
Deductive Molecular Mechanics of sp^3 Carbon Atom

A. L. TCHOUGRÉEFF, A. M. TOKMACHEV

*Karpov Institute of Physical Chemistry, 10 Vorontsovo Pole, Moscow 105064, Russia and
Center for Computational Chemistry at the Keldysh Institute for Applied Mathematics of RAS,
Moscow, Russia*

Received 7 August 2002; accepted 21 March 2002

DOI 10.1002/qua.10618

ABSTRACT: The problem of substantiation of additive systematics such as molecular mechanics is considered. Antisymmetrized product of strictly localized geminals (APSLG) trial wavefunction is used as a starting point for this substantiation. The main force fields of molecular mechanics—those of bonding and bending—are derived for sp^3 carbon from the analysis of the total energy in semiempirical variant of the APSLG scheme. Analytically obtained constants of these force fields are in perfect coincidence with those typically used in the empirical force fields. The formulae for off-diagonal force fields are also obtained and analyzed. The application of the results for understanding the organic stereochemistry is discussed. © 2003 Wiley Periodicals, Inc. *Int J Quantum Chem* 96: 175–184, 2004

Key words: antisymmetrized product of strictly localized geminals; bending; bonding; deductive molecular mechanics; off-diagonal force fields; sp^3 carbon atom

Introduction

Molecular mechanics (MM) [1] is normally considered as a purely empirical scheme based on representation of the total energy as a sum of contributions (force fields) that are explicit functions of molecular geometry. The form of these

force fields is usually taken as one or two first nonzero terms in the Taylor expansion for the energy. The parameters of the force fields are adjusted to reproduce the experimental (or sometimes quantum chemical) data on the heats of formation, equilibrium geometries, and sometimes also vibration frequencies etc. By construction the MM methodology is inapplicable to essentially quantum situations and to the molecules where electronic delocalization or/and correlation are important. Nevertheless, MM remains an extremely useful tool of computational chemistry because of its very low computational costs and high quality of results on molecular geometry and heats of formation, which

Correspondence to: A. Tchougréeff; e-mail: andrei@cc.nifhi.ac.ru

Contract grant sponsor: RAS.

Contract grant number: 6-120.

Contract grant sponsor: Haldor Topsøe A/S.

typically exceed even those obtained by the high-level quantum chemical methods.

Despite its wide applicability, the MM approach lacks any reliable theoretical substantiation. The necessity of such substantiation is caused not only by purely theoretical discomfort but also by strong methodological demand, because the hybrid quantum mechanical/molecular mechanical (QM/MM) schemes combining the advantages of the QM and MM approaches have been actively developing during the last decade. Definition of junction between quantum and classical subsystems on the solid theoretical basis requires some QM description underlying the MM one [2]. The problem of substantiation of MM is also close to that of theoretical understanding of organic stereochemistry. Explanatory concepts in this area [3] are somehow guessed but are not derived from any reliable QM description of molecular electronic structure and thus deserve theoretical substantiation. The same is true for additive systematics for the heats of formation of organic molecules [4].

Reported in the literature were some attempts to understand the applicability of classical additive schemes from the QM principles. The authors of Ref. [5] applied the Perturbative Configuration Interaction on Local Orbitals (PCILO) method [6] as the underlying QM scheme for additive systematics. At the same time, the nonvariational nature of local one-electron states and amplitudes of two-electron configurations employed in Ref. [5] has not allowed derivation of the analytical form of the main force fields—those of stretching and bending. Recently, we proposed an alternative approach leading to a generic MM scheme derived from the QM description [7]. It is based on a semiempirical implementation of the trial wavefunction taken in the form of antisymmetrized product of strictly localized geminals (APSLG) [8]. This implementation possesses some attractive features: correct behavior of the wavefunction at all intrabond separations, variational determination of basis one-electron states, reliable results on heats of formation and molecular structures, and linear scaling of computational costs. It allowed us [7] to construct a series of reliable noniterative (MM-like) schemes. At the same time, these schemes are purely numerical. We present an abridged analytical derivation of generic MM scheme for the case of sp^3 carbon atom, with special attention paid to explicit form of the force fields and to demonstration of their sources.

The article is organized as follows: in the next section we briefly review the APSLG implementation used; then the bond energy functions are considered;

in the following section we thoroughly analyze the bending force field as derived from the APSLG energy functional, after which the consequences of adjustment of the one-electron states due to geometry variation are considered, and, finally, the numerical results are given and discussed.

APSLG-MINDO/3 Method

The APSLG trial wavefunction has the form:

$$|\Phi\rangle = \prod_m g_m^+ |0\rangle, \quad (1)$$

where each bond geminal g_m^+ is constructed as a linear combination of three singlet two-electron configurations (two ionic and one covalent, or Heitler–London) with variable amplitudes:

$$g_m^+ = u_m r_{m\alpha}^+ r_{m\beta}^+ + v_m l_{m\alpha}^+ l_{m\beta}^+ + w_m (r_{m\alpha}^+ l_{m\beta}^+ + l_{m\alpha}^+ r_{m\beta}^+) \quad (2)$$

subject to normalization condition imposed $u_m^2 + v_m^2 + 2w_m^2 = 1$. In the case of lone-pair geminal, only one contribution survives:

$$g_m^+ = r_{m\alpha}^+ r_{m\beta}^+ \quad (3)$$

The basis one-electron states $|r_m\rangle$ and $|l_m\rangle$ assigned to the m -th geminal are taken as strictly local hybrid orbitals (HOs), that is, those obtained by a unitary $SO(4)$ transformation of four atomic orbitals for each “heavy” (nonhydrogen) atom:

$$t_{m\sigma}^+ = \sum_{i \in A} h_{mi}^A a_{i\sigma}^+ \quad (4)$$

The 4×4 hybridization matrices h^A are determined variationally. Geminal amplitudes Eq. (2) are in their turn determined by diagonalizing the effective bond Hamiltonians.

The electronic Hamiltonian used is that of the MINDO/3 type [9] with resonance parameters β^{AB} slightly readjusted [8]. The core–core interaction in the MINDO/3 scheme is not purely Coulomb but is modified by a short-range repulsion term:

$$E_{nn} = \frac{1}{2} \sum_{\alpha \neq \beta} C_{\alpha\beta}, \quad C_{\alpha\beta} = Z_\alpha Z_\beta (\gamma_{\alpha\beta} + D_{\alpha\beta}), \quad (5)$$

where $\gamma_{\alpha\beta}$ is the two-center Coulomb integral.

Bond Energy, Equilibrium Bond Length, and Bond Stretching Constant

The APSLG-MINDO/3 energy can be rewritten in a form close to that of the MM (with different bonding and nonbonding contributions, 1–3 bond interaction contributions) as a sum of increments:

$$\begin{aligned}
 E_A &= \sum_{t_m \in A} [2U_m^{tt} P_m^{tt} + (t_m t_m | t_m t_m)^T \Gamma_m^{tt}] \\
 &\quad + 2 \sum_{\substack{t_k t'_m \in A \\ k < m}} g_{t_k t'_m}^{T_k} P_k^{tt} P_m^{t't'} \\
 E_{R_m L_m}^{bond} &= 2\gamma_{R_m L_m} [\Gamma_m^{rl} - 2P_m^{rr} P_m^{ll}] - 4\beta_{r_m l_m}^{R_m L_m} P_m^{rl} \\
 E_{AB}^{nonbond} &= Q_A Q_B \gamma_{AB}. \quad (6)
 \end{aligned}$$

Here U_m^t is the matrix element of attraction of electron to its own core, $\beta_{r_m l_m}^{R_m L_m}$ is the intrabond resonance integral, Q_A is the Mulliken charge on the atom A , and $g_{t_k t'_m}^{T_k} = 2(t_k t_k | t'_m t'_m)^{T_k} - (t_k t'_m | t'_m t_k)^{T_k}$ is the reduced Coulomb integral.

The energy expression Eq. (6) depends on the electronic structure parameters (ESPs) of two types. First are the elements of density matrices

$$\begin{aligned}
 P_m^{t't'} &= \langle 0 | g_m^+ t_{m\sigma}^+ t'_{m\sigma} g_m^+ | 0 \rangle, \\
 \Gamma_m^{t't'} &= \langle 0 | g_m^+ t_{m\beta}^+ t'_{m\alpha} t'_{m\alpha} t_{m\beta} g_m^+ | 0 \rangle, \\
 P_m^{rr} &= u_m^2 + w_m^2, \quad P_m^{ll} = v_m^2 + w_m^2, \\
 P_m^{rl} &= P_m^{lr} = (u_m + v_m)w_m, \\
 \Gamma_m^{rr} &= u_m^2, \quad \Gamma_m^{ll} = v_m^2, \quad \Gamma_m^{rl} = \Gamma_m^{lr} = w_m^2, \quad (7)
 \end{aligned}$$

where t and t' are either r or l . Second are the matrices transforming the AOs to the $|t\rangle$ one-electron states that describe the hybridization. (The energy depends on the latter through molecular integrals.) It was shown in Ref. [10] by analysis of the bond-related contributions to the energy that there exist two parameters (functions of bond interatomic separation) for each bond ζ_m^{-1} and μ_m describing, respectively, intrabond correlation and bond asymmetry. It was shown that for all chemical bonds these parameters are small. Therefore, it is possible to describe transferability of the elements of density matrices by expression of them as power series on these parameters. It was shown [10] that the contribution of the zero-th order in μ_m can be written as:

$$\Gamma_m^{t't'} = \frac{1}{4} \left(1 - tt' \frac{1}{\Gamma(\zeta_m)} \right), \quad P_m^{tt} = \frac{1}{2}, \quad P_m^{rl} = \frac{\zeta_m}{2\Gamma(\zeta_m)}, \quad (8)$$

where $\Gamma(\zeta_m) = \sqrt{1 + \zeta_m^2}$. At the equilibrium intrabond separations the limit $\zeta_m \rightarrow \infty$ is valid. In this limit

$$\Gamma_m^{t't'} = \frac{1}{4} \left(1 - tt' \frac{1}{\zeta_m} \right), \quad P_m^{rl} = \frac{1}{2} \left(1 - \frac{1}{32\zeta_m^2} \right). \quad (9)$$

Taking the contributions of up to second order in μ_m , the density matrix elements become slightly modified [10]:

$$\begin{aligned}
 \Gamma_m^{rl} &= \Gamma_m^{rl} \left[1 + \mu_m^2 \frac{(2\Gamma(\zeta_m) + 1)(1 - \Gamma(\zeta_m))}{2(\Gamma(\zeta_m) + 1)^2} \right], \\
 P_m^{tt} &= P_m^{tt} \left[1 + t\mu_m \frac{\Gamma(\zeta_m) - 1}{\Gamma(\zeta_m) + 1} \right], \\
 P_m^{rl} &= P_m^{rl} \left[1 + \mu_m^2 \frac{2\Gamma(\zeta_m) + 1 - \Gamma^2(\zeta_m)}{2(\Gamma(\zeta_m) + 1)^2} \right], \quad (10)
 \end{aligned}$$

where subscript 0 corresponds to the estimates by Eq. (8). The formulae Eqs. (9) and (10) allow for the conclusion that the bond order ($2P_m^{rl}$) is transferable up to second order with respect to both parameters ζ_m^{-1} and μ_m ; the bond covalency ($2\Gamma_m^{rl}$) is transferable up to second order with respect to μ_m and up to first order with respect to ζ_m^{-1} ; the bond polarity ($P_m^{rr} - P_m^{ll}$) is transferable up to first order with respect to both ζ_m^{-1} and μ_m . The transferability of bond orders is used in the following considerations. The limit $\zeta_m \rightarrow \infty$, $\mu_m \rightarrow 0$ can be considered as a good approximation in the case of nonpolar bonds (C–H, N–H, C–C, N–C, etc.) for the respective bond lengths close to equilibrium. It gives the following estimates:

$$P_m^{t't'} = \frac{1}{2}, \quad \Gamma_m^{t't'} = \frac{1}{4}. \quad (11)$$

This result can be also obtained from the self-consistent field (SCF) consideration of symmetric bond. Such a picture can be termed "FA," fixed (geminal) amplitudes, to contrast it to the more refined picture of Eqs. (8), (9), and (10), which can be termed as "TA," tuned (geminal) amplitudes. These notations are in the spirit of those of Ref. [7].

From the general APSLG energy expression the equilibrium bond length can be deduced by equat-

ing to zero the first derivative of the bond-related energy. We consider the quite reliable case of non-polar bond Eq. (8). The derivatives of the gemina-related ESPs exactly compensate each other, and the bond energy minimum equation is:

$$\frac{\partial E}{\partial q} = Z_{R_m} Z_{L_m} \frac{\partial D_{R_m L_m}}{\partial q} - 2 \frac{\zeta_m}{\Gamma(\zeta_m)} \frac{\partial \beta_{r_{mlm}}^{R_m L_m}}{\partial q} - \frac{1}{2} \left(1 - \frac{1}{\Gamma(\zeta_m)} \right) \frac{\partial \gamma_{R_m L_m}}{\partial q} = 0. \quad (12)$$

In the limit $\zeta_m \rightarrow \infty$ we recover the equilibrium geometry condition for the FA picture. Note that if q is the intrabond distance the partial derivatives of $\beta_{r_{mlm}}^{R_m L_m}$ and $\gamma_{R_m L_m}$ are positive, which leads to attraction of the bonded atoms. The short-range repulsion is due to a special form of the core–core interaction.

The same concepts can be used to determine the elasticity constant for the bond stretching by taking the second derivative of the energy with respect to the bond length. In the FA picture we get:

$$k_{R_m L_m} = \left(Z_{R_m} Z_{L_m} \frac{d^2 D_{R_m L_m}}{dr_{R_m L_m}^2} - 2 \frac{\partial^2 \beta_{r_{mlm}}^{R_m L_m}}{\partial r_{R_m L_m}^2} - \frac{1}{2} \frac{d^2 \gamma_{R_m L_m}}{dr_{R_m L_m}^2} \right)_{r_{R_m L_m}^0} \quad (13)$$

Numerical results on this constant are given below.

Hybridization and Bending Force Field

This section describes in detail the role of hybridization in construction of the MM force fields. In general we consider a kind of linear response relation between variations of nuclear coordinates and ESPs characterizing HOs. Its general form can be readily obtained by expanding the energy up to the second order with respect to both atomic coordinates q and the ESPs x with subsequent minimization:

$$x - x_0 = -(\nabla_x \nabla_x E)^{-1} \nabla_x \nabla_q E (q - q_0) \quad (14)$$

It leads to the second-order energy expression:

$$E = E_0 + \frac{1}{2} (q - q_0) |\nabla_q \nabla_q E| (q - q_0) - \frac{1}{2} (q - q_0) |\nabla_q \nabla_x E (\nabla_x \nabla_x E)^{-1} \nabla_x \nabla_q E| (q - q_0) \quad (15)$$

To derive the bending force field we need to know the main source of angular dependence of the energy. In the APSLG approximation it is obviously determined by the structure of HOs. In Ref. [10] the one-center hybridization-dependent contribution to the energy was analyzed. It turned out that the one-center energy is hybridization-dependent only if subtle polarization and correlation effects are taken into account. This contradicts the fact that the correct hybridization can be reproduced numerically even by methods without intra-atomic electron–electron repulsion [11]. Therefore, we admit that the main source of hybridization is the resonance energy, which is in agreement with earlier concepts in this area [12].

HYBRIDIZATION DESCRIPTION

To consider the variations of the resonance energy we studied first the mathematical structure of hybridization. Each of the HOs centered at a given atom is a linear combination of the s - and p -AOs (atomic orbitals). One can note, however, that when a molecule is rotated as a whole, the expansion coefficients at the s -AO are invariant, whereas the coefficients at the p -AOs transform as if they were components of a 3-vector. Thus, the HOs can be represented as an object combining a scalar and a vector parts: (s_m, \vec{v}_m) [10]. This mathematical object is a normalized quaternion.

The ends of vectors \vec{v}_m form a hybridization tetrahedron containing full information about hybridization. As mentioned above, the specific hybridization is given by an $SO(4)$ matrix. These matrices form the six-parametric group, and the most direct way to parameterize the latter is to represent them as products of six Jacobi matrices performing rotations in two-dimensional subspaces by the angles $\omega_{sx}^A, \omega_{sy}^A, \omega_{sz}^A, \omega_{yz}^A, \omega_{xz}^A$, and ω_{xy}^A (subscripts refer to the two-dimensional subspace). The first triple of angles corresponds to formation of hybridization tetrahedron (a pseudorotation), whereas the second triple of angles corresponds to rotation of the hybridization tetrahedron as a whole (a quasirotation, where the prefix “quasi” signifies that no physical body actually rotates). The $SO(4)$ group generates all

possible hybridizations, thus being a dynamical group for the system of all possible HOs at a given atom [13].

The above parameterization of the $SO(4)$ group by the Jacobi matrices is inconvenient for analytical considerations because the matrix elements are clumsy combinations of trigonometrical functions. Therefore, we propose a more workable parameterization of the $SO(4)$ group by a pair of normalized quaternions similar to the quaternion parameterization of the $SO(3)$ group of three-dimensional rotations [13]. The quaternions (\mathbf{q} and \mathbf{p}) correspond to fictitious rotations defined by triples of angles:

$$\omega_{\gamma\pm} = \epsilon_{\alpha\beta\gamma} \omega_{\alpha\beta} \pm \omega_{s\gamma} \quad (16)$$

$$h = \begin{pmatrix} q_0 p_0 + q_1 p_1 + q_2 p_2 + q_3 p_3 & q_0 p_1 - q_1 p_0 - q_2 p_3 + q_3 p_2 & q_0 p_2 + q_1 p_3 - q_2 p_0 - q_3 p_1 & q_0 p_3 - q_1 p_2 + q_2 p_1 - q_3 p_0 \\ -q_0 p_1 + q_1 p_0 - q_2 p_3 + q_3 p_2 & q_0 p_0 + q_1 p_1 - q_2 p_2 - q_3 p_3 & -q_0 p_3 + q_1 p_2 + q_2 p_1 - q_3 p_0 & q_0 p_2 + q_1 p_3 + q_2 p_0 + q_3 p_1 \\ -q_0 p_2 + q_1 p_3 + q_2 p_0 - q_3 p_1 & q_0 p_3 + q_1 p_2 + q_2 p_1 + q_3 p_0 & q_0 p_0 - q_1 p_1 + q_2 p_2 - q_3 p_3 & -q_0 p_1 - q_1 p_0 + q_2 p_3 + q_3 p_2 \\ -q_0 p_3 - q_1 p_2 + q_2 p_1 + q_3 p_0 & -q_0 p_2 + q_1 p_3 - q_2 p_0 + q_3 p_1 & q_0 p_1 + q_1 p_0 + q_2 p_3 + q_3 p_2 & q_0 p_0 - q_1 p_1 - q_2 p_2 + q_3 p_3 \end{pmatrix} \quad (18)$$

Due to its algebraic structure, the $SO(4)$ matrix H close to the unity matrix also represents small variation of HOs in a vicinity of a given set of HOs. The first-order contribution, for example, is:

$$\delta^{(1)}H = \begin{pmatrix} 0 & p_1 - q_1 & p_2 - q_2 & p_3 - q_3 \\ q_1 - p_1 & 0 & -q_3 - p_3 & q_2 + p_2 \\ q_2 - p_2 & q_3 + p_3 & 0 & -q_1 - p_1 \\ q_3 - p_3 & -q_2 - p_2 & q_1 + p_1 & 0 \end{pmatrix} \quad (19)$$

and, analogously, the second-order contribution can be extracted. The first-order correction to the HOs in the quaternion form (s, \vec{v}), which appear when small quasirotations and pseudorotations $\delta\vec{\omega}_l$ and $\delta\vec{\omega}_b$ are applied to the system of HOs at a given atom, has a simple form:

$$\begin{aligned} \delta^{(1)}s &= -(\delta\vec{\omega}_b, \vec{v}), \\ \delta^{(1)}\vec{v} &= s\delta\vec{\omega}_b + \delta\vec{\omega}_l \times \vec{v}, \end{aligned} \quad (20)$$

where \times stands for the vector product of 3-vectors.

VALENCE ANGLES AND BENDING ENERGY

It has been known for almost half a century [12] that in general terms stereochemistry (form of the

As in the case of quaternion representation of rotations the components of the quaternions at hand are given by:

$$\begin{aligned} q_0 &= \cos \frac{\omega_+}{2}, & q_1 &= \frac{\omega_{+x}}{\omega_+} \sin \frac{\omega_+}{2}, \\ q_2 &= \frac{\omega_{+y}}{\omega_+} \sin \frac{\omega_+}{2}, & q_3 &= \frac{\omega_{+z}}{\omega_+} \sin \frac{\omega_+}{2}, \\ p_0 &= \cos \frac{\omega_-}{2}, & p_1 &= \frac{\omega_{-x}}{\omega_-} \sin \frac{\omega_-}{2}, \\ p_2 &= \frac{\omega_{-y}}{\omega_-} \sin \frac{\omega_-}{2}, & p_3 &= \frac{\omega_{-z}}{\omega_-} \sin \frac{\omega_-}{2}. \end{aligned} \quad (17)$$

The $SO(4)$ matrix in terms of the quaternion components (the derivation will be published elsewhere) has the form:

coordination polyhedron) is determined by relation between the two-center energy, which favors population of excited and ionized states of an atom under consideration and the excitation and ionization energies themselves, which tend to keep an atom in its ground (unhybridized) neutral state. In the standard MM framework it is assumed that there is a characteristic angle to be maintained for each pair of incident chemical bonds (triple of atoms). These angles are chosen as totally independent parameters. At the same time it is known that only five of six angles (in the case of substituted carbon) are really independent [14]. This contradiction is important and should be taken into account when one tries to derive the MM scheme. Also the parameters characterizing the bending energy in the standard form

$$\frac{1}{2} k_{ABC} (\theta - \theta_0)^2 \quad (21)$$

are considered to be independent parameters. In the framework of the APSLG-derived MM-like scheme, the system of HOs at a given atom (hybridization tetrahedron) has a more rigid structure than does the coordination tetrahedron, because the form of hybridization tetrahedron is determined by only three parameters (pseudorotation angles $\vec{\omega}_b$). It

leads to important, purely algebraic consequences, which are considered below.

We demonstrate the general correspondence between the structure of the APSLG wavefunction and the corresponding energy functional on one hand and the MM force fields on the other hand. To do this we consider a characteristic example—the sp^3 carbon atom. For the sake of simplicity we give explicit formulae only for the simple case of hydride (methane); the generalization of the results to more complex cases is straightforward and will be published elsewhere. We characterize the stereochemistry of the carbon atom (the coordination tetrahedron) by a set of four unit vectors $\vec{e}_{R_m L_m}$, which have the carbon atom as their origin and point to the hydrogen atoms. In the diatomic coordinate frame (DCF), with the z -axis directed along the $\vec{e}_{R_m L_m}$ vector, the resonance integrals can be written as:

$$\beta_{r_m l_m}^{R_m L_m} = \beta_{\sigma\sigma}^{R_m L_m} h_{m\sigma}^{R_m} h_{m\sigma}^{L_m} + \beta_{\zeta\sigma}^{R_m L_m} h_{m\zeta}^{R_m} h_{m\sigma}^{L_m} \quad (22)$$

or

$$\beta_{r_m l_m}^{CH} = \beta_{\sigma\sigma}^{CH} s_m^C + \beta_{\zeta\sigma}^{CH} v_{m\zeta}^C \quad (23)$$

where

$$v_{m\zeta}^{T_m} = (\vec{v}_m^{T_m}, \vec{e}_{R_m L_m}) \quad (24)$$

(in the case of the hydrogen atom, the contribution of the s -orbital always equals unity).

Now we obtain the general equilibrium conditions for the hybridization tetrahedron. Taking the first-order variation of the resonance integrals Eq. (23) and making use of formulae Eq. (20) we get:

$$\delta^{(1)} \beta_{r_m l_m}^{CH} = -\beta_{\sigma\sigma}^{CH} (\delta \vec{\omega}_b, \vec{v}_m^C) + \beta_{\zeta\sigma}^{CH} [s_m^C (\delta \vec{\omega}_b, \vec{e}_{R_m L_m}) + (\delta \vec{\omega}_l \times \vec{v}, \vec{e}_{R_m L_m})] \quad (25)$$

From this equation the expression for pseudotorque \vec{N} and quasitorque \vec{K} can be easily obtained as a sum of contributions to the resonance energy for all four bonds, which are linear in $\delta \vec{\omega}_b$ and $\delta \vec{\omega}_l$, respectively:

$$\begin{aligned} \vec{N} &= 4 \sum_m P_m^r \{ \beta_{\sigma\sigma}^{CH} \vec{v}_m^C - \beta_{\zeta\sigma}^{CH} s_m^C \vec{e}_{R_m L_m} \}, \\ \vec{K} &= 4 \sum_m P_m^r \beta_{\zeta\sigma}^{CH} \vec{e}_{R_m L_m} \times \vec{v}_m^{R_m} \end{aligned} \quad (26)$$

The equilibrium condition for the HOs is then:

$$\vec{N} = \vec{K} = \vec{0}. \quad (27)$$

These conditions govern both the form and orientation of the hybridization tetrahedron. The solution of these equations in the case of symmetric hydride is obvious, because if we direct all vectors \vec{v}_m^C along vectors $\vec{e}_{R_m L_m}$ both equations are satisfied. We note that in the first equation two different contributions (from HOs \vec{v}_m^C and vectors $\vec{e}_{R_m L_m}$) vanish independently and that the second contribution is always zero when directions of HOs coincide with directions of bonds. Tetrahedral coordination and hybridization correspond to the energy minimum in the symmetric case. Meanwhile, corrections to this characteristic case can be considered as perturbations.

We consider now a situation in which parameters $\vec{\omega}_b$ are fixed by some reasons. It can be termed as fixed orbitals (FO) picture following the notations of Ref. [7], because the form of hybridization tetrahedron is fixed in this case. It is worth mentioning that if the form of the hybridization is fixed ($\vec{\omega}_b$), the one-center contributions to the energy are precisely $\vec{\omega}_l$ -independent, irrespective of the above assumptions concerning the geminal-related ESPs. The minimum of the resonance energy that is the only orientation-dependent contribution is achieved for some values of vectors $\vec{e}_{R_m L_m}$ (in the symmetric hydride case they are directed to the vortices of tetrahedron). The angular dependence of the energy (bending) can be described by introducing small rotation vectors $\delta \vec{\varphi}_m$, which after applying them to vectors $\vec{e}_{R_m L_m}$ lead to new (distorted) coordination tetrahedron:

$$\begin{aligned} (\vec{e}_{R_m L_m})' &= \vec{e}_{R_m L_m} + \delta \vec{\varphi}_m \times \vec{e}_{R_m L_m} \\ &+ \frac{1}{2} (\delta \vec{\varphi}_m \otimes \delta \vec{\varphi}_m - \mathcal{I} \delta \vec{\varphi}_m^2) \vec{e}_{R_m L_m} \\ &= \vec{e}_{R_m L_m} + \delta \vec{\varphi}_m \times \vec{e}_{R_m L_m} \\ &+ \frac{1}{2} (\delta \vec{\varphi}_m (\delta \vec{\varphi}_m, \vec{e}_{R_m L_m}) - \delta \vec{\varphi}_m^2 \vec{e}_{R_m L_m}) \end{aligned} \quad (28)$$

These vectors must be inserted in Eq. (24), and the elasticity constant can be obtained by extracting the second-order contribution in vectors $\delta \vec{\varphi}_m$. In the case of hydride:

$$\delta_{\vec{\varphi}_m \vec{\varphi}_m}^{(2)} E = -2\delta_{mm'} P_m^{rl} \beta_{\zeta\sigma}^{CH} \times \{(\vec{e}_{R_m L_m}, \delta\vec{\varphi}_m)(\delta\vec{\varphi}_m, \vec{v}_m^C) - \delta\vec{\varphi}_m^2(\vec{v}_m^C, \vec{e}_{R_m L_m})\}. \quad (29)$$

This equation is remarkable, because it shows that in the FO picture there are no contributions to the bending that can be attributed to interbond interaction. The bending force field is produced by energies of separate chemical bonds.

Typically in the MM framework the increment from the bending is considered a quadratic function of valence angles. The formula for bending Eq. (29) can be rewritten in this form. Variation of the valence angle $\theta_{mm'}$ with $m < m'$ results in rotations of the involved vectors around the axis orthogonal to the both coordination tetrahedron vectors:

$$\delta\vec{\varphi}_m = -\frac{\delta\theta_{mm'}}{2} \frac{\vec{e}_{R_m L_m} \times \vec{e}_{R_{m'} L_{m'}}}{|\vec{e}_{R_m L_m} \times \vec{e}_{R_{m'} L_{m'}}|}; \quad \delta\vec{\varphi}_{m'} = -\delta\vec{\varphi}_m. \quad (30)$$

After substitution of this expression to the second-order expansion Eq. (29) and after significant simplifications based on vector algebra, we obtain that the bending force field constant can be written as:

$$k_{HCH} = \beta_{\zeta\sigma}^{CH} \{P_m^{rl}(\vec{v}_m^C, \vec{e}_{R_m L_m}) + P_{m'}^{rl}(\vec{v}_{m'}^C, \vec{e}_{R_{m'} L_{m'}})\}, \quad (31)$$

which is a sum of two separate single-bond contributions.

According to Eq. (15) some modification of the bending expression Eq. (31) is expected, because of adjustment of hybridization tetrahedron to the geometry (coordination polyhedron) deformation. In the FO picture, which we now consider, this modification can only be due to overall quasirotation of the hybridization tetrahedron. All the angular deformations produced by the four rotation vectors $\delta\vec{\varphi}_m$ form an eight-dimensional space (two components for each vector, because the component collinear to the polyhedron vector can be set to zero). Likewise, the space of quasirotations is three-dimensional and thus cannot accommodate all possible angular deformations. From an algebraic point of view the linear mapping Eq. (14) in this case:

$$\delta\vec{\omega}_l = A\delta\vec{\varphi} = -(\nabla_{\vec{\omega}}^2 E)^{-1} \sum_m (\nabla_{\vec{\omega}_l} \nabla_{\vec{\varphi}_m} E) \delta\vec{\varphi}_m \quad (32)$$

maps an eight-dimensional space to a three-dimensional one, which means after using:

$$\dim \ker A + \dim \text{im } A = 8, \quad (33)$$

that the transformation A has at least a five-dimensional kernel, that is, a five-dimensional subspace of the rotations space spanned by $\{\delta\vec{\varphi}_m | m = 1 \div 4\}$, which results in a zero rotation of the system of the HOs. This kernel can be determined as a subspace orthogonal to the image of the operator A . The structure of the image subspace can be easily determined, because it is obvious that the rotations of a molecule as a whole

$$\Phi = \sum_m \delta\vec{\varphi}_m \quad (34)$$

lead to the equivalent quasirotations of the set of HOs. This result can be directly applied to finding the correction to Eq. (31). As one can see, for whatever values of $\delta\theta_{mm'}$ the resulting deformations are all orthogonal to the overall rotation of the system as a whole (correspond to a kernel of operator A Eq. (32)), because

$$\delta\vec{\varphi}_{m'} + \delta\vec{\varphi}_m = 0, \quad (35)$$

that is, any deformation of valence angles cannot result in quasirotation of HOs and Eq. (31) is not corrected because of adjustment of HOs in the FO picture, and thus the only source of the bending energy is the angular dependence of the resonance integrals $\beta_{r_{mm}^{CH}}$ for the bonds involved in the bending.

ADJUSTMENT OF FORM OF HYBRIDIZATION TETRAHEDRON

Let us now consider the picture that allows for adjustment of both orientation and form of hybridization tetrahedron. It can be termed, "tuned orbitals" (TO), in accordance with the notations of Ref. [7]. To this end we systematically estimate the second derivatives of the energy with respect to variables q and x . Generally, the second-order correction to the energy due to variation of HOs can be written as:

$$\delta_{\omega\omega}^{(2)} E = -4 \sum_m P_m^{rl} \delta_{\omega\omega}^{(2)} \beta_{r_{ml}^{CH}} \quad \text{where}$$

$$\delta_{\omega\omega}^{(2)} \beta_{r_{ml}^{CH}} = \beta_{\sigma\sigma}^{CH} \delta^{(2)} s_m^C + \beta_{\zeta\sigma}^{CH} (\vec{e}_{R_m L_m}, \delta^{(2)} \vec{v}_m^C). \quad (36)$$

This expression significantly simplifies for a symmetric case, where all s_m^C are the same (1/2); the directions of all HOs coincide with those of the bonds, i.e., $\vec{v}_m^C = (\sqrt{3}/2)\vec{e}_{R_m L_m}$; and the bond orders for all bonds and the resonance integrals in the

corresponding DCFs coincide. Taking these relations into account, we show [10] that the second-order energy expansion with respect to small quairotation and pseudorotation angles for the symmetric hydride (methane) takes the form:

$$\delta_{\omega\omega}^{(2)}E = 4P^{rl} \left[\frac{2}{\sqrt{3}} \beta_{\zeta\sigma}^{CH} \delta\tilde{\omega}_l^2 + \left(\beta_{\sigma\sigma}^{CH} + \frac{1}{\sqrt{3}} \beta_{\zeta\sigma}^{CH} \right) \delta\tilde{\omega}_b^2 \right], \quad (37)$$

which is diagonal to two eigenvalues that are triply degenerate.

To write down the necessary mixed second derivatives we consider the contribution to the energy of the second order with respect to the angles and the small variations of the resonance integrals:

$$\delta_{\beta\omega}^{(2)}\beta_{rmlm}^{CH} = -\delta\beta_{\sigma\sigma}^{CHm}(\delta\tilde{\omega}_b, \tilde{v}_m^C) + \delta\beta_{\zeta\sigma}^{CHm} s_m^C(\delta\tilde{\omega}_b, \tilde{e}_{RmLm}) - \delta\beta_{\zeta\sigma}^{CHm}(\delta\tilde{\omega}_l, \tilde{e}_{RmLm} \times \tilde{v}_m^C). \quad (38)$$

Now we consider the deformation of the hybridization tetrahedron due to angular distortions of molecular geometry. The required mixed second derivative can be easily written from Eqs. (28) and (38):

$$\begin{aligned} \nabla_{\tilde{\omega}_b} \nabla_{\tilde{\varphi}_m} E &= 4P_m^{rl} \beta_{\zeta\sigma}^{CH} s_m^C(\tilde{e}_{RmLm} \times) \\ \nabla_{\tilde{\omega}_l} \nabla_{\tilde{\varphi}_m} E &= 4P_m^{rl} \beta_{\zeta\sigma}^{CH}(\tilde{e}_{RmLm} \otimes \tilde{v}_m^C - (\tilde{v}_m^C, \tilde{e}_{RmLm})\mathcal{I}). \end{aligned} \quad (39)$$

The relation Eq. (14) in this case can be written as

$$\delta\tilde{\omega} = -(\nabla_{\tilde{\omega}}^2 E)^{-1} \sum_m (\nabla_{\tilde{\omega}} \nabla_{\tilde{\varphi}_m} E) \delta\tilde{\varphi}_m, \quad (40)$$

where a 6×6 matrix of the second derivatives with respect to angles coming from Eq. (37) is meant by $\nabla_{\tilde{\omega}}^2 E$. This linear operator maps eight-dimensional space of angular molecular deformation to a six-dimensional space of angular deformations of the HOs. Thus, there exists at least a two-dimensional subspace of angular deformations of a tetrahedron that affect neither orientation nor the form of the HOs. These deformations can be called "hybridization-incompatible." To single them out we notice that, as previously, the three-dimensional subspace of overall rotations maps to the three-dimensional subspace of pure HOs rotations. The $\nabla_{\tilde{\omega}}^2 E$ matrix is nondegenerate as is its reverse matrix. Therefore, the mixed second-derivative matrix has a kernel

mapping some of the coordination tetrahedron deformations to zero. It means that the five-dimensional subspace formed by independent distortions of valence angles form a two-dimensional kernel, which maps to the zero deformation of the HOs and a three-dimensional image, which maps to the three-dimensional space of the HO deformations $\delta\tilde{\omega}_b$. We perform direct calculation of these subspaces for the symmetric case, with the coordination tetrahedron vectors pointing to the octants. The following picture of the tetrahedron deformation can be obtained: if a valence angle increases by certain amount $\delta\theta_{1m}$ then the opposite (spiro) angle decreases by the same magnitude. Clearly, in a tetrahedron, one can choose three pairs of spiro angles for $m = 2 \div 4$. On the other hand it can be shown that the symmetric combinations of the spiro angles (when they increase or decrease simultaneously) all fall into the kernel of the $\sum_m (\nabla_{\tilde{\omega}_b} \nabla_{\tilde{\varphi}_m} E)$ operator, thus forming the two-dimensional subspace of the hybridization-incompatible deformations of the tetrahedron.

In the case of the hybridization-compatible rotations of the vectors the result of action of the operator reads:

$$\begin{aligned} \sum_m \nabla_{\tilde{\omega}_b} \nabla_{\tilde{\varphi}_m} E \{ \delta\tilde{\varphi}_m \} \\ = 4P^{rl} \beta_{\zeta\sigma}^{RL} \sqrt{\frac{2}{3}} (\delta\chi_{12}\vec{k} + \delta\chi_{13}\vec{j} + \delta\chi_{14}\vec{i}), \end{aligned} \quad (41)$$

and, therefore, the reaction of the form of hybridization tetrahedron on the angular distortions of molecular geometry can be calculated as:

$$\delta\tilde{\omega}_b = -\frac{\beta_{\zeta\sigma}^{CH}}{\sqrt{2} (\sqrt{3}\beta_{\sigma\sigma}^{CH} + \beta_{\zeta\sigma}^{CH})} (\delta\chi_{12}\vec{k} + \delta\chi_{13}\vec{j} + \delta\chi_{14}\vec{i}), \quad (42)$$

provided the parameters $\delta\chi_{1m}$ describe the hybridization-compatible deformations of the coordination tetrahedron. The above expression has a remarkable feature: the variation of the hybridization angles falls back as compared with the variation of the corresponding valence angles, because the coefficient relating the variation of geometry angle to the hybridization angle is less than unity.

The same considerations also apply to the bond stretching where variation of the resonance parameters can be presented as

$$\delta\beta_{\mu\nu}^{CH_m} = \gamma_{\mu\nu}^{CH} \delta r_{CH_m}. \quad (43)$$

It leads to the mixed second derivatives of the form:

$$\begin{aligned} \frac{\partial^2 \beta_{r_m l_m}^{CH_m}}{\partial \tilde{\omega}_b \partial r_{CH_m}} &= -\gamma_{\sigma\sigma}^{CH} \tilde{v}_m^C + \gamma_{\zeta\sigma}^{CH} S_m^C \tilde{e}_{R_m L_m}^C \\ \frac{\partial^2 \beta_{r_m l_m}^{CH_m}}{\partial \tilde{\omega}_l \partial r_{CH_m}} &= -\gamma_{\zeta\sigma}^{CH} \tilde{e}_{R_m L_m}^C \times \tilde{v}_m^C. \end{aligned} \quad (44)$$

Therefore, in the symmetric case, additional quasisrotation is absent because the vectors \tilde{v}_m^C and $\tilde{e}_{R_m L_m}^C$ are collinear. The response of the form of the hybridization tetrahedron can however be written as:

$$\delta\tilde{\omega}_b = -\frac{\sqrt{3}}{2} \cdot \frac{\gamma_{\sigma\sigma}^{CH} \tilde{v}_m^C - \gamma_{\zeta\sigma}^{CH} S_m^C \tilde{e}_{R_m L_m}^C}{\sqrt{3}\beta_{\sigma\sigma}^{CH} + \beta_{\zeta\sigma}^{CH}} \delta r_{CH_m}. \quad (45)$$

It is well known that solution of the reverse problem in molecular vibrational spectroscopy leads to off-diagonal terms in the potential energy function [14], whereas they are usually absent in the MM schemes. The proposed scheme allows one to solve the question of the off-diagonal terms on purely theoretical basis. So, coupling of stretching for two incident C—H bonds in the methane molecule can be written as:

$$\frac{\partial^2 E}{\partial r_{CH_m} \partial r_{CH_m'}} = \frac{1}{2\sqrt{3}} P^{rl} \frac{(\sqrt{3}\gamma_{\sigma\sigma}^{CH} - \gamma_{\zeta\sigma}^{CH})^2}{\sqrt{3}\beta_{\sigma\sigma}^{CH} + \beta_{\zeta\sigma}^{CH}} \quad (46)$$

and is positive. The same moves result in off-diagonal terms connecting bond stretching with angular distortions at the same atom.

Results and Discussion

In this article we start from the semiempirical APSLG description of molecular electronic structure—produced explicit formulae for different force fields of the sp³ hybridized carbon atom. We discuss the limits of their applicability. Some numerical analysis of schemes either fixing or tuning different ESPs is performed in Ref. [7]. In this article the construction is performed in the linear-response approximation. So we need estimates of numerical reliability of this model. It can be shown that even relatively large elongation of C—H bond of 0.10 Å leads to very small changes of geminal amplitudes not exceeding 0.003. The difference between esti-

mates of $\delta\tilde{\omega}_b$ for this distortion in the FA setting numerically and by formula Eq. (45) is less than 0.2%. In the case of methane distortions conserving the S₄ symmetry, the HOs remain unchanged even for angular deformations corresponding to valence angle 60°. These deformations are hybridization-incompatible and this conclusion derived in the linear approximation remains true even for the above large distortion. The next process studied is the totally hybridization-compatible deformation of the coordination tetrahedron. In this case, even for large distortion $\delta\theta_{12} = 6^\circ$, the difference between numerical and theoretical [by Eq. (42)] estimates of $\delta\tilde{\omega}_b$ is less than 0.03%. Therefore, the use of linear response formulae is valid.

The formulae given above allow us to obtain the parameters of the MM force fields. We give the numerical estimates. The parameter r^0 (“equilibrium” C—H bond length) in the FA picture equals 1.069 Å. The symmetric TA picture Eq. (8) gives the value 1.078 Å, which is closer to the C—H bond length in the methane molecule (1.094 Å). At the same time, both these values seem quite reasonable. In the case of second derivative of the bond energy function, we cannot expect a good coincidence with the same parameters adopted in the MM schemes, for two reasons. First of all, the MM elasticity constants have some implicit contribution of other factors. For example, the structure-related elasticity constant values implicitly refer to a certain range of nonvalence interaction energies. On the other hand, we must admit that the elasticity constant obtained in the current MM derivation strongly depends on the form of specific core–core repulsion D_{CH} adopted in the MINDO/3 method. Within the original method, this repulsion term is parameterized only to reproduce the equilibrium bond lengths (first derivatives of the energy) but not the second derivatives of the energy (MINDO/3 method is not parameterized to reproduce the IR frequencies). Indeed, we obtain the value 8.30 mdin/Å in the framework of the FA picture and 7.77 mdin/Å in the framework of the symmetric TA picture. The standard MM schemes usually use this elasticity constant in the range 4.5 to 4.7 mdin/Å [15–18]. The constant obtained from the solution of reverse spectral problem in the assumption of harmonic potential is equal to 5.31 mdin/Å [14]. At the same time, Ref. [19] gives the value of 7.9 mdin/Å for this constant.

The estimates of other constants based on analysis of the resonance contribution to the energy seem to be more reliable because the bending force

field derived does not depend on any derivatives of molecular integrals on geometric parameters but is determined by coarse features of electronic structure, such as the angle between HO and direction of chemical bond. Equation (31) allows us to estimate the bending elasticity constant for the H—C—H angle. It equals 0.509 m_{din}/deg and is quite close to the constants used in the MM force fields (e.g., 0.549 m_{din}/deg in Ref. [17] and 0.508 m_{din}/deg in Ref. [18]) and to the value 0.493 m_{din}/deg, obtained by solving the reverse spectral problem for methane [14]. It shows validity of the stereochemistry considerations based on the resonance contribution to the energy and confirms our conclusions about the main source of the bending energy. The explicit formula Eq. (46) allows us to estimate the off-diagonal constant *h* coupling the stretching of two incident C—H bonds. The numerical value of the constant obtained from our consideration is 0.120 m_{din}/Å. This value is approximately three times larger than that estimated in Ref. [14] in the harmonic approximation. It is not surprising, because for such a small constant its value strongly depends on the force field used.

Conclusion

In this article we proposed a method of analytical construction of the principal MM force fields—those of stretching and bending. To this end we performed a formal analysis of the energy expression written within the semiempirical implementation of the APSLG method and thoroughly elaborated different types of approximations to its electronic structure parameters, either matrix elements of electron density or hybridization parameters. Both types of parameters can be either fixed or tuned. We developed a linear-response version of the theory with fixed geminal amplitudes (FA) and applied it to a characteristic example, the methane molecule. We constructed bond energy functions and, using the parametrization of the hybridization manifold by a pair of quaternions, gave explicit formulae for different force fields, including off-diagonal ones, which depend on the details of hybridization response to the small geometry variations. The general theoretical consideration is con-

firmed by numerical analysis of force field parameters.

ACKNOWLEDGMENTS

This work was performed with partial financial support from RAS through grant #6-120 dispatched by its Young Researchers' Commission. Financial support for A.M.T. on the part of the Haldor Topsøe A/S is acknowledged, as well. One of the authors (A.L.T.) gratefully acknowledges valuable discussions with Prof. A. A. Levin, Dr. I. A. Misurkin, and Dr. I. V. Pletnev.

References

- Burkert, U.; Allinger, N. *Molecular Mechanics*; ACS: Washington, DC, 1982.
- Tokmachev, A. M.; Tchougréeff, A. L. *Int J Quantum Chem* 2001, 84, 39.
- Gillespie, R. J.; Hargittai, I. *The VSEPR Model of Molecular Geometry*; Allyn & Bacon: New York, 1991.
- Cohen, N.; Benson, S. W. *Chem Rev* 1993, 93, 2419.
- Daudey, J.-P.; Malrieu, J.-P.; Rojas, O. In Chalvet, O.; Daudel, R.; Diner, S.; Malrieu, J.-P., Eds. *Localization and Delocalization in Quantum Chemistry, Vol. 1: Atoms and Molecules in the Ground State*; D. Reidel Publications: Dordrecht, 1975.
- Diner, S.; Malrieu, J.-P.; Claverie, P. *Theoret Chim Acta* 1969, 13, 1.
- Tokmachev, A. M.; Tchougréeff, A. L. *Int J Quantum Chem* 2002, 88, 403.
- Tokmachev, A. M.; Tchougréeff, A. L. *J Comp Chem* 2001, 22, 752.
- Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J Am Chem Soc* 1975, 97, 1302.
- Tchougréeff, A. L.; Tokmachev, A. M. (to be published).
- Hoffmann, R. *J Chem Phys* 1963, 39, 1397.
- Coulson, C. A. *Valence*; Oxford University Press: Oxford, 1961.
- Elliot, J. P.; Dawber, P. G. *Symmetry in Physics, Vols. 1 and 2*; MacMillan Press Ltd: London, 1979.
- Volkenstein, M. V.; Gribov, L. A.; Elyashevich, M. A.; Stepanov, B. I. *Molecular Vibrations*; Nauka: Moscow, 1972 (in Russian).
- Allinger, N. L. *J Am Chem Soc* 1977, 99, 8127.
- Engler, E. M.; Andose, J. D.; Schleyer, P. v. R. *J Am Chem Soc* 1973, 95, 8005.
- Ermer, O.; Lifson, S. *J Am Chem Soc* 1973, 95, 4121.
- Chang, S.; McNally, D.; Shary-Tehrany, S.; Hickey, S. M. J.; Boyd, R. H. *J Am Chem Soc* 1970, 92, 3109.
- Pearson, R. G. *Symmetry Rules for Chemical Reactions*; Wiley-Interscience: New York, 1976.