JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Atomic Charges and the Electrostatic Potential Are Ill-Defined in Degenerate Ground States

Patrick Bultinck,*[,†] Carlos Cardenas,*[,‡,§] Patricio Fuentealba,[‡,§] Paul A. Johnson,[⊥] and Paul W. Ayers[⊥]

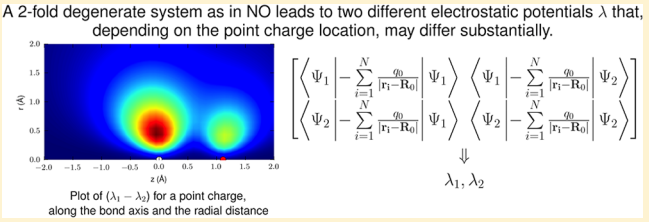
[†]Department of Inorganic and Physical Chemistry, Ghent University, Krijgslaan 281 (S3), 9000 Gent, Belgium

[‡]Departamento de Física, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile

[§]Centro para el desarrollo de la Nanociencias y Nanotecnologia, CEDENNA, Av. Ecuador 3493, Santiago, Chile

[⊥]Department of Chemistry & Chemical Biology, McMaster University, Hamilton, Ontario, Canada L8S 4M1

ABSTRACT: A system in a spatially degenerate ground state responds in a qualitatively different way to positive and negative point charges. This means that the molecular electrostatic potential is ill-defined for degenerate ground states due to the ill-defined nature of the electron density. It also means that it is impossible, in practice, to define fixed atomic charges for molecular mechanics simulations of molecules with (quasi-)degenerate ground states. Atomic-polarizability-based models and electronegativity-equalization-type models for molecular polarization also fail to capture this effect. We demonstrate the ambiguity in the electrostatic potential using several molecules of different degree of degeneracy, quasi-degeneracy, and symmetry.



I. INTRODUCTION

Atomic charge is one of the most powerful descriptive and predictive concepts in chemistry.[1] Many other concepts (e.g., the electronegativity and formal charge/oxidation state) are useful primarily because they provide insight into the population and charges of atoms and functional groups within molecules. Conversely, atomic charges are often used to predict chemical interactions: negatively charged sites tend to be basic/nucleophilic and positively charged sites tend to be acidic/electrophilic. It is, indeed, almost impossible to imagine chemistry without the concept of atomic charge. Most chemists recognize that atomic charges are not quantum-mechanical observables and are therefore not uniquely defined.[1,2] This is of little consequence for practical work: one can use any set of charges that are "chemically reasonable" (e.g., consistent with a molecular-orbital or valence-bond interpretation of electronic structure) and accurately reproduce the molecular electrostatic potential. The molecular electrostatic potential is highly important, especially on molecular surfaces, as it is one of the key reactivity indicators explaining where reactions are most likely to take place, e.g., in biomolecular modeling.[3−6]

Indeed, for the parametrization of force fields in molecular mechanics, it is key to reproduce the electrostatic potential, preferably at as low a computational cost as possible.[7] In this context, atomic charges/populations provide a coarse-grained "condensed" representation of the electrostatic potential whereby one aims to optimally reproduce the potential with as little as possible need to go beyond atomic monopoles. However, what many chemists do not appreciate is that neither the electrostatic potential nor atomic charges/populations are well-defined for molecules in degenerate ground states. While one can often qualitatively assign atomic charges in terms of δ^- or δ^+ values, when (quasi-)degenerate states occur, the value of these charges depends on whether (a) the molecule is interacting with a positive or a negative charge and (b) the location of this charge.[8] The goal of this paper is to further elucidate and demonstrate this effect. In section II the mathematical foundation of this effect is demonstrated. In section IV we show several numerical examples for molecules with different symmetries and different degrees of degeneracy using the methods described in section III. In section V we conclude by emphasizing that conceptual reasoning and molecular-mechanics force fields based on fixed atomic charges, while indubitably useful in practice, are typically theoretically ill-founded.

II. NONEXISTENCE OF THE ELECTROSTATIC POTENTIAL IN SPATIALLY DEGENERATE GROUND STATES

The electrostatic potential of an N-electron molecule with M nuclei located at positions $\{\mathbf{R}_\alpha\}$ and nuclear charges $\{Z_\alpha\}$ at a position $\mathbf{R}_0$ is most often expressed as

$$\Phi(\mathbf{R}_0) = -\int d\mathbf{r} \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_0|} + \sum_{\alpha=1}^{M} \frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{R}_0|} \tag{1}$$

This expression is generally valid although it does not by itself clarify the nature of the electron density $\rho(\mathbf{r})$ to be used. For a nondegenerate wave function, $\rho(\mathbf{r})$ is simply the electron density of the molecule but in cases where (quasi-)degeneracies appear the electron density becomes ill-defined.[8] To show the

Received: June 27, 2013
Published: September 16, 2013

problems arising from the degeneracy, it is useful to derive eq 1 from the energetic response of a molecule, per unit charge, to a point charge of magnitude q_0 at the point $\mathbf{R}_0$

$$\left[\frac{E(\hat{v} + \hat{v}'(\{\mathbf{R}_0, q_0\}); N) - E(\hat{v}; N)}{q_0}\right] \tag{2}$$

Here $E(\hat{v}; N)$ is the energy (including the nuclear repulsion but without external fields) of the system obtained from the Hamiltonian for the number of electrons, N, and the nuclear potential acting on the electrons i, $\hat{v}$:

$$\hat{v} = -\sum_{i=1}^{N}\sum_{\alpha=1}^{M}\frac{Z_\alpha}{|\mathbf{r_i} - \mathbf{R}_\alpha|} \tag{3}$$

The presence of the extra point charge results in a new external potential operator:

$$\hat{v}'(\{\mathbf{R}_0, q_0\}) = \hat{v} - \sum_{i=1}^{N}\frac{q_0}{|\mathbf{r_i} - \mathbf{R}_0|} \tag{4}$$

The notation $\hat{v}'(\{\mathbf{R}_0,q_0\})$ indicates that the operator depends parametrically on the position and magnitude of the point charge. The difference $E(\hat{v} + \hat{v}'(\{\mathbf{R}_0,q_0\}); N) - E(\hat{v}; N)$ can be easily computed using perturbation theory and provided the point charge tends toward zero the first order energy correction to the electronic energy suffices and yields $\Phi(\mathbf{R}_0)$:

$$\begin{aligned}\Phi(\mathbf{R}_0) &= \lim_{q_0\to 0}\left[\frac{E(\hat{v} + \hat{v}'(\{\mathbf{R}_0, q_0\}); N) - E(\hat{v}; N)}{q_0}\right] \\ &= \left[\langle\Psi|\hat{v}'(\{\mathbf{R}_0, q_0\}) - \hat{v}|\Psi\rangle + \sum_{\alpha=1}^{M}\frac{Z_\alpha q_0}{|\mathbf{R}_\alpha - \mathbf{R}_0|}\right]/q_0\end{aligned} \tag{5}$$

For a nondegenerate state, knowledge of the nuclear potential, the electron density of the system, and the perturbing potential suffices and eq 1 is immediately obtained. This establishes the electron density $\rho(\mathbf{r})$ as the first reactivity indicator in the set of derivatives $(\delta^k U/\delta_v^k(\mathbf{r}))_N$.[9]

In developing molecular mechanics force fields and in computing atomic charges to interpret calculations, one often selects point charges, $\{q_\alpha\}$, usually on the atomic centers such that a simple monopole approximation reproduces the electrostatic potential as closely as possible

$$\Phi(\mathbf{R}_0) \approx \sum_{\alpha=1}^{M}\frac{q_\alpha}{|\mathbf{R}_\alpha - \mathbf{R}_0|} \tag{6}$$

Of prime interest in this are electrostatic potential derived atomic charges that minimize the difference between the exactly computed $\Phi(\mathbf{R}_0)$ and the point charge based approximation in eq 6. It is this type of atomic charge[10−12] that is often used in molecular mechanics force fields[13] although clear statistical problems[13] largely limit the possibility to reproduce them using semiempirical methods such as electronegativity equalization.[14,15] There, other definitions of atomic charge that allow a good approximation of the true electrostatic potential, such as Hirshfeld-I,[16] may prove more useful.[17−19]

All the above relies on the existence of the limit in eq 5. Existence of the limit is established when the limit $(\varepsilon \to 0)^+$ (the limit of $\varepsilon \to 0$ from above) equals the limit $(\varepsilon \to 0)^-$. The present paper addresses the problem of degenerate ground states and the associated ambiguous nature of the electrostatic potential and of atomic charges. This is a problem that no specific choice of atomic charge model can alleviate as it is more fundamental than the choice of some atoms-in-molecules model. It is therefore important to stress that the ill-defined nature of atomic charges in the present paper is not due to the conceptual difficulty of defining an atom in the molecule and the need to choose a specific method but is here a consequence of the (quasi-)degeneracy of the electronic states and applies to all methods to describe an atom in the molecule.

The problem is that when the ground state is spatially degenerate, the limit in eq 5 does not exist and one finds that[8]

$$\Phi^+(\mathbf{R}_0) \le \Phi^-(\mathbf{R}_0) \tag{7}$$

where now the notation $\Phi^+(\mathbf{R}_0)$ adds a superscript to distinguish between a positive or negative sign of the point charge. Note that Cardenas et al.[8] have a different sign convention, but eq 7 is consistent with eq 5. Equation 7 means that the response to a positive and negative point charge differs depending on whether the limit is approached from the positive or negative side. As was recently argued,[8] rather than simply using eq 1 with the electron density of the degenerate ground states, one needs to consider degenerate perturbation theory and use the N-electron wave function associated with the lowest eigenvector of the perturbation matrix

$$\begin{bmatrix} \langle\Psi_1|\Delta\hat{v}|\Psi_1\rangle & \langle\Psi_1|\Delta\hat{v}|\Psi_2\rangle & \cdots & \langle\Psi_1|\Delta\hat{v}|\Psi_g\rangle \\ \langle\Psi_2|\Delta\hat{v}|\Psi_1\rangle & \langle\Psi_2|\Delta\hat{v}|\Psi_2\rangle & \cdots & \langle\Psi_2|\Delta\hat{v}|\Psi_g\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\Psi_g|\Delta\hat{v}|\Psi_1\rangle & \langle\Psi_g|\Delta\hat{v}|\Psi_2\rangle & \cdots & \langle\Psi_g|\Delta\hat{v}|\Psi_g\rangle \end{bmatrix} \tag{8}$$

where $\Delta\hat{v}$ denotes $-\sum_{i=1}^{N}(q_0/|\mathbf{r}_i - \mathbf{R}_0|)$ and g is the degree of degeneracy and $\{\Psi_i\}$ are the associated unperturbed N-electron ground state eigenfunctions. This matrix and as a consequence its eigenvectors and eigenvalues, which correspond to the first order energy corrections $E^{(1)}$, depend parametrically on the magnitude and sign of the point charge q_0 and its location $\mathbf{R}_0$. Choosing a positive point charge q_0 and using $\lambda^+_{\min}$ and $\lambda^+_{\max}$ to denote the lowest and highest eigenvalues of the perturbation matrix, one finds that upon diagonalization of the matrix 8, the lowest energy corresponds to $E^{(1)} = \lambda^+_{\min}$. Note that both $\lambda^+_{\min}$ and $\lambda^+_{\max}$ are negative as the electrons and the point charge are of opposite sign. If a point charge -1 is considered, then the same matrix is used but every integral is multiplied by -1 because of the repulsion between the electrons and the point charge. Algebraically, the eigenvalues have the same magnitude but opposite sign. As a consequence, it suffices to discuss the eigenvalues of only one given sign of q_0. In the results and discussion, we therefore limit ourselves to the analysis of $\lambda^+_{\min}$ and $\lambda^+_{\max}$ and their (positive) difference $\lambda^+_{\max} - \lambda^+_{\min}$. The eigenvalues thus show the following (in)equalities:

$$\lambda^+_{\min} \le \lambda^+_{\max} \tag{9}$$

$$\lambda^-_{\min} \le \lambda^-_{\max} \tag{10}$$

$$\lambda^+_{\min} = -\lambda^-_{\max} \tag{11}$$

$$\lambda^+_{\max} = -\lambda^-_{\min} \tag{12}$$

Note that the eigenvalues do depend parametrically on q_0 and $\mathbf{R}_0$. In order not to overload notation, these dependencies, other than on the sign of q_0, are not specified explicitly and in

dx.doi.org/10.1021/ct4005454 | J. Chem. Theory Comput. 2013, 9, 4779−4788

the results always a unit point charge is used and the dependence on $\mathbf{R}_0$ is made explicit in the figures.

The issues due to degeneracy are illustrated graphically in Figure 1. The values λ correspond to the $E^{(1)}$ values on the

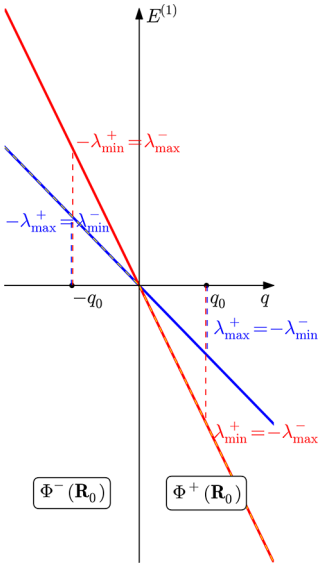


Figure 1. Graphical illustration of the effect of degeneracy on the electrostatic potential.

ordinate for a given value of q_0 on the abscissa and some value of $\mathbf{R}_0$. The relations from eq 12 are clear and the essential feature of Figure 1 is that for the depicted system with two degenerate states the slopes of the red and blue lines differ in such a way that if the lowest energy is sought for any value of q_0, the slope differs depending on the sign of q_0. So the electrostatic potential as a function of q corresponds to the slope of the yellow dashed line, clearly illustrating that the limits below and above $q_0 = 0$ differ and fulfill eq 7. When no degeneracy is present, only one line remains and the limit thus exists. The figure also explains why no unique set of atomic charges can be found in case of degeneracies as this set would have to account for two, often significantly different, slopes (see section IV).

As in the following, Born−Oppenheimer approximated wave functions are used, the focus lies on the electronic part of the electrostatic potential as the nuclear contribution does not differ between the (quasi-)degenerate electronic states. Note also that the simple formula eq 1 still applies but that for every $\{\mathbf{R}_0, q_0\}$ a new density needs to be computed from the eigenvector with the lowest eigenvalue of the perturbation matrix. This is due to the ill-defined nature of the electron density for degenerate states.[8]

As a simple elucidative example, consider how the boron atom responds to the presence of a point charge located on the z axis, at (0, 0, z). If an infinitesimal positive point charge is placed on the z axis, the degeneracy of the $2p_x$, $2p_y$, and $2p_z$ orbitals is broken, with the $2p_z$ orbital becoming the lowest-energy orbital because it allows the electron density to accumulate near the point charge. Conversely, if one puts a negative point charge on the axis, the $2p_z$ orbital goes up in energy and the HOMO becomes a linear combination of the $2p_x$ and $2p_y$ orbitals because this allows the negatively charged electrons to become concentrated in the xy plane, away from the repulsive point charge. As an example, the purely electronic perturbation matrix for a positive point charge at a distance of 2.5 au from the nucleus and along the z axis has the eigenvalues −3.870 au and the expected degenerate pair −3.740 au. Using the RHF/cc-pvtz B^+ 1s, 2s, and 2p orbitals, the eigenvectors follow the given reasoning with the eigenvector associated with λ^{+}_{min} equaling purely $2p_z$. For a negative point charge, the lowest solution is the degenerate pair with eigenvalue 3.740 au with a mixture of $2p_x$ and $2p_y$ as eigenvector and the eigenvector associated with the higher eigenvalue equaling again purely $2p_z$. Note the importance of treating the entire perturbation matrix: for a general point in space, the two degenerate eigenvalues will only be found when including the off-diagonal terms except for a limited set of points contained in the plane defined by two axis of the coordinate system. In order to have a generally working scheme, one always needs to consider the full perturbation matrix.

All the above means that molecular mechanics force fields with fixed point charges are inherently ill-suited for describing systems with (quasi-)degenerate grounds states. (We consider the ground state to be quasi-degenerate whenever the energy difference between a ground state and (an) excited state(s) is comparable to the magnitude of the electrostatic interaction. If one considers that the electrostatic potential from a proton 2.5 Å from a molecule is more than 0.2 au (5.7 eV ≈ 130 kcal/mol), the danger of ignoring the corrections from (quasi-)degeneracy in classical molecular dynamics simulations becomes clear.

To explore the importance of properly considering the effect of degenerate states on the computed electrostatic potential, we performed elaborate calculations for different molecules with different symmetries and degrees of degeneracies (section III). We observe profound differences in energetic responses, eigenvectors of the perturbation matrix, and spatial features of the density response, between a positive and negative test charge (section IV). We also observe that atomic charges change significantly (section IV). We conclude that the concepts of molecular electrostatic potential and atomic partial charges are ill-defined for (quasi-)degenerate systems, but that directly computing the molecular response using (quasi-)degenerate perturbation theory provides a way to recover chemically useful results.

III. METHODS

As test cases we consider the molecules $C_{\infty v}$ NO, D_{6h} $C_6H_6^+$, I_h $B_{12}H_{12}^-$ and C_{2h} $C_4H_6^+$. The basis sets used are consistently of at least double-ζ quality and to account for electron correlation, all calculations were performed at complete active space (CAS) SCF levels of theory. Table 1 shows the respective levels of theory applied.

In some cases, the symmetry would be broken by a Jahn−Teller effect, e.g., in $C_6H_6^+$. In these cases, the geometry optimization consisted of selecting the key geometric

Table 1. Levels of Theory Used for the Different Molecules Considered[a]

molecule	CAS	basis set	degeneracy	geometry
NO	CAS(1,2)	cc-pvtz	2	CAS(1,2)
$C_6H_6^+$	CAS(5,6)	cc-pvdz	2	CAS(5,6)
$B_{12}H_{12}^-$	CAS(7,8)	6-31G	4	MP2 $B_{12}H_{12}^{2-}$
$C_4H_6^+$	CAS(3,4)	6-31G	3QD	CAS(3,4)

[a]QD denotes the degree of quasi-degeneracy considered.

dx.doi.org/10.1021/ct4005454 | J. Chem. Theory Comput. 2013, 9, 4779−4788

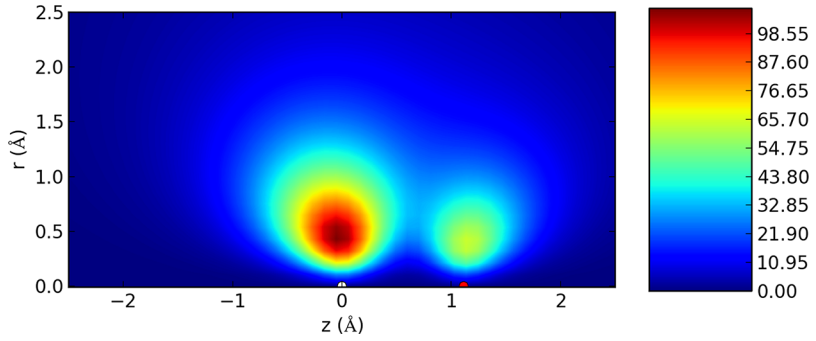


Figure 2. Difference $\lambda^{+}_{max} - \lambda^{+}_{min}$ (kcal/mol) for NO as a function of z (bond axis coordinate; Å) and radial distance r (Å) from the bond. The nitrogen atom is labeled by the white dot, and the oxygen atom, by the red dot.

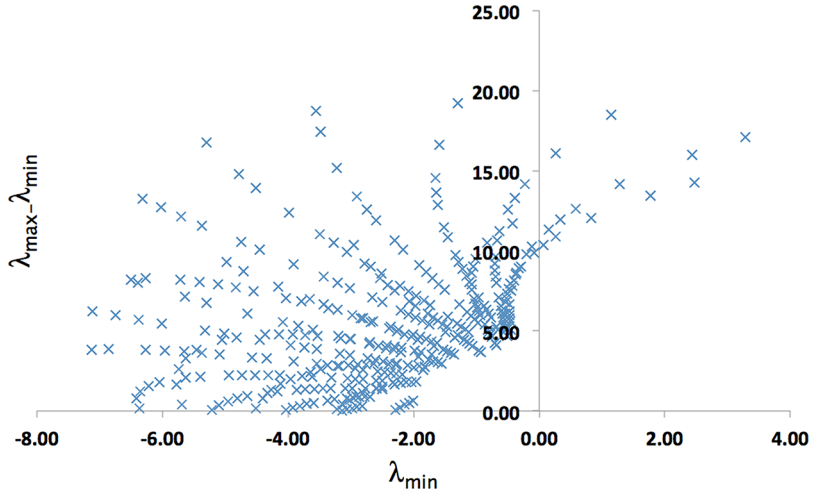


Figure 3. Difference between the minimum and maximum eigenvalues ($\lambda^{+}_{max} - \lambda^{+}_{min}$; kcal/mol) for NO versus the minimum eigenvalue (λ^{+}_{min}; kcal/mol) for a set of point charges located at points used in the CHELPG algorithm for computing electrostatic potential derived charges.

parameters (bond lengths) and generating the geometry by application of the different symmetry operations of the desired point group and scanning the potential energy surface within the constraints set by the point group. The CAS calculations were verified to yield the expected degeneracy and the Slater determinant expansion was used to compute the required density matrices and transition density matrices. The case of $C_4H_6^+$ was included to show that the impact of degeneracy is not limited to highly symmetric molecules forced into a Jahn−Teller avoided geometry but goes far beyond this.

The one electron integrals for the perturbation matrix were obtained using the Obara−Saika[20] scheme. All calculations were done using our own programs except for the CAS and full CI calculations, which were performed using Gamess(US).[21,22] This allows generating all Slater determinant expansions and all determinants with a coefficient above 10^{-15} in absolute value were retained to construct the required density matrices. Validation of the code was performed through comparing the diagonal elements of the perturbation matrix with results of Gamess(US), revealing absolute errors in the order of 10^{-6} au ($<10^{-3}$ kcal/mol) at worst. Concerning the first order reduced (transition) density matrices, the trace was found to deviate from either the number of electrons or zero by less than 10^{-9}. The eigenvalues of the density matrices also match those of Gamess(US) to the same accuracy.

IV. RESULTS AND DISCUSSION

NO. Nitric oxide is an often studied molecule for interaction with biomolecules and is characterized by a 2-fold degenerate ground state. It binds myoglobin in a way reminiscent to CO but with some subtle differences. In a recent molecular dynamics study on the bonding with myoglobin,[23] the importance of the electrostatic interaction was emphasized: "As with carbon monoxide (CO), NO is a versatile and useful local probe for the electrostatic environment of a protein." Moreover, the authors further stress that "As the difference in the kinetics depends on the details of the electronic structures of NO and CO, it would be highly desirable to have force fields that can accurately describe the electrostatics around diatomic molecules such as CO and NO". As explained in the previous section, proper treatment of the electrostatic interaction is therefore of prime importance although this was not fully appreciated previously. Figure 2 shows the difference between the maximum (yet negative) eigenvalue λ^{+}_{max} and the minimum eigenvalue (i.e., λ^{+}_{min}, the most negative). The plot shows the difference between the two eigenvalues for a unit positive point charge located at a point identified by the coordinate z (the bond axis) and with a radial (i.e., orthogonal to the z axis) distance r. Due to the $C_{\infty v}$ symmetry, the 3D data exhibit cylindrical symmetry.

As is clear from Figure 2, this difference is quite significant with values up to roughly 100 kcal/mol. Note that the eigenvalues of the perturbation matrix suffice to show the

differences in electrostatic potential as the nuclear contributions (not included in the eigenvalues) are exactly the same for both electronic states. When changing the sign of the test point charge, the roles of both eigenvalues and eigenvectors reverse. Whereas for a positive test charge, the lowest energy will be obtained using λ^+_{min}, for a negative point charge it will be obtained through $-\lambda^+_{max}$ of the perturbation matrix with a positive point charge. It is also worth mentioning that the largest changes do not occur near the nuclei, but farther away, in regions where the electrostatic potential value is relevant for reactivity prediction.

In order to highlight the effect of the degeneracy, we computed the electrostatic potential for both eigenvectors of the perturbation matrix at the set of points used in the CHELPG scheme[10] for fitting atomic charges. Figure 3 shows that the differences between both eigenvalues of the perturbation matrix may actually grow larger than $|\lambda^+_{min}|$. This very clearly demonstrates that, for force field development, one single set of charges is problematic when degenerate states occur.

Note the importance of proper treatment of the nondiagonal elements of the perturbation matrix. Although between the different states, the differences in electrostatic potential may not be that large (i.e., the difference in diagonal elements), the admittedly smaller off-diagonal elements may significantly enhance this difference. Moreover, the cylindrical symmetry would not be respected upon neglecting the off-diagonal terms.

$C_6H_6^+$. The benzene molecule is a most enigmatic molecule due to its aromatic nature and is thus also a key test molecule for evaluating the performance of new reactivity indices. Benzene itself has no degenerate states but often, as in Fukui functions,[24,25] one needs to compute wave functions for the molecular ions, under constant external potential, meaning (possibly among other constraints) the same molecular geometry. For Fukui functions, one often uses a rather pragmatic approach when considering degeneracy, namely a simple average over the densities of the degenerate states[26−29] (denoted $\langle\rho_{avg}\rangle$). In the present case, density matrices are used; hence, we could analogously use the average density matrix to compute the recently introduced Fukui matrix.[30,31] This means, however, that the off-diagonal terms in the perturbation matrix are not taken into account. From the example described above, this hardly seems to be a proper approach because depending on the location of the test point charge, another set of wave function coefficients may result in the lowest eigenvalue. For the present study, we optimized the geometry of $C_6H_6^+$ at the CAS(5,6)/cc-pVTZ level using as a constraint that the geometry must respect D_{6h} symmetry. The CAS active space contains the 6 lowest π orbitals. Figure 4 shows the results for a positive unit charge for $\lambda^+_{max} - \lambda^+_{min}$, indicating again quite substantial differences between the two eigenvalues.

As stated above, often a simple average is taken of the two density matrices to compute a response but this ignores the essence of having to treat the problem using nondegenerate perturbation theory. Figure 5 shows the difference between the 50−50 average and the lowest eigenvalue for a positive unit test charge, again showing significant errors of up to 25 kcal/mol. Cross sections through other planes reveal similar behavior although in some cases even more pronounced.

Diagonalization of the perturbation matrix also allows computing new density matrices of the correct 0th order wave function for the lowest eigenvalue and thus also allows computing atomic charges. It is important not to confuse these

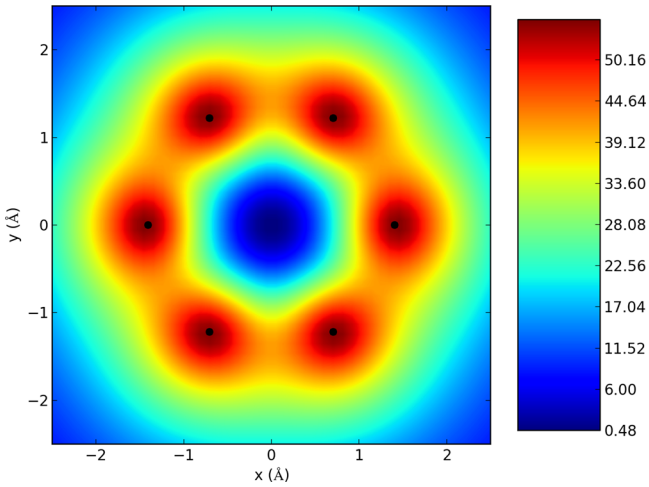


Figure 4. Difference $\lambda^+_{max} - \lambda^+_{min}$ (kcal/mol) for $C_6H_6^+$ in the molecular plane. The carbon atoms are labeled by black dots.

density matrices based on the 0th order wave functions with those that would result from perturbed wave functions. Figure 6 shows the resulting Mulliken atomic charge[32] on the first carbon atom in the molecule using the density matrix from the correct 0th order wave function. For every point in the x, y coordinate system ($z = 0$), the Mulliken charge on carbon atom 1 is shown for a unit positive point charge put at that point. As explained above, the atomic charge is expected to depend on the location and sign of the point charge. With a difference of roughly 0.15 au, the effect is considerable and shows that the use of a uniform set of atomic charges for any interaction is highly questionable. Obviously the nature of the atoms in molecules method used may have an impact, but the effect will be similar as it is determined by the underlying effects of the perturbation on the density matrix and any atoms in molecules method that relies on the density matrix and thus electron density will show these effects.

Similar figures for the other carbon atoms look exactly the same and are merely rotated according to the expected symmetry relation.

$B_{12}H_{12}^-$. Closo-dodecaborate $B_{12}H_{12}^{2-}$ is a well-known borane, characterized by the high I_h symmetry and 4-fold degeneracy of the HOMO orbitals. Using MP2 level calculations and the 6-31G basis set, the geometry was optimized and subsequently used without further refinement for the $B_{12}H_{12}^-$ molecular ion to induce 4-fold degeneracy. The MP2 geometry was found to be quite similar to the B3LYP optimized geometry. Calculations were performed at the CAS(7,8) level to capture a fair amount of electron correlation. As Figure 7 shows, even in the xy plane containing no atoms (the $z = 0$ plane with the z axis a C_5 axis connecting the two poles), an effect of similar magnitude as in the previous example is found for $\lambda^+_{max} - \lambda^+_{min}$. The need for proper treatment of degeneracy is thus clear for both smaller and larger degrees of degeneracy and different plotting planes in molecules.

Exact degeneracy is often a consequence of high symmetry of molecules and one might be tempted to consider the off-diagonal elements in the perturbation matrix to be quite small in magnitude and thus of little importance. Given the ready availability of the integrals needed to compute the entire perturbation matrix, it is advised to make it standard practice to consider the entire perturbation matrix because even small off-diagonal elements may substantially influence the eigenvalues

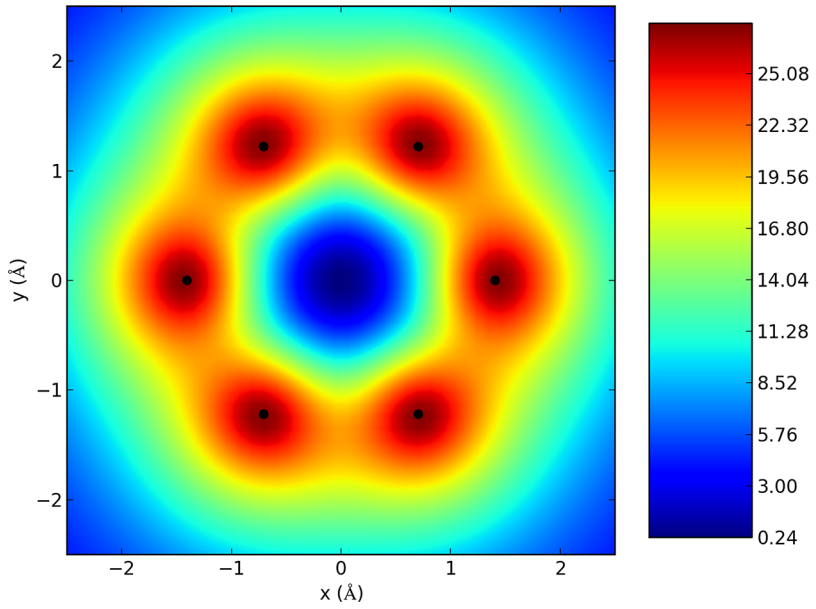


Figure 5. Difference $\langle \rho_{avg} \rangle - \lambda^{+}_{min}$ (kcal/mol) for $C_6H_6^+$ in the molecular plane. The carbon atoms are labeled by black dots.

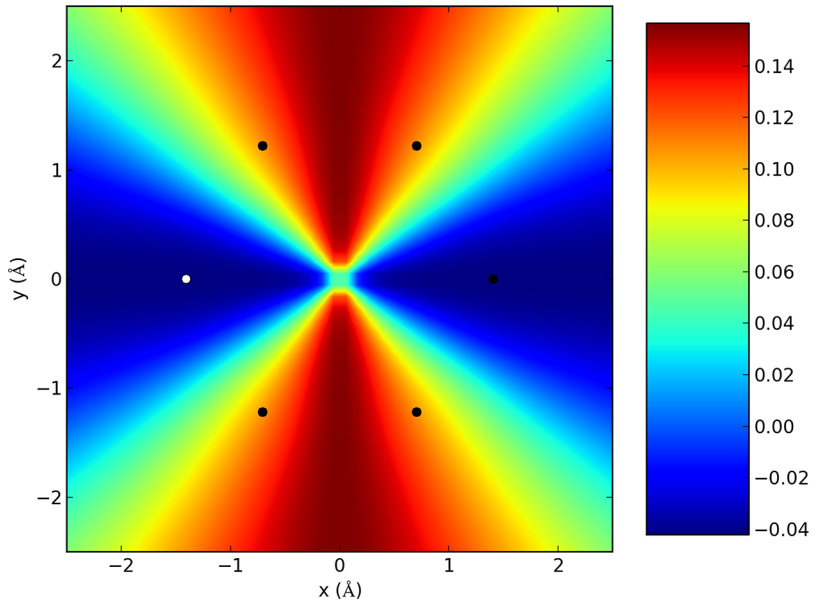


Figure 6. Mulliken atomic charge (au) for the first carbon atom in $C_6H_6^+$ for a positive unit point charge located in the molecular plane. The first carbon atom is indicated with a white dot, the other carbon atoms are labeled by black dots.

when these are only slightly separated as is usually the case for the perturbation considered here.[33] The density of valence eigenstates in $B_{12}H_{12}^-$ is very high, so the response is large, even though the difference between the lowest and highest diagonal element of the perturbation matrix is relatively small. In the realm of reactivity and selectivity, possibly even more important than the eigenvalues alone, the eigenvectors of the perturbation matrix can be seen as normal modes of reactivity. These do not only depend on the molecule itself but also on the nature of the perturbation. Yet, for the same perturbation the different eigenvectors lead to different density matrices and thus different reactivity modes.

$C_4H_6^+$. As mentioned in the introduction, atomic charges and the electrostatic potential become ambiguous also in case of quasi-degenerate states under a point charge perturbation. In order to demonstrate the effect of considering near degeneracy and at the same time the effect of the perturbation on atomic charges, we now consider a system with low lying excited states.

The trans butadiene radical cation $C_4H_6^+$ is known to have several low lying excited states above the ground B_g state[34] and in line with the previous CAS(3,4) study we indeed found two excited states within 0.2 au. We therefore included all three states in our quasi-degenerate perturbation theory treatment and computed the perturbation matrix. The off-diagonal elements in the perturbation matrix are again significant and differences between the highest and lowest eigenvalues for the three lowest states in the perturbation treatment rise up to 0.4 au, as is shown in Figure 8. Moreover, for quasidegenerate cases, any approximation like a simple average of density matrices will be even more error prone as the off-diagonal elements often become significantly larger than in most highly symmetric small molecules. Figure 8 shows the quite impressive

dx.doi.org/10.1021/ct4005454 | J. Chem. Theory Comput. 2013, 9, 4779−4788

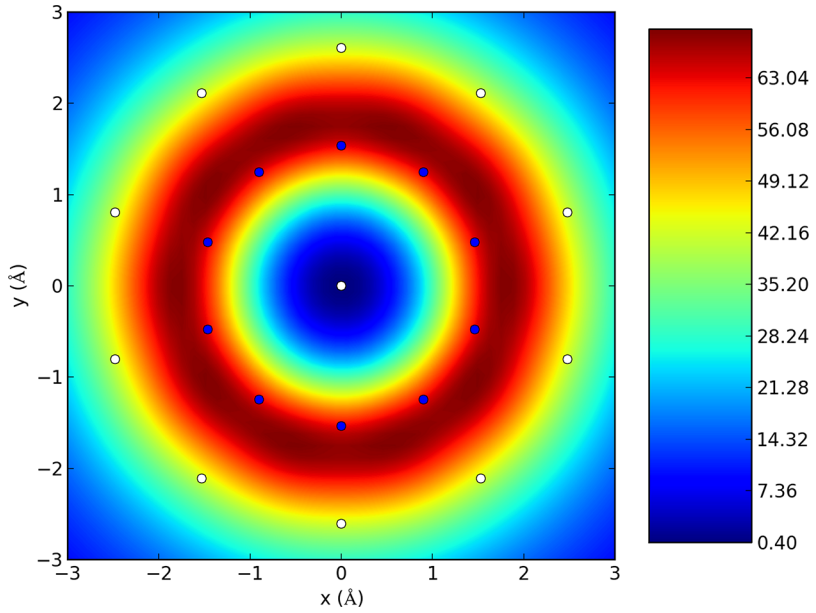


Figure 7. Difference $\lambda^{+}_{max} - \lambda^{+}_{min}$ (au) for $B_{12}H_{12}^{2-}$ in the molecular plane. The boron atoms are labeled by blue dots and hydrogen atoms by white dots. Note that all atoms lie above and below the plane in which the difference is plotted.

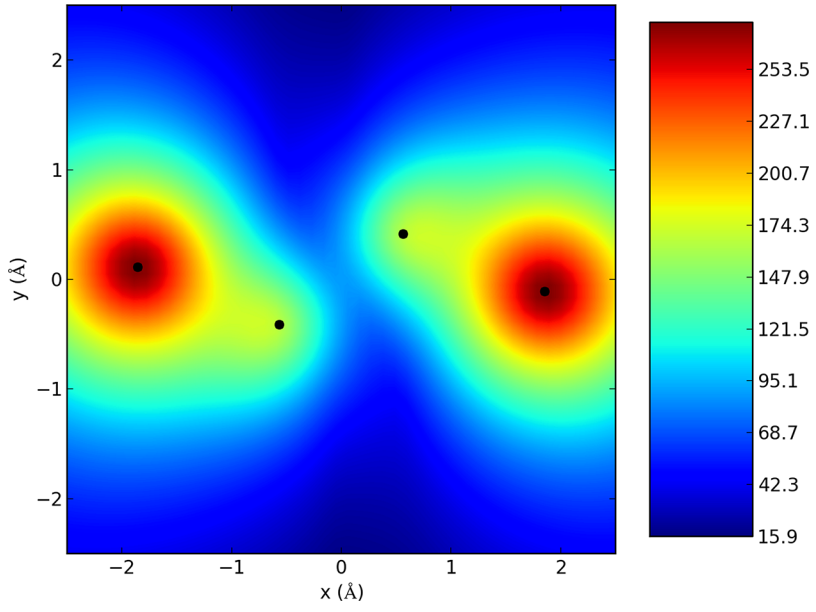


Figure 8. Difference $\lambda^{+}_{max} - \lambda^{+}_{min}$ (kcal/mol) for $C_4H_6^{+}$ in the molecular plane. The carbon atoms are labeled by black dots.

effect of the (quasi-)degeneracy with differences in eigenvalues up to roughly 250 kcal/mol. The large value is not entirely reflected in the diagonal elements of the perturbation matrix which differ by maximally 85 kcal/mol, indicating again the importance of the off-diagonal elements.

Figure 9 again shows the dependence of the Mulliken atomic charge of one of the atoms on the sign of the perturbing point charge. The effect is very significant with changes in charge up to roughly half an electron. This is not due to the specific choice of the Mulliken scheme with its known drawbacks. To show this, electrostatic potential derived charges were computed for different locations of a positive unit test charge as well as for a negative one. Following the CHELPG suggested van der Waals radii and point densities,[10] a set of 556 points on the outer molecular volume was chosen. For each of these points, the degenerate perturbation theory treatment was carried out for a single unit point charge placed at one of these points and the effective density matrix was computed for the lowest energy eigenvector of the perturbation treatment. Using this density matrix for the point charge of given sign in the chosen point, the electrostatic potential was computed on a second grid of points on the van der Waals surface. This is appropriate as it is found that the first degenerate treatment lifts any (quasi-)degeneracies. This is repeated for both positive and negative point charges at each of the 556 grid points. After the electrostatic potential has been computed on the second grid, atomic charges are computed such that the monopole based electrostatic potentials in these points match those computed ab initio in a least-squares sense, i.e. minimizing the difference between the left and right-hand side of eq 6 but adding the

dx.doi.org/10.1021/ct4005454 | J. Chem. Theory Comput. 2013, 9, 4779−4788

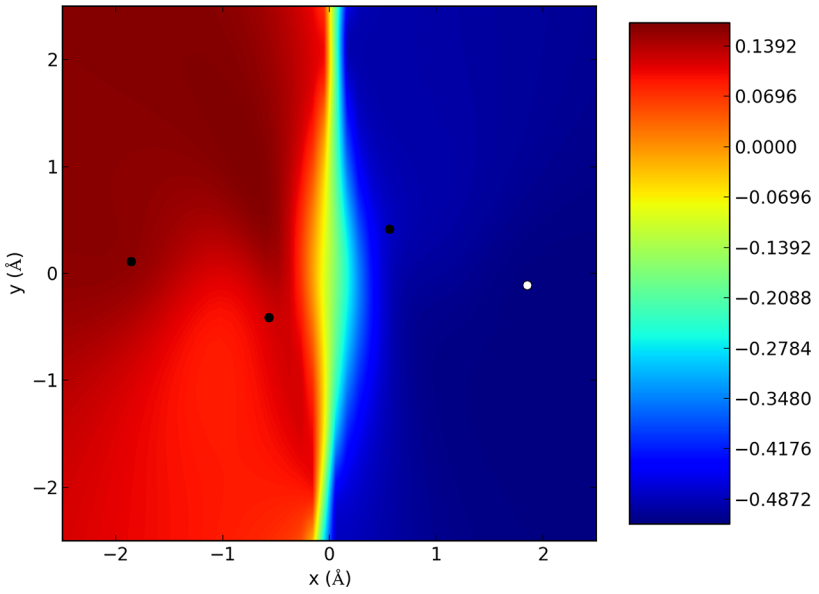


Figure 9. Mulliken atomic charge for the first carbon atom in $C_4H_6^+$ for a positive unit point charge located in the molecular plane. The first carbon atom is indicated with a white dot, the other carbon atoms are labeled by black dots.

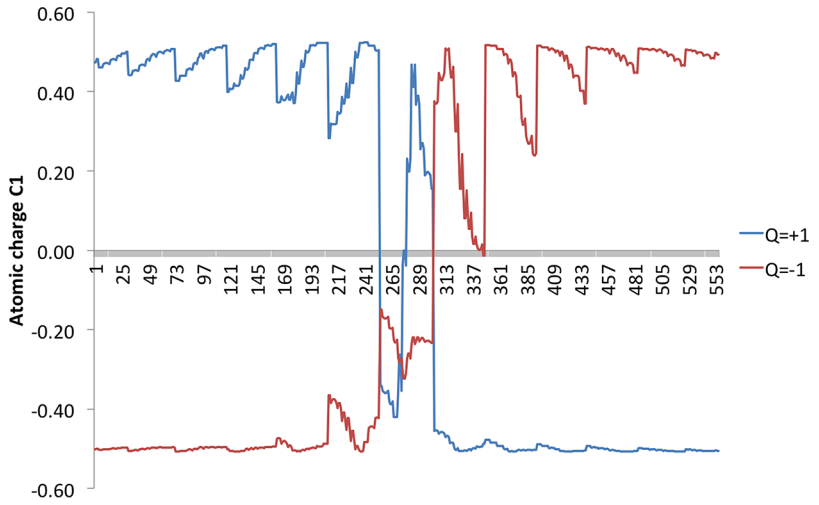


Figure 10. Electrostatic potential fitted atomic charge for the first carbon atom in $C_4H_6^+$ for a positive unit point charge $q_0 = +1$ or negative unit point charge $q_0 = -1$ placed on a grid of 556 points. The abscissa refers to a random ordering of the grid of points where the point charge is placed on the van der Waals surface .

constraint of charge conservation. This minimization is done using singular value decomposition to avoid issues related to the null space of the coefficient matrix. Although the grid of points where the point charge is placed in the degenerate treatment may be chosen different from that used for fitting atomic charges, for computational convenience the same grid is used in the implementation used here. Figure 10 shows the atomic charge on the first carbon atom for both a positive and negative unit point charge placed at one of the 556 points on the grid. As Figure 10 very clearly shows, the electrostatic potential fitted charge differs very significantly whether on the same point either a positive or negative unit point charge is placed. Moreover, the charge also depends strongly on where the point charge is located, causing even changes in sign of the atomic charge. Note that for a nondegenerate state, the same result is obtained independent of where the point charge was placed and independent of its sign. For reference, ignoring quasi-degeneracy over the three states, results in atomic charges of −0.06, −0.28, or 0.08 for respectively state 1, 2, or 3, illustrating the dramatic effect of proper treatment of (quasi-)degeneracy.

The examples shown above stress the importance of using the entire degenerate perturbation matrix instead of any approximation. The differences between the lowest and highest eigenvalue for a given sign of the test charge are significant, and moreover, for a point charge with different sign, their roles interchange. Interpretation of any reactivity modes in space also relies on using the proper density matrix which is based on the eigenvectors of the perturbation matrix. Degeneracy is not limited to small highly symmetric molecules. Especially (quasi-)degeneracy becomes more common with increasing molecular size, such that any study on reactivity should take the possible effects of (quasi-)degeneracy properly into account. Note that although CAS methods were used throughout here, all conclusions hold also for DFT calculations and may significantly impact many other reactivity indicators. At the

DFT level, however, proper treatment of degeneracy is not straightforward and requires one to identify zero-frequency excitations (poles) in the linear response matrix. In a more pragmatic approach, one can proceed in the way described in the present work although admittedly using Kohn−Sham determinants from approximate DFT functionals.

V. CONCLUSION

Atomic charges and the electrostatic potential are ill-defined in case of degenerate or (quasi-)degenerate states. Instead of using only one state, one needs to compute the entire degenerate perturbation theory matrix for every type of perturbation considered. This includes both the location of the point charge and its magnitude and sign. For a K-fold degenerate state, the $K \times K$ perturbation matrix leads in general to K possibly different responses (eigenvalues) and eigenvectors. Considering only the interaction with the electrons, the lowest (most negative) eigenvalue for a positive point charge becomes the highest (most positive) eigenvalue for a negative point charge, implying a change in sign of the eigenvalues. This implies that the density matrix, which is obtained by taking the correct 0th order linear combination of the degenerate states, also depends on the specific sign and location of the point charge. Atomic charges therefore depend on the location, sign, and type of perturbing potential. Properly using degenerate perturbation theory gives quite significant effects, with effective electrostatic potentials differing by amounts comparable to or greater than typical intermolecular interaction energies. Moreover, the variation in atomic charges is also significant, showing that no unique set of atomic charges can be used when modeling intermolecular interactions in, e.g., a force field. Although the magnitude of these variations may differ from method to method, the underlying principles are generally valid.

Given the simplicity of the integrals appearing in the perturbation matrix, we suggest that the (quasi-)degenerate perturbation matrix be used whenever (near) degeneracy sets in (e.g., in large molecules and small molecules with small band gaps), independent of the level of theory used. The interaction with a second molecule may for example be expressed in terms of a perturbation through a set of nuclear centered atomic charges. As a guideline for when to use degenerate perturbation theory, knowledge of the energy separation between the ground and other degenerate or excited states is required. We propose to use the Fermi level $\varepsilon_{\mathrm{Fermi}} = (\varepsilon_{\mathrm{HOMO}} + \varepsilon_{\mathrm{LUMO}})/2$ (where ε refers to the orbital energy) as a reference. We suggest to always examine whether (quasi-)degenerate states exist by considering whether orbitals exist within 0.20−0.25 au from this level and then to carefully examine whether states produced from excitations involving these orbitals need to be treated using degenerate perturbation theory. Force field development and molecular dynamics simulations requiring proper treatment of (quasi-)degeneracy should henceforth consider the important effect of these degeneracies, as they induce differences in electrostatic potentials of the same magnitude as many of the relevant interactions in classical molecular mechanics simulations.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: Patrick.Bultinck@UGent.be.

*E-mail: cardena@macul.ciencias.uchile.cl.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

P.B. acknowledges the Scientific Research Foundation−Flanders for continuous support and a travel grant to Chile and the research board of Ghent University for a research professorship. C.C. acknowledges FONDECYT under grant 11090013. P.W.A. thanks NSERC and the Canada Research Chairs for Support. P.A.J. thanks NSERC for a Vanier CGS fellowship. The authors are grateful to Mr. Guillaume Acke for aiding in the preparation of the figures.

REFERENCES

(1) Bultinck, P.; Popelier, P. L. A. In *Theory of Chemical Reactivity*; Chattaraj, P., Ed.; CRC Press: Boca Raton, FL, 2009; pp 215−227.

(2) Parr, R. G.; Ayers, P. W.; Nalewajski, R. F. *J. Phys. Chem. A* **2005**, *109*, 3957−3959.

(3) Politzer, P.; Laurence, P. R.; Jayasuriya, K. *Environ. Health Perspect.* **1985**, *61*, 191−202.

(4) Narayszabo, G.; Ferenczy, G. G. *Chem. Rev.* **1995**, *95*, 829−847.

(5) Politzer, P.; Murray, J. S. In *Theory of Chemical Reactivity*; Chattaraj, P., Ed.; CRC Press: Boca Raton, FL, 2009; pp 243−254.

(6) Murray, J. S.; Politzer, P. *Comput. Molec. Sci.* **2011**, *1*, 153−163.

(7) Ponder, J. W.; Case, D. A. *Adv. Protein Chem.* **2003**, *66*, 27.

(8) Cardenas, C.; Ayers, P. W.; Cedillo, A. *J. Chem. Phys.* **2011**, *134*, 174103.

(9) Ayers, P. W.; Anderson, J. S. M.; Bartolotti, L. J. *Int. J. Quantum Chem.* **2005**, *101*, 520−534.

(10) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361−373.

(11) Francl, M. M.; Chirlian, L. E. *Reviews in Computational Chemistry, Vol 14*; Wiley-Vch, Inc: New York, 2000; Vol. *14*; pp 1−31.

(12) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129−145.

(13) Francl, M. M.; Carey, C.; Chirlian, L. E.; Gange, D. M. *J. Comput. Chem.* **1996**, *17*, 367−383.

(14) Mortier, W. J.; Vangenechten, K.; Gasteiger, J. *J. Am. Chem. Soc.* **1985**, *107*, 829−835.

(15) Mortier, W. J.; Ghosh, S. K.; Shankar, S. *J. Am. Chem. Soc.* **1986**, *108*, 4315−4320.

(16) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbo-Dorca, R. *J. Chem. Phys.* **2007**, *126*, 144111.

(17) Van Damme, S.; Bultinck, P.; Fias, S. *J. Chem. Theory Comput.* **2009**, *5*, 334−340.

(18) Verstraelen, T.; Bultinck, P.; Van Speybroeck, V.; Ayers, P. W.; Van Neck, D.; Waroquier, M. *J. Chem. Theory Comput.* **2011**, *7*, 1750−1764.

(19) Verstraelen, T.; Pauwels, E.; De Proft, F.; Van Speybroeck, V.; Geerlings, P.; Waroquier, M. *J. Chem. Theory Comput.* **2012**, *8*, 661−676.

(20) Obara, S.; Saika, A. *J. Chem. Phys.* **1986**, *84*, 3963−3974.

(21) Ivanic, J.; Ruedenberg, K. *Theor. Chem. Acc.* **2001**, *106*, 339−351.

(22) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−1363.

(23) Lee, M. W.; Meuwly, M. *J. Phys. Chem. B* **2012**, *116*, 4154−4162.

(24) Parr, R. G.; Yang, W. *J. Am. Chem. Soc.* **1984**, *106*, 4049−4050.

(25) Yang, W.; Parr, R. G.; Pucci, R. *J. Chem. Phys.* **1984**, *81*, 2862−2863.

(26) Meneses, L.; Tiznado, W.; Contreras, R.; Fuentealba, P. *Chem. Phys. Lett.* **2004**, *383*, 181−187.

(27) Chamorro, E.; Perez, P. *J. Chem. Phys.* **2005**, *123*, 114107.

(28) Martinez, J. *Chem. Phys. Lett.* **2009**, *478*, 310−322.

(29) Flores-Moreno, R. *J. Chem. Theor. Comput.* **2010**, *6*, 48−54.
(30) Bultinck, P.; Clarisse, D.; Ayers, P. W.; Carbo-Dorca, R. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6110−6115.
(31) Bultinck, P.; Van Neck, D.; Acke, G.; Ayers, P. W. *Phys. Chem. Chem. Phys.* **2012**, *14*, 2408−2416.
(32) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833−1840.
(33) Paige, C. C. *Linear Algebra Appl.* **1974**, *8*, 1−10.
(34) Cave, R. J.; Perrott, M. G. *J. Chem. Phys.* **1992**, *96*, 3745−3755.