Systematic ab Initio Gradient Calculation of Molecular Geometries, Force Constants, and Dipole Moment Derivatives

Péter Pulay,[1a] Géza Fogarasi,[1a] Frank Pang,[1b] and James E. Boggs*[1b]

Contribution from the Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, and The Institute for General and Inorganic Chemistry, Eötvös L. University, Budapest 8, Pf. 323, Hungary 1445. Received October 18, 1978

Abstract: Systematic ab initio gradient calculation of molecular geometries, force constants, and dipole moment derivatives is described. A new basis set, denoted 4-21, is presented for first-row atoms. It is nearly equivalent to the 4-31G set but requires less computational effort. Completely optimized Hartree-Fock geometries of 18 molecules are compared using several basis sets, with and without polarization functions. The question of the best representation of molecular force fields is discussed, and a set of standardized internal coordinates is suggested for future work. Quadratic and the most important cubic force constants and dipole moment derivatives of first-row hydrides are calculated using the 4-21 basis set, and the results are compared with those from other basis sets, including near-Hartree-Fock ones. Force-field calculations on larger molecules with the 4-21 basis are summarized. A general formulation of the rotational correction to dipole moment derivatives is given.

I. Introduction

The value of ab initio calculations for molecular geometries is well established, particularly by the systematic work of Pople and co-workers.[2] There is less systematic work on the force constants of polyatomic molecules.[3-6] This is partly due to the fact that calculation of the second and higher derivatives of the potential surface is more expensive and more sensitive to numerical errors than the geometry, i.e., the first derivative. Development of the ab initio gradient algorithm[7] has solved this problem in principle. Another fact contributing to the relative scarcity of ab initio force constant calculations is due to the difficulties of comparing theoretical and experimental data. As explained elsewhere,[3] this comparison is best done at the level of potential constants and not at the level of experimental observables like vibrational frequencies. However, good-quality harmonic force fields are rare for larger molecules, and only a few molecules have been satisfactorily analyzed using anharmonic potentials. It is particularly difficult to assess the influence of systematic errors on experimental force fields, with the result that error estimates are often far too optimistic.[8]

There is a recent revival of interest in molecular potential surfaces and vibrational states, due partly to advances in instrumentation, and partly to the realization that vibrational effects are often significant, e.g., for molecular geometries and dipole moments. This research is hampered by the lack of reliable potential constants. Given the present status of the inverse vibrational problem, the only significant advance in sight is the utilization of ab initio (or even semiempirical) information. This is the more natural because the experimental and the theoretical information is largely complementary. Present ab initio methods do not allow the routine determination of the dominant potential constants to an accuracy which comes close to experiment, except for the smallest molecules. On the other hand, this limited accuracy is maintained for coupling and anharmonic force constants which are often very difficult to determine experimentally. Therefore, it seems natural to combine the two sources of information, e.g., by using theoretical values for the off-diagonal force constants and fitting the diagonal terms to experimental vibrational frequencies.[4l,m,9] Another, and slightly preferable, procedure is the scaling of the theoretical force field, as done by Blom[6] and in our semiempirical calculations.[10]

Recently, an efficient ab initio Hartree-Fock gradient program, called TEXAS, has been written by one of us.[11] Using this program, we were able to extend systematic ab initio force field determinations to molecules larger than in previous studies. Complete quadratic force fields and the most important higher force constants have been determined for a number of medium-sized molecules, up to the size of benzene (see section VIII).

In this paper, our procedure for the systematic gradient calculation of molecular geometries, potential constants, and dipole moment derivatives is described. To the uninitiated reader, the amount of space devoted to the proper representation of force fields may seem excessive. However, our systematic calculations have shown the prime importance of this point. The reason for this is that our wave functions are necessarily approximative, both because of basis set incompleteness and because of the limitations inherent in the Hartree-Fock model. Empirical correction of systematic errors in the theoretical force fields is only possible if the latter are represented in a suitable form, i.e., by using suitable nuclear coordinates.

Improvements in the accuracy of our procedure may consist of (1) extending the basis set and (2) introducing electron correlation. As to the first, work is in progress to enhance the efficiency of the program so that polarization functions can be routinely incorporated in the basis sets, even for larger molecules. The second point is more difficult. Although several efficient ways have been found to introduce a limited amount of electron correlation, most of these methods do not allow the simple calculation of the gradient. An exception is Meyer's new SCEP method,[12a] which, similarly to multiconfigurational SCF methods, can be reformulated so as to allow the evaluation of the gradient.[12b] However, even this would be difficult to apply to large molecules. Work on this problem is also progressing in several other laboratories, but for the present it seems useful to accept the Hartree-Fock method and to correct for its deficiencies empirically.

II. Basis Sets

In selecting a basis set for systematic calculations, there are two conflicting requirements: the basis set should be sufficiently complete to permit a good description of the wave function, and at the same time it should be small enough so that the calculations can be extended to larger systems with reasonable cost.

Most of our earlier force constant calculations[4] with the MOLPRO program[13] were performed with atomic basis sets

containing three primitive Gaussian functions in the valence shell. First we used the 7s3p/3s basis set of Csizmadia et al.,[14] and later the more carefully optimized sets of similar size by Roos and Siegbahn.[15] These basis sets, often augmented by bond functions to provide polarization, gave fairly good results.[4] Experiments with even smaller Gaussian sets, e.g., 5s2p/2s, proved, however, to be unsatisfactory. Our results, as well as those of Kutzelnigg,[16] show that contraction of the triple-ζ basis sets to double ζ in the valence shell, and to minimum basis in the atomic core, has practically no effect on the calculated force fields and geometries. Further contraction to a minimum basis has, however, a deleterious effect on the results, particularly on stretching force constants, as the results of Schlegel et al.[5] show.

We have abandoned the 7s3p basis sets because, similarly to GAUSSIAN 70,[17] our new program TEXAS[11] is able to exploit the equality of valence-shell s and p exponents. The most widely used double-ζ basis set with this constraint is the 4-31G set of Ditchfield et al.[18] This set has the further advantage that it has been used for systematic force field calculations by Schlegel et al.[5] and Blom et al.[6] We still decided to use a slightly smaller basis set, denoted as the 4-21 set. It consists of a 4s core, taken over from Ditchfield et al.,[18] and three sp valence shells for first-row atoms, contracted to double-ζ quality. Our reasons for this were the following.

(1) It is more important to decrease the size of the primitive basis set in a gradient program than in customary SCF programs. Integral and, what is particularly important, gradient evaluation is about twice as fast with the 4-21 basis as with the 4-31G.

(2) For molecules not containing lone pairs on central atoms, the 4-21 force constants and geometries are very close to the 4-31G ones.

(3) There is indication that the 4-31G basis set (and perhaps all sp sets over a certain size) is imbalanced, in that it considerably overestimates bond angles on central atoms with lone pairs. This imbalance is less pronounced in our 4-21 basis.

The parameters of the 4-21 set are given in Table I. These values were obtained by energy optimization on first-row hydrides. The values for hydrogen are those of Huzinaga, multiplied by a scale factor of $(1.15)^2$ (see Table I for references and for the details of the optimization).

We adopt the basis sets 5-31G*, 6-31G*, and 6-31G**[19] in those cases where polarization functions are necessary. These sets have a larger sp part; this is desirable if polarization functions are used, as a polarized small sp basis is significantly less accurate but not much cheaper than, say, a 5-31G* basis. The first two of the above sets produce virtually identical geometries and force constants.

Table II shows the 4-21 molecular geometries, dipole moments, and energies, compared to results obtained by 4-31G, as well as larger basis sets, including near-Hartree-Fock results where available. The 5-31G* calculations were performed in this study; this seems to be the most comprehensive series of completely optimized molecular geometries with polarization functions. Most other results were taken from the literature.

The data in Table II show that the 4-21 basis is very successful indeed in describing molecular geometries, particularly considering its modest size. This is also true for force constants, as demonstrated in section VIII.

III. Choice of Internal Coordinates for the Representation of Potential Surfaces

For small displacements from a reference geometry, the molecular potential energy can be expanded in a power series with respect to a *complete* and *nonredundant* set of internal displacement coordinates $\{q_i\}$:

Table I. 4-21 Basis Set Parameters[a]

atom	type	exponent	coeff for s function	coeff for p functions
H	s	5.951 753	0.168 044 1	
		0.900 989	1.0	
	s	0.200 192	1.0	
B	s,p	2.621	1.0	1.0
		0.5314	−2.337	4.113
	s,p	0.1649	1.0	1.0
C	s,p	3.909 14	1.0	1.0
		0.805 908	−2.525	3.677 44
	s,p	0.216 909	1.0	1.0
N	s,p	5.574	1.0	1.0
		1.159	−2.753	3.629
	s,p	0.2822	1.0	1.0
O	s,p	7.357	1.0	1.0
		1.546	−2.986	3.434
	s,p	0.3617	1.0	1.0
F	s,p	9.685 57	1.0	1.0
		2.059 48	−3.028 56	3.396 57
	s,p	0.473 34	1.0	1.0

[a] The 4s atomic cores for the atoms B to F are identical with those in ref 18. The first two functions with the higher exponents are contracted to a single group. Note the inconsistency in the number of significant figures given; it arose because of insufficient communication among us. However, as the calculations reported in this paper have been performed with exactly this basis set, we prefer to quote it in this form. The hydrogen exponents have been taken from S. Huzinaga, *J. Chem. Phys.*, **42**, 1293 (1965), scaled by $(1.15)^2$. The remaining exponents and contraction coefficients have been energy optimized in the molecules BH_3, CH_4, NH_3, H_2O, and HF at their respective experimental equilibrium geometries (see Table II; for BH_3, r_{BH} was 1.187 Å).

$$V = V_0 - \sum_i \phi_i q_i + \frac{1}{2} \sum_{ij} F_{ij} q_i q_j + \frac{1}{6} \sum_{ijk} F_{ijk} q_i q_j q_k + \ldots$$

$$(1)$$

with $\phi_i = -\partial E/\partial q_i$ and $F_{ijk\ldots} = \partial\partial\partial \cdot E/\partial q_i \partial q_j \partial q_k \cdots$ at the reference geometry. The completeness and nonredundance of the coordinates $\{q_i\}$ assures that (1) is unique. Furthermore, care should be taken that the transformation connecting the coordinates q_i with the Cartesians $\mathbf{x}$

$$\mathbf{q} = \mathbf{B}(\mathbf{x} - \mathbf{x}_0) \qquad (2)$$

is not near singular.[20] This is perhaps best expressed by requiring that the matrix $\mathbf{BB}^+$ is not near singular.

Although all regular coordinate systems are equivalent in principle, there are two further practical requirements for the coordinates: (1) They should facilitate the transfer and comparison of force constants between related molecules. (2) They should allow a simple representation of the dominant anharmonic, i.e., cubic and higher terms in (1).

These requirements can be satisfied only by local internal valence coordinates, i.e., by bond lengths, angles, dihedral angles, etc., and their linear combinations. Cartesian coordinates have very significant disadvantages, particularly if the reference geometry does not coincide with the theoretical equilibrium geometry, as is often the case (see ref 3 for a more detailed discussion). The expression "local" means that the coordinates should extend to only a few atoms; among others, this excludes global symmetry coordinates. Requirement (2) prescribes, in essence, that bond stretchings be used as individual coordinates, instead of their linear combinations. This is because the dominant anharmonicity associated with bond stretchings assumes an almost diagonal form only if individual bond lengths are used as coordinates. There are no such unique rules governing the choice of the deformational coordinates,

Table II. Theoretical Geometries, Dipole Moments, and Energies

molecule symmetry	property[a]	4-21[b]	4-31G[c]	5-31G*[d]	6-31G*[e]	6-31G**[d]	near-H-F[f]	expt[g]
H_2	r_{HH}	0.7318	0.730				0.7337[h]	0.7414[h]
$D_{\infty h}$	$-E$	1.122 80	1.126 83				1.133 50[h]	
BH_3	r_{BH}	1.183			1.1887[i]	1.1882[i]		
D_{3h}	$-E$	26.336 28			26.390 01	26.392 89		
CH_4	r_{CH}	1.0815	1.081[j]	1.0835	1.084	1.084[k]	1.083[l]	1.086[m]
T_d	$-E$	40.112 64	40.139 77	40.187 13	40.195 17	40.201 40[k]	40.213 98[l]	
NH_3	r_{NH}	0.9995	0.991[n]	1.003	1.004	1.001	1.000[p]	(1.012)[q]
C_{3v}	α_{HNH}	112.6	115.8[n]	107.1	107.5	107.5	107.2[p]	(106.7)[q]
	θ^r	47.0	35.0	60.3	59.4	59.4	60.1	(61.2)
	μ	1.763	2.105	1.910		1.858	1.660[p]	(1.482)[p]
	$-E$	56.051 64	56.106 69	56.172 64	56.184 34	56.195 45	56.2219[p]	
	ΔH	6.32	1.7	27.5	27.2	23.0	21.3	(24.2)[q]
H_2O	r_{OH}	0.9630	0.951	0.9477	0.948		0.940[s]	0.957[t]
C_{2v}	α_{HOH}	108.06	111.2	105.51	105.5		106.1[s]	104.5[t]
	μ	2.186		2.192			1.948[s]	1.855[t]
	$-E$	75.820 43	75.908 64	75.994 76	76.010 75		76.0649[s]	
HF	r_{HF}	0.936	0.922	0.9124	0.911		0.898[u]	0.917[v]
$C_{\infty v}$	μ	2.186		1.972			1.905[u]	1.827[v]
	$-E$	99.754 73	99.887 29	99.980 99	100.002 91		100.0656[u]	
C_2H_2	r_{CC}	1.1847	1.190	1.185			1.180[w]	1.203[x]
$D_{\infty h}$	r_{CH}	1.0509	1.051	1.057			1.056[w]	1.061[x]
	$-E$	76.664 86	(76.7111)	76.801 50			76.851 22[w]	
C_2H_4	r_{CC}	1.312	1.316	1.316		1.318[k]	y	1.334[z]
D_{2h}	r_{CH}	1.073	1.073	1.076		1.077[k]		1.081[z]
	α_{HCC}	122.0	122.0	121.8		121.7[k]		121.3[z]
	$-E$	77.870 80	(77.9205)	78.015 44		78.0407[k]		
C_2H_6	r_{CC}	1.5406	1.529	1.5272			aa	1.526[bb]
D_{3d}	r_{CH}	1.0832	1.083	1.0854				
	α_{HCC}	110.9	111.2	111.21				
	$-E$	79.066 51	(79.1148)	79.212 76				
HCN	r_{CN}	1.1368	1.140	1.132			1.127[cc]	1.153[dd]
$C_{\infty v}$	r_{CH}	1.0504	1.051	1.059			1.058[cc]	1.066[dd]
	μ	3.077	(3.244)	3.198				(3.00)[dd]
	$-E$	92.666 03	(92.7308)	92.855 84			92.907 68[cc]	
HNC	r_{NC}	1.1615		1.1536				1.169[ee]
$C_{\infty v}$	r_{NH}	0.980		0.986				0.994[ee]
	μ	2.626		2.817				
	$-E$	92.649 41		92.835 56				
H_2CO	r_{CO}	1.208	1.206	1.1836			1.178[ff]	1.203[z,gg]
C_{2v}	r_{CH}	1.084	1.081	1.0917			1.092[ff]	1.099[z,gg]
	α_{HCO}	122.4	121.8	122.17			122.0[ff]	121.8[z,gg]
	μ	2.798	(3.037)	2.665			2.759[ff]	2.34
	$-E$	113.588 72	(113.6920)	113.843 09			113.902 09[ff]	
CO	r_{CO}	1.130	1.128[hh]	1.113			1.101[ii]	1.128[jj]
$C_{\infty v}$	μ	0.595		0.261			0.153[ii]	−0.112[ii]
	$-E$	112.456 60		112.714 65			112.7879[ii]	
N_2	r_{NN}	1.0840	1.085[hh]	1.078			1.065[kk]	1.098[jj]
$D_{\infty h}$	$-E$	108.660 21		108.921 58			108.9956[kk]	
CO_2	r_{CO}	1.158		1.143			1.135[ll]	1.160[mm]
$D_{\infty h}$	$-E$	187.154 75		187.596 56			187.725 36[ll]	
CH_3NH_2[nn]	r_{CHt}	1.089	(1.089[pp])	1.091				
C_s	r_{CHg}	1.081	(1.080[pp])	1.084				
	r_{NH}	1.000	(0.994[pp])	1.002				(1.011)[qq]
	r_{CN}	1.474	(1.452[pp])	1.453				(1.474)[qq]
	$\theta_{CNH_2}{}^r$	45.0	(41[pp])	53.8				(56)[qq]
	α_{HNH}	110.9	(111.1[pp])	106.8				(106)[qq]
	α_{NCHt}	114.6	(114.6[pp])	114.8				(113)[qq]
	α_{NCHg}	108.9	(109.4[pp])	109.2				(108)[qq]
	μ	1.484		1.527				(1.33)[qq]
	$-E$	94.997 05		95.190 28				

Table II (*Continued*)

molecule symmetry	property[a]	4-21[b]	4-31G[c]	5-31G*[d]	6-31G*[e]	6-31G**[d]	near-H-F[f]	expt[g]
CH₃OH[nn']	r_{CO}	1.446	1.430[rr]	1.399				(1.421)[ss]
C_s	r_{OH}	0.962	0.951[rr]	0.947				(0.963)[ss]
	r_{CHt}	1.079	1.076[rr]	1.081				(1.094)[ss]
	r_{CHg}	1.085	1.082[rr]	1.087				(1.094)[ss]
	α^{av}_{OCH}	109.9	109.9[rr]	110.4				(110.4)[ss]
	α_{COH}	110.4	113.2[rr]	109.4				(108.0)[ss]
	tilt	3.7	3.7[rr]	3.25				(3.2)[ss]
	μ		2.206	1.857				
	$-E$		114.769 63	115.011 78				
C₆H₆	r_{CC}	1.3845						(1.397)[tt]
D_{6h}	r_{CH}	1.0721						(1.084)[tt]
	$-E$	230.233 78					(230.7494)[uu]	

[a] r and α denote bond lengths and angles, in Å = 100 pm and degree units, respectively. μ is the dipole moment in debyes (1 D ≃ 3.335 640 × 10⁻³⁰ Cm), E is the energy in E_h units (1 E_h ≃ 4.359 814 aJ), and ΔH is the inversion barrier in kJ/mol. The number of significant figures varies according to the accuracy of the geometry optimization. [b] Results of this study, using the 4-21 basis set of Table I. Integral threshold[11] is 10⁻⁷. [c] Data were taken from W. A. Lathan, W. J. Hehre, L. A. Curtiss, and J. A. Pople, *J. Am. Chem. Soc.*, **93**, 6377 (1971), and from ref 18 unless indicated otherwise. A number of the geometries were also optimized in ref 5 and 6 and, in general, the agreement with the above references is very good. Values in parentheses show that the corresponding quantity has been evaluated at a standard geometry, rather than the geometry given. [d] This paper. Integral threshold[11] is 10⁻⁷. [e] P. C. Hariharan and J. A. Pople, *Mol. Phys.*, **27**, 209 (1974), except as noted. [f] The quality of the calculations included here varies. It can be safely assumed, however, that the geometry parameters have already converged to the Hartree–Fock limit values. It is not implied that the energies given are the lowest ones available. [g] Experimental data which are not true equilibrium values, or thought to be of lower accuracy than the number of significant figures implies, are given in parentheses. [h] See, for instance, ref 3, p 172. [i] C. E. Blom, private communication. [j] Blom et al.[6a] quote 1.082 Å. [k] The basis set used does not agree exactly with that recommended by Hariharan and Pople[19] in that no scale factors were used for the carbon atom, and the hydrogen exponents were multiplied uniformly by 1.15². It is believed that this has only a negligible effect on the results, particularly on the geometries. [l] Reference 4d. [m] P. Pulay, W. Meyer, and J. E. Boggs, *J. Chem. Phys.*, **68**, 5077 (1978). [n] Schlegel et al.[5a] obtain significantly different results, probably because of incomplete force relaxation. We are thankful to Dr. Blom for comfirming the original figures (see footnote c). [p] A. Rauk, L. C. Allen, and E. Clementi, *J. Chem. Phys.*, **52**, 4133 (1970). [q] W. S. Benedict and E. K. Plyler, *Can. J. Phys.*, **35**, 1235 (1957); J. D. Swalen and J. A. Ibers, *J. Chem. Phys.*, **36**, 1914 (1962). [r] The deviation from planarity is better expressed by θ, the angle of an NH bond with the NH'H″ plane, than by the sum of the three HNH angles. The latter ceases to be a regular coordinate near the planar configuration. The corresponding coordinate, the angle of the CN bond with the NH₂ plane, is also used for methylamine. [s] B. J. Rosenberg, W. C. Ermler, and I. Shavitt, *J. Chem. Phys.*, **65**, 4072 (1976). [t] W. S. Benedict, N. Gailer, and E. K. Plyler, *ibid.*, **24**, 1139 (1956); S. A. Clough, Y. Beers, G. P. Klein, and L. S. Rothman, *ibid.*, **59**, 2254 (1973). [u] W. Meyer and P. Rosmus, *ibid.*, **63**, 2356 (1975). [v] D. E. Mann, B. A. Thrush, D. R. Lide, Jr., J. J. Ball, and N. Acquista, *ibid.*, **34**, 420 (1961); J. S. Muenter and W. Klemperer, *ibid.*, **52**, 6033 (1970). [w] Reference 34. [x] W. J. Lafferty and R. J. Thibault, *J. Mol. Spectrosc.*, **14**, 79 (1964); H. Fast and H. L. Welsh, *ibid.*, **41**, 203 (1972). [y] The present 6-31G** results are probably the best fully optimized Hartree–Fock values. [z] J. L. Duncan, *Mol. Phys.*, **28**, 1177 (1974). [aa] Our 5-31G* results confirm the conclusion of R. Ahlrichs, H. Lischka, B. Zurawski, and W. Kutzelnigg, *J. Chem. Phys.*, **63**, 4685 (1975), that the ethane structure of A. Veillard, *Theor. Chim. Acta*, **18**, 21 (1970), is incompletely optimized; in particular, its r_{CC} bond length (1.551 Å) is too large. Our results are virtually identical with the SCF results of Ahlrichs et al.; moreover, these authors find a small contraction (0.002 Å) when electron correlation is taken into account. This brings our calculated value in very good agreement with experiment. [bb] E. Hirota, K. Matsumura, M. Imachi, M. Fujio, and Y. Tsuno, *J. Chem. Phys.*, **66**, 2660 (1977). [cc] Reference 31. [dd] A. E. Douglas and D. Sharma, *J. Chem. Phys.*, **21**, 448 (1953); J. K. Tyler and J. Sheridan, *Trans. Faraday Soc.*, **59**, 2661 (1963). [ee] R. A. Creswell and A. G. Robiette, to be published. [ff] Reference 4d. [gg] K. Yamada, T. Nakawaga, K. Kuchitsu, and Y. Morino, *J. Mol. Spectrosc.*, **38**, 70 (1971). [hh] Reference 2, p 10. [ii] W. M. Huo, *J. Chem. Phys.*, **43**, 624 (1965). [jj] M. Mizushima, "The Theory of Rotating Diatomic Molecules", Wiley, New York, 1975. [kk] P. E. Cade, K. D. Sales, and A. C. Wahl, *J. Chem. Phys.*, **44**, 1973 (1966). [ll] A. D. McLean and M. Yoshimine, *IBM J. Res. Dev., Suppl.*, **11**, 2 (1967). [mm] G. Herzberg and L. Herzberg, *J. Opt. Soc. Am.*, **43**, 1037 (1953). [nn] In methylamine, CH_t and CH_g refer to CH bonds trans and gauche to the lone pair on the nitrogen atom. In methanol, trans and gauche refer to the OH bond. [pp] Reference 5b. Note that the geometry given in this paper represents just the first extrapolation step of the force relaxation method. Therefore these values are given in parentheses. [qq] D. R. Lide, *J. Chem. Phys.*, **27**, 343 (1957); T. Nishikawa, T. Itoh, and K. Shimoda, *ibid.*, **23**, 1735 (1955). [rr] Reference 6c; these values agree with those of Schlegel et al.[5b] except for the tilt angle, which must be in error in the latter reference. The tilt angle is defined as ⅔($\alpha_{OCH_t} - \alpha_{OCH_g}$). [ss] M. C. L. Gerry, R. M. Lees, and G. Winnewisser, *J. Mol. Spectrosc.*, **61**, 231 (1976). [tt] A. Langseth and P. B. Stoicheff, *Can. J. Phys.*, **34**, 350 (1956). [uu] W. C. Ermler and C. W. Kern, *J. Chem. Phys.*, **58**, 3458 (1973).

although local pseudosymmetries often make one set of coordinates strongly preferable to others.

A set of recommended coordinates, applicable in a wide variety of organic molecules, is defined in Figure 1 and Table III. It is hoped that these coordinates will become standard in future work. Indeed, local coordinates like ours find increasing use since first proposed by Duncan.[21] Besides facilitating the construction of potential functions from the force fields of fragments, such a standardized coordinate system helps to avoid awkward transformations.

A few remarks about these internal coordinates are in order.

(1) The hydrogen atom receives preferential treatment. This is reasonable, considering the unique position of hydrogen, both in vibrational spectroscopy (due to its low mass) and in quantum chemistry.

(2) In ring compounds, we recommend the use of deformational symmetry coordinates appropriate to the corresponding planar ring of D_{nh} symmetry, even if the actual symmetry is lower. Redundancies between stretching and bending coordinates are eliminated by this choice, and it facilitates the comparison of ring rigidity. Our ring deformational coordinates involve the whole ring and are therefore nonlocal. However, this is not a serious drawback as most rings must be handled as separate entities anyway. The only cases where different coordinates are required are large saturated rings.

(3) An interesting alternative to the bond-stretching coordinate is the Simons–Parr–Finlan (SPF) expansion vari-

Table III. Recommended Internal Coordinate System[a]

1. bond stretchings:	individual coordinates rather than combinations. possible exceptions: methyl and methylene groups where symmetrized combinations of the CH stretchings may be used
2. X_mA-BY_n and X_mA=BY_n	torsions and twistings: the sum of all possible dihedral angles X_iABY_j
3. methyl:	sym def = $\alpha_1 + \alpha_2 + \alpha_3 - \beta_1 - \beta_2 - \beta_3$
	asym def = $2\alpha_1 - \alpha_2 - \alpha_3$
	asym def' = $\alpha_2 - \alpha_3$
	rocking = $2\beta_1 - \beta_2 - \beta_3$
	rocking' = $\beta_2 - \beta_3$
4. methylene (sp^3):	CH_2 scissoring = $5\alpha + \gamma$
	CXY scissoring = $\alpha + 5\gamma$
	CH_2 rocking = $\beta_1 - \beta_2 + \beta_3 - \beta_4$
	CH_2 wagging = $\beta_1 + \beta_2 - \beta_3 - \beta_4$
	CH_2 twisting = $\beta_1 - \beta_2 - \beta_3 + \beta_4$
5. methylene (sp^2):	sym def = $2\alpha - \beta_1 - \beta_2$
	rocking = $\beta_1 - \beta_2$
	wagging = X out of CH_2 plane
6. methine (sp^2):	CH rocking = $\beta_1 - \beta_2$
	XCY def = $2\alpha - \beta_1 - \beta_2$
	CH wagging = H out of XCY plane
7. methine (sp^3):	CH rocking = $2\beta_1 - \beta_2 - \beta_3$
	CH rocking' = $\beta_2 - \beta_3$
	XCY def = $4\alpha_{XCY} + \alpha_{XCZ} + \alpha_{YCZ}$
	XCZ and YCZ deformations are analogously defined
8. amino:	scissoring = $2\alpha - \beta_1 - \beta_2$
	rocking = $\beta_1 - \beta_2$
	wagging = X out of NH_2 plane
9. imino (sp^3):	NH rocking = $\beta_1 - \beta_2$
	XNY def = $2\alpha - \beta_1 - \beta_2$
	wagging = H out of XNY plane
10. four-ring:	ring def = $\alpha_1 - \alpha_2 + \alpha_3 - \alpha_4$
	puckering = $\tau_1 - \tau_2 + \tau_3 - \tau_4$
11. five-ring:	ring def = $\alpha_1 + a(\alpha_2 + \alpha_5) + b(\alpha_3 + \alpha_4)$
	ring def' = $(a - b)(\alpha_2 - \alpha_5) + (1 - a)(\alpha_3 - \alpha_4)$
	torsion = $b(\tau_1 + \tau_5) + a(\tau_2 + \tau_4) + \tau_3$
	torsion' = $(a - b)(\tau_4 - \tau_2) + (1 - a)(\tau_5 - \tau_1)$
	where $a = \cos 144°$, $b = \cos 72°$
12. six-ring:	trigonal def = $\alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 + \alpha_5 - \alpha_6$
	asym def = $2\alpha_1 - \alpha_2 - \alpha_3 + 2\alpha_4 - \alpha_5 - \alpha_6$
	asym def' = $\alpha_2 - \alpha_3 + \alpha_5 - \alpha_6$
	puckering = $\tau_1 - \tau_2 + \tau_3 - \tau_4 + \tau_5 - \tau_6$
	asym torsion = $\tau_1 - \tau_3 + \tau_4 - \tau_6$
	asym torsion' = $-\tau_1 + 2\tau_2 - \tau_3 - \tau_4 + 2\tau_5 - \tau_6$

[a] See Figure 1 for the definition of bond angles. In the rings, α_2 is, e.g., the angle 1-2-3 and τ_2 is the dihedral angle 1-2-3-4. Note that the coordinates are still subject to some arbitrariness owing to the freedom in the numbering of the atoms for degenerate coordinates. Also, the sign of the out-of-plane coordinates must be specified in each case. Normalization constants are not given here. For definiteness, they are chosen as $N = (\Sigma c_i^2)^{-1/2}$, where the c_i are the coefficients of the individual valence coordinates. In some cases, e.g., for compounds containing divalent oxygen, the definition of coordinates is obvious and is not given. Note the possibility of using nonstandard valence coordinates: the inverse bond length coordinate, or the coordinate replacing the out-of-plane angle, $(e_1) \times (e_2 \cdot e_3)$, where e_1, e_2, and e_3 are unit vectors directed along the bonds.

able[22]

$$\rho = (r - r_e)/r = 1 - r_e/r \qquad (3)$$

This is essentially an expansion in the inverse powers of r, as the second form in (3) shows. It has significant merits: a power series expansion in the r^{-1} coordinate converges much more quickly than in the usual r coordinate. This is due to the fact that the origin of the strong anharmonicity in bond stretching is the coulomb repulsion of the atomic cores;[3] for this, r^{-1} is a natural coordinate. Although the higher terms in the potential function are less important when using the r^{-1} coordinate than in the usual case, physical anharmonicity is merely shifted to the kinetic energy part.

The SPF variable is dimensionless, and the force constants expressed by it correspond to the "descaled" constants advocated by Mills.[8] This is, however, one of its weaknesses because it contains the equilibrium bond length explicitly. This means that the same potential curve produces different force constants depending on the accepted value of r_e. Therefore the use of the r^{-1} coordinate is recommended, rather than ρ; to make it dimensionless, define it as 1 Å/r.

There have been suggestions to use compliance constants, i.e., the elements of the inverse force constant matrix, instead of the force constants.[23] These have the advantage over force constants that an element, say C_{ij}, depends only on the definition of coordinates q_i and q_j but is independent of the definition of the other coordinates. According to Jones,[24] this may imply better transferability. The latter statement is certainly true if any internal coordinate set is allowed, including highly unusual ones. However, there is no evidence that compliance constants are more transferable than force constants if the latter are evaluated in a carefully chosen, physically meaningful valence coordinate system. Moreover, compliance constants have significant disadvantages, the foremost from our point of view being that there is no direct method for their quantum-chemical evaluation. The simplest procedure is probably the calculation of the full quadratic force constant matrix and its inversion; the alternative procedure of optimizing the geometry with one coordinate given a fixed displacement is usually too expensive. As a consequence, individual elements of the compliance constant matrix cannot be evaluated. A second disadvantage is that the treatment of

anharmonicity is unsolved in the compliance matrix formalism. For these reasons, the compliance formalism is not recommended.

IV. Evaluation of Force Constants

The final result from our quantum mechanical calculations is the gradient, usually in terms of internal coordinates. To determine the force constants, the forces must be differentiated further numerically, e.g., for the quadratic constants

$$F_{ij} \approx [\phi_j(\text{ref}) - \phi_j(q_i = \Delta_i)]/\Delta_i$$
$$\approx [\phi_i(\text{ref}) - \phi_i(q_j = \Delta_j)]/\Delta_j \quad (4)$$

with ϕ_k denoting the force acting along q_k. This scheme yields accurate values only if cubic anharmonicity can be neglected, e.g., by symmetry, or when very small displacements Δ_k are used; the latter is, however, dangerous numerically. Significantly more reliable values are obtained if the gradient is evaluated for two displacements along each coordinate at $\pm\Delta_k$. Besides eliminating the effect of cubic anharmonicity on the quadratic constants, this construction allows the calculation of all diagonal and semidiagonal cubic force constants:

$$F_{ij} = [\phi_j(q_i = -\Delta_i) - \phi_j(q_i = +\Delta_i)]/2\Delta_i$$
$$= [\phi_i(q_j = -\Delta_j) - \phi_i(q_j = +\Delta_j)]/2\Delta_j \quad (5)$$

$$F_{iij} = [2\phi_j(\text{ref}) - \phi_j(q_i = +\Delta_i) - \phi_j(q_i = -\Delta_i)]/\Delta_i^2 \quad (6)$$

When two-sided distortions are used, about half of the SCF iteration time can be saved by extrapolating the starting Fock matrix for, say, the negative displacement, F_-, from the Fock matrices at the reference geometry and at the positive displacement, according to $F_- \approx 2F_0 - F_+$. For larger molecules, one-sided displacements can be used with cubic force constants taken over from smaller molecules to correct for cubic anharmonicity. This speeds up the calculation of quadratic force constants by a factor of nearly 2.

Calculation of the dominant cubic terms in the potential function is quite important. First, these constants allow the transformation of the quadratic force field to a slightly different reference geometry (see later). Second, for a satisfactory description of molecular vibrations, anharmonicity must be considered. It is our experience that in rigid molecules only the diagonal and semidiagonal cubic terms need be considered if a well-chosen internal coordinate system is used. Cubic terms of the form F_{ijk}, $i \neq j \neq k$, are negligible.

Although they appear formally in the same order of perturbation theory as the cubic terms, it seems that most quartic potential constants are unimportant. Diagonal quartic terms may be, however, significant. These can be obtained from the three energy and gradient values at the reference geometry and at $\pm\Delta_i$ by the method of Schlegel et al.,[5a] although this method is sensitive to numerical errors. For stretching coordinates, all higher order constants can be estimated from the quadratic and cubic term by assuming a Morse potential curve. Alternatively, a cubic potential in the R^{-1} variable contains most of the higher terms expressed in the usual R variable. The Morse potential predicts for the quartic force constant expressed in R

$$F_4 = 7F_3^2/9F_2 \quad (7)$$

while a cubic potential in R^{-1} gives

$$F_4 = -12(F_3R_e + 3F_2)/R_e^2 \quad (8)$$

These two expressions have a comparable accuracy, about 10% in most cases.

V. Choice of the Reference Geometry

The choice of the reference geometry is probably the most difficult problem confronting systematic ab initio force con-
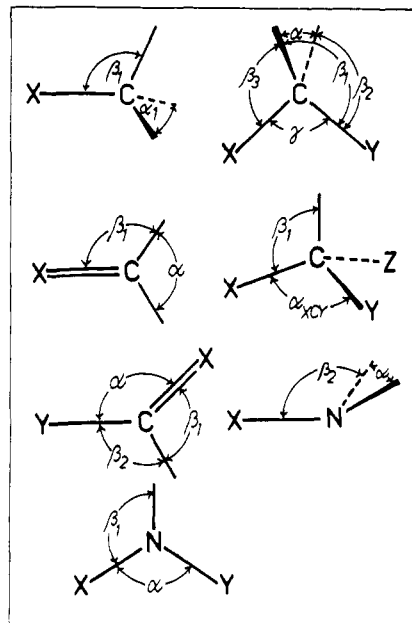

Figure 1. Definition of internal coordinates in methyl, methylene (sp³), methylene (sp²), methine (sp³), methine (sp²), amino, and imino groups.

stant calculations. Both the quadratic force constant matrix and the kinetic energy matrix depend on the molecular geometry; the former is particularly significant for bond-stretching coordinates. There are two basic possibilities for choosing the reference geometry: one can accept either the theoretical geometry or the experimental one. The first has the advantage of not containing empirical elements. However, as shown in ref 3, systematic errors in the Hartree–Fock reference geometry lead to large errors in the theoretical force constants, particularly in the stretching constants. Therefore we have adopted the experimental geometry as the reference geometry in previous work.[4] In series of larger molecules, however, a difficulty arises in this approach: the experimentally derived geometry parameters often show an artificial irregular trend, owing to experimental errors and to differences in the definition of the data (e.g., r_e, r_0, or r_g values). Such a noise in the experimental geometries is easy to detect by comparison with the theoretical geometries.

For the above reason, we advocate the use of empirically corrected theoretical geometries, as practiced by Blom and Altona.[6b] The corrections are established in small molecules and apply to a given basis set. For hydrocarbons and oxo compounds, corrections are applied only to bond lengths, as bond angles are usually well reproduced for these compounds. However, in compounds containing divalent oxygen and trivalent nitrogen, it may be necessary to correct the bond angles as well if the basis set does not contain polarization functions. Table IV contains a few correction values for bond lengths. As Table IV shows, the sign of the correction is different for CC single and double bonds. For intermediate bond lengths, linearly interpolated values are used.

VI. Generation of Distorted Molecular Geometries and the Evaluation of Dipole Moment and Polarizability Derivatives

Hand calculation of the nuclear coordinates for distorted molecular geometries is very strenuous, particularly if distortions along the exact curvilinear coordinates are required. For systematic force constant calculations, an automatic method is needed to calculate the Cartesians x for arbitrary small values of the internal displacement coordinates q. This is part of the program system TEXAS. Inversion of eq 2

Table IV. Some Empirical Corrections to the Theoretical 4-21 Bond Lengths

bond	correction[a]	ref molecule[b]
C—H	+0.005	CH_4[c]
C—H aldehyde	+0.016	H_2CO
C—C[d]	−0.014$_5$	C_2H_6
C=C[d]	+0.022	C_2H_4
C=O	−0.005	H_2CO

[a] Reference bond length = theoretical bond length + correction (in Å = 0.1 nm units). [b] See Table II. [c] CH bond lengths in other compounds indicate that a slightly larger correction, about 0.007 Å, is more appropriate in sp²- and sp-hybridized carbon compounds. Note also that sp basis sets do not reproduce the significant lengthening of the aldehydic CH bond. This necessitates a separate correction for the latter. [d] In the range between 1.31 and 1.56 Å, we recommend the following formula for the correction: $\Delta r = 0.16(1.45 - r_{th})$, where r_{th} is the theoretical 4-21 CC bond length.

yields

$$x^1 = x_0 + Aq \tag{9}$$

where x_0 is the column vector of the Cartesians at the reference geometry, and

$$A = mB^+(BmB^+)^{-1} \tag{10}$$

Here **m** is any nonsingular $3N$ by $3N$ matrix (N is the number of nuclei). Equation 10 is of first-order accuracy; accurate Cartesians are obtained by using it iteratively. To this end, evaluate the internal coordinate values q^1 corresponding to x^1 and calculate

$$x^2 = A(q - q^1) + x^1; x^3 = A(q - q^2) + x^2 \tag{11,12}$$

etc. This procedure converges well even if **A** is not recalculated, if the distortions are not too large.

Dipole moment derivatives, necessary for infrared intensity calculations, are obtained by numerical differentiation

$$\partial\mu/\partial q_i = [\mu(q_i = \Delta_i) - \mu(q_i = -\Delta_i)]/2\Delta_i \tag{13}$$

One should realize, however, that (13) holds only if the Cartesian distortions used to calculate the dipole moments satisfy the Eckart conditions,[25] unless the molecule has no permanent dipole moment. The Cartesians obtained from (12) satisfy the Eckart conditions for a particular isotopic species if the matrix **m** is composed of triplets of inverse atomic masses in the diagonal. For a different isotopic species, a rotational correction must be applied. As we have not found any simple and general method in the literature for doing this, our procedure is described in the Appendix.

Calculation of the polarizability derivatives, necessary for the evaluation of Raman intensities, may be accomplished in a similar fashion, by determining the polarizability components at distorted nuclear conformations, e.g., by applying a small but finite electric field. However, this method is rather cumbersome; moreover, the satisfactory reproduction of polarizability derivatives (and, to a lesser extent, that of the dipole moment derivatives) requires better basis sets than force-field calculations (i.e., polarization functions are mandatory). Therefore we suggest evaluating these quantities in a separate calculation, using an alternative method, recently proposed by McIver and Komornicki.[26] Briefly, this method consists of reversing the order of differentiation and calculating dipole moment and polarizability derivatives as numerical first and second derivatives of the energy gradient with respect to an external electric field. In this method, significant economy can be achieved, as the basis set integrals need be evaluated only once, at the reference geometry. Moreover, a minor modification of the program allows the simultaneous calculation of

Table V. Force Constants and Dipole Moment Derivatives of Borane[a]

constant	4-21	9s2p/5s[b]
F_{rr}	4.087	4.26
$F_{rr'}$	0.026	
$F_{ad,r}$	−0.102	
$F_{ad,ad}$	0.428	0.424
F_{ww}	0.109	0.126
F_{rrr}	−20.49	
$F_{rrr'}$	0.00	
$F_{r,ad,ad}$	−0.06	
$F_{r',ad,ad}$	−0.11	
F_{rww}	−0.05	
$|\partial\mu/\partial r|$[c]	1.264	
$|\partial\mu/\partial ad|$	0.524	
$|\partial\mu/\partial w|$	0.654	

[a] Energy in aJ, bond lengths in Å = 0.1 nm, angles in radians. The angle coordinates are defined as follows: $ad = (2\alpha - \alpha' - \alpha'')/6^{1/2}$; $ad' = (\alpha' - \alpha'')/2^{1/2}$; the angle opposite to r is α; that opposite to r' is α'; etc. $w = (\theta + \theta' + \theta'')/3^{1/2}$, where θ is the angle of r with the $r'r''$ plane. Reference geometry: $r_e = 1.1896$ Å. [b] M. Gelus and W. Kutzelnigg, *Theor. Chim. Acta,* **28,** 103 (1973); note that an inconsistency in the definition of the out-of-plane coordinate in this paper has been corrected here. [c] Dipole moments in D units (1 D ≃ 3.335 64 × 10⁻³⁰ Cm). The sign of the dipole moment derivatives is consistent with B⁺H⁻ partial charges.

the gradient from a number of SCF wave functions which differ only in the imposed electric field. Note that special attention must be paid to the transformation of the Cartesian forces to internal coordinates,[7] as the net torque of the former does not generally vanish in the presence of an external field. Thus it becomes mandatory to use the inverse atomic masses for the matrix **m** in the calculation of the transformation matrix (which is the transpose of **A** in eq 11).

VII. Calculation of Equilibrium and Saddle-Point Geometries and Reaction Paths

In this section, the results of earlier work[3,4,7] are briefly recapitulated and some new development is added. Calculation of molecular geometries can be significantly speeded up, both in terms of computer time and human effort, if the gradient is available. However, geometry prediction requires, besides the gradient, also an estimate of the Hessian, i.e., of the quadratic force constant matrix. Simple steepest descent or similar methods are too expensive for ab initio work. If an estimate, F_0, of the force constant matrix is available, the force relaxation method[7] may be used. In this Newton–Raphson-type method, the correction to the geometry is given by

$$\Delta q = F_0^{-1}\phi \tag{14}$$

where ϕ is the force vector in the ith step. Equation 14 clearly shows the advantages of a well-chosen internal coordinate system over, say, Cartesians, as only the former allow the estimation of a reasonable force constant matrix.

For strained systems, unusual molecules, and saddle points, it is difficult to estimate the **F** matrix, even in internal coordinates. These systems often have large coupling force constants, so that a simple diagonal **F** may be a poor approximation. Our experience in the past few years shows that it is generally worthwhile to perform first some calculations for a good estimate of the **F** matrix in these cases. The gain from the faster convergence of (14) usually outweighs the additional expense. One may perform a complete force-field calculation with a simpler basis set (Schlegel[27] even advocates the use of CNDO force constants in conjunction with the ab initio gradient) and scale it with suitable scale factors. Alternatively, calculations may be performed on fragments, and the force constants may be united to a composite matrix.

Table VI. Force Constants and Dipole Moment Derivatives of Methane[a]

constant	4-21	4-31G[b]	7s3p/3s/1[c]	near-H–F[d]
F_{rr}	5.590	5.584	5.780	5.521
$F_{rr'}$	0.044	0.043	0.036	0.049
$F_{ed,ed}$	0.699	0.674	0.708	0.642
$F_{r,fd}$	0.130	0.131[e]	0.118	0.110
$F_{fd,fd}$	0.674	0.667	0.668	0.612
F_{rrr}	−32.74	−32.65	−33.4	−31.03
$F_{rrr'}$	0.04	0.12		0.09
$F_{r,ed,ed}$	−0.14			−0.11
$F_{r,fd,fd}$	−0.15			−0.10
$F_{fd,fd,fd}$	0.07			0.05
$\partial\mu/\partial r$	0.795[f]	0.989[f,g]	0.952[f]	0.846[f]
$\partial\mu/\partial fd$	0.451[f]	0.501[f,g]	0.363[f]	0.360[f]

[a] See Table V for units. Deformational coordinates are defined as follows: $ed = (2\alpha_{12} + 2\alpha_{34} - \alpha_{13} - \alpha_{14} - \alpha_{23} - \alpha_{24})/12^{1/2}$; $fd = (\alpha_{12} - \alpha_{34})/2^{1/2}$ and their symmetry-related degenerate pairs. The stretching coordinate r refers to r_1 in this scheme with obvious notation. The reference geometry adopted is $r_{CH} = 1.090$ Å. [b] Reference 5a, corrected for the present reference geometry. [c] Reference 4b, corrected for the present reference geometry. [d] Harmonic data are taken from calculation no. 15 of ref 4d, anharmonic data and the correction for the reference geometry from P. Pulay, W. Meyer, and J. E. Boggs, *J. Chem. Phys.*, **68**, 5077 (1978); this paper also discusses correlation effects. [e] As a recalculation has shown, the value given in Table IV of ref 5a is in error by a factor of 2. Correctly, $F_{34} = 0.262$, hence $F_{r,fd} = 0.131$. [f] The dipole moment derivatives are consistent with C^+H^- partial charges for the bond stretching, and with C^-H^+ for the *fd* deformation. [g] This work.

The force relaxation method converges in a single step if the potential surface is strictly quadratic and if the force-constant matrix is accurately known. In practice neither of these conditions is likely to be fulfilled exactly. If the geometry correction Δq is not small then it is worthwhile to include the cubic part of the potential energy in the minimization. If the complete force field has been determined from positive and negative displacements, then all diagonal and semidiagonal cubic force constants may be included in an iterative fashion: one solves

$$\Delta q = F_0^{-1}[\phi + \psi(\Delta q)] \tag{15}$$

iteratively until self-consistency is obtained. The anharmonic correction to the forces is given by

$$\psi(\Delta q)_k = \sum_i \Delta q_i(F_{kki}\Delta q_k + \tfrac{1}{2}F_{iik}\Delta q_i) + F_{kkk}\Delta q_k^2 \tag{16}$$

If both the geometry and the force field are to be determined, then this method yields the final geometry essentially in one step. However, it is too expensive if the force field is not needed for other purposes. Even in these cases, the inclusion of the dominant anharmonic terms (possibly transferred from smaller molecules) accelerates the convergence of the force relaxation method.

In general, eq 14 converges toward a stationary point on the potential surface. This point is a local minimum if the (quadratic) force constant matrix is positive definite, and a saddle point if it has one negative eigenvalue. Saddle points are of special importance, as their determination is very cumbersome in the ordinary energy hypersurface method. On the other hand, the force relaxation method converges as well for saddle points as for minima if a proper force-constant matrix is used. More complex methods, e.g., the minimization of the length of the gradient vector,[28] have been suggested, but these suffer from slow convergence and the appearance of false minima. Our experience in calculating the transition state for the methyl isocyanide rearrangement reaction[29] shows that the most practical method is to determine the complete quadratic (and possibly also part of the cubic) force field at a qualitatively correct reference geometry using a small basis set. This force field is then used in conjunction with the gradient evaluated with better basis sets; some of the most important force constants may be recalculated by using the better basis, or improved with the variable metric technique (see below). Botschwina[30] has recently discussed a similar, potentially very useful algorithm for molecular geometry determination. He

calculates one-dimensional cross sections of the multidimensional potential surface by a sophisticated, e.g., highly correlated, wave function (usually without the gradient) with cross terms taken from simpler Hartree–Fock calculations using the gradient. With the proper choice of the coordinates, the accuracy of the composite potential surface may approach that of the correlated surface.

It is possible to gather information about the force field during the iterative determination of molecular geometries. This is the basis of the variable metric minimization technique, first used for molecular geometry determination by McIver and Komornicki[31] (for an alternative formulation by Meyer, see ref 3).

For constrained geometry optimization, the same procedure can be used as in the unconstrained case, but the diagonal force constants, corresponding to the frozen coordinates, should be assigned an infinite value. The modified compliance matrix, F^{-1*}, is obtained by omitting the rows and columns of F which correspond to the frozen coordinates, inverting the matrix, and restoring the original dimension by zeroes.

VIII. Results for Force Constants and Dipole Moment Derivatives

Tables V–IX contain the 4-21 force constants and dipole moment derivatives of first-period hydrides, compared to results obtained with other basis sets: 4-31G, 7s3p/3s/1, and near-Hartree–Fock results. The constants are presented in the internal coordinate representation advocated in section III. The complete quadratic force field is given, as well as a few selected cubic constants: diagonal and semidiagonal stretchings, diagonal deformational constants, and terms of the form $F_{r\alpha\alpha}$; the latter are sometimes significant in Fermi resonances. No experimental figures are shown, as our objective is to compare different basis sets. Moreover, a fair comparison to experiment must include a lengthy discussion of the reliability of the experimental values, and depends sensitively on the reference geometry adopted. More meaningful would be a comparison of Hartree–Fock and highly correlated potential constants: this is available for methane, water, and hydrogen fluoride (see the notes on Tables VI, VIII, and IX).

The conclusions from Tables V–IX can be summarized as follows.

(1) The 4-21 basis set is very similar to the 4-31G set. The stretching F_{rr} constants agree within about 1%, except for hydrogen fluoride. Stretch–stretch coupling terms are insignificant in these compounds; they agree to within ±0.015

Table VII. Force Constants and Dipole Moment Derivatives of Ammonia[a]

constant	4-21	4-31G[b]	7s3p/3s/1[c]	5-31G*	near-H-F[d]
F_{rr}	7.242	7.328	7.704	7.479	7.380
$F_{rr'}$	-0.010	-0.020	-0.060	-0.022	-0.002
$F_{r,sd}$	0.397	0.374	0.350	0.279	0.275
$F_{r,ad}$	-0.220	-0.209	-0.169	-0.152	-0.149
$F_{sd,sd}$	0.629	0.615	0.606	0.669	0.593
$F_{ad,ad}$	0.864	0.828	0.805	0.800	0.743
F_{rrr}	-47.33	-48.08	-49.54	-47.90	
$F_{rrr'}$	-0.14	0.02		0.13	
$F_{r,sd,sd}$	-0.03			0.08	
$F_{r,ad,ad}$	-0.17			-0.12	
$F_{r',ad,ad}$	-0.28			-0.19	
$F_{sd,sd,sd}$	-0.52			-0.18	
$F_{ad,ad,ad}$	-0.10			-0.06	
$\partial\mu_z/\partial r$[e]	0.012		-0.170	0.040	-0.010
$\partial\mu_x/\partial r$[e]	-0.207		-0.299	-0.007	0.171
$\partial\mu_z/\partial sd$[e]	-1.919		-2.017	-1.872	-1.819
$\partial\mu_x/\partial ad$[e]	0.365		0.402	0.440	0.439

[a] See Table V for units. Internal coordinates are defined as follows: stretchings, r, r', r'' (angle α_{23} lies opposite to r); deformations, $sd = (\alpha_{23} + \alpha_{12} + \alpha_{13})/3^{1/2}$, $ad = (2\alpha_{23} - \alpha_{13} - \alpha_{12})/6^{1/2}$, $ad' = (\alpha_{13} - \alpha_{23})/2^{1/2}$. Note that the sum of the three out-of-plane angles is a better coordinate for the description of the symmetric deformation, particularly for large amplitudes. It is, however, not yet accepted generally. The following reference geometry is adopted: $r_{NH} = 1.0116$ Å, $\alpha_{HNH} = 106.67°$. [b] Reference 5a, corrected for the present reference geometry. [c] Reference 4b. [d] Reference 4e. [e] The z axis is the C_3 symmetry axis. The present results refer to the $^{14}N^1H_3$ isotopic species. The signs are defined in ref 4e.

Table VIII. Force Constants and Dipole Moment Derivatives of Water[a]

constant	4-21	4-31G[b,e]	4-31G[c,e]	7s3p/3s/1[d]	near-H-F[f]
F_{rr}	8.723	8.689	8.684	9.372	8.694
$F_{rr'}$	-0.081	-0.067	-0.066	-0.156	-0.052
$F_{r\alpha}$	0.381	0.370	0.355	0.280	0.240
$F_{\alpha\alpha}$	0.918	0.878	0.880	0.757	0.783
F_{rrr}	-59.4	-60.		-62.7	-59.7
$F_{rrr'}$	-0.22	-0.14			0.00
$F_{r\alpha\alpha}$	-0.37	-0.43			-0.32
$F_{\alpha\alpha\alpha}$	-0.87	-0.81		-0.75	-0.76
$\partial\mu_z/\partial r$	0.100			-0.098	0.483
$\partial\mu_y/\partial r$[g]	0.220			0.186	0.937
$\partial\mu_z/\partial\alpha$	0.874			0.833	1.003

[a] Reference geometry: $r_{OH} = 0.9573$ Å, $\alpha_{HOH} = 104.52°$. See Table V for units. [b] Reference 5a. [c] Reference 6c. [d] Reference 4a. [e] In order to transform the force constants to the present reference geometry, the anharmonic constants given by Schlegel et al.[5a] were used, except for the quartic deformational constant, which is probably seriously in error numerically in that work. A value $F_{\alpha\alpha\alpha\alpha} = -1$ aJ/rad^4, taken from the reference in footnote f, was used instead. Also, because of the significant difference in the reference bond angle, the constants $F_{rr,\alpha} \simeq -0.55$ aJ/Å^2·rad, and $F_{rrr\alpha} \simeq 0.67$ aJ/Å^3·rad both taken from footnote f had to be included. [f] SCF results of B. J. Rosenberg, W. C. Ermler, and I. Shavitt, *J. Chem. Phys.*, **65**, 4072 (1976), transformed to the present reference geometry. The effect of electron correlation on the force constants is thoroughly discussed by these authors; similar results have been obtained by W. Meyer, Proceedings of the SRC Atlas Symposium No. 4, "Quantum Chemistry—The State of Art", V. R. Saunders and J. Brown, Eds., Atlas Computer Laboratory, Chilton, England, 1975, p 97. Several previous calculations also gave essentially identical results at the SCF level: W. C. Ermler and C. W. Kern, *J. Chem. Phys.*, **55**, 485 (1971); B. J. Krohn, W. C. Ermler, and C. W. Kern, *ibid.*, **60**, 22 (1974); T. H. Dunning, Jr., R. M. Pitzer, and S. Aung, *ibid.*, **57**, 5044 (1972); and the above paper of Meyer. [g] In order to be compatible with the results of Rosenberg et al.,[f] this derivative refers to an isotopic species where atoms O and H' (i.e., the hydrogen atom not involved in bond r) have infinite masses. The z axis coincides with the C_2 symmetry axis, and the molecule lies in the yz plane, with r pointing in the first quadrant (from O to H). At the present reference geometry, the derivative appropriate to the $^1H_2^{16}O$ species is obtained by adding 0.029 52 Å^{-1} μ_0, where μ_0 is the static dipole moment (positive for this orientation of the molecule).

aJ/Å^2 (obviously, a percentage error figure has little meaning for these constants). The deformational constants of the 4-21 set are ~3% higher than those of the 4-31G set. As the theoretical values show systematic deviations of ~20% anyway, and must be scaled to the experimental frequencies, a systematic deviation of this order is negligible and can be absorbed in the scale factors. Agreement is almost perfect for the stretching-deformation couplings. Cubic constants show generally good agreement; for the most important F_{rrr} type the values predicted by the two basis sets are virtually identical.

(2) As compared to near-Hartree-Fock results, both the 4-21 and 4-31G results are qualitatively correct; particularly reassuring is the correct sign of the coupling force constants. However, both bases systematically overestimate most force constants, particularly the diagonal deformational ones and the stretching-deformation couplings. These deviations are larger in the lone-pair compounds, showing the importance of polarization functions. For ammonia, our 5-31G* results are also shown; note that, contrary to the case of methane, this basis is still quite far from the Hartree-Fock limit.

(3) 4-21 dipole moment derivatives are near the Hartree-Fock limit in methane and for the deformational coordinates in ammonia and water. However, the stretching derivatives are not well reproduced, particularly in ammonia and in water. Fortunately, the stretching derivatives are probably more transferable than the deformational ones. We conclude that the 4-21 values may be useful for the prediction of infrared intensities, particularly in hydrocarbons.

Table IX. Force Constants and Dipole Moment Derivative of Hydrogen Fluoride[a]

constant	4-21	4-31G[b]	7s3p/3s/1[c]	near-H-F[d]
F_{rr}	10.380	9.874	10.751	9.839
F_{rrr}	−70.09	−70.11	−75.06	−71.84
$\partial\mu/\partial r^{e}$	0.939		0.941	1.930

[a] See Table V for the units. Reference geometry: $r_{FH} = 0.917$ Å. [b] Reference 5a. [c] Reference 4b. [d] W. Meyer and P. Rosmus, *J. Chem. Phys.*, **63**, 2356 (1975), transformed to the present reference geometry. These authors also give highly correlated potential curves. [e] $H^{+}F^{-}$.

The above conclusions, reached in simple hydrides, are confirmed by our 4-21 calculations on more complex molecules, including C_2 hydrocarbons, formaldehyde, methylamine, hydrogen cyanide, methyl cyanide and isocyanide, methyl alcohol, formic acid and its methyl ester, acetaldehyde, glyoxal, acrolein, butadiene, isoprene, formamide, acetamide and their *N*-methyl derivatives, and benzene. This material is being analyzed. For the simpler molecules on the above list, 4-31G force fields are also available;[5b,e,6b,c,e,32-34] the agreement between the two sets is generally very good. There is also a growing number of molecules for which force fields have been determined using extended basis sets: formaldehyde,[4f] hydrogen cyanide,[3,35] carbon dioxide,[36] ethylene,[37] and the symmetric vibrations of acetylene.[38] A comparison of these results with our limited 4-21 basis set results shows the same qualitative agreement and the same systematic deviations as for the hydride series. A noteworthy exception is carbon dioxide, where an sp basis set is clearly inadequate (note that our 4-21 results are virtually identical with the sp basis set results in ref 36). For hydrogen cyanide[35] and acetylene,[38] correlated potential surfaces have also been determined. These studies, as well as those on the hydrides, show that correlation effects are not dramatic for these molecules. Coupling force constants seem particularly insensitive to correlation effects. This is reassuring, in that a simple scaling procedure can empirically correct for most of the deficiencies of the Hartree-Fock model.

Finally, it is an interesting question how a small basis set like ours is able to reproduce the complete cubic force field. This is shown in Table X for methane. Although methane is probably a favorable case, the good agreement between the 4-21 results and those obtained with a larger basis set looks certainly optimistic. Note that some previous attempts of this kind are not conclusive for the smaller constants because of numerical problems.

Acknowledgments. This work has been supported in part by grants from the U.S. National Science Foundation and the Hungarian Institute for Cultural Relations as part of the cooperative research program of The University of Texas at Austin and the Central Research Institute for Chemistry of the Hungarian Academy of Sciences. It has also been supported in part by a grant from The Robert A. Welch Foundation.

Appendix. Rotational Correction to the Theoretical Dipole Moment Derivatives

As mentioned in section VI, the calculated dipole moment derivatives of polar molecules can be used directly for the calculation of infrared intensities only if the nuclear distortions (relative to the chosen reference geometry) satisfy the Eckart conditions.[25] Such distorted geometries can be generated by means of eq 10-12. The latter are, as the Eckart conditions themselves, mass dependent, so that different Cartesian distortions must be used for different isotopic species. Moreover, in the quantum-chemical literature, Cartesian distortions are

Table X. Cubic Force Field of Methane[a]

constant	4-21	9s5p1d/5s1p[b]	expt[c]
F_{11}	5.722	5.673	5.422
F_{22}	0.699	0.646	0.585
F_{33}	5.546	5.472	5.370
F_{34}	0.259	0.220	0.221
F_{44}	0.674	0.619	0.549
F_{111}	−16.25	−15.08	−15.27
F_{122}	−0.28	−0.23	−0.32
F_{133}	−16.42	−15.47	−15.66
F_{134}	−0.15	−0.07	
F_{144}	−0.31	−0.20	−0.11
F_{222}	0.07	0.10	0.10
F_{233}	−0.42	−0.37	
F_{234}	−0.24	−0.16	
F_{244}	−0.52	−0.39	−0.31
F_{333}	16.49	15.65	15.84
F_{334}	−0.22	−0.27	
F_{344}	0.12	0.10	
F_{444}	0.50	0.43	0.35

[a] In symmetry coordinate representation. See footnote b for definitions and units. [b] P. Pulay, W. Meyer, and J. E. Boggs, *J. Chem. Phys.*, **68**, 5077 (1978). [c] D. L. Gray and A. G. Robiette, *Mol. Phys.*, in press. Only six of the cubic constants could be determined from the experimental data; the others were fixed at the ab initio values (column 2).

often used which do not correspond to any existing isotopic species (see, e.g., Table VIII). The first step in calculating the rotational correction to the dipole moment derivatives consists of determining the correctly oriented set of distorted nuclear Cartesians, through eq 10-12. Both these and the original set of distorted coordinates are translated so that the center of mass is at the origin of the Cartesian frame. The 3×3 orthogonal matrix which rotates the old Cartesian position vectors $\{r_i, i = 1, \ldots, N\}$ to the new ones r_i' is expressed by the three parameters of an antisymmetric matrix $\mathbf{K}$ (see, e.g., ref 39):

$$r_i' = \mathbf{U}r_i = (\mathbf{I} - \mathbf{K})^{-1}(\mathbf{I} + \mathbf{K})r_i \qquad (17)$$

where

$$\mathbf{K} = \begin{pmatrix} 0 & k_1 & k_2 \\ -k_1 & 0 & k_3 \\ -k_2 & -k_3 & 0 \end{pmatrix}$$

From this

$$\mathbf{K}(r_i' + r_i) = r_i' - r_i \qquad (18)$$

and the three parameters in $\mathbf{K}$ are given by a system of $3N$ linear equations:

$$(y_i' + y_i)k_1 + (z_i' + z_i)k_2 = x_i' - x_i$$
$$-(x_i' + x_i)k_1 + (z_i' + z_i)k_3 = y_i' - y_i \qquad (19)$$
$$-(x_i' + x_i)k_2 - (y_i' + y_i)k_3 = z_i' - z_i$$

For numerical reasons, this system of equations is solved by the method of least squares: denoting the coefficient matrix in (19) by $\mathbf{J}$, the 3-vector $\mathbf{k}$ is obtained from

$$\mathbf{k} = (\mathbf{J}^{+}\mathbf{J})^{-1}\mathbf{J}^{+}(x' - x) \qquad (20)$$

The rotated dipole moment, μ', is simply expressed by the original dipole as

$$\mu' = \mathbf{U}\mu = (\mathbf{I} - \mathbf{K})^{-1}(\mathbf{I} + \mathbf{K})\mu = [2(\mathbf{I} - \mathbf{K})^{-1} - \mathbf{I}]\mu \qquad (21)$$

References and Notes

(1) (a) Department of General and Inorganic Chemistry, Eötvös L. University, Budapest; (b) The University of Texas at Austin.
(2) J. A. Pople in "Modern Theoretical Chemistry", Vol. 4, H. F. Schaefer III,

Ed., Plenum Press, New York, 1977, pp 1–27.

(3) P. Pulay in ref 2, pp 153–185.

(4) (a) P. Pulay, *Mol. Phys.*, **18**, 473–480 (1970); (b) *ibid.*, **21**, 329–339 (1971); (c) P. Pulay and W. Meyer, *J. Mol. Spectrosc.*, **40**, 59–70 (1971); (d) W. Meyer and P. Pulay, *J. Chem. Phys.*, **56**, 2109–2116 (1972); (e) P. Pulay and W. Meyer, *ibid.*, **57**, 3337–3340 (1972); (f) W. Meyer and P. Pulay, *Theor. Chim. Acta*, **32**, 253–264 (1974); (g) P. Pulay and W. Meyer, *Mol. Phys.*, **27**, 473–490 (1974); (h) W. Sawodny and P. Pulay, *J. Mol. Spectrosc.*, **51**, 135–141 (1974); (i) P. Pulay and F. Török, *J. Mol. Struct.*, **29**, 239–246 (1975); (j) P. Pulay, A. Ruoff, and W. Sawodny, *Mol. Phys.*, **30**, 1123–1131 (1975); (k) K. Molt, W. Sawodny, P. Pulay, and G. Fogarasi, *ibid.*, **32**, 169–176 (1976); (l) P. Botschwina, *Chem. Phys. Lett.*, **29**, 98–101, 580–583 (1974); (m) W. Bleicher and P. Botschwina, *Mol. Phys.*, **30**, 1029–1036 (1974); (n) P. Botschwina, K. Pecul, and H. Preuss, *Z. Naturforsch. A*, **30**, 1015–1017 (1975); (o) P. Botschwina, W. Meyer, and A. Semkow, *Chem. Phys.*, **15**, 25–34 (1976); (p) P. Botschwina, E. Nachbaur, and B. M. Rode, *Chem. Phys. Lett.*, **41**, 486–489 (1976); (q) P. Botschwina, *Mol. Phys.*, **32**, 729–733 (1976); (r) P. Botschwina, K. Srinivasan, and W. Meyer, *ibid.*, **35**, 1177–1181 (1978).

(5) (a) H. B. Schlegel, S. Wolfe, and F. Bernardi, *J. Chem. Phys.*, **63**, 3632–3638 (1975); (b) *ibid.*, **67**, 4181–4193 (1977); (c) H. B. Schlegel, S. Wolfe, and K. Mislow, *J. Chem. Soc., Chem. Commun.*, 246–247 (1975); (d) H. B. Schlegel, S. Wolfe, and F. Bernardi, *J. Chem. Phys.*, **67**, 4194–4198 (1977); (e) H. B. Schlegel, private communication (ethane, ethylene, formaldehyde, cyclopropane, thioformaldehyde).

(6) (a) C. E. Blom, P. J. Slingerland, and C. Altona, *Mol. Phys.*, **31**, 1359–1376 (1976); (b) C. E. Blom and C. Altona, *ibid.*, **31**, 1377–1391 (1976); (c) C. E. Blom, L. P. Otto, and C. Altona, *ibid.*, **32**, 1137–1149 (1976); (d) C. E. Blom and C. Altona, *ibid.*, **33**, 875–885 (1977); (e) *ibid.*, **34**, 177–192 (1977); (f) C. E. Blom, B. Altona, and A. Oskam, *ibid.*, **34**, 557–571 (1976); (g) C. E. Blom and A. Müller, *J. Mol. Spectrosc.*, **70**, 449–458 (1978).

(7) P. Pulay, *Mol. Phys.*, **17**, 197–204 (1969).

(8) I. M. Mills, *Spec. Period. Rep., Theor. Chem.*, **1**, 110–159 (1974).

(9) G. Fogarasi, P. Pulay, K. Molt, and W. Sawodny, *Mol. Phys.*, **33**, 1565–1570 (1977).

(10) (a) F. Török, A. Hegedüs, K. Kósa, and P. Pulay, *J. Mol. Struct.*, **32**, 93–99 (1976); (b) Yu. N. Panchenko, P. Pulay, and F. Török, *ibid.*, **34**, 283–289 (1976); (c) G. Fogarasi and P. Pulay, *ibid.*, **39**, 275–280 (1977).

(11) P. Pulay, *Theor. Chim. Acta*, **50**, 299–312 (1979).

(12) (a) W. Meyer, *J. Chem. Phys.*, **64**, 2901–2907 (1976); (b) W. Meyer and P. Pulay, to be published.

(13) W. Meyer and P. Pulay, MOLPRO Description, München, Stuttgart, 1969.

(14) I. G. Csizmadia, M. C. Harrison, J. W. Moskowitz, and B. T. Sutcliffe, *Theor.*

Chim. Acta, **6**, 191–216 (1966).

(15) B. Roos and P. Siegbahn, *Theor. Chim. Acta*, **17**, 209–215 (1970).

(16) W. Kutzelnigg, private communication.

(17) W. J. Hehre, W. A. Lathan, R. Ditchfield, M. D. Newton, and J. A. Pople, GAUSSIAN 70, Quantum Chemistry Program Exchange, Indiana University, Bloomington, Ind., 1972, Program No. 236.

(18) R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.*, **54**, 724–728 (1971).

(19) W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.*, **56**, 2257–2261 (1972); P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta*, **28**, 213–222 (1973).

(20) Examples of near-singular internal coordinates are the sum of the three XAX angles in a nearly planar AX₃ molecule, or the dihedral angle of a near-linear XYZU chain.

(21) J. L. Duncan, *Spectrochim. Acta*, **20**, 1197–1221 (1964).

(22) G. Simons, R. G. Parr, and J. M. Finlan, *J. Chem. Phys.*, **59**, 3229–3234 (1973); P. F. Fougere and R. K. Nesbet, *ibid.*, **44**, 285–298 (1966).

(23) J. C. Decius, *J. Chem. Phys.*, **38**, 241–248 (1963).

(24) L. H. Jones and R. R. Ryan, *J. Chem. Phys.*, **52**, 2003–2004 (1970).

(25) C. Eckart, *Phys. Rev.*, **47**, 552–558 (1935).

(26) This method is referred to briefly by M. Dupuis and H. F. King, *J. Chem. Phys.*, **68**, 3998–4004 (1978).

(27) H. B. Schlegel, private communication.

(28) J. W. McIver, Jr., and A. Komornicki, *J. Am. Chem. Soc.*, **94**, 2625–2633 (1972).

(29) P. Pulay and H. F. Schaefer III, to be published.

(30) P. Botschwina, *Chem. Phys.*, **28**, 231–241 (1978).

(31) J. W. McIver and A. Komornicki, *Chem. Phys. Lett.*, **10**, 303–306 (1971).

(32) Both 4-21 and 4-31G results are cited by J. L. Duncan, D. C. McKean, F. Tullini, G. D. Nivellini, and J. Perez Pena, *J. Mol. Spectrosc.*, **69**, 123–140 (1978).

(33) L. S. Bartell, S. Fitzwater, and W. J. Hehre, *J. Chem. Phys.*, **63**, 4750–4758 (1975).

(34) L. A. Curtiss and J. A. Pople, *J. Mol. Spectrosc.*, **48**, 413–426 (1973).

(35) U. Wahlgren, J. Pacansky, and P. S. Bagus, *J. Chem. Phys.*, **63**, 2874–2881 (1975).

(36) J. Pacansky, U. Wahlgren, and P. S. Bagus, *Theor. Chim. Acta*, **41**, 301–309 (1976).

(37) G. Fogarasi and P. Pulay, *Acta Chim. Acad. Sci. Hung.*, submitted.

(38) P. S. Bagus, J. Pacansky, and U. Wahlgren, *J. Chem. Phys.*, **67**, 618–623 (1977).

(39) P. Pulay, *Acta Chim. Acad. Sci. Hung.*, **52**, 49–56 (1967).

Theoretical ab Initio SCF Investigation of the Photochemical Behavior of the Three-Membered Rings. 5. Oxaziridine[1]

B. Bigot,* D. Roux, A. Sevin, and A. Devaquet

Contribution from the Laboratoire de Chimie Organique Théorique,[2] Université Pierre et Marie Curie, 4 Place Jussieu, 75230, Paris Cedex 05, France. Received November 1, 1978

Abstract: In a first step, MO natural correlation concepts are used in order to understand the potential energy curve profile of the ground and first excited states of the oxaziridine system when different reaction paths are simulated. In a second step, the conclusions of this analysis are corroborated by calculating these potential energy curves by an ab initio SCF CI method. On the one hand, the oxaziridine-nitrone transformations are examined in relation with the cis-trans nitrone isomerization which involves the same intermediate. On the other hand, the formamide-oxaziridine isomerization is studied in relation with the nitrene plus formaldehyde formation. The theoretical results afford a rationale for the spin multiplicity selectivity as well as for an understanding of the origins of various stereospecific transformations experimentally observed. Some predictions relative to the influence of substitutions or solvation are reported.

Since the pioneering works of Emmons[3] and Calvin et al.,[4a] it is now clearly established[5] that oxaziridines may intervene as intermediates in the photochemical reactions of nitrones.[6] Indeed, isolable oxaziridines can be photochemically formed from nitrones and then thermally isomerized either back to nitrones or to other products such as amides. Various authors[4b,7] have proved that oxaziridines can also be converted photochemically to amides. It was observed that the last transformation possesses regiospecificity[3,7a,8] and stereospecificity[8] characters, as well as the photochemical transformation which links nitrone and oxaziridine.[9]

In 1965, Tanaka et al.[10] determined that this nitrone photocyclization involves a singlet state while the photochemically induced cis-trans nitrone isomerization involves a triplet state. Lastly, Bjorgo et al.[11] proved that oxaziridine can also isomerize photochemically back to nitrone and that this process is, in some cases, stereospecific.

Notwithstanding the quality and the number of these experimental results and the theoretical works[9b,12] they called forth, no complete understanding of all these monomolecular reactions has been proposed. For example, what are the factors which govern the NO/CO rupture dichotomy, what is the