P. Pulay, Mol. Phys. 17, 197 (1969)

However, experience indicates that even with the best wavefunctions available today the Hellmann–Feynman force differs considerably from the energy-derivative force. As Salem and Wilson have shown [12], in the case of a discrepancy the energy derivative is much more reliable. It is like the energy itself, stationary with respect to changes in the wavefunction. Therefore, the error in the wavefunction enters the energy derivative *only in second order*, but it enters the Hellmann–Feynman force in first order.

Several numerical tests have been made ([13–15], see also Part II of this series) which all show how extraordinarily sensitive the Hellmann–Feynman forces are. Kern and Karplus [14] calculated the Hellmann–Feynman forces for hydrogen fluoride with the quite good wavefunction of Clementi (16 Slater-type orbitals) and obtained 0·78 A.U. = 6·42 mdyn for the rest force on fluorine at the experimental H–F distance. This force would shift the interatomic distance by about 0·7 Å. Thus the Hellmann–Feynman force is practically of no value in determining equilibrium configurations unless very good wavefunctions are used. The better the wavefunction is, however, the more accurate is the energy derivative compared with the Hellmann–Feynman force, because of its second-order dependence on the error in the wavefunction. Considering the quickly increasing computational labour as the wavefunction is made more and more accurate, it is much more advantageous to work with a medium-accuracy function and the energy derivative than with a very accurate function and the Hellmann–Feynman force.

If the Hellmann–Feynman theorem does not hold, another difficulty arises if one tries to calculate force constants from the Hellmann–Feynman forces. Because the sum of the forces on the atoms does not necessarily vanish for the Hellmann–Feynman force as it does for the energy derivative force, the force constants will depend on how the internal coordinates are defined with respect to the external ones (translations and rotations of the molecule as a whole), e.g. force constants of a diatomic molecule will depend on whether the force on atom A or B was used to calculate it.

We conclude that quantities derived from the Hellmann–Feynman force are generally unreliable and should not be used, unless it is assured that the Hellmann–Feynman theorem holds. Examples show that this is not true for present-day wavefunctions. An exception would be the use of floating basis functions [10].

4. Transformation from cartesian to internal forces

Let N be the number of atoms and n the number of internal coordinates in a symmetry species. $\mathbf{x}=(x_1, x_2, \ldots, x_{3N})^+$ denotes the cartesian coordinates, $\mathbf{f}=[-(\partial E/\partial x_1), \ldots, -(\partial E/\partial x_{3N})]^+$ the corresponding forces $\mathbf{q}=(q_1, q_2, \ldots, q_n)^+$ the internal coordinates and $\boldsymbol{\varphi}=[-(\partial E/\partial q_1), \ldots, -(\partial E/\partial q_n)]^+$ the internal forces. The internal displacements $d\mathbf{q}$ are related to the cartesian ones by the well-known $\mathbf{B}$ matrix of Wilson and Eliashevich [16]:

$$d\mathbf{q}=\mathbf{B}\,d\mathbf{x}.$$

Thus for the forces:

$$\mathbf{f}=\mathbf{B}^+\boldsymbol{\varphi}. \tag{8}$$

From (8), $\boldsymbol{\varphi}$ cannot be expressed directly, $\mathbf{B}^+$ being a rectangular matrix and having no inverse. We can, however, construct a set of matrices, all denoted symbolically by $\mathbf{B}^{+-1}$ for which

$$\mathbf{B}^{+-1}\mathbf{B}^+ \quad \mathbf{I},$$

where **I** is the n by n unit matrix. Multiplying (8) by $\mathbf{B}^{+-1}$ from the left, we obtain the internal forces:

$$\boldsymbol{\varphi}=\mathbf{B}^{+-1}\mathbf{f}.$$

The matrices $\mathbf{B}^{+-1}$ can be obtained by:

$$\mathbf{B}^{+-1}=(\mathbf{BmB}^{+})^{-1}\mathbf{Bm},$$

where $\mathbf{m}$ is any $3N$ by $3N$ matrix which has the full symmetry of the molecule so that ($\mathbf{BmB}^{+}$) is non-singular†. The fact that an infinite set of matrices $\mathbf{B}^{+-1}$ exists is a consequence of the non-uniqueness of the transformation from internal coordinates to cartesians.

5. DETERMINATION OF THE EQUILIBRIUM GEOMETRY

In determining the equilibrium nuclear configuration we can restrict ourselves to the totally symmetric (A_1) symmetry species if the symmetry of the equilibrium configuration is known. Let $\mathbf{q}_0$ be the starting vector of internal coordinates and $\mathbf{F}_0$ a guess for the force constant matrix. To obtain a better approximation $\mathbf{q}_1$ we calculate the forces $\boldsymbol{\varphi}_0$ and form:

$$\mathbf{q}_1=\mathbf{q}_0+\Delta\mathbf{q}=\mathbf{q}_0+\mathbf{F}_0^{-1}\boldsymbol{\varphi}_0.$$

It is evident that if $\mathbf{F}_0$ were equal to the calculated force constant matrix and if the forces $\boldsymbol{\varphi}$ depended linearly on the coordinates q, then $\mathbf{q}_1$ would give the final equilibrium geometry. If this is not entirely true, $\mathbf{q}_1$ will differ from its equilibrium value. Nevertheless, if the guesses $\mathbf{F}_0$ and $\mathbf{q}_0$ are not too far from the correct values, a considerable improvement can be expected in $\mathbf{q}_1$ compared with $\mathbf{q}_0$. The iteration process:

$$\mathbf{q}_{i+1}=\mathbf{q}_i+\mathbf{F}_0^{-1}\boldsymbol{\varphi}_i,$$

will then converge towards equilibrium. Experience indicates that three or four steps are usually sufficient to achieve the final configuration. Suitable guesses $\mathbf{q}_0$ and $\mathbf{F}_0$ are almost always available. Note that

(*a*) this process does not depend on how many coordinates fall in the A_1 species.
(*b*) the final equilibrium configuration does not depend on F_0. F_0 determines only the rate of convergence.

The possibility of determining economically the equilibrium geometry is one of the most attractive features of the force method.

6. WHICH FORCE CONSTANTS SHOULD BE CALCULATED *ab initio*?

Let us distinguish between the following types of force constants in internal valence coordinates:

---

† There are two reasonable choices for $\mathbf{m}$. First, we can choose $\mathbf{m}=\mathbf{M}^{-1}$ where $\mathbf{M}$ is a diagonal matrix consisting of triplets of the atomic masses. As Crawford and Fletcher have shown [17], the centre of gravity of the molecule as well as its orientation will be left unchanged on a displacement dq with this choice. However, for our purposes it seems to be more appropriate if some of the cartesians are kept unchanged. The number of cartesians which can be kept constant simultaneously is equal to the number of translations and rotations in the symmetry species; for a whole molecule it is six (five for a linear molecule). $\mathbf{m}$ is then a diagonal matrix with zeros in the diagonal for those cartesians which have to be kept constant and with arbitrary non-zero diagonal elements, usually 1's, elsewhere.

P. Pulay in: Applications of Electronic Structure Theory, Vol. 3, H.F. Schaefer III, Ed., Plenum Press 1977, p. 135

Table 1. Geometry and Force Constants of the Hydrogen Molecule Near the Hartree–Fock Limit (7s4p2d/2b Gaussian Basis Set)^a

R	1.3	1.35	1.4	1.45	1.5
E	−1.1318488	−1.1332234	−1.1334654	−1.1327485	−1.1312196
dE/dR	0.0401356	0.0155367	−0.0052808	−0.0229076	−0.0378342

Property	Points used	Fitting	From force	From energy
R_e	All	Quartic	1.386507	1.3865
$(dE/dR)_{1.4}$	All	Quartic	−0.0052808	−0.005283
$(d^2E/dR^2)_{1.4}$	All	Quartic	0.382641	0.38267
$(d^3E/dR^3)_{1.4}$	All	Quartic	−1.27294	−1.2824
$(d^4E/dR^4)_{1.4}$	All	Quartic	4.3248	4.288
$(d^2E/dR^2)_{1.4}$	1.35, 1.45	Linear	0.38444	—
$(d^3E/dR^3)_{1.4}$	1.35, 1.4, 1.45	Quadratic	−1.2763	—

[a] In atomic units.

ment that for a given exponent and n, l quantum numbers, all functions with the possible m quantum numbers should be included. This requirement can be reduced to a subgroup of the rotational group only if the molecule preserves a certain symmetry for all geometries investigated. For example, in a planar π system different exponents may be used for p_σ and p_π functions as long as the system remains strictly planar. Note that in a Gaussian lobe basis set, rotational invariance is only approximately fulfilled, even if the basis set contains all functions to a given l quantum number and exponent. This may cause a very small net torque.(56)

Recently application of the gradient method to semiempirical calculations of the CNDO type has become popular.(17,57–61) It is appropriate to note that gradient calculation offers even more advantages in semiempirical than in *ab initio* work. Integral evaluation, which must be redone for gradient calculation, is only a minor time consumer in CNDO, resulting in gradients at virtually no extra cost.(57,59)

2.6. Transformation of Cartesian Forces and Force Constants to Internal Coordinates

Programming considerations dictate the calculation of the forces first in Cartesian coordinates. However, the relevant quantities are the forces and force constants in internal coordinates, typically in internal valence coordinates, i.e., bond lengths and angles, out-of-plane angles are dihedral angles, or a linear combination of them. This section deals with this transformation.(8)

Let $\mathbf{X} = (X_1, \ldots, X_{3N})^+$ denote the set of Cartesian displacement coordinates from a reference configuration, and $\mathbf{q} = (q_1, \ldots, q_M)^+$ the chosen set of internal displacement coordinates. The forces are defined in internal coordinates as $\varphi_i = (\partial E/\partial q_i)_0$ with all q_j, $j \neq i$, fixed. It is clear from this definition that *the internal forces are unambiguously defined only if the internal coordinates form a complete but nonredundant set*, requiring $M = 3N - 6$ ($3N - 5$ for linear molecules). In general, it makes sense to speak of an internal force φ_i only if all coordinates q_j are defined and independent.

Let us write the energy as a quadratic function of $\mathbf{X}$ and $\mathbf{q}$:

$$E = E_0 - \sum_a f_a X_a + \tfrac{1}{2} \sum_{a,b} K_{ab} X_a X_b = E_0 - \mathbf{f}^+\mathbf{X} + \tfrac{1}{2}\mathbf{X}^+\mathbf{K}\mathbf{X} \tag{30}$$

and

$$E = E_0 - \sum_i \varphi_i q_i + \tfrac{1}{2} \sum_{i,j} F_{ij} q_i q_j = E_0 - \boldsymbol{\varphi}^+\mathbf{q} + \tfrac{1}{2}\mathbf{q}^+\mathbf{F}\mathbf{q}$$

where $\mathbf{f}$ and $\boldsymbol{\varphi}$ are the column vectors of the forces, and $\mathbf{K}$ and $\mathbf{F}$ are the harmonic force constant matrices in Cartesian and internal coordinates, respectively. Deviating from the usual procedure, let us retain the quadratic terms in the expression of the internal coordinates by the Cartesian ones:

$$q_i = \mathbf{B}_i\mathbf{X} + \tfrac{1}{2}\mathbf{X}^+\mathbf{C}^i\mathbf{X} \tag{32}$$

where $\mathbf{B}_i$ is the ith row of the matrix which relates $\mathbf{q}$ and $\mathbf{X}$ to first order,(62) $\mathbf{q} = \mathbf{BX}$; $\mathbf{B}$ can be easily determined.(63) Substituting (32) into (31) and comparing the coefficients of like terms we obtain

$$\mathbf{f} = \mathbf{B}^+\boldsymbol{\varphi}$$

and

$$\mathbf{K} = \mathbf{B}^+\mathbf{F}\mathbf{B} - \sum_i \varphi_i \mathbf{C}^i$$

We cannot directly express $\boldsymbol{\varphi}$ and $\mathbf{F}$ from (33) and (34) because $\mathbf{B}^+$, being a $3N$ by M matrix, has no inverse. However, a set of $M \times 3N$ matrices can be defined, all symbolically denoted by $\mathbf{B}^{\pm 1}$, having the property

$$\mathbf{B}^{\pm 1}\mathbf{B}^+ = \mathbf{I}_M \tag{35}$$

Indeed, if $\mathbf{m}$ is any matrix for which $(\mathbf{BmB}^+)$ is not singular* then

$$\mathbf{B}^{\pm 1} \quad (\mathbf{BmB}^+)^{-1}\mathbf{Bm}$$

satisfies (35). Multiplying (33) and (34) by $\mathbf{B}^{\pm 1}$ we get

$$\boldsymbol{\varphi} = \mathbf{B}^{\pm 1}\mathbf{f}$$

and

$$\mathbf{F} = \mathbf{B}^{\pm 1}\mathbf{K}\mathbf{B}^{-1} + \sum \varphi_i \mathbf{B}^{\pm 1}\mathbf{C}^i\mathbf{B}^{-1}$$

where $\mathbf{B}^{-1}$ is the transpose of (36).

*If $\mathbf{BmB}^+$ is singular for every m then the transformation to internal coordinates is singular and the chosen set of internal coordinates is not appropriate. This is the case, e.g., if we try to describe the out of plane distortion of the BF_3 molecule by the sum of the three FBF angles.

By explicitly including the six translational and rotational coordinates in the set $\mathbf{q}$, it can be shown that any of the matrices (36) yields the same internal forces if the sum and torque of the Cartesian forces vanish. They also yield the same internal force constants if $\mathbf{K}$ obeys the translational and rotational invariance condition. The simplest choice of $\mathbf{m}$ is obviously the unit matrix $\mathbf{I}_{3N}$. Note that ***B** must be calculated at the actual nuclear configuration.* This condition corresponds to the use of curvilinear internal coordinates instead of rectilinear ones. In general, curvilinear coordinates are preferable for two reasons: (a) the anharmonic force field is much more diagonal in curvilinear valence coordinates; and (b) serious difficulties are encountered in the calculation of force constants in rectilinear coordinates at nonequilibrium reference geometries (see Section 3.2.1).

The second term on the right-hand side of (38) is usually neglected because it is in the equilibrium geometry that the second derivatives of the energy are of main interest. However, as shown in Section 3, in quantum chemical calculations it may be preferable to determine the force constants at the experimental geometry, where the *theoretical* φ_i are usually not zero. In these cases the full equation (38) must be used. For example, consider a diatomic molecule AB which is initially oriented along the z axis, with forces $f_A = -f_B$ acting on the nuclei. The Cartesian force constant for the coordinate X_A is clearly nonvanishing: $K_{XX} = -f_A/(Z_A - Z_B)$. Introducing the new coordinates R_{AB} and β, the angle of the molecule with the z axis in the xz plane, the first term in (38) yields a spurious nonzero rotational force constant $F_{\beta\beta} = R_{AB}^2 K_{XX}$. The second term, however, exactly compensates for this.* Note that Cartesian and internal force constants may lead to different harmonic vibrational frequencies if they are not calculated at the theoretical equilibrium geometry. This represents an inherent ambiguity in the definition of harmonic vibrational frequencies at a nonequilibrium geometry. The fact that a diatomic molecule may have nonzero rotational frequency in Cartesian coordinates, but not in valence coordinates, is a strong argument in favor of the latter.

3. Applications

3.1. Molecular Geometries and Reaction Paths

The gradient greatly facilitates the determination of molecular geometry. If a reasonable approximation F_0 to the *theoretical* force constant matrix is

---

*Recently Thomsen and Swanstrøm(64) have attributed the difficulties in the transformation of Cartesian force constants (the apparent violation of rotational invariance) to cubic and quartic terms in the potential energy. This is not correct because these terms cannot contribute to the quadratic terms in internal coordinates, provided the transformation is smooth and nonsingular. Their difficulties are due to the neglect of the second term in (38).

By explicitly including the six translational and rotational coordinates in the set $\mathbf{q}$, it can be shown that any of the matrices (36) yields the same internal forces if the sum and torque of the Cartesian forces vanish. They also yield the same internal force constants if $\mathbf{K}$ obeys the translational and rotational invariance condition. The simplest choice of $\mathbf{m}$ is obviously the unit matrix $\mathbf{I}_{3N}$. Note that ***B** must be calculated at the actual nuclear configuration.* This condition corresponds to the use of curvilinear internal coordinates instead of rectilinear ones. In general, curvilinear coordinates are preferable for two reasons: (a) the anharmonic force field is much more diagonal in curvilinear valence coordinates; and (b) serious difficulties are encountered in the calculation of force constants in rectilinear coordinates at nonequilibrium reference geometries (see Section 3.2.1).

The second term on the right-hand side of (38) is usually neglected because it is in the equilibrium geometry that the second derivatives of the energy are of main interest. However, as shown in Section 3, in quantum chemical calculations it may be preferable to determine the force constants at the experimental geometry, where the *theoretical* φ_i are usually not zero. In these cases the full equation (38) must be used. For example, consider a diatomic molecule AB which is initially oriented along the z axis, with forces $f_A = -f_B$ acting on the nuclei. The Cartesian force constant for the coordinate X_A is clearly nonvanishing: $K_{XX} = -f_A/(Z_A - Z_B)$. Introducing the new coordinates R_{AB} and β, the angle of the molecule with the z axis in the xz plane, the first term in (38) yields a spurious nonzero rotational force constant $F_{\beta\beta} = R_{AB}^2 K_{XX}$. The second term, however, exactly compensates for this.* Note that Cartesian and internal force constants may lead to different harmonic vibrational frequencies if they are not calculated at the theoretical equilibrium geometry. This represents an inherent ambiguity in the definition of harmonic vibrational frequencies at a nonequilibrium geometry. The fact that a diatomic molecule may have nonzero rotational frequency in Cartesian coordinates, but not in valence coordinates, is a strong argument in favor of the latter.

3. *Applications*

3.1. Molecular Geometries and Reaction Paths

The gradient greatly facilitates the determination of molecular geometry. If a reasonable approximation F_0 to the *theoretical* force constant matrix is

---

*Recently Thomsen and Swanstrøm(64) have attributed the difficulties in the transformation of Cartesian force constants (the apparent violation of rotational invariance) to cubic and quartic terms in the potential energy. This is not correct because these terms cannot contribute to the quadratic terms in internal coordinates, provided the transformation is smooth and nonsingular. Their difficulties are due to the neglect of the second term in (38).