Can the electronegativity equalization method predict spectroscopic properties?

T. Verstraelen [a,*], P. Bultinck [b]

[a] Center for Molecular Modeling (CMM), Ghent University, 9000 Gent, Belgium
[b] Ghent University, Department of Inorganic and Physical Chemistry, Krijgslaan 281 (S3), 9000 Gent, Belgium

HIGHLIGHTS

- IR spectroscopic quantities derived from the electronegativity equalization method.
- EEM cannot reproduce DFT reference data for these spectroscopic quantities.
- Impact on development of polarizable and reactive force-field models is discussed.

GRAPHICAL ABSTRACT

Quantity		Optimal RRMSE of the EEM model	
Atomic charge	$q_{gs} = \tilde{q}_{gs} + \chi_{mol}\bar{\bar{\eta}}^{-1}d$	20%	✔
Molecular dipole	$D_j = \sum_{i=1}^{N} R_{ij}q_{gs,i},$	30%	✔
Dipole derivatives	$\frac{\partial D_j}{\partial \alpha} = \sum_{i=1}^{N} \delta_{\alpha,R_{ij}}q_i + R_{ij}\frac{\partial q_{gs,i}}{\partial \alpha}$	72%	✘
Cartesian Hessian	$\frac{\partial^2 E_{gs}}{\partial \alpha \partial \beta} = \frac{\partial q_{gs}^{T}}{\partial \beta}\frac{\partial \bar{\bar{\eta}}}{\partial \alpha}q_{gs} + \frac{1}{2}q_{gs}^{T}\frac{\partial^2 \bar{\bar{\eta}}}{\partial \alpha \partial \beta}q_{gs}$	99%	✘
Long-range part (>5Å) of the Cartesian Hessian		91%	✘

ARTICLE INFO

ABSTRACT

The electronegativity equalization method is classically used as a method allowing the fast generation of atomic charges using a set of calibrated parameters and provided knowledge of the molecular structure. Recently, it has started being used for the calculation of other reactivity descriptors and for the development of polarizable and reactive force fields. For such applications, it is of interest to know whether the method, through the inclusion of the molecular geometry in the Taylor expansion of the energy, would also allow sufficiently accurate predictions of spectroscopic data. In this work, relevant quantities for IR spectroscopy are considered, namely the dipole derivatives and the Cartesian Hessian. Despite careful calibration of parameters for this specific task, it is shown that the current models yield insufficiently accurate results.

© 2013 Elsevier B.V. All rights reserved.

Introduction

Electronegativity equalization is a pervasive concept in chemistry that is fundamental to our understanding of molecular charge distributions [1,2]. One of its modern incarnations is the Electronegativity Equalization Method (EEM), which allows a rapid computation of *ab initio quality* atomic charges from just the molecular geometry and a set of atomic electronegativity and hardness parameters [3–5]. Over the past 27 years, EEM strongly influenced related areas, such as in-silico screening [6], chemical reactivity

descriptors [7–9] and empirical models for molecular potential energy surfaces, hereafter referred to as (polarizable and possibly reactive) force fields [10–16]. Extensive calibration studies have demonstrated the accuracy of atomic charges and the parameter transferability of EEM [17–22,14,23] and its recent generalizations, the Split-Charge Equilibration (SQE) [15,24–26] and Atom-Condensed Kohn-Sham approximated to second order (ACKS2) [16].

EEM is essentially a semi-empirical form of density functional theory (DFT) [27,28], in which the molecular electron density is modeled with a minimal set of variables (the atomic charges) and the energy of a system with a specific molecular geometry is expanded in a second order Taylor series of the charges. From that perspective, one may be tempted to derive other quantities from

* Corresponding author. Tel.: +32 9 264 65 56.
  E-mail address: Toon.Verstraelen@UGent.be (T. Verstraelen).

EEM beyond just the atomic charges of the ground state [7–9]. In this work, we explore the ability of EEM to reproduce key quantities in the field of infrared spectroscopy, namely the Cartesian Hessian (i.e. force constants or second order derivative of the energy toward nuclear displacements) and the molecular dipole derivatives (first order derivative of the molecular dipole moment toward nuclear displacements). Such insights are essential for the development of force-field models and their applications in which IR spectra are derived from molecular dynamics simulations. It should be noted that such attempts have been made before by Mortier and co-workers [29,30] although limited to mainly vibrations involving bonded atoms. Here the ambition is to go beyond this and consider the entire vibrational structure. For comparison, similar assessments for traditional fixed-charge models are also considered.

The paper is structured as follows. Analytic expressions for the key spectroscopic quantities are derived in Section 2. Section 3 describes all computational aspects related to the model assessment, i.e. the test set of molecular reference computations, the cost functions used to calibrate EEM parameters and the calibration algorithm. All results are presented in Section 4 and discussed in Section 5. Finally, Section 6 summarizes the main conclusions.

Key spectroscopic quantities in EEM

In this section, we start from the EEM energy expression. Justifications and a detailed description of the underlying physics can be found elsewhere [4,12,31,16]. The molecular energy for a system with N atoms is approximated to second order in the atomic charges,

$$E(\{q\}) = \boldsymbol{\chi}^{\mathsf{T}} \mathbf{q} + \frac{1}{2} \mathbf{q}^{\mathsf{T}} \bar{\bar{\boldsymbol{\eta}}} \mathbf{q}, \tag{1}$$

where $\mathbf{q}$ is a column vector with N atomic charges, $\boldsymbol{\chi}$ is a column vector with N atomic electronegativity parameters and $\bar{\bar{\boldsymbol{\eta}}}$ is a symmetric N by N matrix containing all second order coefficients, known as the hardness matrix. The diagonal elements of $\bar{\bar{\boldsymbol{\eta}}}$ are atomic hardness parameters, while the off-diagonal elements correspond to electrostatic interactions between the atoms. In line with previous work, the electrostatics are based on Gaussian charge distributions, whose widths are derived from covalent radii [24,25]. Note that our atomic hardness parameters do not include the factor 1/2 of the Taylor series as in Parr's operational definition of the atomic hardness [32].

The EEM *ground state* charges are found by minimizing this energy with a constraint on the total charge q_{mol},

$$E_{\mathrm{gs}} = E(\{q\}_{\mathrm{gs}}) = \min_{\{q\}, \mathbf{d}^{\mathsf{T}}\mathbf{q} = q_{\mathrm{mol}}} E(\{q\}), \tag{2}$$

where $\mathbf{d}$ is a column vector filled with N times the number 1. This ground state is defined as the stationary point of the following Lagrangian:

$$L = E - \chi_{\mathrm{mol}}(\mathbf{d}^{\mathsf{T}}\mathbf{q} - q_{\mathrm{mol}}). \tag{3}$$

To facilitate the remainder of the derivations, we introduce the vector with *unconstrained* ground state charges:

$$\tilde{\mathbf{q}}_{\mathrm{gs}} = -\bar{\bar{\boldsymbol{\eta}}}^{-1}\boldsymbol{\chi} \tag{4}$$

Using this notation, the stationary point of the Lagrangian takes the following form:

$$\chi_{\mathrm{mol}} = \frac{q_{\mathrm{mol}} - \mathbf{d}^{\mathsf{T}}\tilde{\mathbf{q}}_{\mathrm{gs}}}{\mathbf{d}^{\mathsf{T}}\bar{\bar{\boldsymbol{\eta}}}^{-1}\mathbf{d}} \quad \text{and} \quad \mathbf{q}_{\mathrm{gs}} = \tilde{\mathbf{q}}_{\mathrm{gs}} + \chi_{\mathrm{mol}}\bar{\bar{\boldsymbol{\eta}}}^{-1}\mathbf{d} \tag{5}$$

We will now treat the derivatives of the energy and the charges towards nuclear displacements. For the sake of compactness, nuclear

Cartesian coordinates are labeled with Greek letters α and β instead of explicitly writing R_{ij}. (At most two are needed as we consider at most second order derivatives. Note that R_{ij} is component j of the Cartesian coordinates of atom i and *not* an interatomic distance between two atoms i and j.) In EEM, the electronegativity parameters are geometry-independent; such derivatives are not considered below. The only geometry-dependence in the model is found in the off-diagonal elements of $\bar{\bar{\boldsymbol{\eta}}}$.

The forces acting on the atoms consist of first-order derivatives of the energy of the following form:

$$\frac{\partial E_{\mathrm{gs}}}{\partial \alpha} = \underbrace{\left(\boldsymbol{\chi}^{\mathsf{T}} + \mathbf{q}_{\mathrm{gs}}^{\mathsf{T}}\bar{\bar{\boldsymbol{\eta}}}\right) \frac{\partial \mathbf{q}_{\mathrm{gs}}}{\partial \alpha}}_{=0} + \frac{1}{2}\mathbf{q}_{\mathrm{gs}}^{\mathsf{T}} \frac{\partial \bar{\bar{\boldsymbol{\eta}}}}{\partial \alpha} \mathbf{q}_{\mathrm{gs}} \tag{6}$$

In analogy with the Helmann-Feynman theorem in quantum mechanics, the first term is zero because the charges are obtained from a variational principle. However, in the derivation of the force constants, the response of the charges does appear. Therefore, let us first derive the *charge response* and subsequently consider the force constants. The response of the charges with respect to a change in geometry can be found after careful application of the chain rule to Eq. (5):

$$\frac{\partial \mathbf{q}_{\mathrm{gs}}}{\partial \alpha} = \frac{\partial \tilde{\mathbf{q}}_{\mathrm{gs}}}{\partial \alpha} + \frac{\partial \chi_{\mathrm{mol}}}{\partial \alpha} \bar{\bar{\boldsymbol{\eta}}}^{-1}\mathbf{d} + \chi_{\mathrm{mol}} \frac{\partial \bar{\bar{\boldsymbol{\eta}}}^{-1}}{\partial \alpha} \mathbf{d} \tag{7}$$

with

$$\frac{\partial \bar{\bar{\boldsymbol{\eta}}}^{-1}}{\partial \alpha} = -\bar{\bar{\boldsymbol{\eta}}}^{-1} \frac{\partial \bar{\bar{\boldsymbol{\eta}}}}{\partial \alpha} \bar{\bar{\boldsymbol{\eta}}}^{-1}, \tag{8}$$

$$\frac{\partial \tilde{\mathbf{q}}_{\mathrm{gs}}}{\partial \alpha} = -\bar{\bar{\boldsymbol{\eta}}}^{-1} \frac{\partial \bar{\bar{\boldsymbol{\eta}}}}{\partial \alpha} \tilde{\mathbf{q}}_{\mathrm{gs}}, \tag{9}$$

$$\frac{\partial \chi_{\mathrm{mol}}}{\partial \alpha} = \frac{q_{\mathrm{mol}} - \mathbf{d}^{\mathsf{T}}\tilde{\mathbf{q}}_{\mathrm{gs}}}{\left(\mathbf{d}^{\mathsf{T}}\bar{\bar{\boldsymbol{\eta}}}^{-1}\mathbf{d}\right)^2} \left(\mathbf{d}^{\mathsf{T}} \frac{\partial \bar{\bar{\boldsymbol{\eta}}}^{-1}}{\partial \alpha} \mathbf{d}\right) - \frac{1}{\mathbf{d}^{\mathsf{T}}\bar{\bar{\boldsymbol{\eta}}}^{-1}\mathbf{d}} \frac{\partial \mathbf{d}^{\mathsf{T}}\tilde{\mathbf{q}}_{\mathrm{gs}}}{\partial \alpha}. \tag{10}$$

The Cartesian Hessian can be found through the application of the chain rule to Eq. (6):

$$\frac{\partial^2 E_{\mathrm{gs}}}{\partial \alpha \partial \beta} = \frac{\partial \mathbf{q}_{\mathrm{gs}}^{\mathsf{T}}}{\partial \beta} \frac{\partial \bar{\bar{\boldsymbol{\eta}}}}{\partial \alpha} \mathbf{q}_{\mathrm{gs}} + \frac{1}{2}\mathbf{q}_{\mathrm{gs}}^{\mathsf{T}} \frac{\partial^2 \bar{\bar{\boldsymbol{\eta}}}}{\partial \alpha \partial \beta} \mathbf{q}_{\mathrm{gs}} \tag{11}$$

In EEM, the molecule is modeled as a superposition of spherical atomic charge distributions. Hence, the Cartesian components of the molecular dipole moment simply become

$$D_j = \sum_{i=1}^{N} R_{ij} q_{\mathrm{gs},i}, \tag{12}$$

where q_i is the charge of atom i. The dipole derivatives are obtained by differentiating D_j once with respect to a nuclear displacement:

$$\frac{\partial D_j}{\partial \alpha} = \sum_{i=1}^{N} \delta_{\alpha, R_{ij}} q_i + R_{ij} \frac{\partial q_{\mathrm{gs},i}}{\partial \alpha}. \tag{13}$$

In this paper, the EEM model will also be compared to a model with fixed atomic charges. Such a comparison is relevant as many force-field models treat the atomic charges as geometry-independent parameters. The above results for the Cartesian Hessian and the dipole derivatives are still valid for fixed-charge models, except that the terms containing derivatives of the charges can be dropped.

Assessment protocol

To test the accuracy of EEM for the key spectroscopic quantities of interest, DFT reference computations were carried out for a set