A non-local representation of the effective potential due to a molecular fragment

Renato Colle[1,2], Alessandro Curioni[2], and Oriano Salvetti[3]

[1] Dipartimento di Chimica Applicata, Facoltá di Ingegneria, Universitá di Bologna, Viale Risorgimento 2, I-40136 Bologna, Italy

[2] Scuola Normale Superiore, I-56100 Pisa, Italy

[3] Istituto di Chimica Quantistica ed Energetica Molecolare del CNR, I-56100 Pisa, Italy

Summary. A non-local representation of the effective potential due to a molecular fragment is proposed here. Using this technique one can reproduce both Coulomb and exchange operators with kernels made up by molecular orbitals localized on a given molecular fragment. Such an approach seems particularly effective for large molecules with well-defined chemical fragments since in this case the kernel orbitals can be prepared through separate calculations on each fragment. The performance of the method is illustrated through calculations on specific molecular examples.

Key words: Non-local representation – Coulomb operator – Molecular fragments

1 Introduction

The ability to predict the electronic structure of an extended chemical system in a quantitative way, as a function of its geometrical parameters, represents an ultimate goal in quantum chemistry. A way to approach this goal – without introducing empirical parameters, but still reducing drastically the dimension of the computational effort – is based on the observation that many of the chemical properties of a molecule depend on a few "valence" electrons moving in the effective field of the nuclei and of the other "core" electrons. Furthermore, a typical large system can be described as an aggregate of molecular fragments or functional groups such that the effective field, which governs the motion of the few "valence" electrons, consists of the sum of the fields due to the various fragments.

 These ideas are interesting from a computational point of view only if two main conditions can be satisfied. The first is the ability to obtain transferable quantities and, in particular, transferable potentials for each of the fragments that can be consistently defined inside the extended system. The second is the ability to

represent these effective potentials in a simple, economical, but sufficiently accurate way.

In regard to the first point, there are many studies ranging from small molecules to large systems [1–2] on the transferability of localized molecular orbitals and Fock matrix elements. The main requirements of this process that can be considered the natural extension of the "frozen core" approximation are *1*) a consistent definition of the fragments inside the large system – a definition that must reflect the nature and the structure of bonds and lone pairs – and *2*) an efficient method of localizing the orbitals obtained from calculations on the separate fragments. This last problem can be easily solved by using one of the various methods [3–7] proposed for localizing the canonical orbitals through unitary transformations. Note that the resulting orbitals can also be used as an efficient tool for performing large CI calculations since they allow one to reduce the steep dependence of the computational effort on the molecular size [8–10].

As for the problem of reproducing the effective potential which governs the behaviour of the valence electrons, we observe that there is a large quantity of literature – quoted under the names of "Pseudopotentials" [11–16] or "effective/pseudo/model Hamiltonians" [17–29], see Refs. [30, 31] for a general overview – which is based on the assumption that the valence electrons move inside the effective field produced by non-overlapping, spherically symmetric charge distributions centred around the atomic nuclei. The resulting effective potential is the sum of atomic contributions: $\hat{V}_{eff} = \Sigma_A \hat{V}_A$, each one representing, at a level of complexity that depends on the chosen functional form, an atomic core potential with its various components (Coulomb, exchange . . .).

Different representations have been proposed for these effective potentials. The simplest is a local form of the type $\hat{V}_A = V_A(r)$, where $V_A(r)$ is usually a linear combination of gaussian or error functions modulated by powers of the radial coordinate. This form, which represents a crude approximation if applied to the entire potential [13], is widely used for the Coulomb part of the potential [20–25]. More sophisticated representations are based on semi-local forms [11–16], of the type $\hat{V}_A = \Sigma_{lm}|lm; A\rangle V_l^A(r)\langle lm; A|$, or non-local forms [17–19] such as $\hat{V}_A = \Sigma_{pq}|g_{pA}\rangle C_{pq}^A \langle g_{qA}|$, where $\{|lm; A\rangle\}$ and $\{|g_{pA}\rangle\}$ are respectively harmonic and gaussians functions centred on the atom A.

Finally, in the most complex model Hamiltonians [20–24] different representations are used for the different parts of the atomic core potential, more specifically, a local representation is used for the Coulomb part and a non-local one for the exchange part as well as for the pseudopotential owing to the orthogonality constraints between core and valence oribitals. Note that, although characterized by different functional forms, these pseudopotentials are in general parametrized from *ab initio* calculations. This means that the exponents and the expansion coefficients, which characterize the chosen functional form, are determined through a least-square fitting to *ab initio* atomic core potentials.

Although the results obtained using these model Hamiltonians are very often quite significant even in medium size molecules, the fact that the proposed effective potentials are sums of atomic core contributions represents a severe limitation since it does not allow one to take advantage of clear chemical information. Thus, for example, in a conjugated organic polymer one knows that the nature and the structure of the sigma skeleton is practically the same along the chain and, furthermore, very similar to that in the monomeric unit. Therefore, if one is interested in properties which depend on the π-electrons only, the most natural

choice is to derive the corresponding orbitals as eigenfunctions of a Fock operator in which the effective potential is that produced by the whole sigma skeleton given as the sum of identical monomeric contributions. An accurate and simple representation of such a potential cannot be obtained through a superposition of atomic core potentials, but it requires a different and strictly molecular type of approach. The ability to obtain an accurate and simple representation of the effective potential due to each monomeric unit (i.e. to a well-defined molecular fragment) would therefore represent a relevant improvement in the definition and the usefulness of model Hamiltonians since it would allow one to construct efficiently the whole effective potential of any large chemical system.

The idea of introducing "group potentials" in molecular calculations is strictly related to that of partitioning the electrons into groups, an idea which was formally developed by McWeeny [32, 33] in the so-called "Generalized Product Function" approach. This idea is also present in a paper of Morokuma et al. [34] where the two lone pair electrons of the NH_3 molecule are treated in the effective potential of the N_{1s} core and of the three N–H bonds. This effective potential, however, is represented in its Coulomb part as a sum of atomic contributions.

More recently – see Ref. [35] – the authors have proposed a simple method for representing an exchange operator whose kernel consists of a set of molecular orbitals localized on a well-defined chemical fragment. This kernel is represented as a non-local form consisting of a weighted sum of projection operators constructed from a set of orthonormal orbitals, each of which is centred on the fragment itself and weighted by the mean value of the exchange operator on that orbital. The quality of this representation depends on the number and type of functions used in the expansion and can be improved, should the problem require it, simply by extending the basis set.

The central observation on which this method is based is the following: quantum mechanical calculations, based on the LCAO method, require only the knowledge of the representation of the total Hamiltonian inside the space of the Slater determinants that can be built up using the chosen set of basis functions. This means that one needs to know the spectral resolution of a few important operators (such as the Coulomb or the exchange operator) only inside the space defined by these functions.

In molecular calculations the basis sets are in general polycentric which thereby implies that the matrix elements representing the operators are defined also between functions centred on different fragments. However, if the operator is "short range" – i.e. different from zero only inside a finite volume, such as the exchange operator – the absolute value of its matrix elements decreases rapidly as the distance between the functions that are implied in the matrix elements increases. It follows that one can reasonably represent the kernel of each operator using only a reduced number of functions centred on the fragment where the kernel orbitals are localized.

The aim of this paper is to show that a representation, similar to that proposed by us in Ref. [35] for the exchange operator, can be used also for the "short-range" part of a Coulomb operator the kernel of which consists of molecular orbitals localized on a well defined chemical fragment. This "short-range" part can be obtained by subtracting from the Coulomb operator the potential due to the first few multipole moments of the charge distribution which defines the kernel of the operator.

In this paper we also propose a procedure to obtain the optimal basis functions necessary for representing this part of the Coulomb operator. Our procedure is

based on the minimization of the "distance" between the exact and the approximate operator in a reduced space defined by functions centred both on the fragment itself and on adjacent fragments. This minimization is achieved by increasing the number of functions used for the representation of the operator and by looking simultaneously for the orbital exponents that minimize this distance. Note that we refer to the definition of distance between two operators that is given in Ref. [30] and according to which our approximate operator can be named a "pseudo"-Coulomb operator.

We will show that, through a judicious choice of the type and number of basis functions centred on the fragment itself, it is possible to obtain an accurate and economical representation of the "short-range" part of the Coulomb operator the kernel of which consists of orbitals localized on the fragment. This fact, together with the results obtained in Ref. [35] for the exchange, allows one to reproduce easily the whole effective potential due to the electronic charge distribution localized on a molecular fragment. In this respect our approach gives a way to construct a pseudo Hamiltonian that neither introduces empirical parameters nor imposes any given functional form upon the pseudo operators, but simply takes advantage of the available accurate calculations on separate molecular fragments.

Furthermore, we observe that this technique, proposed for representing a fragment potential, is strictly related to the type and the dimension of the basis set chosen for spanning the valence space. This approach, in fact, does not introduce any specific or definite form of effective potential for any given fragment, but it suggests a way of representing the fragment potential in the specific space chosen for the "valence only" calculations. It follows that the quality of the representation can be properly calibrated on the characteristics and the type of the used valence basis set and, therefore, problems related to the change or the increase of the valence basis set – problems frequently encountered when using standard pseudopotential approaches and leading to various types of artefacts in the calculations – are avoided.

Finally, we want to remark that our method is qualitatively different from the approach known as "Model Potential" [20–24] for the following two main reasons:

i) we propose an approximate representation of "short-range" operators the kernels of which consist of molecular orbitals localized on a given chemical fragment and not simply of operators with kernels made up by atomic core orbitals.

ii) The representation of the "short-range" part of the Coulomb operator does not depend on the choice of a particular functional form but, having a generalized non-local form analogous to that proposed in Ref. [35] for the exchange, allows one to treat on an equal footing the exchange and the Coulomb operator of each molecular fragment and to take into account the characteristics of the valence basis set. For these reasons our approach is also qualitatively different from the "Effective Fragment Potential Method" proposed by Morokuma et al. [34].

In Sect. 2, along the line of Ref. [35], we describe the method proposed for representing the "short-range" part of a Coulomb operator. In Sect. 3 we describe the procedure to obtain the optimal basis functions necessary for this type of representation and, in general, for representing efficiently any "short range" operator. Finally, we present the results of test calculations on diatomic and polyatomic molecules considering also the case in which the fragment potential consists of a Coulomb plus an exchange part and analyzing also the degree of accuracy that can be reached through our approach.

2 Method

Let us consider first a large system having a group of delocalized (valence) electrons and a set of $\mathscr{M}$ distinct chemical fragments such that each one has a definite number of electrons localized on it. If the total wave function is an antisymmetrized product of group functions [32, 33] that are mutually related by the strong orthogonality condition, one can write the valence energy E_v (i.e. the energy of the valence electrons in the effective field of the nuclei and of the other localized electrons) in the following way:

$$E_v = \int \hat{h}'(1)\rho_{1v}(\tilde{x}_1, \tilde{x}'_1)_{x'_1 = x_1} d\tilde{x}_1 + \tfrac{1}{2}\int \rho_{2v}(\tilde{x}_1, \tilde{x}_2)\frac{1}{|\vec{r}_1 - \vec{r}_2|} d\tilde{x}_1 \tilde{x}_2 \qquad (1)$$

where ρ_{1v} and ρ_{2v} are the first- and second-order density matrices obtained from the valence orbitals $\tilde{x}_i = \vec{r}_i s_i$ and:

$$\hat{h}' = \hat{T} + \hat{V}_{val} + \hat{V}_{eff} \qquad (2)$$

with

$$\hat{V}_{eff} = \sum_{\mu}^{\mathscr{M}} \{[\hat{I}(\rho_{1\mu}) - \hat{V}_{\mu}] - \hat{\mathscr{K}}(\rho_{1\mu})\} = \sum_{\mu}^{\mathscr{M}} [\hat{\mathscr{I}}(\rho_{1\mu}) - \hat{\mathscr{K}}(\rho_{1\mu})] \qquad (3)$$

$$\hat{V}_{val} = \hat{V}_{en} + \sum_{\mu}^{\mathscr{M}} \hat{V}_{\mu}. \qquad (4)$$

In Eqs. (3–4) $\hat{V}_{\mu}$ represents the "long-range" components of the potential produced by the charge distribution $\rho_{1\mu}$ on the fragment μ. It is precisely this "long-range" component which must be subtracted from the Coulomb operator $\hat{I}(\rho_{1\mu})$ to obtain its "short-range" part $\hat{\mathscr{I}}(\rho_{1\mu})$. The kernel of these two operators, like that of the exchange operator $\hat{\mathscr{K}}(\rho_{1\mu})$, depends on the density matrix $\rho_{1\mu}$ relative to the fragment μ, while $\hat{V}_{en}$ is the standard electron-nuclei attraction potential.

If one expands ρ_{1v} in terms of a finite set of valence functions $\{v_\alpha\}$ – as done routinely in standard molecular calculations – one needs to represent the operators defined in Eqs. (2–4) only inside the space $\{v_\alpha\}$. In Ref. [35] we have shown that the kernel of an exchange operator $\hat{\mathscr{K}}(\rho_{1\mu})$, consisting of molecular orbitals localized on the fragment μ, can be represented as follows:

$$\mathscr{K}_{\mu}(\tilde{x}, \tilde{x}') = \sum_{ij}^{\mathscr{N}_{\mu}} \chi_i^{(\mu)}(\vec{r})\eta_i(s)k_{ij}^{(\mu)}\eta_j^*(s')\chi_j^{(\mu)*}(\vec{r}') \qquad (5)$$

with

$$k_{ij}^{(\mu)} = \int \chi_i^{(\mu)*}(\vec{r})\eta_i^*(s)\frac{\rho_{1\mu}(\tilde{x}, \tilde{x}')}{|\vec{r} - \vec{r}'|}\eta_j(s')\chi_j^{(\mu)}(\vec{r}')d\tilde{x}\,d\tilde{x}' \qquad (6)$$

and $\{\chi_i^{(\mu)}\eta_i\}$ a set of $\mathscr{N}_{\mu}$ orthonormal spin-orbitals centred on μ. Expansion (5) is exact in the limit of a complete set of orthonormal orbitals, however, because of the localized character of $\rho_{1\mu}$ and of the "short-range" nature of the operator $\hat{\mathscr{K}}(\rho_{1\mu})$, one can obtain a representation suitable for molecular calculations even using a finite and relatively small number of orbitals, as long as they are properly chosen.

Exactly the same considerations can be applied to the representation of the "short-range" component $\mathscr{I}(\rho_{1\mu})$ of the Coulomb operator whose kernel without spin:

$$\mathscr{I}_\mu(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}')\left[\int \frac{\rho_{1\mu}(\vec{x}_1, \vec{x}_1)}{|\vec{r} - \vec{r}_1|}d\vec{x}_1 - \hat{V}_\mu(\vec{r})\right] \tag{7}$$

can be represented as follows:

$$\mathscr{I}_\mu(\vec{r}, \vec{r}') = \sum_{ij}^{\mathscr{N}_\mu}\chi_i^{(\mu)}(\vec{r})\,\varepsilon_{ij}^{(\mu)}\,\chi_j^{(\mu)*}(\vec{r}') \tag{8}$$

with

$$\varepsilon_{ij}^{(\mu)} = \int \chi_i^{(\mu)*}(\vec{r})\left[\int \frac{\rho_{1\mu}(\vec{x}_1, \vec{x}_1)}{|\vec{r} - \vec{r}_1|}d\vec{x}_1 - \hat{V}_\mu(\vec{r})\right]\chi_j^{(\mu)}(\vec{r})d\vec{r}. \tag{9}$$

Obviously, Eq. (8) as well as Eq. (5) for the exchange operator has to be modified when non-orthogonal orbitals $\{\varphi_i^{(\mu)}\}$ are used. In this case, by taking into account the overlap matrix elements $S_{\alpha\beta} = \langle\varphi_\alpha^{(\mu)}|\varphi_\beta^{(\mu)}\rangle$, one gets:

$$\mathscr{I}_\mu(\vec{r}, \vec{r}') = \sum_{ij\alpha\beta}^{\mathscr{N}_\mu} \varphi_i^{(\mu)}(\vec{r})[S_{i\alpha}^{-1}\varepsilon_{\alpha\beta}^{(\mu)}S_{\beta j}^{-1}]\varphi_j^{(\mu)*}(\vec{r}') = \sum_{ij}^{\mathscr{N}_\mu}\varphi_i^{(\mu)}(\vec{r})\bar{\varepsilon}_{ij}^{(\mu)}\varphi_j^{(\mu)*}(\vec{r}'). \tag{10}$$

Using simultaneously Eqs. (5) and (8), one obtains for the matrix elements of the effective potential in the space of the valence functions $\{v_\alpha\}$ the following expression:

$$\langle v_\alpha\eta_\alpha|\hat{V}_{eff}|v_\beta\eta_\beta\rangle = \sum_\mu^{\mathscr{M}}\langle v_\alpha\eta_\alpha|\mathbf{X}^{(\mu)}\rangle\mathbf{V}^{(\mu)}\langle\mathbf{X}^{(\mu)}|v_\beta\eta_\beta\rangle \tag{11}$$

with

$$\langle v_\alpha\eta_\alpha|\mathbf{X}^{(\mu)}\rangle\mathbf{V}^{(\mu)}\langle\mathbf{X}^{(\mu)}|v_\beta\eta_\beta\rangle = \sum_{ij}^{\mathscr{N}_\mu}\langle v_\alpha\eta_\alpha|\chi_i^{(\mu)}\eta_i\rangle o_{ij}^{(\mu)}\langle\chi_j^{(\mu)}\eta_j|v_\beta\eta_\beta\rangle \tag{12}$$

$$o_{ij}^{(\mu)} = \varepsilon_{ij}^{(\mu)} - k_{ij}^{(\mu)} \tag{13}$$

and $\mathbf{X}^{(\mu)}$ a vector of $\mathscr{N}_\mu$ orthonormal spin-orbitals centred on the fragment μ. The construction of these matrix elements requires only the evaluation of overlap integrals between the valence functions and the orbitals used for the expansion of $\hat{V}_{eff}$, a process that can be performed in a simple and economical way. Note that $\{v_\alpha\}$ may also include functions which are centred on the fragment μ and are used for the expansion of the molecular orbitals $\{\chi_i^{(\mu)}\}$. In this case the matrix elements of $\hat{V}_{eff}$ between each pair of these functions coincide with the exact ones.

Two main problems need to be solved in order to obtain a working method. The first one is the definition of a simple and efficient criterion for the choice of basis functions appropriate for expanding the orthonormal orbitals $\{\chi_i^{(\mu)}\}$. In Sect. 3 we suggest a procedure for the choice of functions which one should add to the original basis set of the fragment μ in order to improve the quality of the representation without increasing appreciably the number of basis functions.

The second problem is an appropriate definition of the "long-range" potential $\hat{V}_\mu$ that must be subtracted from the Coulomb operator $\hat{I}(\rho_{1\mu})$ in order to obtain its "short range" component. In principle, one could simply make a "one center"

expansion of $\rho_{1\mu}$ in a coordinate system centred on the fragment μ, evaluate the first few multipole moments and subtract the corresponding potential from $\hat{I}(\rho_{1\mu})$. However, since we are dealing with charge distributions localized on fragments which are constituted by several atoms, it seems more convenient to distribute a set of point charges on a few relevant positions (atoms or bond centres) and require that the resulting charge distribution have the same total charge and dipole moment as calculated with the wave function of the fragment μ. In order to obtain this result one can use, for example, the "extended Mulliken electron population analysis" proposed by Huzinaga et al. [36], which requires the evaluation of standard monoelectronic integrals. Furthermore, if a better correspondence is required between exact and approximate potential, one can either use a larger number of point charges distributed in such a way as to reproduce also higher multipole moments of the charge distribution on the fragment μ or, using a smaller number of point charges, minimize the errors on the first few multipole moments.

3 Results and discussion

In order to give an example of the quality of the results obtainable using this type of representation, we have considered three molecular systems inside which one can clearly identify chemical fragments. Each fragment is characterized by a set of localized orbitals which represent the electronic charge distribution governing the dynamics of the valence electrons. The systems considered are the following: Li_2, bipyrrole ($C_8N_2H_8$), an aromatic molecule obtained from the condensation of two pyrroles, and SiH_4.

Li$_2$ has been chosen since, in this case, each fragment is clearly defined as a Li ion with a pair of core electrons localized on it. The corresponding Coulomb operator $I(1s^2)$ has a purely atomic character and we will show that a non-local representation of the type proposed in Eq. (8), with a basis set appropriate to the problem, allows one to reproduce almost exactly various matrix elements of the Coulomb operator which are important in the energy calculation. Such an example indicates that this representation constitutes an alternative to those based on linear combinations of gaussian [20–24] or error [25] functions multiplied by powers of the radial coordinate.

Bipyrrole has been chosen since it was used in Ref. [35] for testing the non-local representation proposed for the exchange operator. This molecule, at its equilibrium geometry, is characterized by a clear separation between π and σ electrons, and therefore by a well-defined sigma skeleton inside which three distinct chemical fragments can be identified – two pyrrole rings (without a C–H bond) and one C–C bond. It follows that the effective potential, due to the sigma skeleton, can be represented as the sum of three distinct contributions pertaining to the charge distributions localized on the three molecular fragments.

SiH$_4$ has been chosen for testing the quality of the non-local representation when applied to the entire effective potential operator, i.e. to $\hat{G}(\rho) = \hat{\mathscr{J}}(\rho) - \hat{\mathscr{K}}(\rho)$. We have considered two different examples. One in which the charge density ρ is that of the core electrons of Si and a second in which ρ is that of the electrons localized on the molecular fragment SiH$_3$. In the first case we have calculated, at the Hartree–Fock (HF) level, the electronic energy as a function of the Si–H bond distance for symmetric and simultaneous changes of the four Si–H bonds from the equilibrium position. In the other case, the energy relative to the (H$_3$Si)–H bond

has been calculated as a function of the bond distance while keeping the geometry of the SiH$_3$ fragment fixed.

We want to remark that in the examples reported below we have considered only the problem of correctly representing a given effective potential which consists of a "short-range" Coulomb part plus, in the last example, an exchange potential. This means that we have not introduced in our test calculations any pseudopotential owing to the orthogonality constraints between core and valence orbitals. These last have been taken strictly orthogonal to the core orbitals in all the examples considered and this has been achieved either by exploiting the symmetry of the problem, as in the case of the bipyrrole, or by including in the valence space also the components of the core orbitals. The reasons of this choice derive from the fact that the orthogonality problem is qualitatively different from that of representing an effective Coulomb or exchange potential and, therefore, the two problems have to be treated separately in order to clearly understand the relative importance of the various effects that can be present.

Let us consider first the case of Li$_2$ at its equilibrium geometry. In this case, and according to the frozen core approximation, the effective Coulomb potential which governs the dynamics of the valence electrons is the sum of two operators each of which has the $1s$ of a Li atom as its kernel orbital. This Coulomb potential can be thereby expressed as $\hat{I}(1\sigma_g^2 + 1\sigma_u^2) \simeq \hat{I}(1s_1^2 + 1s_2^2) = \hat{I}(1s_1^2) + (1s_2^2)$ with the "long-range" component of each operator that is a pure Coulomb potential since $\lim_{r \to \infty} \hat{I}(1s^2(r)) = -\frac{2}{r}$. It follows that the "short range" operator that we want to represent through Eq. (8) is given by: $\hat{\mathscr{I}}(1s_i^2) = [\hat{I}(1s_i^2) + \frac{2}{r_i}]$ with $i = 1, 2$.

In order to obtain a high quality representation of $\hat{\mathscr{I}}(1s^2)$ by means of the expansion (8), but without using too many functions, we start from the atomic basis set centred on each Li atom and used for the molecular calculations. To this basis set [37], which consists of $10s/4p_z$ gaussian functions $\{g_i\}$ contracted to $3s/2p_z$, we add one more s-type function and search for the orbital exponent that minimizes the "distance" between exact and approximate operator over a space of M valence functions $\{v_\alpha\}$. This means that we minimize the following distance:

$$\sigma = \|\hat{\mathscr{I}}_{ex}(1s_1^2) - \hat{\mathscr{I}}_{app}(1s_1^2)\| = \left(\sum_{\alpha\beta}^{\mathscr{M}} |\langle v_\alpha | \hat{\mathscr{I}}_{ex}(1s_1^2) - \hat{\mathscr{I}}_{app}(1s_1^2) | v_\beta \rangle|^2 \right)^{1/2} \quad (14)$$

$$\hat{\mathscr{I}}_{app}(1s_1^2) = \sum_{ij\alpha\beta}^{5+1} |g_i\rangle S_{i\alpha}^{-1} \varepsilon_{\alpha\beta} S_{\beta j}^{-1} \langle g_j|; \quad \varepsilon_{\alpha\beta} = \langle g_\alpha | \hat{\mathscr{I}}_{ex}(1s_1^2) | g\beta \rangle;$$

$$S_{i\alpha} = \langle g_i | g_\alpha \rangle \quad (15)$$

by changing the orbital exponent of the 6th gaussian function added to the atomic basis set. Note that the valence space, inside which the minimization is performed, is that spanned by the sum of the two atomic basis sets, thereby implying $M = 10$ in Eqs. (14).

The results obtained are summarized in Table 1 where we report the behaviour of the "distance" between exact and approximate operators as a function of the orbital exponent (α) of the s-type gaussian added to the primitive basis set of the fragment. We observe that the minimum of σ is obtained for $\alpha \sim 30$, a value that allows the reproduction of the matrix elements of $\hat{\mathscr{I}}(1s_1^2 + 1s_2^2)$ over the molecular orbitals ($1\sigma_g$, $1\sigma_u$, $2\sigma_g$) with an accuracy that is of the order of about 3.10^{-4} a.u.

The same type of analysis has been performed in the case of the bipyrrole molecule. In this case we have restricted the variational calculation to the π-space

Table 1. Li_2 molecule: behaviour of the distance (σ) between exact and approximate representations of the "short-range" part of the Coulomb operator: $\hat{\mathscr{F}}(1s_1 + 1s_2)$, and behaviour of the modulus of various matrix elements of the approximate operator as functions of the orbital exponent (α) of the gaussian added to the atomic basis set. Note that $\alpha = 0$ means that no function has been added. All the quantities are given in atomic units

α	σ	$|\langle 1\sigma_g|\mathscr{F}|1\sigma_g\rangle|$	$|\langle 2\sigma_g|\mathscr{F}|2\sigma_g\rangle|$	$|\langle 1\sigma_u|\mathscr{F}|1\sigma_u\rangle|$
100.0	0.0563	2.0646	0.0562	2.0754
50.0	0.0403	2.0681	0.0562	2.0712
30.0	0.0274	2.0685	0.0552	2.0709
10.0	0.0419	2.0709	0.0553	2.0684
7.0	0.0598	2.0682	0.0563	2.0711
2.0	0.1164	2.0265	0.0600	2.1134
0.0	0.1180	2.0425	0.0609	2.0967
exact	–	2.0688	0.0553	2.0706

and used, for the sigma skeleton, orbitals obtained from calculations on the separate fragments. The procedure used is similar to that expounded in Ref. [35] and is based on the following four steps:

a) Localization of the sigma orbitals obtained from a SCF calculation on the isolated pyrrole.

b) Minimization of the distance between exact and pseudo-Coulomb operators having kernels which consist of the sigma orbitals of one pyrrole.

c) Construction of the total pseudo-Coulomb operator for bipyrrole through adding those relative to the two pyrroles (prepared as explained in the previous steps) and a third kernel consisting of the C–C bond orbital obtained from a HF calculation on the bipyrrole.

d) Calculation of the valence energy of the bipyrrole using the total pseudo-Coulomb operator.

The calculations have been performed using a minimal basis set of STO-5G gaussian functions [38–39], which has been increased for the π-orbitals by including two diffuse p_x functions centred on each heavy atom ($\alpha_C = 0.1$–0.05; $\alpha_N = 0.12$–0.06). To represent the "short-range" part of the Coulomb operator of each fragment, via Eq. (8), we have used the set of p_x functions centred on that fragment plus one p_x function for each heavy atom. Note that the ratio of the orbital exponent (α_1) of the function added on each C to that of the function added on N has been kept fixed and equal to 0.833. This orbital exponent (α_1) has been changed to minimize the "distance" between exact and approximate operator inside the space of the valence functions of the bipyrrole. The "long-range" potential that has been subtracted from the Coulomb operator of each fragment is that produced by a set of point charges having the same total charge and dipole moment as calculated with the sigma orbitals of that fragment. The values and locations of these charges have been obtained by following the procedure suggested in Ref. [36]. We see, from the results reported in Table 2, that a satisfactory value can be obtained both for the π-energy and for the transition energies to the first singlet and first triplet excited states by adding

Table 2. Bipyrrole molecule: behavior of the distance (σ) between exact and approximate representations of the "short-range" part of the Coulomb operator and behaviour of the energy difference and of the transition energies to the first singlet (Δ_s) and to the first triplet (Δ_t) excited state as functions of the orbital exponent (α_1) of the gaussian added to the basis set of the fragment. Note that $\alpha_1 = 0$ means that no function has been added. All the quantities are given in atomic units

α_1	σ	$\Delta E_{app.-ex.}$	Δ_s	Δ_t
0.0	0.462124	− 0.0483	0.2140	0.1460
0.005	0.461583	− 0.0451	0.2135	0.1456
0.010	0.460989	− 0.0337	0.2097	0.1429
0.020	0.460887	− 0.0159	0.2093	0.1412
0.030	0.460765	− 0.0095	0.2108	0.1408
0.040	0.460755	− 0.0078	0.2129	0.1409
0.060	0.460801	− 0.0103	0.2148	0.1419
0.070	0.460819	− 0.0125	0.2161	0.1425
0.080	0.460823	− 0.0147	0.2172	0.1430
0.090	0.460925	− 0.0167	0.2182	0.1435
exact	–	–	0.2177	0.1427

only one more function per atomic center to the primitive basis set of each fragment.

In order to improve the quality of these results, we have added to the previously optimized basis set another p_x function on both the C atoms and the N and afterwards searched for the orbital exponent (α_2) that minimizes the distance between exact and approximate operators. From this procedure we have obtained a minimum distance ($\sigma = 0.4605$) for a value ($\alpha_2 = 0.03$) that gives a difference between "exact" and "approximate" energies ($\Delta E_{app-ex} = -0.0021$ a.u.) that is about four times smaller in absolute value than that obtained by adding a single function. Furthermore, the new values of the transition energies ($\Delta_s = 0.2147$ a.u. and $\Delta_t = 0.1421$ a.u.) differ from the exact ones by less than 0.1 eV. We observe that by using this representation one can correctly reproduce also the eigenvalues of the Fock operator relative to the HOMO and LUMO orbitals of the two symmetries (b_g, a_u) involved in the calculation. This is shown by the following results:

$$\varepsilon_{3b_g}^{ex} = -0.3742 \text{ a.u.}; \quad \varepsilon_{3b_g}^{app} = -0.3741 \text{ a.u.}$$

$$\varepsilon_{4b_g}^{ex} = 0.0968 \text{ a.u.}; \quad \varepsilon_{4b_g}^{app} = 0.0986 \text{ a.u.}$$

$$\varepsilon_{3a_u}^{ex} = -0.2827 \text{ a.u.}; \quad \varepsilon_{3a_u}^{app} = -0.2774 \text{ a.u.}$$

$$\varepsilon_{4a_u}^{ex} = 0.1360 \text{ a.u.}; \quad \varepsilon_{4a_u}^{app} = 0.1375 \text{ a.u.}$$

As a final test of our method, we have considered the SiH_4 molecule and constructed a non-local representation of the total effective potential $\hat{G}(\rho) = \hat{\mathscr{I}}(\rho) - \hat{\mathscr{K}}(\rho)$. We have considered two cases. In the first, ρ is the density of the core electrons of Si, and in the second, ρ is that of the 16 electrons of SiH_3 fragment. In the first case, we have represented $\hat{G}(\rho)$, via Eqs. (5–8), starting from the functions of the molecular basis set [40–42] that are centred on Si and then adding to them functions of different types with orbital exponents which minimize the distance between exact and approximate operator in the valence space.

Table 3. SiH_4 molecule: behaviour of the distance (σ) between exact and approximate representations of the "short-range" part of the effective operator and behaviour of the corresponding error on the energy as functions of the number and type of gaussians with orbital exponent α added to the basis set of the fragment. All the quantities are given in atomic units

Functions	α	σ	$\Delta E_{app.-ex.}$
$1s(\alpha)$	1.0	0.0590	-0.07844
$1s + 1s(\alpha)$	38.0	0.0519	-0.06865
$2s + 1p(\alpha)$	2.1	0.0355	-0.00474
$2s + 1p + 1d(\alpha)$	3.2	0.0170	0.00126

Table 4. SiH_4 molecule: comparison between "exact" and approximate values of the total energy given as function of the Si–H bond distance for symmetric and simultaneous changes of the four Si–H bonds. All the quantities are given in atomic units

$R(Si-H)$	$E_{ex.}$	$E_{app.}$	$\Delta E_{app.-ex.}$
2.740	-291.25011	-291.25509	-0.00498
2.760	-291.25084	-291.25572	-0.00488
2.780	-291.25120	-291.25599	-0.00479
2.790	-291.25127	-291.25603	-0.00476
2.800	-291.25122	-291.25596	-0.00474
2.820	-291.25091	-291.25562	-0.00471
2.840	-291.25028	-291.25497	-0.00469
2.860	-291.24935	-291.25404	-0.00469
2.880	-291.24813	-291.25284	-0.00471

In Table 3 we report the values of the minimum distance and of the error on the energy obtained by enlarging, step by step, the basis set used for the representation. Note that in the following examples the reference or "exact" energies are those obtained from standard Hartree–Fock calculations performed in the frozen core approximation, i.e. by keeping the fragment orbitals frozen. As one can see from the values reported in Table 3 the convergency to the "exact" energy value is quite significant and the rate of convergency increases appreciably when functions of higher angular momentum are included in the basis set.

Since, however, one is especially interested in the dependence of the energy on the geometrical parameters, more than in the absolute energy values, we have reported in Table 4 the total energy curve, i.e. the electronic energy (core and valence) plus the nuclear repulsion term calculated for symmetric changes of the four SiH_4 bonds around the equilibrium position. The curve obtained using our approach is compared in Fig. 1 with the "exact" one plotted with respect to its minimum. One can observe that the two curves are significantly parallel along a wide range of internuclear distances, even if for the representation of the effective potential we have not used the largest basis set, i.e. that giving the best energy value, but a smaller one: the $(2s + 1p)$ of Table 3 with orbital exponents optimized at the equilibrium geometry.

Another measure of the quality of these results is given by the calculation of a few quantities which characterize the approximate curve, i.e. equilibrium position
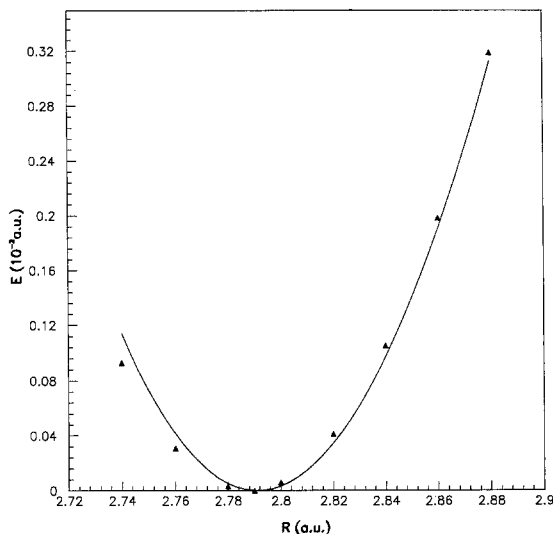
Fig. 1. "Exact" (—) and approximate (▲) energy curves given as functions of the Si–H bond distance for symmetric changes of the four Si–H bonds around the equilibrium position. Both the curves are calculated with respect to their energy minimum. The energy (E) is given in 10^{-2} a.u. and the bond distances (R) in a.u.

and value of the second derivative at the equilibrium position. The satisfactory values obtained for these quantities – which are compared in Table 5 with those deriving from a SCF calculation performed with and without the frozen core approximation – indicate that an approach of this type can be used also for predicting spectroscopic constants.

In the other example concerning the SiH_4 molecule we have applied our method to the representation of the "short-range" part of the effective potential produced by the electronic charge distribution localized on the SiH_3 fragment. In this example, a proper definition of the "long-range" potential which one has to subtract from $\hat{I}(\rho)$ in Eq. (7) is important, since the electronic charge distribution of the fragment presents also relevant quadrupole components that should be taken into account for obtaining a "short-range" effective operator. In order to achieve this result in the simplest possible way, we have used the minimum number of charges (one for each atomic center) which respect the symmetry of the problem, reproduce exactly the total charge and give dipole and quadrupole components sufficiently accurate. In practice, we have used the effective charge on Si and the ratio between this charge and that on each hydrogen as variational parameters in the process of minimization of the distance between exact and approximate "short-range" operators. Such a process has been performed at the equilibrium geometry and under the constraint of conservation of the total charge. In Table 6 we present the absolute values of dipole and quadrupole components relative to the electronic charge distribution of the fragment and to the distributions of positive charges derived either as suggested in Ref. [36] or as proposed in this example. We see that our charge distribution allows a better compensation of the quadrupole components while introduces only a small error on the dipole.

The basis set used for the representation of the "short-range" effective operator $\hat{G}(\rho)$ through Eqs. (5–8) is the $(2s + 1p)$ of the previous example with orbital exponents optimized for the specific problem. The behaviour of the error on the total energy for different values of the (H_3Si)–H bond distance is reported in

Table 5. SiH$_4$ molecule: equilibrium distance (R_{eq}.) and second derivative (K) at the equilibrium position of the curves obtained from an exact SCF calculation, a "frozen core" calculation and from our method. All the quantities are given in atomic units

	R_{eq}.	K
Exact	2.792	0.847
Frozen-core	2.791	0.848
Method	2.789	0.870

Table 6. (H$_3$Si)–H molecule: absolute values of total charge (q), x-component of the dipole moment (D_x) and xx-component of the quadrupole moment (Q_{xx}) relative to the electronic charge distribution (ρ_{el}), to the point charge distribution (ρ_H) obtained through the procedure of Ref. [36] and to that (ρ_t) used in the text. All the quantities are given in atomic units

	ρ_{el}	ρ_H	ρ_t
q	16.000	16.000	16.000
D_x	2.497	2.497	2.593
Q_{xx}	17.367	10.178	12.538

Table 7. (H$_3$Si)–H molecule: behaviour of the difference between exact and approximate values of the total energy given as function of the length (R) of the (H$_3$Si)–H bond. All the quantities are in atomic units

R	$\Delta E_{app.-ex.}$
2.740	-0.00416
2.760	-0.00415
2.780	-0.00416
2.790	-0.00416
2.800	-0.00418
2.820	-0.00420
2.840	-0.00423
2.860	-0.00428
2.880	-0.00433

Table 7 and exact and approximate curves are compared in Fig. 2. The values of equilibrium position and second derivative of the approximate curve are respectively $R_{eq} = 2.782$ a.u. and $K = 0.210$ a.u. which compare very well with the "exact" ones: $R_{eq} = 2.780$ a.u. and $K = 0.220$ a.u. The fact that the difference between approximate and exact energy values increases at large bond distances depends
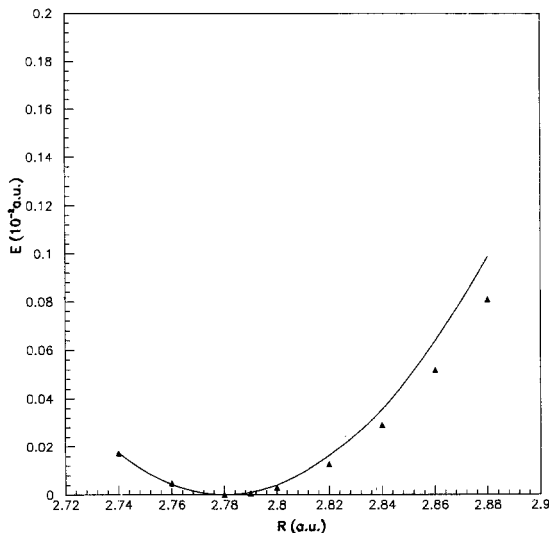
Fig. 2. "Exact" (—) and approximate (▲) energy curves given as functions of the (H$_3$Si)–H bond distance and calculated with respect to their energy minimum. The energy (E) is given in 10^{-2} a.u. and the bond distances (R) in a.u.

essentially on the type of point charge distribution chosen for eliminating the "long-range" component of the potential. A better reproduction of the exact energy curve could be easily obtained also in this region by using a more sophisticated distribution of point charges.

To conclude, we summarize the following considerations:

– The method allows one to obtain an accurate and economical representation not only of the "short-range" part of a Coulomb operator with kernel orbitals localized on a given molecular fragment, but also of the whole effective potential due to the electrons localized on the fragment. It follows that the representation of the effective potential governing the dynamics of the valence electrons in a large chemical system can be obtained as the sum of the representations of the effective potentials due to the various fragments.

– This type of approach has been proposed for representing "short-range" operators, and therefore it requires the minimization of the distance between exact and approximate operators only in the space of the valence functions centred on positions which are contained inside the volume of definition of the "short-range" operator.

References

1. O'Leary H, Duke BJ, Eilers JE (1975) Adv Quant Chem 9:1
2. Chalvet O, Daudel R, Diner S, Malrieu JP (eds) (1976) Localization and delocalization in quantum chemistry. Reidel, Dordrecht, Vol I e II
3. Boys SF (1960) Rev Mod Phys 32:296
4. Edmistone C, Ruedenberg K (1963) Rev Mod Phys 35:457

5. Newton MD, Lathan WA, Hehre WJ, Pople JA (1969) J Chem Phys 51:3927
6. Billingsley FP, Bloor JE (1971) J Chem Phys 55:5178
7. Zerner M (1975) J Chem Phys 62:2788
8. Saebo S (1992) Int J Quantum Chem 42:217
9. Panas I, Almlöf J (1992) Int J Quantum Chem 42:1073
10. Roszak S, Hariharan PC, Kaufman JJ, Koskiw S (1990) J Com Chem 11:1076
11. Kahn LR, Baybut P, Truhlar DG (1976) J Chem Phys 65:3826
12. Durand P, Barthelat JC (1978) Gazz Chim Ital 108:225
13. Dixon RN, Robertson IL (1978) in: Theoretical chemistry Vol 3, Specialist Periodical Reports, The Chemical Society, London – see for a review
14. Bachelet GB, Haman DR, Schlüter M (1982) Phys Rev B26:4199
15. Krauss M, Stevens WJ (1984) Ann Rev Phys Chem 35:357
16. Hay PJ, Wadt WR (1985) J Chem Phys 82:299
17. Bredas JL, Change RR, Silbey R, Nicolas G, Durand P (1981) J Chem Phys 75:255
18. Bredas JL, Themans B, André JM (1983) J Chem Phys 78:6137
19. André JM, Delhalle J, Bredas JL (1991) Quantum chemistry aided design of organic polymers. World Scientific Lectures and Courses Notes in Chemistry, Vol 2 and references therein
20. Bonifacic U, Huzinaga S (1974) J Chem Phys 60:2779
21. Sakai Y (1981) J Chem Phys 75:1303; Sakai Y, Huzinaga S (1982) J Chem Phys 76:2537,2552
22. Andzelm J, Radzio E, Barandiarán Z, Seijo L (1985) J Chem Phys 83:4565
23. Barandiaran Z, Seijo L (1986) J Chem Phys 84:1941
24. Huzinaga S, Seijo L, Barandiaran Z, Klobukowski M (1987) J Chem Phys 86:2132
25. Montagnani R, Salvetti O (1984) Theor Chim Acta 64:371; 65:159
26. Colle R, Fortunelli A, Salvetti O (1984) J Chem Phys 80:2654
27. Colle R, Fortunelli A, Salvetti O (1986) Mol Phys 57:1305
28. Colle R, Fortunelli A, Salvetti O (1986) Chem Phys Lett 127:84
29. Colle R, Fortunelli A, Salvetti O (1987) J de Chimie Physique 84:667
30. Durand P, Malrieu JP (1987) in: Lawley KB (ed) Ab initio methods in quantum chemistry. Wiley, NY
31. Pelisser M, Daudey JP, Malrieu JP, Jeung GH (1987) in: Veillard E (ed) Quantum chemistry: The challenge of transition metals and coordination chemistry. Reidel, Dordrecht
32. McWeeny R (1956) Proc R Soc A253:242
33. McWeeny R (1960) Rev Mod Phys 32:335
34. Ohta K, Yoshioka Y, Morokuma K, Kitaura K (1983) Chem Phys Lett 101:12
35. Colle R, Salvetti O (1991) Theoret Chim Acta 80:63
36. Huzinaga S, Sakai Y, Miyoshi E, Narita S (1990) J Chem Phys 93:3319
37. Dill JD, Pople JA (1975) J Chem Phys 62:2921
38. Hehre WJ, Steward RF, Pople JA (1969) J Chem Phys 51:2657
39. Hehre WJ, Ditchfield R, Stewart RF, Pople JA (1970) J Chem Phys 52:2769
40. McLean AD, Chandler GS (1980) J Chem Phys 72:5639
41. Clark T, Chandrasekhar J, Spitznagel GW, Schleyer PvR (1983) J Comp Chem 4:294
42. Frisch MJ, Pople A, Binkley JS (1984) J Chem Phys 80:3265