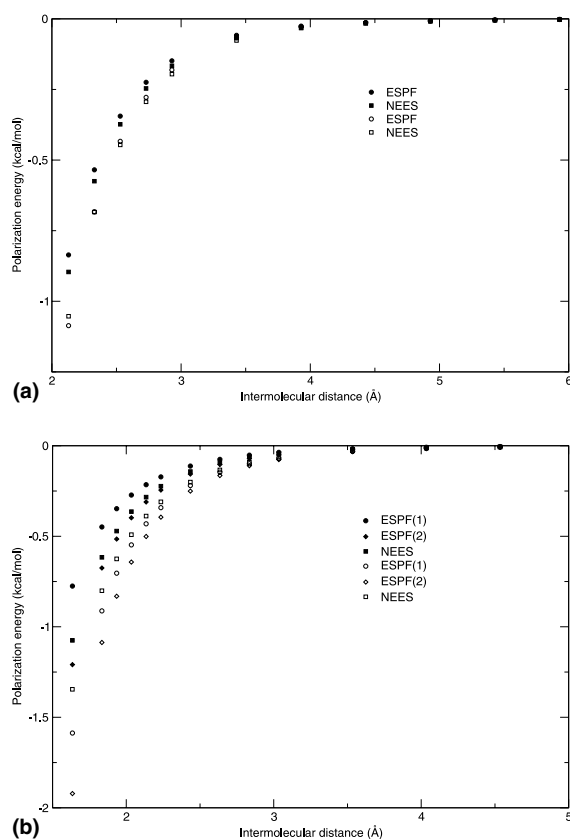


Fig. 3. Polarization energies for the (a) formamide–water and (b) water–water complexes calculated by the two operators (ESPF and NEES). ESPF(1) was obtained from charges, ESPF(2) from charges and dipoles. The black symbols correspond to the polarization energies obtained with quantum chemically treated formamide (a) and proton acceptor water (b) molecules, respectively, while the empty symbols designate results obtained after inverting the roles of the two subsystems, i.e. quantum mechanically treated water (a) and proton donor water (b) molecules.

One should be aware of the fact that an (over)simplified representation of the charge density operator, e.g. in terms of atomic charges, may lead to considerable errors in the total polarization energies. The most obvious shortcomings of this representation are the missing out-of-plane polarization of the planar molecules and the poor representation of the highly polarizable lone-pair regions. In this respect the NEES operator is of superior quality, since it does not contain any ‘built-in’ restriction limiting the nature of the charge reorganization. This is quite well illustrated by the comparison of the formamide–water and water–water complexes. For the quantum mechanically treated water molecule, that is polarized by the classical charge distribution of the formamide, the agreement between the NEES and ESPF polarization energies is very good. In this complex, the two molecules are coplanar and the water acts as a proton donor. The resulting polarization takes place essentially by an O–H charge transfer, which can be quite well reproduced by the point charge operator. The situation is still acceptable for the formamide polarized by the water. Again, owing to the coplanarity of the FW complex, most of the polarization of the formamide can be described in terms of charge flows. However, one expects that the deformation of the electron cloud in the H-acceptor oxygen lone pair region might be incomplete at this level. In the case of the water dimer (WW), the proton acceptor water molecule is polarized in the out-of-plane direction. The charge-flow representation is unable to describe such a deformation, therefore the inclusion of atomic dipoles in the ESPF operator seems to be mandatory in this case. In fact, the inclusion of atomic dipole moments lead to an appreciable increase of the polarization energy for this configuration. Although the proton donor water is not exposed to an out-of-plane electric field, the charge-flow-only description is insufficient to describe the polarization of the lone pair region, which explains that a non-negligible error was found for this case too.

In the case of the FW complex the difference of the ESPF and NEES polarization energies are almost negligible, less than 0.1 kcal/mol. For the water dimer, the deviations are slightly larger, at

most of 0.2 kcal/mol at an intermolecular separation of 2 Å. It is interesting to note that while the ESPF polarization energy of the acceptor water gets nearer to the NEES value by adding dipoles to the charges, an opposite behavior can be observed for the proton donor water.

In conclusion, it seems that the interchangeability with respect to the total interaction energy of the quantum and the classical descriptions can be ensured using the ESPF interaction operator proposed in this paper, especially if classical point multipoles are not too close to the quantum atoms.

3.2. Validation of the ESPF interaction operator

In QM/MM calculations, it is important to polarize efficiently the electronic wave function in order to get accurate reaction energies and realistic molecular geometries. Therefore we attempted to verify the reliability of the ESPF operator method.

3.2.1. Geometry optimization of the complexes

Using again the two model systems described above, the geometry of the quantum molecule is optimized at the HF/6-31G** level in the field of the point charges used to describe the classical molecule. Whatever model is chosen to fit the multipoles (type of grid, length of the expansion, etc.), the agreement between a calculation using the ESPF and NEES interaction operators is very satisfying. Geometry differences are always less than 1 pm for the intra- or inter-molecular distances and less than 0.3° for the valence angles. At a given geometry, energy differences are about one tenth of the calculated interaction energy in the worst cases and can reach one hundredth of the interaction energy (in the case of the water dimer for example, when the fitted multipoles are charges and dipoles). Although these results are satisfying, the simple binary complexes are not really representative of real QM/MM situations, where a quantum subsystem is usually surrounded by numerous classical atoms. A typical QM/MM study is presented in the following paragraph.

3.2.2. Solvation of the formamide molecule

A formamide molecule is placed at the center of a cubic box made of 216 water molecules, repre-

sented by the TIP3P model (the box edge measures 18.8 Å in order to simulate a density of 1). The geometry of the formamide molecule is optimized in the field of the classical charges (−0.834 electrons for each oxygen atom and +0.417 for each hydrogen atom). Positions of classical atoms are kept frozen. No van der Waals potential is added to the pure electrostatic QM/MM interactions. Hartree–Fock calculations are performed using the 6-31G** basis set. Three types of grid are used, the first one is obtained using the Merz–Kollman–Singh scheme, the second one using the PNT scheme and the last one is called a ‘molecular’ grid because every classical atom is used as a grid point. Here again, the same kind of calculation, using the NEES interaction operator, serves as the reference.

If the fitted multipoles are limited to the order 0 (charges), final geometries are all in good agreement (variations of bond lengths are less than 1 pm). For a given geometry, energy differences are small: 0.007 kcal/mol for the PNT grid, 0.038 kcal/mol for the MKS grid and 0.155 kcal/mol for the ‘molecular’ grid. This last result can be explained because this grid is not sufficiently regular: the variation of the electrostatic potential are not represented well enough by the values it takes on these grid points.

It can also be interesting to study the effect of describing the electronic charge distribution with charges and dipoles while the solvent is kept the same. Optimized geometries are still in excellent agreement with the reference calculation. At the same geometry, the energy differences are also small: −0.079 kcal/mol for the PNT grid, −0.090 kcal/mol for the MKS grid, and 0.006 kcal/mol for the ‘molecular’ grid. If we compare these values to the previous one, the variations are, respectively, 0.086, 0.128 and 0.149 kcal/mol. The PNT grid seems to give the most stable behavior, this is the main reason why we suggest choose it as the default grid in all future calculations using the ESPF operator.

4. Conclusions

A new electrostatic interaction operator (ESPF) is presented that is able to correctly take into

account the interaction energy between a molecular charge distribution and an arbitrary external potential, that we are supposed to know only on the selected expansion centers, usually the atomic sites. A major feature of this approach is the effective continuity and interchangeability between representations of the quantum system and its external environment. Its implementation in QM/MM codes is straightforward and presents a number of advantages for calculating electrostatic QM/MM interaction terms.

Besides the treatment of periodic environment, where the benefits of the ESPF operator are obvious owing to the possibility of using directly the Madelung potential, the present approach can be promising for the treatment of non-periodical environments too. For point charge environments consisting of only a few hundred to thousand atoms, the computer time requirement of the ESPF calculation is approximately identical to that of the traditional (NEES) one. This is so, because the ESPF method replaces the nuclear attraction integrals over the environment charges in the NEES calculation by the electrostatic potential integrals over the points of the grid and the additional evaluation of the electrostatic potential of the point multipoles of the environment at the expansion centers. As far as the size of the environment increases, the advantages of the ESPF method become more and more pronounced, since the number of quantum chemical integrals remains always constant. If the classical electrostatic potential is treated by some linear scaling technique, like the fast multipole method, it is easy to obtain an overall linear scaling of the computer time with respect to the size of the global QM/MM system.

5. Acknowledgements

The authors thank Drs Christophe Chipot and Xavier Assfeld (Nancy) for useful discussions and Dr. Stéphanie Chalmet (Nancy) for providing the

simulation box of the hydrated formamide molecule.

References

- [1] J. Almlöf, U. Wahlgren, *Theoret. Chim. Acta* 28 (1973) 161.
- [2] J.O. Noell, K. Morokuma, *J. Phys. Chem.* 80 (1976) 2675.
- [3] E.V. Stefanovich, T.N. Truong, *J. Phys. Chem. B* 102 (1998) 3018.
- [4] J.E. Jaffe, A.C. Hess, *J. Chem. Phys.* 105 (1996) 10983.
- [5] H.C. Anderson, S. Hammes-Schiffer, *J. Chem. Phys.* 101 (1994) 375.
- [6] V.R. Saunders, C. Freyria-Fava, R. Dovesi, L. Salasco, C. Roetti, *Mol. Phys.* 77 (1992) 629.
- [7] S.E. Derenzo, M.K. Klintenberg, M.J. Weber, *J. Chem. Phys.* 112 (2000) 2074.
- [8] H.C. Longuet-Higgins, *Proc. Roy. Soc. A* 235 (1956) 537.
- [9] D. Bakowies, W. Thiel, *J. Phys. Chem.* 100 (1996) 10580.
- [10] M.J. Field, P.A. Bash, M. Karplus, *J. Comp. Chem.* 11 (1990) 700.
- [11] A. Warshel, M. Levitt, *J. Mol. Biol.* 103 (1976) 227.
- [12] R.W.F. Bader, *Atoms in Molecules – A Quantum Theory*, University of Oxford Press, Oxford, 1990.
- [13] A.J. Stone, *Chem. Phys. Lett.* 83 (1981) 233.
- [14] W.A. Sokalski, R.A. Poirier, *Chem. Phys. Lett.* 98 (1983) 86.
- [15] A.D. Reed, L.A. Curtiss, F. Weinhold, *Chem. Rev.* 88 (1988) 899.
- [16] M.M. Francl, L.E. Chirlian, *Rev. Comp. Chem* 14 (2000) 1.
- [17] J.G. Ángyán, C. Chipot, *Int. J. Quantum Chem.* 52 (1994) 17.
- [18] C. Chipot, J.G. Ángyán, G.G. Ferenczy, H.A. Scheraga, *J. Phys. Chem.* 97 (1993) 6628.
- [19] P.E. Blöchl, *J. Chem. Phys.* 103 (1995) 7422.
- [20] T.K. Woo, P.E. Blöchl, T. Ziegler, *J. Mol. Struct. Theochem.* 506 (2000) 313.
- [21] E.F. Bertaut, *J. Phys.* 39 (1978) 1331.
- [22] A.J. Stone, *The Theory of Intermolecular Forces*, Oxford University Press, Oxford, 1997.
- [23] C. Hättig, B.A. Heß, *Mol. Phys.* 81 (1994) 813.
- [24] M.J. Frisch et al., *GAUSSIAN 98 Revision A.9*, (1998).
- [25] B.H. Besler, K.M. Merz Jr, P.A. Kollman, *J. Comp. Chem.* 11 (1990) 431.
- [26] U.C. Singh, P.A. Kollman, *J. Comp. Chem.* 5 (1984) 129.
- [27] L.E. Chirlian, M.M. Francl, *J. Comp. Chem.* 8 (1987) 894.
- [28] C.M. Breneman, K.B. Wiberg, *J. Comp. Chem.* 11 (1990) 361.
- [29] C. Chipot, J.G. Ángyán, *Q.C.P.E. Bull.* 14 (1994) 4.