Charge Equilibration for Molecular Dynamics Simulations

Anthony K. Rappe and William A. Goddard III

BioDesign, Inc., Pasadena, California 91101, Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, and Materials Simulation Center, Beckman Institute (139-24), California Institute of Technology, Pasadena, California 91125 (Received: October 4, 1989)

We report here an approach for predicting charge distributions in molecules for use in molecular dynamics simulations. The input data are experimental atomic ionization potentials, electron affinities, and atomic radii. An atomic chemical potential is constructed by using these quantities plus shielded electrostatic interactions between all charges. Including only terms through second order in \( \phi \) leads to

\[
E_A(Q) = E_A(0) + \sum \left( \frac{\partial E}{\partial Q} \right)_{Q=0} + \frac{1}{2} \sum m_i \phi_i^2,
\]

so that

\[
\frac{\partial E}{\partial Q} = \frac{1}{2} (IP + EA) = \chi_{AA}
\]

and

\[
\frac{\partial^2 E}{\partial Q^2} = IP - EA
\]

where IP and EA denote the ionization potential and electron affinity and \( \chi_{AA} \) is referred to as the electronegativity.

To understand the physical significance of the second-derivative quantity \( \partial^2 E/\partial Q^2 \), consider the simple case of a neutral atom with a singly occupied orbital, \( \phi_A \), that is empty for the positive ion and doubly occupied for the negative ion. The difference between the IP and EA for this system is

\[
IP - EA = J^R_{AA}
\]

where \( J^R_{AA} \) is the Coulomb repulsion between two electrons in the \( \phi_A \) orbital (the self-Coulomb integral). We refer to this atomic repulsion quantity as \( J^R_{AA} \) as the ideompotential (self-Coulomb) for less awkward reference to it in later discussions. Of course, the

---

1. Introduction

Knowledge of the charge distribution within molecules is essential for determining the electrostatic energies (including hydrogen bonding) in molecular mechanics and molecular dynamics calculations. Unfortunately, reliable charge distributions are known only for a few organic molecules. Thus, currently there is no effective approach to estimate the charges for inorganic systems (ceramics, zeolites, high-\( T_c \) superconductors), and current estimates of charges for polymers and large organic systems are quite uncertain. For biological molecules, the 20 standard amino acids and four standard bases have been assigned charges that are expected to be reasonably accurate; however, charges are not available for nonstandard amino acids, unusual bases, and various cofactors and substrates.

An additional serious problem is that current approaches to molecular mechanics and molecular dynamics use fixed charges that cannot readjust to match the electrostatic environment. Since the charges are not allowed to respond to the environment, the tradition is to incorporate a dielectric constant in the interaction potential, leading to additional uncertainties in the calculations.

We propose here a general scheme for predicting charges of large molecules based only on geometry and experimental atomic properties. The charge equilibration (QEq) approach allows the charges to respond to changes in the environment, including those in applied fields, and can be applied to any material (polymer, ceramic, semiconductor, biological, metallic).

In section II, we derive the basic equations for the charge equilibration approach. The scaling parameter \( \chi \) relating atom size to crystal atomic radii is determined in section III by comparing theory and experiment for the alkali-metal halide diatomic molecules. In section IV, we discuss hydrogen atoms, which require an extension of the simple scheme of section II. Finally, in sections V and VI we apply the QEq method to a number of molecules and compare our results with experiment or ab initio theory. The concepts involved in the QEq approach rest upon earlier ideas of Pauling, Mulliken, Margrave, Parr, Pearson, Mortier, and others. Section VI summarizes the relationship between QEq and some of these earlier ideas.

In section VII, we mention some possible extensions utilizing the ability of QEq to allow polarization of the charge distribution.

II. Charge Equilibration

A. Charge Dependence of Atomic Energy. In order to estimate the equilibrium charges in a molecule, we first consider how the energy of an isolated atom changes as a function of charge. Using a neutral reference point, we can write the energy of atom A as

\[
E_A(Q) = E_A(0) + \sum \left( \frac{\partial E}{\partial Q} \right)_{Q=0} + \frac{1}{2} \sum m_i \phi_i^2.
\]
optimum shape of the orbital changes upon adding an additional electron, and an accurate description of the electron affinity requires configuration interaction so that the \( J_{AA} \) derived from (4) may differ somewhat from the \( J_{AA} \) calculated with a Hartree–Fock wave function.

Using (2) and (4) leads to

\[
E_A(Q) = E_{AO} + \chi_A^2 Q_A + \frac{1}{2} J_{AA} Q_A^2
\]

where the \( \chi_A^2 \) and \( J_{AA} \) can be derived directly from atomic data. However, the atomic IP and EA must be corrected for exchange interactions present in atoms but absent in molecules. The atomic states contain unpaired spins, whereas the molecules for which we use \( \chi_A \) and \( J_{AA} \) generally have all spins paired.) This leads to the generalized Mulliken–Pauling electronegativities and idepotentials in Table 1.

The idepotential is roughly proportional to the inverse size of the atom, and indeed, one can define a characteristic atomic size \( R_A^x \) by

\[
R_A^x = 14.4 / R_A^x \quad \text{or} \quad R_A^x = 14.4 / J_{AA}
\]

where the conversion factor 14.4 allows \( R_A^x \) to be in angstroms and \( J_{AA} \) to be in electronvolts. This equation leads to \( R_A^x = 0.84 \) Å, \( R_A^x = 1.42 \) Å, \( R_A^x = 1.22 \) Å, \( R_A^x = 1.08 \) Å, \( R_A^x = 2.06 \) Å, \( R_A^x = 3.68 \) Å, and \( R_A^x = 3.38 \) Å. Comparing with bond distances of diatomics \( R_{HH} = 0.74 \) Å, \( R_{CC} = 1.23 \) Å, \( R_{NN} = 1.10 \) Å, \( R_{OO} = 1.21 \) Å, \( R_{SS} = 2.20 \) Å, \( R_{BB} = 1.63 \) Å, and \( R_{BB} = 3.08 \) Å, we see that this characteristic atomic distance corresponds roughly with the homopolar bond distance.

Use of a quadratic relation such as (1') is expected to be valid only in a restricted region. In particular, the \( \chi_A \) and \( J_{AA} \) are clearly invalid outside the range corresponding to emptying or filling the valence shell of electrons. Thus we restrict the ranges to

\[-7 < Q_{Li} < +1 \quad -4 < Q_{C} < +4 \quad -2 < Q_{O} < +6 \] (5)

etc. and take \( E_A(Q) = \infty \) outside these ranges.

B. Electrostatic Balance. In order to calculate the optimum charge distribution, we need to evaluate the interatomic electrostatic energy, \( \sum_{A < B} Q_A Q_B r_{AB} \), where \( J_{AB} \) is the Coulomb interaction between unit charges on centers A and B ( \( J_{AB} \) depends on \( R_{AB} \), the distance between A and B). This leads to a total electrostatic energy of

\[
E(Q_1, ..., Q_N) = \sum_A (E_{AO} + \chi_A^2 Q_A + \frac{1}{2} J_{AA} Q_A^2) + \sum_{A < B} Q_A Q_B r_{AB}
\]

which we rewrite as

\[
E_0(Q_1, ..., Q_N) = \sum_A (E_{AO} + \chi_A^2 Q_A) + \frac{1}{2} \sum_{A < B} Q_A Q_B r_{AB} \] (6')

(suggesting that \( J_{AA}(R) \rightarrow J_{AA}^B \) as \( R \rightarrow 0 \)).

Taking the derivative of \( E \) with respect to \( Q_A \) leads to an atomic-scale chemical potential of the form

\[
\chi_A(Q_1, ..., Q_N) = \frac{\partial E}{\partial Q_A} = \chi_A^2 + \sum_{B} J_{AB} Q_B
\]

or

\[
\chi_A(Q_1, ..., Q_N) = \chi_A^2 + J_{AA} Q_A + \sum_{B \neq A} J_{AB} Q_B
\]

where \( \chi_A \) is a function of the charges on all the atoms. For equilibrium, we require that the atomic chemical potentials be equal, leading to \( N - 1 \) conditions

\[
\chi_1 = \chi_2 = ... = \chi_N
\]

Adding the condition on total charge

\[
Q_{tot} = \sum_{i=1}^{N} Q_i
\]

leads to a total of \( N \) simultaneous equations for the equilibrium self-consistent charges that are solved once for a given structure. These QEq equations can be written as

\[
CD = -D
\]

where

\[
D_i = Q_{tot} - \sum_{B \neq i} Q_B
\]

and

\[
C_{ij} = J_{ij} \quad \text{for} \quad i \neq j
\]

The inequalities in (5) are implemented in our programs as follows. We first solve \( (10)-(12) \) for the charges and check the inequalities in (5). If any atom is outside its range, we fix its charge at the boundary. Defining \( D \) for the nonfixed atoms as

\[
D_i = \chi_i^2 - \chi_i^2 \quad \text{for} \quad i \neq 1
\]

we solve the reduced set of equations. We find that this procedure works reliably for all cases considered.

C. Shielding Corrections. In order to solve the QEq equations (10), we first must specify the form for the Coulomb potential \( J_{AB} \) between unit charges on centers A and B separated by a distance \( R \). For large separations

\[
J_{AB}(R) = 14.4 / R
\]

(14)

(where 14.4 converts units so that \( R \) is in angstroms and \( J \) is in electronvolts). However, for distances where the charge distributions on centers A and B overlap, the simple Coulomb law (14) is no longer valid. Indeed, as \( R \rightarrow 0 \), (14) leads to

\[
J_{AB}(R) \rightarrow \infty
\]

whereas it should lead to a finite value related to \( J_{AA} \) and \( J_{BB} \) as illustrated in Figure 1. This overlap or shielding correction to (14) will be quite large for bonded atoms.

There are a number of ways of evaluating the shielding of the two charge distributions. We have chosen to express the shielding as the Coulomb integral between atomic densities. We could obtain the atomic densities from accurate (spherically averaged) Hartree–Fock (HF) or local-density calculations on atoms. However, in the current implementation of QEq, we describe the atomic density in terms of a single Slater orbital. For an atom whose outer valence orbital is \( n_s \), \( n_p \), or \( n_d \), we construct a normalized \( n_s \) Slater orbital of the form

\[
\phi_{n_s} = N_s e^{-r^2} e^{-\mu} \]

(15)
dipole moments of the alkali-metal halide diatomics. The diatomic alkali-metal halide molecules MX, where M

where \( A \) is the covalent radius in atomic units (\( \text{a.u.} \)).

Consequently, we choose the valence orbital exponent \( \xi_{A} \) for atom A by the relation

\[ \xi_{A} = \lambda(2n + 1)/(2R_{A}) \]

where \( R_{A} \) is the covalent radius in atomic units (\( a_{0} = 0.52917 \text{ Å} \)) for atom A, which we select from experimental crystal structure data (see Table I). An adjustable parameter \( \lambda \) is included in (17) to account for the difference between an average atom size as given by (16) and the crystal covalent radius \( R_{A} \). We require that the same \( \lambda \) be used for all atom of the periodic table and in section III determine \( \lambda \) by comparing the predicted and experimental dipole moments of the alkali-metal halide diatomics. The diatomic Coulomb integral \( J_{AB} \) involving these Slater functions is evaluated exactly for \( \xi_{A} \) and \( \xi_{B} \) at the various distances.

III. Alkali-Metal Halides

In order to determine the scaling factor \( \lambda \) that adjusts atomic radii to Coulomb shielding distance, we considered the 12 alkali-metal halide molecules MX, where M = Na, K, Rb, or Cs and X = Cl, Br, or I. For these systems, (8) and (9) reduce to

\[ Q_{A} = \frac{x_{A}^{2} - x_{A}^{2}}{J_{MM} + J_{AX} - 2J_{MX}} \]

We require that the calculated \( Q_{M} \) lead to the experimental dipole moment

\[ \mu_{MX} = \left( \frac{1}{4.80324} \right) Q_{M} R_{MX} \]

where \( R_{MX} \) is the experimental bond distance (the constant \( 4.80324 \) allows \( Q \) to be in electron units, \( R \) in angstroms, and \( \mu \) in debyes). The only variable here is the scaling parameter \( \lambda \). The best value of \( \lambda \) is 0.4913, which leads to an average error of 0.0018 e (see Table II). Rounding off to \( \lambda = 1/2 \) leads also to an average error of 0.0018 e, and hence (17) becomes

\[ \xi_{A} = \frac{(2n + 1)}{(4R_{A})} \]

We did not use M = Li and X = F in the fits because the errors were larger for these first-row elements. However, the results for these eight cases are also listed in Table II. Including these cases, the average error increases to 0.15 e.

IV. Hydrogen

The Mulliken-like definition \(^{11}\) for electronegativity leads for hydrogen to \( x_{H}^{M} = \frac{1}{2} \left( \text{IP} + \text{EA} \right) \) eV, which is not consistent with the Pauling \(^{12}\) or other empirical values for electronegativities. With \( x_{H}^{M} \), the hydrogen is more electronegative than C (\( x_{C}^{M} = 5.34 \)) or N (\( x_{N}^{M} = 6.90 \)), whereas the Pauling scale (based on chemical experience) has hydrogen much more electronegative than C (\( x_{C}^{P} = 2.1 \)) and slightly more electronegative than boron (\( x_{B}^{P} = 2.0 \)). As discussed in ref 9, the problem with \( x_{H}^{M} \) is that the effective EA for H is much smaller than the atomic value because the H orbital involved in a bond cannot expand to the value achieved in a free H\(^{+}\) ion. Consequently, we redefine \( x_{H}^{M} \) and \( J_{HH} \) for hydrogen, allowing EA\(_{H} \) to be a variable.

From an examination of the charges on H in the molecules LiH, CH\(_{4}\), NH\(_{3}\), H\(_{2}\)O, and HF, we find that an accurate description of \( Q_{H} \) is obtained if the effective charge parameter \( \xi_{H} \) is allowed to be charge-dependent:

\[ \xi_{H}(Q_{H}) = \xi_{H}^{0} + Q_{H} \]

Here \( \xi_{H}^{0} = 1.0698 \) is based on (17) where \( R_{H} = 0.371 \text{ Å} \). The idempotential \( J_{HH} \) becomes charge-dependent:

\[ J_{HH}(Q_{H}) = (1 + Q_{H}/\xi_{H}^{0})J_{HH} \]


TABLE III: Charges on Hydrogen

<table>
<thead>
<tr>
<th>compd</th>
<th>excl</th>
<th>QEq</th>
<th>HF</th>
<th>QEqHF</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF</td>
<td>0.415</td>
<td>0.462</td>
<td>0.462</td>
<td>0.457</td>
</tr>
<tr>
<td>H2O</td>
<td>0.325</td>
<td>0.353</td>
<td>0.398</td>
<td>0.346</td>
</tr>
<tr>
<td>NH3</td>
<td>0.267</td>
<td>0.243</td>
<td>0.338</td>
<td>0.233</td>
</tr>
<tr>
<td>CH4</td>
<td>0.150</td>
<td>0.149</td>
<td>0.124</td>
<td>0.124</td>
</tr>
<tr>
<td>LiH</td>
<td>-0.768</td>
<td>-0.767</td>
<td>-0.682</td>
<td>-0.679</td>
</tr>
</tbody>
</table>

* From eq 23 for xH(Q) and eq 7 for other atoms. Fitted to experimental charges. 

To determine the parameters xH and qH, we considered the five cases in Table III and compared with experiment (where the experimental charges are based on the lowest moments) or accurate theory. A least-squares fit leads to

\[ x_H = 4.5280 \text{ eV} \quad j_{HH}(0) = 13.8904 \text{ eV} \] (22)

and a good fit to the experimentally derived charges (see Table III, first two columns). Thus

\[ E_H + E_{HH} + x_H Q_{HH} + \frac{1}{2} j_{HH} Q_{HH}^2 (1 + Q_{HH}/1.0698) \] (23)

Instead of determining the parameters in (20) with experimentally derived charges, we could use the charges calculated from electrostatic potentials of HF wave functions. This might be more appropriate for comparing with charges from HF calculations. This leads to

\[ x_H = 4.7174 \text{ eV} \quad j_{HH} = 13.4725 \text{ eV} \] (24)

and other results as in the last column of Table III.

Equations 20 and 21 are well-behaved in the range of QH

\[-1.0 < Q_H < +1.0 \] (5')

corresponding to (5). (They would lead to unphysical results for sufficiently negative values of QH.)

In solving (10) using (21), we use (20) with an estimated QH and iterate until all QHs are self-consistent. This converges rapidly for all cases tried (six to ten iterations for an initial guess of zero for all cases discussed).

V. Results

To test the utility of the charge equilibration approach, calculations were carried out on a representative set of molecules for which ab initio partial charges or experimentally derived charges are available. The calculations were carried out at the experimental geometries14-16 using the electronegativities, idem-

(14) All polyatomic geometries are from ref 15 or 16. Diatomic geometries are from ref 10.

Figure 2. (a) Predicted charges for Ala-His-Ala. The N and O termini are charged as appropriate for a peptide. Comparisons with charges from AMBER are given in parentheses. (b) Same as (a) except that His is protonated.
VI. Comparison with Other Methods

The idea that the Mulliken electronegativity, \( \chi = 1/2(IP + EA) \), is equal to the chemical potential \( \mu = -\delta E/\delta Q \) was suggested by Iczkowski and Margrave.\(^2\) The relation between electronegativity and quantum-mechanical wave functions was established by Parr et al.,\(^2\) who showed that \( \mu = \partial E(\rho)/\partial \rho \), where \( E(\rho) \) is the density functional for the energy.

The charge expansion of the total energy (1) was suggested by Iczkowski and Margrave.\(^7\)

It was Parr and Pearson\(^8\) who identified the physical significance of the second-order coefficient \( IP - EA \) as an atomic hardness

\[ \eta_A^N = \frac{1}{2}(IP - EA) = \frac{1}{2}p_A^N \]

We agree with them that the quantity \( IP - EA \) is an important chemical quantity. However, the identification with hardness in acids and bases is less obvious. Consequently, we refer\(^9\) to this atomic Coulomb repulsion quantity \( p^N_{AB} = IP - EA = 2\eta_A^N \) as the idempotential (self-Coulomb interaction) for less awkward reference in discussions.

Popular approaches for estimating charges in molecules have been the partial equalization of orbital electronegativities (PEOE) of Gasteiger\(^1\) and the Del Re\(^1\) scheme. These methods depend upon the topology (bond connections) but not on geometry. From Table IV, we see that these schemes generally lead to the proper sign but underestimate the \( Q_H \) of alkylamines by a factor of 3–6.

A simplification of QEq theory would be to replace the Coulomb interaction \( J_{AB}(R) \) by a shielded Coulombic term

\[ J_{AB}(R) = 14.4/(R_{AB}^3) \]  

We show in Table III the effect of using this approximation (using (21) with the fit of \( \chi_H \) and \( J_{HH} \) in experiment). Using \( \epsilon = 1.0 \) often leads to net charges opposite those expected from electronegativities and hence to very unreasonable charge distributions. Using dielectric constants around \( \epsilon = 2.0 \) leads to the best self-consistent charges. For \( \epsilon = 14.0 \), we obtain values that are generally the right sign but a factor of 2–3 small. The results are very sensitive to geometry and to \( \epsilon \), resulting often in singular points in the variation of \( Q \) with \( R \). Thus, we cannot recommend the simple Coulomb potential with dielectric constant approximation (25).

Mortier et al.\(^1\) derived equations analogous to (10)–(12) from density functional theory and suggested that these equations lead to reasonable charges if the standard atomic electronegativities are modified for the molecular environment. The major difference from QEq is that Mortier used the unshielded Coulomb potential (25) with \( \epsilon = 14.4 \). Mortier’s electronegativity equalization method\(^6\) (EEM) starts with the Sanderson\(^1\) values for electronegativities (which are not in electronvolts) and the Parr-Pearson\(^8\) values\(^9\) of the Mulliken hardnesses (in electronvolts) and modifies them so that the predicted charges best fit the Mulliken charges from STO-3G calculations on small molecules. Using the Mortier-modified Sanderson electronegativities (\( \chi_H = 3.832 \), \( \chi_C = 4.053 \), \( \chi_N = 5.002 \), \( \chi_O = 5.565 \)) and Parr-Pearson hardnesses (\( \eta_H = 6.836 \), \( \eta_C = 5.671 \), \( \eta_N = 6.158 \), \( \eta_O = 6.777 \)) with \( \epsilon = 14.4 \) leads to the EEM results in Table III. The correct sign is generally obtained, but the magnitudes are low by factors of 3–6. Mortier has applied this approach to the prediction of charges for a number of ceramic crystals.\(^2\)

Allinger and co-workers\(^2\) have developed the induced dipole moment and energy (IDME) method for treating electrostatic effects in molecules in terms of bond dipole moments and induced bond dipole moments. This approach is geometry-dependent but

assumes that electrostatic interactions can be built upon a molecular connectivity framework. Thus the extension to salts is not apparent. Partial charges reported by Allinger and co-workers are provided in Table IV in the column labeled IDME. The charges calculated by this approach appear to underestimate charge transfer from hydrogen by a factor of 2 (e.g., 0.07 on H in CH₄ and C₂H₆ compared with 0.15 and 0.17 for experimentally derived charges).

VII. Properties

In addition to calculating electrostatic energies and multipole moments, the self-consistent charges can be used to evaluate the other properties such as infrared or Raman intensities. For example, if we express the α component of the dipole moment as

\[ \mu_\alpha = \sum_i \alpha R_{i\alpha} \]  

then the dipole derivative can be written as

\[ \frac{\partial \mu_\alpha}{\partial R_{\beta}} = Q_{\alpha\beta} = \sum_i \frac{\partial \alpha}{\partial R_{\beta}} R_{i\alpha} \]  

Using (10)

\[ \sum_k A_{ik} Q_k = -B_i \]

we can write

\[ \sum_{i \neq k} A_{ik} \frac{\partial Q_i}{\partial R_{\beta}} + \sum_k A_{ik} \frac{\partial \alpha}{\partial R_{\beta}} = 0 \]

Equation 28 is solved to obtain the \( \frac{\partial Q_i}{\partial R_{\beta}} \) that are substituted into (27), which is transformed to normal modes to yield the dipole intensity of each mode. Similar formulas can be derived for Raman intensities and other charge-related quantities.

VIII. Summary

The charge distributions from charge equilibration (QEQ) lead to good agreement with experiment. The QEQ approach uses only readily available experimental data (atomic IP and EA, atomic radius) and thus can be applied to any combination of atoms. (The relevant χ and J values have been tabulated for all elements through Lw₂,24) The results for simple examples of typical organic, inorganic, biological, and polymer systems seem reasonable, and we believe that this approach will prove valuable in simulating biological, polymer, and inorganic materials.

Acknowledgment. These studies were initiated while A.K.R. was on sabbatical at BioDesign, and we thank the BioDesign staff for useful discussions. We also thank BioDesign for the use of BIOGRAF in carrying out these calculations. Partial funding for this research (W.A.G.) was provided by a grant from the Air Force Office of Scientific Research (No. AFOSR-88-0051).