Perspective

Perspective on "Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules. I. Theory"

Pulay P (1969) Mol Phys 17: 197–204

H. Bernhard Schlegel

Department of Chemistry, Wayne State University, Detroit, MI 48202, USA

Abstract. This article provides an outline of the title paper by Peter Pulay and discusses some of the methodology that grew from it, and the impact that it has had on the development of computational chemistry.

Key words: Gradients – Force constants – Hessians – Energy derivatives – Geometry optimization

1 The paper

The title indicates that this paper is about the calculation of vibrational force constants and the geometry optimization of polyatomic molecules; however, its primary impact on computational chemistry comes from the methodology for calculating analytic first derivatives with respect to molecular coordinates at the Hartree–Fock (HF) level of theory. Applications of first and higher derivatives of the energies obtained by molecular orbital (MO) calculations have revolutionized computational chemistry, allowing molecular structures and properties to be computed efficiently and reliably [1–5]. Almost all electronic structure codes compute analytic first derivatives of the energy, and Pulay's paper was the first to describe a practical calculational approach.

In the linear combination of atomic orbitals to form MOs (LCAO-MO) approach, the HF energy can be expressed in terms of the one-electron density matrix $\mathbf{D}$ and integrals over the basis functions:

$$
\begin{aligned}
E = \langle\psi|H|\psi\rangle &= 2\Sigma_j\big(\phi_j|H_1|\phi_j\big) + \Sigma_{jk}\big(\phi_j^2|1/r_{jk}|\phi_k^2\big) \\
&= 2\Sigma_{mn}(m|H_1|n)D_{mn} \\
&\quad + \Sigma_{mnrs}(mn|rs)(2D_{mn}D_{rs} - D_{mr}D_{ns}) \ ,
\end{aligned}
\tag{1}
$$

where H_1 is the one-electron part of the Hamiltonian, ϕ are the MOs and m, n, r, and s refer to the basis functions (shown for a spin-restricted, closed-shell system with real orbitals). The first derivative of the energy is

$$
dE/dq_i = \langle\psi|dH/dq_i|\psi\rangle + 2\langle d\psi/dq_i|H|\psi\rangle \ .
\tag{2}
$$

The first term is the Hellmann–Feynman contribution and is readily calculated as an expectation value of a one-electron operator. The second part can be called the wavefunction derivative term (it has also come to be known as the Pulay term). Typical MO wavefunctions are constructed from basis functions that are centered on the atoms and follow them rigidly. Such wavefunctions do not obey the Hellmann–Feynman theorem and the wavefunction derivative term is not zero. Many molecular properties can also be written as energy derivatives [1–4]. For some properties, it is advantageous to make the wavefunction depend on the electric or magnetic field; for these cases the wavefunction derivative term must also be calculated.

The wavefunction derivative term depends on the derivatives of the one- and two-electron integrals over the basis functions, and on the derivative of the density matrix. The self-consistent-field (SCF) process finds the density matrix that minimizes the energy under the constraint of orthonormality of the MOs; hence, only the changes required to maintain orthonormality are needed for the density matrix derivative contributions to the wavefunction derivative term. The constraint that the molecular orbitals are orthonormal is equivalent to the requirement that the density matrix be idempotent:

$$
\mathbf{D}\,\mathbf{S}\,\mathbf{D} = \mathbf{D} \ ,
\tag{3}
$$

where $\mathbf{S}$ is the overlap matrix. Equation (3) can be differentiated to obtain the change in $\mathbf{D}$ that maintains idempotency

$$
d\mathbf{D}/dq_i\,\mathbf{S}\,\mathbf{D} + \mathbf{D}\,d\mathbf{S}/dq_i\,\mathbf{D} + \mathbf{D}\,\mathbf{S}\,d\mathbf{D}/dq_i = d\mathbf{D}/dq_i \ .
\tag{4}
$$

A simple solution that satisfies Eq. (4) is

$$
d\mathbf{D}/dq_i = -\mathbf{D}\,d\mathbf{S}/dq_i\,\mathbf{D} \ .
\tag{5}
$$

The derivative of the energy, including both the Hellmann–Feynman and wavefunction terms, is

$$dE/dq_i = 2\sum_{mn}(m|dH_1/dq_i|n)D_{mn}$$
$$+ 2\sum_{mn}[(dm/dq_i|H_1|n) + (m|H_1|dn/dq_i)]D_{mn}$$
$$+ \sum_{mnrs}[(dm/dq_i n|rs) + (m\,dn/dq_i|rs)$$
$$+ (mn|dr/dq_i s) + (mn|r\,ds/dq)]$$
$$\times (2D_{mn}D_{rs} - D_{mr}D_{ns})$$
$$+ 2\sum_{mn}(m|H_1|n)dD_{mn}/dq_i + \sum_{mnrs}(mn|rs)$$
$$\times (2dD_{mn}/dq_i\,D_{rs} + 2D_{mn}dD_{rs}/dq_i$$
$$- dD_{mr}/dq_i\,D_{ns} - D_{mr}dD_{ns}/dq_i)\ . \tag{6}$$

Insertion of Eq. (5) and recognition that the terms multiplying the density derivative constitute the Fock matrix, $\mathbf{F}$, yields the following compact expression for the analytic first derivative of the HF energy

$$dE/dq_i = 2\sum_{mn} d(m|H_1|n)/dq_i\,D_{mn}$$
$$+ \sum_{mnrs} d(mn|rs)/dq_i\,(2D_{mn}D_{rs} - D_{mr}D_{ns})$$
$$- 2\sum_{mn} dS_{mn}/dq_i\,W_{mn}\ , \tag{7}$$

where $\mathbf{W} = \mathbf{D}\,\mathbf{F}\,\mathbf{D}$ (borrowing the notation of later papers on energy derivatives).

Pulay's paper goes on to calculate force constants (second derivatives of the energy with respect to geometrical parameters) by numerically differentiating the analytical first derivatives. At the time this was by far the best compromise between efficiency and accuracy in computing the force constants. The second use of geometric derivatives described in Pulay's paper is for geometry optimization. Optimization is best carried out in internal coordinates, but the calculation of analytic derivatives is most practical in Cartesian coordinates. The forces are the negative of the first derivatives of the energy. The relation between Cartesian and internal displacements ($d\mathbf{x}$ and $d\mathbf{q}$) and forces ($\mathbf{f}$ and ϕ) can be written in terms of the Wilson $\mathbf{B}$ matrix [6]

$$d\mathbf{q} = \mathbf{B}\,d\mathbf{x},\ \mathbf{f} = \mathbf{B^t}\phi \tag{8}$$

The $\mathbf{B}$ matrix is rectangular; however, a suitable left inverse can be constructed to transform Cartesian forces into internal coordinates.

$$\phi = (\mathbf{B}\,\mathbf{m}\,\mathbf{B^t})^{-1}\,\mathbf{B}\,\mathbf{m}\,\mathbf{f} \tag{9}$$

The internal forces and a suitable approximation to the force constants, $\mathbf{F}_0$, can be used to relax the molecule to its equilibrium geometry.

$$\mathbf{q}_{new} = \mathbf{q}_{old} + \mathbf{F}_0^{-1}\phi_{old} \tag{10}$$

2 Perspective

Pulay's paper is an early landmark in the explosive growth in computational chemistry that we have seen in the past quarter century. The method for calculating first derivatives as outlined in the article forms the basis for the subsequent development of first, second and higher energy derivatives for many different theoretical methods (for reviews, see Refs. [1–5]). The advances brought about by energy derivative methods have enabled theoretical calculations to become practical and efficient

methods for determining molecular structures, exploring potential-energy surfaces and computing molecular properties.

Pulay demonstrated that analytic first derivatives with respect to geometric parameters can be calculated easily and efficiently for HF energies. Derivatives of correlated methods followed a number of years after SCF derivatives [4, 5]. Extensions of the SCF derivatives to density functional theory methods were straightforward. In the three decades since Pulay's article, hundreds of papers on energy derivatives have been published, and all can trace their roots back to his paper. Energy derivatives have become so useful for calculating molecular structures and properties that, almost universally, first derivatives are formulated and coded soon after a new theoretical method is developed for the energy.

Pulay's paper opened the way for analytic second and higher derivatives of the SCF energy. Earlier papers had suggested that this might be prohibitively expensive [7], but the development of an efficient method to solve the couple perturbed HF (CPHF) equations, made the calculation of SCF second derivatives practical [8]. As a consequence, vibrational force constants and frequencies could be calculated routinely and efficiently. Third and fourth geometric derivatives of the SCF energy followed after a few years [9–12]. The solution of the CPHF equations (in their full or reduced Z-vector form [13]) also made post-SCF first derivatives practical and cost-effective.

Analytic first derivatives with respect to geometrical parameters have proven to be extremely useful for exploring potential-energy surfaces [14]. Almost all studies using electronic structure methods involve geometry optimization at some level of theory. With the possible exception of diatomics, energy derivative methods are much less expensive than energy-only methods for obtaining equilibrium geometries. Calculations on molecules with hundreds and thousands of atoms are now possible, and energy derivatives provide an enormous amount of information about the potenital-energy surface at very little additional cost. Pulay's paper also outlined the transformation of first derivatives from Cartesian to internal coordinates. Optimization in internal coordinates is significantly more efficient than in Cartesian coordinates, and it is now the standard approach. Energy derivatives are indispensable for exploring other aspects of potential-energy surfaces. For example, they are essential for finding transition states and following reaction paths [14]. Recently it has become practical to compute classical trajectories using energy derivatives directly from the electronic structure calculations without first fitting a global potential-energy surface (for a review, see Ref. [15]).

Many molecular properties, such as vibrational frequencies, IR and Raman intensities, NMR shielding constants, etc., can be formulated in terms of second and higher derivatives with respect to geometry and applied fields [1–4]. Such calculations are now practical and routine using analytic derivatives at the SCF level and a few correlated methods. For some levels of theory, analytic second and higher derivatives are not yet available or are too complicated to code. In these cases, the 'force

method' described by Pulay's paper is still the method of choice, i.e. differentiating once analytically and the remaining times numerically.

References

1. Pulay P (1995) In: Yarkony DR (ed) Modern electronic structure theory. World Scientific, Singapore, p 1191, and references therein
2. Helgaker T, Jørgensen P (1988) Adv Quantum Chem 19: 183
3. Dykstra CE (1988) Ab initio calculation of the structures and properties of molecules. Elsevier, Amsterdam
4. Jørgensen P, Simons J (eds) (1986) Geometrical derivatives of energy surfaces and molecular properties. Reidel, Dordrecht
5. Shepard R (1995) In: Yarkony DR (ed) Modern electronic structure theory. World Scientific, Singapore, p 345, and references therein
6. Wilson EB, Decius JC, Cross PC (1955) Molecular vibrations. McGraw-Hill, New York
7. (a) Thompsen K, Swanstron P (1973) Mol Phys 26: 735; Thompsen K, Swanstron P (1973) Mol Phys 26: 751
8. Pople JA, Krishnan R, Schlegel HB, Binkley JS (1979) Int J Quantum Chem Symp 13: 225
9. Pulay P (1983) J Chem Phys 78: 5043
10. Gaw JF, Yamaguchi Y, Schaefer HF III, Handy NC (1986) J Chem Phys 85: 5132
11. Augspurger JD, Dykstra CE (1991) J Phys Chem 95: 9230
12. Colwell, SM, Jayatilaka D, Maslen, PE, Amos RD, Handy NC (1991) Int J Quantum Chem 40: 179
13. Handy NC, Schaefer HF III (1984) J Chem Phys 81: 5031
14. Schlegel HB (1995) In: Yarkony DR (ed) Modern electronic structure theory. World Scientific, Singapore, p 459, and references therein
15. Bolton K, Hase WL, Peshlherbe GH (1998) In: Thompson DL (ed) Modern methods for multidimensional dynamics computation in chemistry. World Scientific, Singapore, pp 143