

---

# Remarks on the Use of the Apparent Surface Charges (ASC) Methods in Solvation Problems: Iterative versus Matrix-Inversion Procedures and the Renormalization of the Apparent Charges

---

**R. CAMMI\***

*Dipartimento di Chimica, Università di Parma,  
Viale delle Scienze, 1, 43100 Parma, Italy*

**J. TOMASI**

*Dipartimento di Chimica, Università di Pisa,  
Via del Risorgimento 35, 56100 Pisa, Italy*

*Received 21 March 1994; accepted 13 December 1994*

## ABSTRACT

---

We present a formal and numerical comparison between the iterative and matrix-inversion approaches of the polarizable continuum model. The formal analysis shows completely the equivalence of the two approaches. Numerical equivalence is also recovered, introducing in both methods the proper boundary conditions on the apparent charge distribution. © 1995 by John Wiley & Sons, Inc.

---

## Introduction

**I**n this article we discuss some methodological and computational aspects of the polarizable continuum model (PCM). The PCM is a method to describe solvent effects within the framework of

the representation of the solvent as a continuum medium, using a formulation of the electrostatic problem in terms of apparent surface charges. Since its first formulation in 1981,<sup>1</sup> the PCM has been extended to deal with a large spectrum of chemical applications, also including terms of nonelectrostatic origin and other features like the consideration of nonisotropic continuous media (for a review, see Tomasi et al.<sup>2</sup>). The basic element of the model consists of a method for evaluating the

\* Author to whom all correspondence should be addressed.

electrostatic solute–solvent interaction energy, in a formulation which enables the direct solution of the problem at both the *ab initio* quantum level and the semiempirical and classical levels of formulation of the theory. This formulation, which has greatly contributed to the success of the PCM and of other similar methods in recent years (see Tomasi and Persico<sup>3</sup> for a recent general review), has remained substantially unchanged since 1981.

In this article we reconsider this subject for the following reasons:

1. The original version is not suited to obtain the analytical derivatives of the energy with respect both to internal<sup>4</sup> (e.g., geometry) and to external<sup>5</sup> (e.g., dielectric permittivity) parameters.
2. There have been some misunderstandings of the original method that deserve clarification.<sup>6,7</sup>
3. The evolution of computers makes profitable some changes in computational strategy; in particular, the increase of fast memory has shifted the balance between two alternative strategies, as will be shown later.

The formulation we consider in this article has been used in other continuum methods. It is based on solution of the electrostatic problem in terms of a system of linear equations, written in matrix form.

In the original version of the PCM program, the same problem was solved with the aid of iterative expressions, without using matrices. Both approaches may be considered as applications to the specific problem of solving a mathematical formulation, called the boundary elements method (BEM), which is used in computational engineering and physics.<sup>8</sup>

The matricial BEM formalism has been exposed in detail by the Sakurai group.<sup>6,7</sup> More concise formulations may be found in articles by Rashin and Namboodiri,<sup>9</sup> Drummond,<sup>10</sup> Zahuar and co-workers,<sup>11–14</sup> Wang and Ford,<sup>15</sup> and Grant et al.<sup>16</sup> In many cases, the matrix formulation strictly follows the original PCM formulation, with a few changes (which, in some cases, deserve further comment).

We are going to reformulate this formalism by introducing some simplifications to get a computationally more efficient version, and we implement it for *ab initio* calculations. The matrix-BEM meth-

ods have been used until now for classical or semiempirical calculations,<sup>6,7,10–15</sup> with one exception.<sup>16</sup>

In the next section we report the essential features of the PCM model in order to introduce the notation and make it easier to appreciate the changes. Then we report a reformulation of the matrix-PCM procedure, with some changes with respect to the formulation given by Hoshi et al.<sup>6,7</sup> Later we discuss the relationship between the effective solvent Hamiltonian and the derivation of the free energy functional; then we consider the surface-charge renormalization, a concept which is essential in the analysis of the formalism. Finally, we report some numerical results.

## General Features of the PCM Method

The PCM method is aimed at the simultaneous solution of a quantum mechanical problem, the determination of the solute wave function and a classical electrostatic problem, the evaluation of the interaction energy between solute and solvent, also including mutual polarization effects.

The quantum problem imposes the presence of a cavity, which for electrostatic reasons must be realistic. The boundary condition due to the presence of the cavity must be satisfied with a high degree of accuracy; otherwise the results are of little numerical meaning. To satisfy these stringent requirements, the PCM resorts to the definition of an apparent charge distribution  $\sigma(\mathbf{s})$  spread on the cavity surface  $\Sigma$  and to the corresponding electrostatic potential  $V_\sigma$ , which is the solvent reaction potential. The charge distribution  $\sigma(\mathbf{s})$  (and  $V_\sigma$ ) depends on the solute charge distribution and on the solute geometry. Conversely, the solute charge distribution  $\gamma(\mathbf{r})$  (electrons and nuclei) depends on  $\sigma(\mathbf{s})$ . In fact,  $\sigma(\mathbf{s})$  has the following implicit definition:

$$\sigma(\mathbf{s}) = -\frac{\varepsilon - 1}{4\pi\varepsilon} \mathbf{E} \cdot \hat{\mathbf{n}} = -\frac{\varepsilon - 1}{4\pi\varepsilon} (\mathbf{E}^\gamma + \mathbf{E}^\sigma) \cdot \hat{\mathbf{n}} \quad (1)$$

where  $\mathbf{E} = \mathbf{E}^\gamma + \mathbf{E}^\sigma$  is the total electric field, which has two sources, the solute charge distribution  $\gamma(\mathbf{r})$  and  $\sigma(\mathbf{s})$ , both computed on the inner side of the surface cavity;  $\hat{\mathbf{n}}$  is the outward normal unit vector to the cavity and  $\varepsilon$  is the solvent dielectric constant.

Equation (1) may also be written in terms of directional derivatives of the electrostatic poten-

tial. This formulation is the one we have used in coding the various versions of the PCM program:

$$\sigma(\mathbf{s}) = -\frac{\varepsilon - 1}{4\pi\varepsilon} \mathbf{E} \cdot \hat{\mathbf{n}} = \frac{\varepsilon - 1}{4\pi\varepsilon} (\nabla V_{\text{tot}}(\mathbf{r})) \cdot \hat{\mathbf{n}} \quad (2)$$

with

$$\begin{aligned} V_{\text{tot}}(\mathbf{r}) &= V_{\gamma}(\mathbf{r}) + V_{\sigma}(\mathbf{r}) \\ &= \int \frac{\gamma(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d\mathbf{r}' + \int \frac{\sigma(\mathbf{s})}{|\mathbf{s} - \mathbf{r}|} ds \end{aligned} \quad (3)$$

$V_{\sigma}$  is introduced in the Schroedinger equation of the solute system,  $M$ , as a component of the effective Hamiltonian:

$$H = H^{\circ} + V_{\sigma} \quad (4)$$

The solution of the Schroedinger equation

$$(H^{\circ} + V_{\sigma})|\Psi'\rangle = E'|\Psi'\rangle \quad (5)$$

gives as output the total (electron and nuclei) charge distribution  $\gamma(\mathbf{r})$  on which  $V_{\sigma}$  depends.

To solve this implicit problem, we have adopted, in the original formulation, an iterative procedure based on a discretization of the problem. The boundary conditions imposed by the presence of an irregular cavity make it difficult to get the direct solution of the integral-differential equation [eq. (5)].

We have introduced a finite set of surface charges  $q_i$ , each associated with a small portion (tessera) of the cavity surface  $\Sigma$ , with area  $a_i$ :

$$q_i = \sigma(\mathbf{s}_i)a_i \quad (6)$$

where  $\sigma(\mathbf{s}_i)$  is the value of the apparent charge distribution at a representative point  $\mathbf{s}_i$  within the tessera  $i$ . The reaction potential assumes the discrete form

$$V_{\sigma}(\mathbf{r}) = \sum_i V_{q_i}(\mathbf{r}) = \sum_i q_i^{\text{tot}} \int \frac{\delta(\mathbf{s}_i - \mathbf{s}')}{|\mathbf{s}' - \mathbf{r}|} d\mathbf{s}' \quad (7)$$

Once this discretization is done, a first guess of the  $q_i$ 's is obtained by using  $V_{\gamma}^{\circ}$  as provisional value for  $V_{\text{tot}}$  and then by performing an iterative calculation. At the cycle  $m + 1$  we have

$$\begin{aligned} q_i^{m+1} &= q_i^{\circ} + \kappa_{\varepsilon} q_i^m 2\pi(1 - \xi) + \kappa_{\varepsilon} a_i \frac{1}{|\delta_i|} \sum_{k \neq i} q_k^m \\ &\quad \times (|\mathbf{s}_k - \delta_k - \mathbf{s}_i|^{-1} - |\mathbf{s}_k - \mathbf{s}_i|^{-1}) \end{aligned} \quad (8a)$$

where  $\kappa_{\varepsilon} = \varepsilon - 1/4\pi\varepsilon$  is a factor depending on the dielectric constant and  $\delta_k$  is a vector normal to the surface at  $\mathbf{s}_i$  (pointing inward) introduced to evaluate numerically the directional derivative of  $V_{q_i}$ . The first two terms of eq. (8) give the self-contribution of tessera  $i$  to the polarization of the surface charges. The term  $\xi$  is a correction for the curvature of the convex tessera, introduced in the first article,<sup>1</sup> without a detailed derivation:

$$\xi = \sqrt{a_i/4\pi R_i^2} \quad (8b)$$

where  $R_i$  is the radius of the sphere to which  $a_i$  belongs. The formal derivation of  $\xi$  may be now found in two articles, by Hoshi et al.<sup>6</sup> and by Wang and Ford.<sup>15</sup>

An approximate formula, based on a partial closure of eq. (8a) with  $m \rightarrow \infty$ , computationally more efficient, is the object of a separate article.<sup>17</sup> The final charges  $q_i^f$  thus obtained must satisfy another boundary condition, namely

$$\sum_i q_i^f = -\frac{\varepsilon - 1}{\varepsilon} Q \quad (9)$$

where  $Q$  is the total charge of the solute  $M$ . This condition is satisfied by imposing a renormalization of the  $q_i^f$ 's. The potential corresponding to these charges is introduced in eq. (5), and a modified value of  $\gamma(\mathbf{r})$  is obtained. The cycle continues by applying again eq. (7). We may call this procedure the iterative BEM-PCM method.

An alternative approach is offered by a matrix formulation. Basically it consists of rewriting eqs. (2) and (3) under the form of a set of linear equations in which the unknown is given by the  $q_i$  charges. The iterative-matrix BEM procedure consists of getting a solution of the linear equations at each step of the self-consistent field (SCF) iterative cycle. This procedure has been used by Grant et al.<sup>6</sup> in *ab initio* calculations. In a separate article,<sup>17</sup> we have examined advantages and disadvantages of this approach.

Another version of the matrix formulation is more interesting. Here the iterative procedure by inverting the matrix involves only geometric elements and  $\kappa_{\varepsilon}$ . We shall call this approach matrix-inversion BEM procedure. The real advantage of the matrix-inversion BEM procedure, when properly applied, lies in that it may replace all the iterative SCF cycle with one only Hartree-Fock (HF)-like calculation, as Hoshi et al.<sup>6,7</sup> have shown.

Actually, another SCF calculation is necessary if one has to make a comparison with in vacuo

calculations [e.g., to get the electrostatic component of the solvation free energy  $\Delta G_{\text{el}} = G(\text{sol}) - G(\text{vac})$ ]. Therefore, taking into account that three to four SCF steps are in general sufficient for routine applications and that the matrix-BEM requires the inversion of a large matrix, the balance between the two approaches depends on the size of the system and on the computer configuration.

The formal properties of the matrix-inversion BEM procedure are, however, the starting point to obtain analytical derivatives within the framework of the PCM method. These properties have been exploited in other recent articles,<sup>4,5</sup> on the evaluation of analytical first and second derivatives of  $G(\text{sol})$ .

### A Matrix-Inversion BEM Reformulation of the PCM Approach

The discretization of the continuous charge distribution, introduced via eq. (6), is combined with a separation of the solute charge distribution  $\gamma(r)$  into nuclear and electronic contributions, the latter being expressed as an expansion over a finite atomic orbital (AO) basis set  $\{\chi\}$ :

$$\gamma(\mathbf{r}) = \gamma^N(\mathbf{r}) + \rho(\mathbf{r}) = \sum_m Z_m \delta(\mathbf{R}_m - \mathbf{r}) + \sum_{\mu\nu} P_{\mu\nu} \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r}) \quad (10)$$

By combining eqs. (1), (2), (4), (5), and (10), one obtains a set of linear equations which may be put in the following matrix form:

$$\mathbf{D}\mathbf{A}^{-1}\mathbf{q} = -(\mathbf{E}_n^e + \mathbf{E}_n^N) \quad (11)$$

where  $\mathbf{D}$  is a square matrix, with size equal to the number of the surface tesseras, and elements

$$D_{ii} = \kappa_\epsilon^{-1} - 2\pi(1 - \xi) \quad (12)$$

$$D_{ij} = a_j \frac{(\mathbf{s}_i - \mathbf{s}_j)}{|\mathbf{s}_i - \mathbf{s}_j|^3} \cdot \hat{\mathbf{n}}_i = a_j \frac{\partial}{\partial \hat{\mathbf{n}}_i} \frac{1}{|\mathbf{s}_i - \mathbf{s}_j|} \quad (13)$$

and  $\mathbf{A}$  is a square diagonal matrix (same dimension as  $\mathbf{D}$ ) with elements equal to the tessera areas

$$A_{ii} = a_i \quad (14)$$

The column matrix  $\mathbf{q}$  collects the surface charges.

On the right side of eq. (11) we have two column matrices reporting the values of the normal

components of the solute electric field [partitioned according to eq. (10)] computed at the discrete set of positions  $\{\mathbf{s}_i\}$ . Their elements are the following:

$$E_{\mathbf{n}_i}^e = -\frac{\partial}{\partial \hat{\mathbf{n}}_i} (V_i^e) \quad (15)$$

$$E_{\mathbf{n}_i}^N = -\frac{\partial}{\partial \hat{\mathbf{n}}_i} (V_i^N) \quad (16)$$

where

$$V_i^e = \text{tr} \mathbf{P} \mathbf{V}^i \quad (17)$$

$$V_i^N = \sum_m^N Z_m \frac{1}{|\mathbf{R}_m - \mathbf{s}_i|} \quad (18)$$

and  $\mathbf{P}$  and  $\mathbf{V}$  are square matrices, of the same dimension as the basis set  $\{\chi\}$ . The first is the charge and bond order matrix,  $\mathbf{P} = 2\mathbf{C}\tilde{\mathbf{C}}$  for a closed shell system, and the second has the following elements:

$$V_{\mu\nu}^i = \left\langle \chi_\mu \left| \frac{1}{|\mathbf{r} - \mathbf{s}_i|} \right| \chi_\nu \right\rangle \quad (19)$$

The partition of  $\gamma(r)$  allows a partition of the apparent charges into two components:

$$\mathbf{q} = \mathbf{q}^e + \mathbf{q}^N \quad (20)$$

with elements

$$q_i^e = \sum_j^t - \left[ D_{ij}^{-1} E_{\mathbf{n}_j}^e \right] a_j \quad (21)$$

$$q_i^N = \sum_j^t - \left[ D_{ij}^{-1} E_{\mathbf{n}_j}^N \right] a_j \quad (22)$$

### Expression of the Interaction Potential between Solute Charge Distribution and Surface Charge

The solute-solvent interaction energy

$$U_{\gamma\sigma} = \iint \frac{\gamma(\mathbf{r})\sigma(\mathbf{s})}{|\mathbf{r} - \mathbf{s}|} d\mathbf{r} d\mathbf{s} \quad (23)$$

may be divided into four components:

$$U_{\gamma\sigma} = U_{ee} + U_{eN} + U_{Ne} + U_{NN} \quad (24)$$

where the first index refers to the components of  $\gamma$ , the second one to the component of  $\sigma$ . These

four terms may be written in the following form:

$$U_{ee} = \text{tr} \mathbf{P} \mathbf{X}(\mathbf{P}) \quad (25)$$

where the  $\mathbf{X}$  matrix has as elements:

$$X_{\mu\nu} = \sum_i V_{\mu\nu}^i q_i^e \quad (26)$$

where

$$q_i^e = \sum_{\mu\nu} P_{\mu\nu} q_{\nu\mu}^e(i) \quad (27)$$

with

$$q_{\mu\nu}^e(i) = \sum_j D_{ij}^{-1} \left( \frac{\partial}{\partial \hat{\mathbf{n}}_j} V_{\mu\nu}^j \right) a_i \quad (28)$$

$$U_{eN} = \text{tr} \mathbf{P} \mathbf{J} \quad (29)$$

with

$$J_{\mu\nu} = \sum_i V_{\mu\nu}^i q_i^N \quad (30)$$

$$U_{Ne} = \text{tr} \mathbf{P} \mathbf{Y} \quad (31)$$

with

$$Y_{\mu\nu} = \sum_i V_i^N q_{\mu\nu}^e(i) \quad (32)$$

$$U_{NN} = \sum_i V_i^N q_i^N \quad (33)$$

In the next section we shall make use of the following relationship:

$$\langle \Psi' | V_\sigma | \Psi' \rangle = U_{ee} + U_{eN} \quad (34)$$

## The Expression of the Free Energy $G$ in Terms of the HF Equations

The electrostatic component of the Gibbs free energy of the solute-solvent system, at fixed position of the solute nuclei, referred both to noninteracting electrons, nuclei, and to the unpolarized dielectric, may be put in the following form<sup>18,19</sup>:

$$G(\text{sol}) = \langle \Psi' | H^0 | \Psi' \rangle + \frac{1}{2} U_{\gamma\sigma} + V_{NN} \quad (35a)$$

or

$$G(\text{sol}) = \langle \Psi' | H^0 | \Psi' \rangle + \frac{1}{2} \langle \Psi' | V_\sigma | \Psi' \rangle + \frac{1}{2} (U_{Ne} + U_{NN}) + V_{NN} \quad (35b)$$

where  $V_{NN}$  is the nuclear repulsion energy of the solute  $M$ . To expedite notation, we shall indicate in the following  $G(\text{sol})$  with the simple notation  $G$ .

Using the definitions given in the previous section and also using the usual matrix expression of the electronic energy *in vacuo*  $E = \langle \Psi' | H^0 | \Psi' \rangle = \text{tr} \mathbf{P} \mathbf{H} + \frac{1}{2} \text{tr} \mathbf{P} \mathbf{G}$ , we have

$$G = [\text{tr} \mathbf{P} \mathbf{H} + \frac{1}{2} \text{tr} \mathbf{P} \mathbf{G}(\mathbf{P})] + \frac{1}{2} [\text{tr} \mathbf{P} \mathbf{X}(\mathbf{P}) + \text{tr} \mathbf{P} \mathbf{J} + \text{tr} \mathbf{P} \mathbf{Y} + U_{NN}] + V_{NN} \quad (36)$$

which may be rewritten in a more condensed form as

$$G = \text{tr} \mathbf{P} \mathbf{h}' + \frac{1}{2} \text{tr} \mathbf{P} \mathbf{G}'(\mathbf{P}) + V'_{NN} \quad (37)$$

with

$$\mathbf{h}' = \mathbf{h} + \frac{1}{2}(\mathbf{J} + \mathbf{Y}) \quad (38)$$

$$\mathbf{G}'(\mathbf{P}) = \mathbf{G}(\mathbf{P}) + \mathbf{X}(\mathbf{P}) \quad (39)$$

$$V'_{NN} = V_{NN} + \frac{1}{2} U_{NN} \quad (40)$$

The  $\mathbf{J}$  and  $\mathbf{Y}$  matrices are the  $\mathbf{J}_1$  and  $\mathbf{J}_2$  matrices in the Hoshi et al.<sup>6</sup> formulation. Though of different physical meaning, they are numerically identical, when the renormalizations of the set  $\mathbf{q}^N$  and  $\mathbf{q}^e$  charges [see eq. (9)] are properly introduced. A formal derivation of their equivalence may be obtained by exploiting the symmetry of the Green operator (see Blaive and Metzger,<sup>20</sup> Constanciel,<sup>21</sup> and Newton and Friedman<sup>22</sup>).

By combining the stationarity condition  $\delta G = 0$  for an arbitrary variation of the molecular orbital (MO) coefficients with the auxiliary conditions of orthonormality, one easily arrives at the matrix HF-like equation:

$$\mathbf{F}' \mathbf{C} = \mathbf{SCE}' \quad (41)$$

with

$$\mathbf{F}' = \mathbf{h}' + \mathbf{G}' \quad (42)$$

The extension to open-shell cases is immediate.

The  $G$  functional for which we have a stationary point via eq. (41) corresponds to the variational functional  $J$  of Sanhueza et al.,<sup>23</sup> which has been introduced to obtain a variational solution of the Schroedinger equation

$$(H^0 + V(\Psi')) |\Psi'\rangle = E' |\Psi'\rangle \quad (43)$$

when the operator  $V(\Psi')$  may be put in the form  $\hat{A}(\Psi'^* \Psi')^m$ , with  $m = 1$ , as is the case for the solvent reaction potential operators. Therefore, the variational solution of eq. (43), at the HF level and

in the fixed nuclei approximation, is done by eq. (41) (see also Tapia<sup>24</sup>).

The situation is different in the iterative PCM process. Here we start again from the Schroedinger equation [eq. (5)]:

$$(H^o + V_\sigma)|\Psi'\rangle = E'|\Psi'\rangle \quad (44)$$

but  $V_\sigma$  is fixed at each step of the iterative procedure. The usual expression of the variational energy functional is valid, and the procedure gives the free energy under the following form:

$$G = \langle \Psi' | H^o + V_\sigma | \Psi' \rangle - \frac{1}{2} \langle \Psi' | V_\sigma | \Psi' \rangle + \frac{1}{2} U_{Ne} + V'_{NN} \quad (45)$$

which is exactly the expression given in eq. (35). The iterative procedure and the one passing through eq. (41) give the same numerical results when the other parameters of the calculations are kept constant. This point has been misunderstood by some authors, and we consider it necessary to provide numerical evidence of the complete equivalence of the two approaches (see Numerical Results and Discussion).

## Surface Charges Renormalization

During the calculation with the PCM, both in the iterative and in the matrix-BEM formulation, it frequently happens that the total surface charge  $Q_\sigma^{\text{tot}}$  does not satisfy a basic requirement of the electrostatic model, namely that

$$Q_\sigma^{\text{tot}}(\text{th.}) = \int_\Sigma \sigma(\mathbf{s}) d\mathbf{s} = \sum_i q_i^{\text{tot}} = -\frac{\varepsilon - 1}{\varepsilon} Q_M \quad (46)$$

where  $Q_M$  is the charge of the solute and  $Q_\sigma^{\text{tot}}(\text{th.})$  is the theoretical value of the apparent charge distribution  $\sigma$ . We have already mentioned that an empirical correction was introduced in the original PCM formulation [see eq. (9)]. We may now proceed a step further.

It is easy to show that analogous relations must apply to the separate nuclear and electronic components:

$$Q_\sigma^N(\text{th.}) = \sum_i q_i^N = -\frac{\varepsilon - 1}{\varepsilon} Q_M^N \quad (47)$$

$$Q_\sigma^e(\text{th.}) = \sum_i q_i^e = -\frac{\varepsilon - 1}{\varepsilon} Q_M^e \quad (48)$$

where  $Q_M^N$  and  $Q_M^e$  are the nuclear and electronic charge of  $M$ . These relations are not respected in actual calculations for two reasons. The first has a numerical origin, because in the PCM we have replaced an integration with a finite summation assuming  $\sigma$  constant within each tessera, and because the evaluation of  $q_i^x$  ( $x = N$  or  $e$ ) proceeds via a finite difference calculation of the  $E_n^x$  components [see eq. (8)]. The second reason has a physical origin. The electronic distribution  $\rho(\mathbf{r})$  has tails spreading out to infinity, and the cavity surface cuts out the outer portions of  $\rho(\mathbf{r})$ . The first effect acts on both components  $Q_\sigma^N$  and  $Q_\sigma^e$ , the second on  $Q_\sigma^e$  alone.

The simplest way to renormalize surface charges consists of introducing two separate factors,  $f^N$  and  $f^e$ , to satisfy eqs. (46) and (47):

$$f^N Q_\sigma^N(\text{calc.}) = -\frac{\varepsilon - 1}{\varepsilon} Q_M^N \quad (49)$$

$$f^e Q_\sigma^e(\text{calc.}) = -\frac{\varepsilon - 1}{\varepsilon} Q_M^e \quad (50)$$

where  $Q_\sigma^x(\text{calc.})$  ( $x = e, N$ ) is the calculated value of the apparent charge distribution  $\sigma^x$ . The factor  $f^e$  may be rewritten as

$$f^e = f^t f^N \quad (51)$$

Numerical tests performed on various systems (also with a complex cavity shape and  $Q_M = \pm 2$ ) show that  $f^t > 1$  always (see Numerical Results and Discussion).

We now consider different options for the surface charge renormalization. In the first option, the sets of charges  $q_i^N$  and  $q_i^e$  are multiplied by the two factors  $f^N$  and  $f^e$ , respectively. The free energy functional [eq. (36)] with the explicit introduction of the two factors  $f^N$  and  $f^e$  assumes the following form:

$$G = \text{tr} \tilde{\mathbf{P}} \tilde{\mathbf{h}}' + \frac{1}{2} \text{tr} \tilde{\mathbf{P}} \tilde{\mathbf{G}}'(\mathbf{P}) + \tilde{V}'_{NN} \quad (52)$$

where

$$\tilde{\mathbf{h}}' = \mathbf{h} + \frac{1}{2}(f^N \mathbf{J} + f^e \mathbf{Y}) \quad (53)$$

$$\tilde{\mathbf{G}}'(\mathbf{P}) = \mathbf{G}(\mathbf{P}) + f^e \mathbf{X}(\mathbf{P}) \quad (54)$$

$$\tilde{V}'_{NN} = V_{NN} + \frac{1}{2} \tilde{U}_{NN} \quad (55)$$

with

$$\tilde{U}_{NN} = f^N \sum_i V_i^N q_i^N \quad (56)$$

The corresponding Fock equation will be related to the following Fock matrix:

$$\tilde{\mathbf{F}}' = \tilde{\mathbf{h}}' + \tilde{\mathbf{G}}' \quad (57)$$

An alternative option exploits the formal equivalence between  $\mathbf{J}$  and  $\mathbf{Y}$  matrices. We rewrite  $\tilde{\mathbf{h}}'$  [eq. (53)] in the following form:

$$\tilde{\mathbf{h}}' = \mathbf{h} + f^N \mathbf{J} \quad (58)$$

The factor  $f^N$  can be computed before performing any SCF calculation because it does not depend on  $\mathbf{P}$ . The  $f^e$  factor is still present in the expression of  $\tilde{\mathbf{G}}'$  [eq. (54)]. A third computationally more effective option consists of adopting the device proposed in the original PCM article.<sup>1</sup> The renormalization procedure is applied by looking at each surface tessera separately, taking into account the correspondent total charge  $q_i = q_i^e + q_i^N$ . If we call  $q_i^+$  and  $q_i^-$  two surface charges having, respectively, positive ( $q_i > 0$ ) and negative ( $q_i < 0$ ) signs, the renormalization procedure may be written in the following form:

$$\tilde{q}_i^e = q_i^e(\text{calc.}) + \frac{1}{2}g_+ q_i^+ \quad (59a)$$

$$\tilde{q}_i^e = q_i^e(\text{calc.}) + \frac{1}{2}g_- q_i^- \quad (59b)$$

with

$$g_{\pm} = \frac{Q_{\sigma}^{\text{tot}}(\text{th.}) - Q_{\sigma}^{\text{tot}}(\text{calc.})}{Q_{\pm}(\text{calc.})} \quad (60)$$

where  $Q_{\sigma}^{\text{tot}}(\text{calc.}) = Q_{\sigma}^N(\text{calc.}) + Q_{\sigma}^e(\text{calc.})$  and  $Q_{\pm}(\text{calc.})$  is the sum of the positive (or negative) surface charges before renormalization ( $Q_{\pm} = \sum_i q_i^{\pm}$ ). The  $\tilde{q}_i^e$  charges are then introduced in eq. (26) to have a better description of the renormalization effects on the  $\tilde{\mathbf{X}}$  matrix (and then of  $\tilde{\mathbf{G}}$ ).

More refined renormalization procedures could be introduced. For instance, instead of using a

unique  $f^e = f^i f^N$  factor for all the  $q_i^e$  charges, a distinct  $f_i^i$  factor for each tessera, based on the local evaluation of the electronic tail for that tessera, could be introduced, or a direct correction for each  $q_{\mu\nu}^e(i)$  value could be evaluated. These refinements have not been introduced in the routine version of the program. The elaboration of a computationally efficient strategy will be considered in a further stage of our research.

## Numerical Results and Discussion

In Table I some numerical results for  $\text{H}_2\text{CO}$  obtained using the iterative and the matrix-inversion BEM-PCM procedures are compared. The calculations refer to a cavity formed by the union of four spheres centered on the nuclei with radii  $R_x = 1.2R_{\text{vdw}}$ . The dielectric constant is  $\epsilon = 78.50$ . In the iterative formulation, we have performed six steps, starting from a previous SCF calculation in vacuo. We have reported the total number of iterations (cycles) in the SCF procedure necessary to reach a convergence in the density matrix within  $10^{-5}$ . The table reports values of  $G$ ,  $\Delta G_{\text{el}} = G(\text{sol}) - G(\text{vac})$ , the dipole moment, and the Mulliken charges as representative values of energetic and structural interest.

For the matrix-inversion procedure, we report the number of SCF iterations in the unique step, starting again from a SCF calculation in vacuo. The matrix inversion calculations have been performed with the computational procedure (d) defined later in this section.

The central processing unit (CPU) calculation times on an IBM RISC 6000/560 are the following for the iterative formulation: 24.0 s (3-21G); 46.5 s (6-31G\*\*). For the matrix inversion, they are 30.0 s (3-21G), 57.4 s (6-31G\*\*). The calculations were

**TABLE I.** Comparison between the Iterative and Matrix-inversion BEM-PCM for  $\text{H}_2\text{CO}$  (at Experimental Geometry) within a Cavity Made by Sphere with Radii  $R_x = 1.2R_{\text{vdw}}$  and Surrounding by a Medium with Dielectric Constant  $\epsilon = 78.5$ .

Basis Set	Method	Steps	SCF Cycles	G (au)	$\Delta G_{\text{el}}$ (kcal/mol)	$\mu$ (Debye)	$Q_C$	$Q_O$	$Q_H$
3-21G	Iterative	6	22	-113.225337	-6.226	3.319	5.857	8.575	0.784
	Matrix inversion	—	12	-113.225339	-6.227	3.320	5.857	8.575	0.784
6-31G**	Iterative	6	22	-113.874393	-6.387	3.491	5.751	8.540	0.855
	Matrix inversion	—	13	-113.874397	-6.389	3.489	5.735	8.543	0.862

Numerical result refers to the total free energy,  $G$ , to the solvation free energy,  $\Delta G_{\text{el}}$ , to the dipole moment,  $\mu$ , and to the Mulliken charges,  $Q_x$  ( $x = C, O, H$ ).

performed by using two different quantum-chemical packages: MONSTERGAUSS-85<sup>25</sup> for the iterative calculations and HONDO-8<sup>26</sup> for the matrix-inversion calculations. The use of different packages affects the comparison of computing times—for example, the vacuo SCF calculation with the 6-31G\*\* basis set (14 iterations in both cases) requires 10.7 s on MONSTERGAUSS and 7.7 in HONDO-8. In this case, the relative computational efficiency  $t(\text{iter})/t(\text{matrix inversion})$  is near to 1/1.7. Table I gives numerical evidence to the complete equivalence of the two approaches when an appropriate apparent charge renormalization is performed.

Table II shows a selection of  $f^N$ ,  $f^e$ ,  $f^t = f^e/f^N$ , and  $\text{trPJ}/\text{trPY} = U_{eN}/U_{Ne}$  values, computed with different basis sets and by using cavities defined as the union of overlapping spheres with radii  $R_k = 1.2R_{\text{vdw}}(k)$ . The data show that  $f^t > 1$  and that the effect of the basis set on  $f^e$  is not negligible. Of course, the absolute values of these correction factors strongly depend on the volume of the cavity. To show this effect more clearly, we have reported in Table III the  $f^N$ ,  $f^e$ , and  $f^t$  values referred to a spherical cavity. The first line of each column corresponds to a cavity having the same value as the corresponding entry of Table II. The comparison of these values in the two tables shows that the shape of the cavity has the same effects on the renormalization factors: We recall that the cavity shape effect is larger on the evaluation of the solute-solvent interaction energy and related properties.

The comparison between the  $\text{trPJ}/\text{trPY}$  and  $f^e/f^N$  ratios is reported in Tables II and III. As stated earlier, the formal equivalence between **J** and **Y** matrices is partially lost because of the different violation of the boundary condition on the  $Q_M^N$  and  $Q_M^e$  total surface charges.

The occurrence of  $f^N \neq 1$  values mainly depends on two reasons. The first one is the use of a finite difference method to evaluate the normal component of the electric field on the cavity surface, which is here described by the distance increment  $\delta$ :

$$\mathbf{E}^N \cdot \mathbf{n} \cong \sum_{\alpha} Z_{\alpha} \left[ \frac{1}{|\mathbf{R}_{\alpha} - \mathbf{s}_i - \delta|} - \frac{1}{|\mathbf{R}_{\alpha} - \mathbf{s}_i|} \right]$$

The second one is related to the essence of the BEM method (i.e., to the assumption that the sampling point faithfully reproduces the function in the selected surface element). A critical point lies in the definition of the diagonal elements of the **D** matrix [eq. (12)] and in particular in the definition of the correction factor  $\xi$  [eq. (8b)]. When working out the original article,<sup>1</sup> we performed a formal analysis of this point, with the conclusion that the leading term in an expansion over the solid angle  $\eta$  subtending the tessera is of the order of  $-\frac{1}{4}\eta^3$ . This correction will be introduced, if we consider it suitable, in further versions of the method. Numerical tests performed on the nuclear component alone show that these two factors are of comparable importance in determining the occurrence of  $f^N \neq 1$ .

As a final numerical example, we show in Table IV the results obtained for the same systems with different versions of the matrix-inversion BEM procedure, characterized by

- the use of eq. (36) of the free energy functional, without introducing any renormalization factors, and with the explicit use of **J** and **Y** matrices; the Fock matrix is given by eq. (42);
- the use of eq. (52) of the free energy functional with the explicit use of  $f^e$  and  $f^N$

**TABLE II.** Numerical Values of the Renormalization Factors  $f^N$  and  $f^e$  and Comparison of Their Ratio  $f^e/f^N$  with the Ratio between the Energetical Quantity  $U_{eN}$  and  $U_{Ne}$ :  $U_{eN}/U_{Ne}$ .

Solute	$f^N$	$f^e$	STO-3G		$f^e$	3-21G		$f^e$	6-31G**	
			$f^e/f^N$	$U_{eN}/U_{Ne}$		$f^e/f^N$	$U_{eN}/U_{Ne}$		$f^e/f^N$	$U_{eN}/U_{Ne}$
HF	1.001986	1.003600	1.001613	1.001349	1.003201	1.001212	1.001101	1.005457	1.003463	1.003154
H <sub>2</sub> O	1.011762	1.014823	1.003025	1.002677	1.015435	1.003630	1.003333	1.020530	1.008666	1.007923
N <sub>2</sub> H <sub>2</sub>	1.018112	1.020731	1.001983	1.001803	1.022531	1.003749	1.003385	1.033043	1.014068	1.012118
NO <sub>2</sub> <sup>-</sup>	1.0119117	1.012345	1.000424	1.000434	1.013308	1.001341	1.001376	1.018461	1.005999	1.006468

The data are relative to several solutes, with the cavity made by the union of sphere centered on the atoms, with radii  $R_x = 1.2R_{\text{vdw}}$  and with the dielectric constant of the medium  $\epsilon = 78.5$ .



TABLE III.

Numerical Values of the Renormalization Factors  $f^N$  and  $f^e$  and Comparison of Their Ratio  $f^e/f^N$  with the Ratio between the Energetical Quantity  $U_{eN}$  and  $U_{Ne}$ :  $U_{eN}/U_{Ne}$ .

Solute	$R$ (Ang.)	STO-3G				3-21G				6-31G**		
		$f^N$	$f^e$	$f^e/f^N$	$U_{eN}/U_{Ne}$	$f^e$	$f^e/f^N$	$U_{eN}/U_{Ne}$	$f^e$	$f^e/f^N$	$U_{eN}/U_{Ne}$	
HF	1.739	0.99548	0.99740	1.00193	1.00167	0.99766	1.00219	1.002321	1.00122	1.00577	1.00615	
	2.087	0.99552	0.99587	1.00036	1.00028	0.99569	1.00017	1.000016	0.99605	1.00053	1.00055	
	2.435	0.99554	0.99562	1.00007	1.00006	0.99557	1.00002	1.00002	0.99559	1.00005	1.00005	
H <sub>2</sub> O	1.878	0.99552	0.99981	1.00431	1.00377	1.00185	1.00636	1.00563	1.01113	1.01568	1.01369	
	2.254	0.99553	0.99648	1.00095	1.00084	0.99618	1.00065	1.00058	0.99669	1.00116	1.00107	
	2.629	0.99556	0.99574	1.00019	1.00018	0.99565	1.00010	1.00009	0.99569	1.00014	1.00013	
N <sub>2</sub> H <sub>2</sub>	2.225	0.99630	1.00163	1.00534	1.00510	1.02855	1.03237	1.03000	1.09477	1.09883	1.08884	
	2.670	0.99572	0.99645	1.00073	1.00067	0.99652	1.00080	1.00077	0.99791	1.00219	1.00209	
	2.115	0.99562	0.99571	1.00009	1.00009	0.99567	1.00005	1.00005	0.99574	1.00013	1.00012	
NO <sub>2</sub> <sup>-</sup>	2.167	0.99606	0.99991	1.00386	1.00410	1.00689	1.01087	1.01149	1.04991	1.05406	1.05450	
	2.600	0.99564	0.99588	1.00019	1.00021	0.99616	1.00047	1.00050	0.99780	1.00211	1.00218	
	3.038	0.99561	0.99561	1.00000	1.00000	0.99562	1.00000	1.00001	0.99572	1.00011	1.00011	

The data are relative to several solutes, with spherical cavities, of radius  $R$ , and with dielectric constant of the medium  $\epsilon = 78.5$ .

renormalization factors [eqs. (53) and (54)]; the Fock matrix is given by eq. (57);

(c) the use of eq. (58) for the  $\tilde{\mathcal{H}}'$  term of the Fock matrix [eq. (57)], without using of the  $\mathbf{Y}$  matrix and with the use of the second option of the renormalization;

(d) the use of eq. (58) for the  $\tilde{\mathcal{H}}'$  term of the Fock matrix [eq. (57)] and of the third option of the renormalization. Surface charges appearing in  $\mathbf{X}$  are modified according to the rule given in eqs. (59a) and (59b); and  $f^N$ ,

used to define  $\tilde{\mathbf{h}}$  [eq. (58)], is computed as in eq. (49).

In both versions (c) and (d) for the calculations of the free energy we explicitly use both  $\mathcal{J}$  and  $\mathcal{Y}$  matrices, computed with the appropriate renormalized surface charges.

The numerical results indicate a remarkable dependence on the renormalization procedure. Without renormalization, case (a), the results are unacceptable. For example, the 6-31G\*\* value of the

TABLE IV.

Comparison of the Four Versions of the BEM-PCM for H<sub>2</sub>O (at Experimental Geometry) with a Standard Cavity ( $R_x = 1.2R_{vdw}$ ) and  $\epsilon = 78.5$ .

Basis Set	Method	$G$ (au)	$\Delta G_{el}$ (kcal/mol)	$\mu$ (Debye)	$Q_O$	$Q_H$
STO-3G	(a)	-74.968897	-3.670	1.70	8.35	0.83
	(b)	-74.969082	-3.786	1.70	8.35	0.83
	(c)	-74.969717	-4.184	1.82	8.40	0.80
	(d)	-74.969694	-4.170	1.82	8.40	0.80
3-21G	(a)	-75.590720	-3.324	2.27	8.67	0.67
	(b)	-75.590856	-3.409	2.27	8.66	0.67
	(c)	-75.598671	-8.917	2.67	8.80	0.60
	(d)	-75.598627	-8.317	2.67	8.80	0.60
6-31G**	(a)	-76.014283	5.576	1.94	8.47	0.76
	(b)	-76.013831	5.826	1.93	8.47	0.77
	(c)	-76.034343	-8.853	2.45	8.71	0.64
	(d)	-76.034220	-7.199	2.45	8.71	0.64

Numerical results refers to the total free energy,  $G$ , to the solvation free energy,  $\Delta G_{el}$ , to the dipole moment,  $\mu$ , and to the Mulliken charges,  $Q_x$  ( $x = O, H$ ).

solvation energy,  $\Delta G_{el}$ , is positive, in contrast with physical intuition. The computationally simple renormalization procedure considered in case (b) is ineffective.

The renormalization procedures used in cases (c) and (d) are practically equivalent. We have already shown (see Table I) that procedure (d) gives exactly the same results as the iterative PCM procedure. The good predictive value of this last procedure has been amply tested.<sup>5</sup> This strong dependence of the values on the renormalization is specific of the matrix-inversion BEM method, and it has not been found in the iterative PCM procedure or in a recent matrix-BEM iterative procedure.<sup>17</sup> The reason for this difference in sensitivity with respect to renormalization is probably that quantum problem is solved here in a single step, while in iterative methods the same problem is solved by using at each step a fixed, and renormalized, set of apparent charges.

An accurate selection of the renormalization algorithm is the price to be paid to get a direct minimization of the free energy functional, a step necessary to get analytical derivatives. We consider the problem of getting a satisfactory renormalization procedure not yet completely solved. The good agreement between procedures (c) and (d) indicates that future refinements will be minor.

## Acknowledgments

The authors acknowledge the 'Ministero dell'Università e della Ricerca Scientifica e Tecnologica' (MURST), the 'Consiglio Nazionale delle Ricerche' (CNR), and the European Economic Community (EEC) for financial support.

## References

1. S. Miertus, E. Scrocco, and J. Tomasi, *Chem. Phys.*, **55**, 117 (1981).
2. J. Tomasi, R. Bonaccorsi, R. Cammi, and F. J. Olivares del Valle, *J. Mol. Struct. (Theochem)*, **234**, 401 (1991).
3. J. Tomasi and M. Persico, *Chem. Rev.*, **94**, 2027 (1994).
4. R. Cammi and J. Tomasi, *J. Chem. Phys.*, **100**, 7495 (1994).
5. R. Cammi and J. Tomasi, *J. Chem. Phys.*, **101**, 3888 (1994).
6. H. Hoshi, M. Sakurai, Y. Inoue, and R. Chūjō, *J. Chem. Phys.*, **87**, 1107 (1987).
7. H. Hoshi, M. Sakurai, Y. Inoue, and R. Chūjō, *J. Mol. Struct. (Theochem)*, **180**, 267 (1988).
8. D. E. Beskos, Ed., *Boundary Methods in Mechanics*, North-Holland, Amsterdam, 1987.
9. A. A. Rashin, and K. Namboodiri, *J. Chem. Phys.*, **91**, 6003 (1987).
10. L. Drummond, *J. Chem. Phys.*, **88**, 5014 (1988).
11. R. J. Zahuar and R. S. Morgan, *J. Comp. Chem.*, **9**, 171 (1988).
12. R. J. Zahuar and R. S. Morgan, *J. Comp. Chem.*, **11**, 603 (1990).
13. R. J. Zahuar and R. S. Morgan, *J. Comp. Chem.*, **12**, 575 (1991).
14. T. Fox, N. Rosch, and R. J. Zahuar, *J. Comp. Chem.*, **14**, 253 (1993).
15. B. Wang and G. Ford, *J. Chem. Phys.*, **97**, 4162 (1992).
16. J. A. Grant, R. L. Williams, and H. A. Scheraga, *Biopolym.*, **30**, 929 (1990).
17. R. Cammi, L. Coitino, and J. Tomasi, *J. Comp. Chem.*, **16**, 20 (1995).
18. S. Yomosa, *J. Phys. Soc. Jap.*, **36**, 1655 (1974).
19. S. Miertus and J. Tomasi, *Chem. Phys.*, **65**, 239 (1982).
20. B. Blaive and J. Metzger, *Phys. Rev.*, **A30**, 1600 (1984).
21. R. Costanciel, *Theor. Chim. Acta*, **69**, 503 (1986).
22. M. D. Newton and H. L. Friedman, *J. Chem. Phys.*, **88**, 4460 (1988).
23. J. E. Sanhueza, O. Tapia, W. G. Laidlaw, and M. Trsic, *J. Chem. Phys.*, **70**, 3096 (1979).
24. O. Tapia, In *Molecular Interactions*, Vol. 3, H. Ratajczak and W. J. Orville-Thomas, Eds, John Wiley & Sons, New York, 1982, p. 47.
25. M. Peterson and R. Poirier, MONSTERGAUSS, Department of Chemistry, University of Toronto, Toronto, Ontario, Canada.
26. M. Dupuis, A. Farazdel, S. P. Karna, and S. A. Maluenes, HONDO-8, IBM Corporation, Scientific and Engineering Computations, Kingston, NY.