Hessian-biased force fields from combining theory and experiment

Siddharth Dasgupta and William A. Goddard III
Arthur Amos Noyes Laboratory of Chemical Physics,[a] *California Institute of Technology, Pasadena, California 91125*

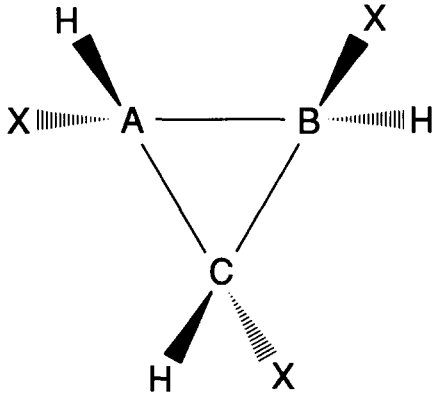
(Received 3 January 1989; accepted 16 February 1989)

We describe a new approach of combining experimental and theoretical information to develop accurate valence force fields. We combine the Hessian from *ab initio* calculations with the structural and spectroscopic data from experiment to generate a new Hessian that is used for extracting force fields. Emphasis is placed on obtaining accurate geometries and accurate frequencies. This technique is illustrated by calculations on formaldehyde, formate anion, thioformaldehyde, carbonyl chloride, and carbonyl fluoride.

I. INTRODUCTION

The great progress in quantum chemistry and in the power of readily available computers over the last decade has led to accurate first principles calculations on molecules containing up to ten or so atoms, and less accurate calculations for up to 100 atoms. Rapid progress will continue here; however, it is hard to imagine accurate first principles calculations for hemoglobin, DNA, amorphous poly ether ether ketone (PEEK), or ZSM-5 zeolite. To predict the structures and dynamics of such systems, it is necessary to average out the electronic information contained in the wave function to obtain a force field in which the energy is determined solely by the positions of the nuclei. To be useful for these purposes, the force field should be described as a superposition of terms, each involving two to four atoms (e.g., two-body terms in bond stretch or van der Waals interactions, three-body angle-bend, four-body torsion or inversion) and for which the parameters (e.g., equilibrium geometries, force constants) are transferable from molecule to molecule.

The most common approach to developing such a force field has been to start with the experimental geometries and vibrational frequencies for a set of related molecules and to find force field parameters that provide the best fit to these data. This has been useful; however, the experimental vibrational data do not provide sufficient constraints to ensure that the second derivative matrix is well fit. To illustrate the problem, consider the general nine-atom molecule in (1.1)



(1.1)

A simple spectroscopic quality force field would involve a minimum of 18 parameters to describe the nine different bond stretch terms, 90 parameters to describe the 18 different bond angle terms, and at least 24 parameters to describe the 24 different torsion terms. However, the 21 vibrational modes provide only 21 pieces of data to determine second derivatives, and the experimental geometry provides only 21 constraints on the first derivatives (forces). These 42 experimental constraints are far short of determining the 132 parameters needed here. Of course, with experiments on various isotopes and sequences of molecules all described with the same force field parameters, one could eventually overdetermine these parameters. Our conclusion is that in generating accurate force fields one *must use* additional information beyond that available from spectroscopic studies.

The solution is to use the power of modern quantum chemistry programs[1,2] which can now accurately calculate optimum geometries and obtain directly the full second derivative matrix (Hessian) at any geometry. Indeed, Schaefer and co-workers[2] and Pulay *et al.*[3] have shown that with a moderate basis set (e.g., double ζ with polarization) at the Hartree–Fock level, quantum calculations can predict the correct relative infrared intensities and in most cases the correct ordering of the normal modes. Unfortunately, the actual harmonic frequencies are usually overestimated by $\sim 10\%$–20%. This is illustrated in Table I which shows the results for H_2CO from accurate HF calculations. A problem here is that restrictions in the HF wave function (that lead to dissociation to incorrect atomic limits) generally provide too short a distance and too high a force constant. To correct for

TABLE I. Comparison of experimental geometry and vibrational frequencies with Hartree–Fock calculations on formaldehyde (H_2CO).

	Expt.	Hartree–Fock	
		Opt. Geom.	Exp. Geom.
Geometry			
$R_{C=O}$ (Å)	1.2030	1.1887	(1.2030)
R_{C-H} (Å)	1.0990	1.0943	(1.0990)
θ_{H-C-H} (deg)	116.5	116.2	(116.5)
Vibrational frequencies (cm^{-1})			
CH_2 wag	1167	1335.3	1344.2
CH_2 rock	1249	1367.6	1366.8
CH_2 scis.	1500	1656.0	1645.0
C=O str.	1746	2007.9	1925.5
CH_2 *s*-str.	2783	3149.2	3107.8
CH_2 *a*-str.	2843	3226.7	3189.6

[a] Contribution No. 7898.

 0021-9606/89/127207-09$02.10

these deficiencies several scaling schemes have been suggested[4] whereby the force constants obtained by *ab initio* calculations are scaled down to fit the experimental frequency spectrum. This method has been found useful for several compounds.[5] We propose an alternative scheme within this same philosophy which we have found useful and report results on a series of related systems.

II. HESSIAN-BIASED FORCE FIELDS

Expanding the energy of a molecule about some reference geometry leads to

$$E(X_1 \cdots Z_N) = E_0 + \sum_{i=1}^{3N} \left(\frac{\partial E}{\partial R_i}\right)_0 (\delta R_i) + \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 E}{\partial R_i \partial R_j}\right)_0 (\delta R_i)(\delta R_j) + \cdots, \quad (2.1)$$

where

$$F_i = -\frac{\partial E}{\partial R_i} \quad (2.2)$$

is the force on the ith component and

$$H_{ij} = \frac{\partial^2 E}{\partial R_i \partial R_j} \quad (2.3)$$

is the Hessian. The vibrational states are calculated by transforming H_{ij} to the mass weighted form

$$\mathscr{H}_{ij} = H_{ij}(M_i M_j)^{-1/2} \quad (2.4)$$

and diagonalizing $\mathscr{H}$

$$\mathscr{H}\,\mathbf{U} = \mathbf{U}\lambda$$

to obtain (squares of) the frequencies. The columns of $\mathbf{U}$ are the vibrational modes.

The Hessian $\mathbf{H}$ has a wealth of information that should be useful in extracting a force field. Thus, for N atoms there are $(3N-6)(3N-5)/2$ independent elements. Adding the $3N-6$ constraints for the forces to be zero leads to $3(N-1)(3N-6)/2$ total constraints on the force field. Thus, for our example in (1.1) there are 252 constraints, more than sufficient to determine the force field discussed above (132 parameters) and even much more elaborate force fields.

The problem with the use of the theoretical Hessian is that the frequencies are poor. However, the experimental frequencies provide only $(3N-6)$ pieces of data, far short of those needed to determine the Hessian. We propose the following approach to combine the information from the theoretical Hessian and from experiment. From the theory we obtain

$$\mathscr{H}^t \mathbf{U}^t = \mathbf{U}^t \lambda^t \quad (2.5)$$

or

$$\mathscr{H}^t = \mathbf{U}^t \lambda^t \tilde{\mathbf{U}}^t$$

(where $\tilde{\mathbf{U}}$ indicates transpose). We then replace λ^t with the experimental frequencies to determine the λ^x and define a new *experimentally biased Hessian* as

$$\mathscr{H}^{xt} = \mathbf{U}^t \lambda^x \tilde{\mathbf{U}}^t. \quad (2.6)$$

This amounts to assuming that the *form* of the normal modes remains the same while the *eigenvalues* are adjusted to fit experiment. Sellers *et al.* and Ha *et al.*[6] have made similar arguments but have not reported results using an analytically calculated *ab initio* Hessian scaled to fit the experiment. By construction, solving for the eigenvalues of $\mathscr{H}^{xt}$ will give the experimental frequencies exactly. Since the theoretical modes may not be in the same sequence as the experimental modes, it is essential to assign the experimental frequencies to the proper modes.

To obtain the force field, we use the following method:

(1) At the *experimental* geometry we calculate the forces $\mathbf{F}$ and Hessian $\mathbf{H}^t$ for the Hartree–Fock wave function using modern quantum chemistry codes (we used GAUSSIAN 86[1]).

(2) We solve Eq. (2.5) for the theoretical vibrational modes U^t.

(3) We use the experimental λ^x and the theoretical U^t to calculate $\mathscr{H}^{xt}$.

(4) Given any starting set of parameters the force field is used to calculate $\mathscr{H}^{FF}$ and the parameters are varied to minimize the difference between $\mathscr{H}^{FF}$ and $\mathscr{H}^{xt}$ while requiring that the forces F_i^{FF} be zero. This fitting can be done in either the Cartesian coordinate system or in the Hessian eigenvalue coordinate system.

III. ELECTRONIC STRUCTURE CALCULATIONS

We utilized the GAUSSIAN 86 package[1] developed by the Pople group for the *ab initio* electronic structure calculations. We used the Dunning–Huzinaga contracted Gaussian basis set (double zeta with polarization)[7] for formaldehyde and formate anion, 6-311G** for thioformaldehyde and 6-311G*[8] for carbonyl chloride and carbonyl fluoride. From these Hartree–Fock calculations we obtained the analytic Hessian in Cartesian coordinates H^t. For the examples in this paper, we used the experimental geometry from the most accurate determination of the geometry. This was from infrared spectroscopy[9] for the formate ion and from microwave spectroscopy for the other molecules,[10] except for formaldehyde for which the suggested equilibrium structure was used.[10(b)]

A. Force fields

For these studies we represent the potential energy of the molecule by the following analytic expression:

$$E = \sum_{\text{bonds}} E_b + \sum_{\text{angles}} E_\theta + \sum_{\text{torsions}} E_\phi + \sum_{\text{inversions}} E_\chi + \sum_{\text{nonbond}} E_{\text{nb}} + \sum_{\text{cross terms}} E_{\text{cross}}, \quad (3.1)$$

where the terms used are summarized below.

1. Bonds

The simplest description for the interaction of two bonded atoms is a spring leading to a harmonic expression of the form

$$E_b = \frac{1}{2} K_b (r - r_e)^2, \quad (3.2)$$

where K_b is the force constant and r_e is the location of the

minimum of the potential energy. While this potential function is adequate near the minimum and generally suffices for describing vibrational frequencies, it is not appropriate for processes involving large displacements (e.g., dissociation). We use instead the Morse potential, which has the form

$$E_b = D_e[e^{\alpha(r-r_e)} - 1]^2, \tag{3.3}$$

where D_e is the bond dissociation energy and α is related to the force constant

$$\alpha = \left(\frac{K_e}{2D_e}\right)^{1/2}.$$

The D_e in our Morse potential should be the energy to break a bond while retaining all angles fixed (the snap bond energy); however, in the current calculations we based D_e on experimental (adiabatic) values. We obtained force fields using both Eqs. (3.2) and (3.3).

2. Angles

Given two bonds IJ and JK sharing a common atom, the angle interactions are often described using a harmonic potential of the form

$$E_\theta = \frac{1}{2} K_\theta (\theta - \theta_e)^2, \tag{3.4}$$

where θ is the angle between bond IJ and JK, K_θ is the force constant and θ_e is the equilibrium angle. A problem with Eq. (3.4) is that it does not obey the relations

$$E(\pi - \theta) = E(\pi + \theta)$$
$$E(-\theta) = E(\theta) \tag{3.5}$$

required by symmetry. Consequently, we prefer the alternate form

$$E_\theta = \tfrac{1}{2} C(\cos\theta - \cos\theta_e)^2 \tag{3.6a}$$

(where $C = K_\theta/\sin^2\theta_e$) which does satisfy Eq. (3.5). For $\theta_e = \pi$ or 0, the form (3.6a) would lead to $\frac{1}{8}C\theta^4$ near θ_e and we instead replace (3.6a) by

$$E_\theta = K_e(\cos\theta - \cos\theta_e). \tag{3.6b}$$

For a planar molecule having three bonds (IJ, KJ, and LJ) the three angles satisfy $\theta_{IJK} + \theta_{KJL} + \theta_{LJI} = 360°$. However, in our fits we do *not* constrain the sum of the θ_e's to be 360°.

An alternative to Eqs. (3.5) or (3.6) is the Urey–Bradley interaction

$$E_{UB} = \tfrac{1}{2} K(R_{IK} - R^0_{IK})^2 \tag{3.7}$$

or

$$E_{UB} = Ae^{-BR_{IK}}$$

(I and K are both bonded to J) which can allow for anharmonicity in the angle bending potential.

3. Torsions

Given bonds IJ, JK, and KL we would describe the energy dependence on the dihedral angle ϕ as

$$E_\phi = \sum_n \tfrac{1}{2} K_{\phi,n} \cos(n\phi) \tag{3.8}$$

(sine terms are not allowed by symmetry). In the cases reported herein, torsional energy terms are not relevant.

4. Inversions

Given an atom I bonded to three other atoms J, K and L, we might include an inversion to describe the umbrella motion (to maintain planarity about I when I is sp^2 hybridized or to adjust the barrier to inversion for pyramidal molecules). A common choice for the potential is the harmonic expression

$$E_\chi = \frac{1}{2} K_\chi (\chi - \chi_e)^2, \tag{3.9}$$

where χ is the angle between the IL bond and the IJK plane and K_χ is the force constant. We prefer the alternate harmonic expression

$$E_\chi = \frac{1}{2} C(\cos\chi - \cos\chi_e)^2, \tag{3.10a}$$

where $C = K_\chi/\sin^2\chi_e$. As in Eq. (3.6) for $\chi_e = 0$ (or π), we use the form

$$E_\chi = K_\chi(\cos\chi - \cos\chi_e) \tag{3.10b}$$

to have a harmonic restraining force. Since any one of the three atoms J, K or L could have been used as the special atom in defining χ, we average over all three cases in order to avoid bias.

5. Nonbonded interactions

The nonbond terms include both van der Waals interactions and the electrostatic interactions arising from charges on the atoms (including hydrogen bonding interactions). Such interactions are generally not included for nearest and next-nearest neighbors (1–2 and 1–3 interactions), since they are considered to be included in the above terms. Since our current studies involve only 1–2 and 1–3 interactions, we will not include these terms.

6. Cross terms

Since the optimum bond angle θ_e depends on the bond distance r_e, we should also include terms coupling the value of θ with the distances in the associated bonds. For the calculations herein, we include bond–bond

$$E_{\text{bond 1–bond 2}} = K_{r_1r_2}(r_1 - r_{1e})(r_2 - r_{2e}) \tag{3.11}$$

and bond-angle terms. When Eq. (3.6) is used, the bond angle term is

$$E_{\text{bond–angle}} = K_{rc}(r - r_e)(\cos\theta - \cos\theta_e) \tag{3.12}$$

whereas when Eq. (3.4) is used, the bond-angle term is

$$E_{\text{bond–angle}} = K_{r\theta}(r - r_e)(\theta - \theta_e) \tag{3.13}$$

matching (3.12) and (3.13) at $\theta = \theta_e$ leads to

$$K_{rc} = -K_{r\theta}/\sin\theta_e$$

if $\theta_e \neq 0, \pi$.

B. Levels

In this paper we developed the force fields at two levels, both simple. The first level eschews cross terms and is suitable for geometry minimizations and dynamics. The second level includes cross terms leading to a significant improvement in the accuracy of the vibrational spectrum. Additional cross terms would lead to further improvement in the vibra-

tional frequencies. However, such increases in the number of cross terms make it more difficult to obtain transferability. Consequently, we have been parsimonious in allowing only the most important interaction terms. Implicit in this approach is the assumption that the physical description of the normal modes afforded by the *ab initio* HF calculation is essentially correct.

IV. PARAMETER OPTIMIZATION

With a particular set of force field parameters, we evaluate $E_i' = \partial E/\partial q_i$, and $E_{ij}'' = \partial^2 E/\partial q_i \partial q_j$, for all $3N$ Cartesian coordinates q_i. We consider the total error in the fit for this particular set of parameters as

$$S = \omega_F \sum_{\alpha} (E_i')^2 + \omega_H \sum_{\beta} (E_\beta'' - H_\beta)^2,$$

where N is the number of atoms in the molecule, $\alpha = 1, \ldots, 3N$, and $\beta = 1, \ldots, 3N(3N+1)/2$. Here ω_F, ω_H are appropriate weighting factors. To obtain equal average errors in these two terms would require

$$\frac{\omega_F}{\omega_H} = \frac{3N(3N+1)}{2(3N)} = \tfrac{1}{2}(3N+1)$$

or 6.5 for four-atom molecules. We have found $\omega_F/\omega_H = 100$ leads to very accurate geometries and reasonable force constants. The objective function S is minimized by systematically varying the force field parameters, using the Nelder–Mead downhill simplex algorithm.[11] We found that due to the strong coupling between the force constants (K_b, K_θ, etc.) and the geometry parameters (r_e, θ_e, etc.), it is better to optimize them alternately, as suggested by Huige *et al.*[12]

The process is begun by using the experimental geometry for the molecule and a starting set of default parameters. In the first pass, the geometry parameters are optimized to minimize the residual forces at the experimental geometry while the force constants are held fixed. At this stage the second derivative terms in the objective function are neglected. In the next pass, the geometry parameters are held fixed and the force constants are optimized to minimize the objective function which now contains all the first and second derivative terms. At the end of this pass the forces on the molecules are not optimum as the geometry parameters were not optimized in this pass. This procedure is repeated iteratively until the optimum is reached (with root mean square forces on the molecule below a cutoff of 0.01 kcal/mol). In our optimizations we used a hierarchical scheme of first optimizing the purely diagonal force field and only then adding the cross terms.

V. RESULTS

Four different force fields were determined with differing analytical expressions and number of terms. In all cases, inversion was treated using Eq. (3.10):

(a) MC: In MC force field (denoting Morse, Cosine) we used the Morse term (3.3) for bonds, the cos θ angle interactions (3.6) but no cross terms. The values of D_e for the Morse terms were taken from experiment[13] and were not optimized.

(b) MCX: The MCX force field includes cross terms (hence the X) for bond–angle (3.12) and bond–bond (3.11) interactions, in addition to the MC terms.

(c) HT: The HT force field (denoting harmonic bond and theta angle) uses the traditional harmonic bond stretch (3.2) and θ angle expansion terms (3.4), but no cross terms.

(d) HTX: The HTX force field has cross terms (3.11) and (3.13) added to HT.

The force fields for all the molecules are summarized in Table II.

TABLE II. Hessian biased force field (HBFF) parameters.[a]

Parameters		MCX[b]	MC[c]	HTX[d]	HT[e]
Formaldehyde (H_2CO)					
C–O	r_e	1.2022	1.2030	1.2017	1.2030
	K_b	1809.13	1601.30	1799.22	1600.74
	D_e	175.0	175.0		
C–H	r_e	1.0980	1.0990	1.0972	1.0990
	K_b	632.45	624.72	628.83	624.75
	D_e	87.0	87.0		
O–C–H	θ_e	122.46	123.16	122.98	124.48
	K_θ	111.38	103.98	113.97	107.59
	$K_{r_1\theta}$	−71.47		60.60	
	$K_{r_2\theta}$	−18.57		15.74	
	$K_{r_1r_2}$	63.96		63.93	
H–C–H	θ_e	117.89	119.50	118.93	122.01
	K_θ	55.05	48.00	57.01	53.30
	$K_{r_1\theta}$	−22.84		20.37	
	$K_{r_2\theta}$	−22.84		20.37	
	$K_{r_1r_2}$	10.73		10.68	
C–X–X–X	K_χ	50.31	48.79	49.10	45.90
Thioformaldehyde (H_2CS)					
C–S	r_e	1.6099	1.6110	1.6021	1.6110
	K_b	930.36	779.41	928.70	778.05
	D_e	129.3	129.3		
C–H	r_e	1.0925	1.0930	1.0896	1.0930
	K_b	710.14	708.40	708.05	708.50
	D_e	87.0	87.0		
S–C–H	θ_e	122.19	121.63	126.57	123.02
	K_θ	78.45	77.04	79.56	75.05
	$K_{r_1\theta}$	−51.75		47.78	
	$K_{r_2\theta}$	−9.13		7.24	
	$K_{r_1r_2}$	15.09		14.99	
H–C–H	θ_e	117.77	117.02	123.45	118.97
	K_θ	56.08	51.21	58.15	53.17
	$K_{r_1\theta}$	−19.87		17.23	
	$K_{r_2\theta}$	−19.87		17.23	
	$K_{r_1r_2}$	13.95		14.06	
C–X–X–X	K_χ	36.87	37.70	30.41	35.64
Carbonyl chloride (Cl_2CO)					
C–O	r_e	1.1655	1.1660	1.1653	1.1660
	K_b	2029.93	1853.61	2026.31	1853.60
	D_e	175.0	175.0		
C–Cl	r_e	1.7378	1.7460	1.7359	1.7460
	K_b	423.25	259.77	412.09	259.68
	D_e	74.0	74.0		
Cl–C–O	θ_e	128.91	128.37	130.31	130.61
	K_θ	102.68	81.11	117.69	94.20
	$K_{r_1\theta}$	−48.10		32.94	
	$K_{r_2\theta}$	−29.51		23.45	
	$K_{r_1r_2}$	163.85		167.63	

TABLE II (*continued*).

Parameters		MCX[b]	MC[c]	HTX[d]	HT[e]
Cl–C–Cl	θ_e	116.4	115.24	117.80	117.64
	K_θ	92.85	85.05	107.50	92.70
	$K_{r_1\theta}$	− 13.05		15.93	
	$K_{r_2\theta}$	− 13.05		15.93	
	$K_{r_1r_2}$	75.86		93.03	
C–X–X–X	K_χ	58.06	60.88	54.30	56.10
Carbonyl fluoride (F_2CO)					
C–O	r_e	1.1694	1.1700	1.1678	1.1700
	K_b	2141.51	1919.84	2142.27	1917.20
	D_e	175.0	175.0		
C–F	r_e	1.3022	1.3150	1.3030	1.3150
	K_b	998.39	686.75	914.38	687.21
	D_e	105.0	105.0		
O–C–F	θ_e	131.16	126.21	130.57	127.52
	K_θ	144.20	114.97	177.00	114.02
	$K_{r_1\theta}$	− 47.09		60.47	
	$K_{r_2\theta}$	− 102.70		85.83	
	$K_{r_1r_2}$	198.41		188.28	
F–C–F	θ_e	113.12	107.61	112.25	108.93
	K_θ	145.47	107.87	176.11	112.27
	$K_{r_1\theta}$	− 81.35		80.28	
	$K_{r_2\theta}$	− 81.35		80.28	
	$K_{r_1r_2}$	146.97		140.48	
C–X–X–X	K_χ	69.43	84.07	69.90	81.10
Formate anion (HCO_2^-)					
C–O	r_e	1.2409	1.2460	1.2359	1.2460
	K_b	1350.26	1079.71	1326.85	1079.58
	D_e	115.4	115.4		
C–H	r_e	1.0901	1.1060	1.0740	1.1060
	K_b	707.82	633.74	642.68	633.74
	D_e	87.0	87.0		
O–C–H	θ_e	125.45	116.56	134.75	128.67
	K_θ	72.02	85.17	90.48	84.50
	$K_{r_1\theta}$	− 67.27		47.65	
	$K_{r_2\theta}$	− 44.44		35.36	
	$K_{r_1r_2}$	69.30		73.51	
O–C–O	θ_e	130.35	126.16	134.45	131.98
	K_θ	149.80	176.29	190.95	175.79
	$K_{r_1\theta}$	− 3.50		23.42	
	$K_{r_2\theta}$	− 3.50		23.42	
	$K_{r_1r_2}$	193.05		241.85	
C–X–X–X	K_χ	58.84	72.46	41.36	52.02

[a] Units: bonds in Å, angles in degrees; K_b and $K_{r_1r_2}$ in kcal/mol Å^2; K_θ in kcal/mol rad^2; $K_{r_1\theta}$ $K_{r_2\theta}$ in kcal/mol Å rad; and D_e in kcal/mol.
[b] MCX: MC + cross terms.
[c] MC: Morse bonds, Cosine angle expansion.
[d] HTX: HT + cross terms.
[e] HT: Harmonic bonds, theta angle expansion.

A. Geometry

With all force fields the optimized geometry matches the experimental geometry to 0.0001 Å and 0.01° for all cases as summarized in Table III. This is a consequence of the method of development where we place a strong premium on reproducing the experimental geometry as accurately as possible. We do not find it necessary to add cross terms in order to reproduce the experimental structure.

B. Frequencies

With the purely diagonal force fields (no cross terms) some of the vibrational frequencies show large deviations from the experimental values. In the case of formaldehyde, Table IV, the CH_2 scissoring mode is too low (by 132 cm^{-1}) as is the CH_2 rocking mode (by 27 cm^{-1}). With HT the splitting between the CH_2 stretching modes is too large (by 22 cm^{-1}). In general, molecules that contain $-CH_2$ and $-CH_3$ groups lead to splittings between the various C–H stretching modes that are difficult to reproduce without cross terms. The two common alternatives have been the linear cross terms as in Eqs. (3.11)–(3.13) or a harmonic force between the two outer atoms as in Eq. (3.7). In both MTX and HTX these errors are corrected.

TABLE III. Comparison of optimum geometries from HBFF with experiment. All bonds in Å and angles in degrees.

Parameter	Expt.[a]	MCX	MC	HTX	TH
Formaldehyde					
C=O	1.2030	1.2030	1.2030	1.2030	1.2030
C–H	1.0990	1.0990	1.0990	1.0990	1.0990
H–C–H	116.5	116.5	116.5	116.5	116.5
Thioformaldehyde					
C=S	1.6110	1.6110	1.6110	1.6110	1.6110
C–H	1.0930	1.0930	1.0930	1.0930	1.0930
H–C–H	116.9	116.9	116.9	116.9	116.9
Formate anion					
C–O	1.2460	1.2460	1.2460	1.2460	1.2460
C–H	1.1060	1.1060	1.1060	1.1060	1.1060
O–C–O	126.3	126.3	126.3	126.3	126.3
Carbonyl chloride					
C=O	1.1660	1.1660	1.1660	1.1660	1.1660
C–Cl	1.7460	1.7460	1.7460	1.7460	1.7460
Cl–C–Cl	111.3	111.3	111.3	111.3	111.3
Carbonyl fluoride					
C=O	1.1700	1.1700	1.1700	1.1700	1.1700
C–F	1.3150	1.3150	1.3150	1.3150	1.3150
F–C–F	107.6	107.6	107.6	107.6	107.6

[a] For formaldehyde, Ref. 10(b), for formate anion, Ref. 9 and for the others, Ref. 10(a).

1. H_2CO

The best empirical force field of H_2CO from a traditional fit to spectroscopic frequencies[14] is compared in Table V

TABLE IV. Vibrational frequencies and isotope shifts for H_2CO. Frequencies and shifts in cm^{-1}.

Modes	Expt.	MCX	MC	HTX	HT
CH_2 wag	1167	1167.2	1166.8	1166.8	1166.8
CH_2 rock	1249	1249.7	1218.4	1249.9	1211.0
CH_2 scis.	1500	1499.9	1364.5	1499.3	1366.9
C=O str.	1746	1746.4	1752.3	1746.3	1753.9
CH_2 *s*-str.	2783	2783.0	2789.5	2782.9	2789.6
CH_2 *a*-str.	2843	2842.7	2871.2	2842.9	2871.2
$\lvert\Delta\nu\rvert_{avg}$		0.3	34.5	0.4	35.6
Isotopic shifts (H_2CO–D_2CO)					
CH_2 wag	229	231.9	231.8	231.8	231.8
CH_2 rock	259	264.4	261.5	264.5	259.8
CH_2 scis.	394	404.2	339.6	404.0	339.4
C=O str.	46	61.4	154.2	61.5	155.5
CH_2 *s*-str.	727	743.7	686.3	743.7	686.2
CH_2 *a*-str.	683	720.2	719.2	720.3	719.5

TABLE V. Comparison for H_2CO of the HTX parameters for HBFF with the parameters for GVFF (Ref. 14). For units see Table II.

	Parameters		HTX	GVFF
Bonds	C–O	r_e	1.2017	1.2030
		K_b	1799.22	1857.13
	C–H	r_e	1.0972	1.0990
		K_b	628.83	706.12
Angles	O–C–H	θ_e	122.98	121.75
		K_θ	113.97	73.58
		$K_{r_1\theta}$	60.60	23.96
		$K_{r_2\theta}$	15.74	9.10
		$K_{r_1r_2}$	63.93	75.19
	H–C–H	θ_e	118.93	116.50
		K_θ	57.01	54.62
		$K_{r_1\theta}$	20.37	6.39
		$K_{r_2\theta}$	20.37	6.39
		$K_{r_1r_2}$	10.68	7.98
Inversion	C–X–X–X	K_χ	49.10	57.52

with the HTX force field. The diagonal stretching force constants are comparable in magnitude and the signs of all the cross terms are also in agreement. The slight differences in the numbers could be due to slightly differing geometries for the equilibrium structure and also because our force field does not include all the possible cross terms in this molecule whereas the experimental force field does include them. The $K_{\text{C-H, C=O}}$ cross term is quite large in both force fields indicating a strong coupling of the two bonds.

As a test of the validity of our force fields we report the isotopic frequency shifts calculated for D_2CO along with the experimental values in Table IV. All the isotopic shifts are calculated to within 5.5% (MCX and HTX) except for the C=O stretching mode. This discrepancy (21 cm^{-1}) in the isotope shift for C=O stretch has been observed even in the more elaborate scaled quantum mechanical force field (21 cm^{-1}).[15]

In molecules containing hydrogen atoms it is well known[14] that anharmonicity effects are important, especially in normal modes involving the H. In order to decrease this effect we used the vibrational spectra of deuterated formaldehyde to evaluate the same force field with the results in Table VI. The geometry is the same as that used for the undeuterated formaldehyde molecule (even though C–D bonds should be somewhat shorter than C–H bonds). The force constants differ by small amounts indicating that anharmonicity effects are not insignificant in formaldehyde. However, the isotopic shift pattern does not differ appreciably from that calculated by the previous force field. When available, we recommend fitting to the deuterated analogs to yield force field parameters most consistent with the optimum structure. For formaldehyde the problem with the isotope shift of the C=O stretching mode remains, indicating that the C=O mode is not as well described at the current level of force field.

Because of the concern that linear cross terms might lead to false minima away from the equilibrium geometry, we performed quenched molecular dynamics[16] for 2 ps at an elevated temperature of 600 K. In this process, normal dynamics is performed repeatedly for 100 cycles with a 1 fs time step and then the structure is minimized. In quenched dynamics runs on formaldehyde with both off-diagonal force fields (MTX and HTX) we did not observe any local minima.

TABLE VI. HBFF parameters for H_2CO derived using D_2CO frequencies.

	Parameters	MCX	MC	HTX	HT
C–O	r_e	1.1993	1.2030	1.1959	1.2030
	K_b	1874.50	1625.69	1822.89	1625.91
	D_e	175.0	175.0		
C–D	r_e	1.0950	1.0990	1.0910	1.0990
	K_b	661.42	636.94	646.33	636.93
	D_e	87.0	87.0		
O–C–D	θ_e	125.15	123.37	128.41	121.89
	K_θ	100.48	103.64	113.97	109.63
	$K_{r_1\theta}$	− 75.54		60.52	
	$K_{r_2\theta}$	− 16.68		13.61	
	$K_{r_1r_2}$	69.40		70.81	
D–C–D	θ_e	122.68	120.46	129.11	116.83
	K_θ	52.04	41.37	59.12	45.75
	$K_{r_1\theta}$	− 21.97		18.85	
	$K_{r_2\theta}$	− 21.97		18.85	
	$K_{r_1r_2}$	6.96		7.63	
C–X–X–X	K_χ	45.21	48.66	37.55	51.83
Geometry Parameters	Expt.				
C=O	1.2030	1.2030	1.2030	1.2030	1.2030
C–D	1.0990	1.0990	1.0990	1.0990	1.0990
D–C–D	116.5	116.5	116.5	116.5	116.5
Frequencies					
Mode	Expt.				
CD_2 wag	938	937.9	938.0	938.2	938.0
CD_2 rock	990	978.3	958.9	984.8	960.2
CD_2 scis.	1106	1114.8	1000.8	1105.1	999.9
C=O str.	1700	1699.3	1606.4	1700.8	1606.3
CD_2 *s*-str.	2056	2055.8	2120.6	2056.4	2120.6
CD_2 *a*-str.	2160	2159.9	2172.5	2160.0	2172.6
$\lvert\Delta\nu\rvert_{avg}$		3.6	51.2	1.2	51.1

In Table VII we show a comparison with the results of some standard force fields (MMP2, AMBER) for H_2CO. These force fields were developed for studies of structures of wide classes of molecules and should not be expected to be accurate for vibrational frequencies.

2. H_2CS

In thioformaldehyde the purely diagonal force fields MC and HT again fail to reproduce the CH_2 rocking and

TABLE VII. Comparison of geometry and frequencies of H_2CO obtained with MCX, MMP2, and AMBER force fields.

Parameter	Expt.	MCX	AMBER	MMP2
$r_{C=O}$	1.2030	1.2030	1.2290	1.2080
r_{C-H}	1.0990	1.0990	1.0800	1.1130
θ_{H-C-H}	116.5	116.5	120.0	119.1
CH_2 wag	1167	1167.2	878.0	1730.0
CH_2 rock	1249	1249.7	997.0	841.5
CH_2 scis.	1500	1499.9	1256.0	1146.6
C=O str.	1746	1746.4	1655.0	1681.1
CH_2 *s*-str.	2783	2783.0	2893.0	2860.6
CH_2 *a*-str.	2843	2842.7	2998.0	2954.1

TABLE VIII. Calculated vibrational frequencies and isotope shifts for thioformaldehyde. Frequencies and shifts in cm^{-1}.

Modes	Expt.	MCX	MC	HTX	HT
CH_2 wag	980	980.4	980.3	980.5	980.3
CH_2 rock	988	987.9	969.2	984.7	955.4
C═S str.	1052	1052.1	937.0	1055.8	936.6
CH_2 scis.	1447	1447.8	1448.7	1447.0	1453.5
CH_2 *s*-str.	2968	2968.3	2953.1	2968.4	2953.3
CH_2 *a*-str.	3017	3017.1	3052.8	3017.0	3052.9
$\|\Delta\nu\|_{avg}$		0.3	31.1	1.3	34.2
Isotopic shifts H_2CS–D_2CS					
CH_2 wag	207	208.3	208.3	218.7	208.4
CH_2 rock	209	235.7	232.2	247.3	228.8
C═S str.	120	129.2	126.5	139.2	125.2
CH_2 scis.	280	285.2	275.9	298.6	278.6
CH_2 *s*-str.	813	816.2	799.0	857.1	799.0
CH_2 *a*-str.	737	767.7	773.8	805.9	774.0

TABLE IX. Calculated vibrational frequencies for carbonyl chloride (Cl_2CO) and fluoride (F_2CO). Frequencies in cm^{-1}.

Modes	Expt.	MCX	MC	HTX	HT
Carbonyl chloride					
CCl_2 scis.	285	245.0	220.9	250.3	220.9
C–O bending	440	393.6	317.9	382.7	317.9
CCl_2 *s*-str.	567	576.8	425.3	578.3	425.3
C–O wag	580	578.0	578.1	582.1	578.1
CCl_2 *a*-str.	849	849.6	838.6	850.4	838.6
C═O str.	1827	1825.6	1832.6	1825.8	1832.6
$\|\Delta\nu\|_{avg}$		16.7	57.6	18.0	57.1
Carbonyl fluoride					
CF_2 def.	584	522.7	417.7	540.7	422.1
CO def.	626	626.7	483.8	631.0	482.2
Op. def.	774	773.4	771.4	773.3	772.1
CF_2 *s*-str.	965	974.2	773.6	971.0	773.6
CF_2 *a*-str.	1249	1250.5	1272.5	1249.8	1271.6
C═O str.	1928	1929.2	1946.8	1927.8	1946.8
$\|\Delta\nu\|_{avg}$		12.4	90.8	9.3	90.0

C═S stretching modes as well as the splitting between the symmetric and antisymmetric CH_2 stretching modes. Addition of the bond–bond and the bond–angle cross terms removes these deficiencies. Table VIII contains the isotopic shifts calculated for dideutero–thioformaldehyde along with the experimental values. For H_2CS the purely diagonal force fields are much more accurate in calculating the isotopic shifts than in the case of formaldehyde.

3. F_2CO, Cl_2CO

For both carbonyl fluoride and carbonyl chloride the purely diagonal force fields fail to account satisfactorily for the two in-plane angle deformation modes as well as the splitting between the symmetric and antisymmetric stretching modes, Table IX. The out-of-plane deformation and the C═O stretching modes are described with less error. The off-diagonal terms help correct the splitting in the stretching modes and in the case of carbonyl fluoride also the F–C–O deformation. There remains some residual error in the lowest frequency CF_2 deformation mode. This could be due to the fact that our off-diagonal force fields contain only a minimum complement of cross terms.

4. Formate

For the formate anion there is no data on the free ion, hence we used the experimental data for the sodium salt (one of the most ionic formates[17]). The geometry[9] is based on the x-ray diffraction structure of $NaHCO_2$ crystal[18] with the C–H bond length corrected to be consistent with the neutron diffraction value. The experimental vibrational frequencies are also from Ref. 9. For the polycrystalline $NaHCO_2$ there are splittings between the values of the modes in the infrared and Raman spectra probably due to interactions between individual ions. The largest splitting (24.3 cm^{-1}) is in the b_2 ν_{as} C═O stretching mode. Consequently, the arithmetic mean of the infrared and Raman values have been used.[9] The mode assignments have been confirmed by a large set of isotopic shift data.[9]

We were concerned that the double ζ basis set with only polarization functions (DZd) used for all other molecules might not be adequate[19] to describe the formate anion. Consequently, we calculated formate anion with two basis sets, the standard basis (DZd) and the same augmented with diffuse functions. The results are essentially the same as indicated in Table X for the MCX force field. With the newly calculated MCX force field, as before, the geometry is calculated quite accurately. The O–C–O bending mode is still the poorest fit, the discrepancy now being 148 cm^{-1} from the experimental values.

There is considerable amount of isotopic data that is available for the formate anion.[9] Table X tabulates the calculated isotope shifts (with MCX) in comparison with the experimental data. The degree of accuracy in reproducing the isotopic shift data is reassuring in that even though the absolute frequencies are not calculated exactly, the basic normal mode description must be quite reasonable.

In previously published theoretical work on the formate anion,[20] neither the force field parameters nor the actual computed frequencies were reported, so it is not possible to compare the two force fields directly. However with the same number of terms (9), our value of $|\Delta\nu|_{avg} = 102.6$ for the overall fit is a significant improvement over that (155.6) reported.[20] Addition of cross terms improves the frequencies in both MCX and HTX but residual errors remain. However, our off-diagonal force fields with only 14 terms have significantly better $|\Delta\nu|_{avg}$ than all but the largest force field reported[20] which has 40 parameters.

For force fields MC and MCX the D_e for the C═O bond Morse potential was estimated as the average of the D_e in carbon dioxide and in formic acid. The D_e for the C═O bond in formic acid was estimated by subtracting the heats of formation of the OH and CHO radical fragments from the heat of formation of formic acid.[13] The D_e for the C–H bond was the same as in formaldehyde since no accurate estimate is available for this bond. However, upon gentle heating, sodium formate is known to decompose to sodium oxalate and hydrogen[21] indicating that this bond must be quite weak.

The vibrational frequencies calculated with the different

TABLE X. Calculated vibrational frequencies and isotope shifts for formate anion. Frequencies and shifts in cm^{-1}.

Modes	Expt.	MCX	MC	HTX	HT
O–C–O bend	772.8	644.2	637.1	680.3	603.1
CH wag	1070.7	1064.4	1064.4	1063.4	1064.3
CH bend	1358.8	1355.9	1215.2	1327.0	1220.7
C=O *s*-str.	1367.6	1386.7	1177.6	1390.7	1165.9
C=O *a*-str.	1591.1	1594.2	1711.6	1606.7	1718.0
C–H str.	2829.6	2828.5	2848.9	2829.8	2848.2
$\lvert\Delta\nu\rvert_{avg}$		26.9	102.6	28.4	110.2

Isotopic shifts with MCX

	$\Delta(CHO_2-CDO_2)$		$\Delta(CHO_2-{}^{13}CDO_2)$		$\Delta(CHO_2-{}^{13}CHO_2)$	
	Calc.	Expt.	Calc.	Expt.	Calc.	Expt.
O–C–O def.	6.3	7.7	5.8	6.2	6.2	7.0
C–H wag	154.2	153.5	18.9	19.2	16.3	15.1
C–H bend	340.6	353.8	0.2	0.5	1.9	1.0
C=O sym–str.	25.4	29.3	18.8	16.7	23.5	21.1
C=O anti–str.	24.2	18.8	43.0	48.6	40.9	45.0
C–H str.	749.5	699	13.5	13.9	7.3	12.5

force fields are tabulated in Table X. The purely diagonal force fields MC and HT show large deviations from the experimental frequencies in almost all modes except the C–H stretching mode. In MCX and HTX, the modes showing the largest deviations are the lowest frequency O–C–O bend and the C=O symmetric stretch. This might be indicative that additional coupling terms [e.g., (C–H)–(O–C–O)] are needed to properly describe this molecule. The other significant source of error might be associated with the net negative charge delocalized over the heavy atoms, as well as the absence of the Na cation which should modify directly the deformation mode. Our current expression does not account directly for any electrostatic effects.

VI. CONCLUSION

In this work we demonstrated that for a series of similar tetra-atomic planar molecules, diagonal force fields can be derived from *ab initio* theory and experiments to give a good description of the geometry. The vibrational spectrum is not described adequately without angle-stretch cross terms. We also show that addition of a few cross-terms can improve significantly the vibrational mode description. The main criteria used for developing these force fields are that (1) there be a limited set of parameters so that they may be transferable while (2) the form of the force field be consistent with an accurate description of the global potential energy surface. The largest errors are made in the case of the formate anion, possibly due to the neglect of crystal effects (and a net charge on the molecule).

We stress the importance of describing the correct geometry by incorporating the forces at the experimental geometry in the fit. Contrary to an earlier assertion,[20] we have demonstrated that the cross terms are not necessary to ensure accurate structures. Moreover, linear cross terms lead to the possibility of false local minimas in the potential energy surface away from the equilibrium geometry. This would be a particularly serious deficiency for molecular dynamics and reactions. However, with the set of molecules described herein we have not yet encountered such local minimas. The addition of the cubic terms as advocated in Ref. 20 necessarily leads to a negative infinity in the energy as the bond is strained, making it useless for molecular dynamics.

For the bond stretching terms, anharmonicity effects are important only in modes involving the hydrogens and, hence a simple harmonic potential is adequate near the equilibrium geometry. We have chosen to emphasize the Morse potential for the simple reason that it is dissociation consistent, leading to a better description of the global potential surface required for reactive dynamics.

The angle bending potential is certainly not harmonic as can be demonstrated by the failure of harmonic force fields to adequately describe as simple a molecule as H_2CO. We find that linear cross terms fit the vibrational spectra adequately.

ACKNOWLEDGMENTS

We would like to thank Dr. Gilles Ohanessian for helpful discussions. Funding for this research was provided by a grant from the Air Force Office of Scientific Research (No. AFOSR-88-0051) and by a grant from Imperial Chemical Industries, Cleveland, England.

[1] M. J. Frisch, J. S. Binkley, H. B. Schlegel, K. Raghavachari, C. F. Melius, R. L. Martin, J. J. P. Stewart, F. W. Bobrowicz, C. M. Rohlfing, L. R. Kahn, D. J. Defrees, R. Seeger, R. A. Whiteside, D. J. Fox, E. M. Fleuder, and J. A. Pople, Carnegie-Mellon Quantum Chemistry Publishing Unit, Pittsburgh, PA, 1984.

[2] Y. Yamaguchi, M. Frisch, J. Gaw, and H. F. Schaefer III, J. Chem. Phys. **84**, 2262 (1986).

[3] P. Pulay, G. Fogarasi, F. Pang, and J. E. Boggs, J. Am. Chem. Soc. **101**, 2550 (1979).

[4] (a) G. Fogarasi and P. Pulay, J. Mol. Struct. **39**, 275 (1977); (b) C. E. Blom and C. Altona, Mol. Phys. **31**, 1377 (1976); (c) for review see G. Fogarasi and P. Pulay, in *Vibrational Spectra and Structure*, edited by J. R. Durig (Elsevier, New York, 1985), Vol. 14.

[5] B. A. Hess, Jr., L. J. Schaad, P. Cársky, and R. Zahradník, Chem. Rev. **86**, 709 (1986).

[6] (a) H. L. Sellers, V. J. Klimkowski, and L. Schäfer, Chem. Phys. Lett. **58**, 541 (1978); (b) H. L. Sellers and L. Schäfer, J. Mol. Struct. **51**, 117 (1979); (c) T.-K. Ha, R. Meyer, and Hs. H. Günthard, Chem. Phys. Lett. **59**, 17 (1978).

[7] T. H. Dunning and P. J. Hay, in *Modern Theoretical Chemistry*, edited by H. F. Schaefer III (Plenum, New York, 1976), Vol. 3.

[8](a) R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, J. Chem. Phys. **72**, 650 (1980); (b) A. D. McLean and G. S. Chandler, *ibid.* **72**, 5639 (1980).
[9]K. G. Kidd and H. H. Mantsch, J. Mol. Spectrosc. **85**, 375 (1981).
[10](a) M. D. Harmony, V. W. Laurie, R. L. Kuczkowski, R. H. Schwendeman, D. A. Ramsey, F. J. Lovas, W. J. Lafferty, and A. G. Maki, J. Phys. Chem. Ref. Data **8**, 619 (1979); (b) K. Yamada, T. Nakagawa, K. Kuchitsu, and Y. Morino, J. Mol. Spectrosc. **38**, 70 (1971).
[11]J. A. Nelder and R. Mead, Comp. J. **7**, 308 (1965).
[12]C. J. M. Huige, A. F. Hezemans, and K. Rasmussen, J. Comp. Chem. **8**, 204 (1987).
[13]D. F. McMillen and D. M. Golden, Annu. Rev. Phys. Chem. **33**, 493 (1982).
[14]J. L. Duncan and P. D. Mallinson, Chem. Phys. Lett. **23**, 597 (1973).
[15]P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs, and A. Vargha, J. Am. Chem. Soc. **105**, 7037 (1983).
[16]Using the BioGraf™ molecular simulation program from Molecular Simulations Inc. (BioDesign), Pasadena, CA 91101.
[17]J. D. Donaldson, J. F. Knifton, and S. D. Ross, Spectrochim. Acta Part A **20**, 847 (1964).
[18]P. L. Markila, S. J. Rettig, and J. Trotter, Acta Crystallogr. B **31**, 2927 (1975).
[19]T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. v. R. Schleyer, J. Comp. Chem. **4**, 294 (1983).
[20]J. R. Maple, U. Dinur, and A. T. Hagler, Proc. Natl. Acad. Sci. USA **85**, 5350 (1988).
[21]A. I. Vogel, *A Textbook of Macro and Semimicro Qualitative Inorganic Analysis*, 4th ed. (Longmans, London, 1959), p. 401.