# The Propagation of Basis-Set Error and Geometry Optimization in *Ab Initio* Calculations. II. Correlation Between the Balance of Gaussian Basis Sets and Calculated Molecular Properties

**Paul G. Mezey and Ernst -C. Hass**

*Department of Chemistry, University of Saskatchewan, Saskatoon, Canada S7N OWO*
*Received 12 August 1982; accepted 7 March, 1983*

Basis-set balance parameters, defined in terms of various projections of an abstract force vector in the space spanned by the logarithms of orbital exponents, are evaluated for a sample of 100 Gaussian basis sets. These basis sets are taken from a random Gaussian distribution of bases, centered on the best energy, fully variational uniform quality (UQ) atomic orbital (AO) basis sets. With each basis geometry optimization has been carried out for model molecule dimethyl sulfoxide, the wavefunction of which molecule is exceptionally sensitive to basis-set errors. Correlations between the balance of basis sets and calculated molecular properties are analyzed.

## INTRODUCTION

The optimum theoretical geometries of molecules, as well as theoretical reaction paths and potential energy hypersurfaces, calculated using *ab initio* self-consistent field (SCF) molecular orbital (MO) techniques, are always subject to certain constraints. The most obvious, and the most crucial of such constraints, beyond the choice of the *ab initio* method itself, is the choice of basis functions used in the expansion of molecular orbitals. Usually, the molecular basis set is a combination of finite *atomic* basis sets,[1-4] optimized for energy or for a certain molecular property (e.g., for experimental geometry in a series of molecules). Energy-optimized basis sets are very special when used within the framework of any variational method, such as the Hartree–Fock–Roothaan method, since the entire determination of the molecular wavefunction is based on energy optimization of the linear (MO) coefficients. In energy-optimized basis sets the parameters of the basis functions, e.g., orbital exponents of Gaussian-type orbitals, play a similar role as the linear coefficients of the MO expansions, and the calculated wavefunction is consistently variational.

The "best" wavefunction for a given set of constraints is the variational optimum that implies energy optimization. That is, quality *is* associated with energy optimization. Most other criteria lack the universality and mathematical clarity of the energy criterion (variational theorem), the very criterion which is already used for the determination of the linear (MO) coefficients. A consistent application of the variational theorem implies energy optimization of orbital exponents. The alternative, at its logical extreme, would be to drop the energy optimization for the linear MO coefficients too, and to replace it with some "calibration procedure," e.g., calibration of MO coefficients for experimental geometry. Such methods, however, would have little to do with quantum mechanics. A calibration method, however, when restricted to near optimum orbital exponents, can give excellent results for molecular geometry and other properties within the general quantum chemical framework.

Deviations from optimality of basis functions influence the calculated molecular properties, energies, theoretical geometries, and charge distributions. In an earlier study[4] the interrelation between calculated conformational energies and the removal of some of the core atomic orbital (AO) functions from the MO basis, followed by a partial reoptimization of core-deficient basis sets, has been investigated. In Part I of the present study[5]

we have reported the results of a more detailed statistical analysis of the propagation of basis-set error and its effect on variations in the calculated theoretical geometries. A sample of 100 different Gaussian basis functions has been generated by choosing orbital exponents randomly from a Gaussian distribution centered on the exponents of the fully variational (energy-optimized) uniform quality (UQ) AO basis sets.[5] The latter (optimum) reference basis sets have been obtained by minimizing the norm of the energy gradient $\mathbf{g}$ in the logarithmic orbital exponent space:

$$\sigma = \left(\sum_i g_i^2\right)^{1/2} = \left[\sum_i (\partial E/\partial \ln\alpha_i)^2\right]^{1/2}$$
$$= \text{minimum} \quad (1)$$

where $\alpha_i$ are the orbital exponents. The gradient $\mathbf{g}$ is calculated analytically.

The molecule chosen for our study, dimethylsulfoxide (DMSO), has shown an exceptional basis-set dependence of the calculated theoretical geometry, especially that of the S=O bond length. The theoretical S=O bond-length values are usually much larger than the experimental value, if $sp$ bases are used.

We have carried out geometry optimizations with all 100 basis sets. For the optimized geometries, the means, standard deviations, skewness, and kurtosis values, as well as the pair correlations of the distributions for the optimum S=O bond lengths, $E_{tot}$ (total), $E_{nuc}$ (nuclear), $E_{el}$ (electronic), $E_{HOMO}$ (highest-occupied molecular orbital), $E_{LUMO}$ (lowest-unoccupied molecular orbital) values, and sulfur and oxygen formal charges, have been calculated and analyzed. The most important conclusions obtained in Part I are

(i) No reasonable geometry and adequate approximation to the variationally optimum total energy can be obtained simultaneously without the inclusion of $d$ orbitals on sulfur.

(ii) None of the 100 random MO basis sets gives better total energy than the uniform quality reference basis set (−552.93993 a.u.), which finding strongly supports the use of uniform quality Gaussian bases in molecular calculations.

(iii) The "quantum chemical Le Châtelier principle," i.e. the general tendency of mutual compensation between the variations of the electronic energy and nuclear repulsion energy terms, which reduces

variations in the total energy,[6-8] has been confirmed.

(iv) $E_{HOMO}$ and $E_{LUMO}$ energies and their variations are more basis-set dependent than have been expected.

In this study we use the same concepts and conventions as those described in Part I, ref. 5, and references therein. In the following, we briefly review these concepts and definitions. In the abstract logarithmic orbital-exponent space, where the $i$th coordinate axis corresponds to $\ln\alpha_i$, each point represents a basis set which is unique up to trivial permutations of the exponents. The use of the $\ln\alpha_i$ rather than the $\alpha_i$ values solves two problems: (a) $\alpha_i$ values within a typical exponent set differ by several orders of magnitude, and the logarithmic transformation ensures that comparable numerical differences of $\ln\alpha_i$ values have comparable effects on the performance of the basis, and (b) in the $\ln\alpha$ space the constraint that all exponents must be positive is automatically satisfied. For any molecule studied, a variety of energy functionals with various constraints can be defined over this space, e.g., the one for a given fixed nuclear geometry. In our study we consider a more important energy functional, which is defined by the *local optimum geometry for each point* of the $\ln\alpha$ space, i.e. by the geometries optimized independently for each basis set. The variational theorem implies that for any *fixed size* of basis sets, i.e. for any fixed dimension of the $\ln\alpha$ space, the best basis set is the one which is obtained at the minimum point of this energy functional. The functional itself can be thought of as a hypersurface over the $\ln\alpha$ space, and any given basis can be represented by a point on this hypersurface. There are obvious analogies between our problem and the classical mechanical problem of a mass point whose motion is constrained by a potential function. One can define an abstract force in the $\ln\alpha$ space "acting" on any point representing a basis as the energy gradient vector. Just as in the case of the mass point, the basis is "stationary," or is in balance, when the net force acting upon it is zero, i.e. when the gradient vector vanishes. The physical significance of such a balanced basis is clear as implied by the variation principle: The molecular wavefunction obtained with such a balanced basis is consistently and fully variational within the LCAO constraints since just as the linear MO coefficients the orbital exponents are also variationally optimum. Note, however, that the analysis is restricted to the usual linear combination of atomic orbitals (LCAO) approxi-

484                                                                 Mezey and Hass

mation; i.e., the basis functions are attached to the nuclei, and no independent variation of the positioning of basis functions is considered. For a detailed specification of the size and type of the Gaussian basis sets used in our study, see Part I (ref. 5).

In the present article (Part II of the study) we report an analysis of the correlations between the *balance* of Gaussian basis sets, defined in terms of quality measure $\sigma$ and various calculated *molecular properties*. The balance of a particular basis set is characterized by abstract orbital-exponent forces calculated analytically, using a direct analogy with the classical mechanical concept of balance as applied to point masses. It is expected that this analysis will provide information for a reliable evaluation of the importance of balanced basis sets and also some additional information on the propagation of basis-set error within the *ab initio* Hartree–Fock–Roothaan framework.

## RESULTS AND DISCUSSION

The method used in generating the 100 random Gaussian basis sets and the orbital exponents of the reference basis sets (fully variational uniform quality AO bases for S, C, O, and H) have been described in ref. 5. In the present study we use the same model compound, DMSO, for which the calculated molecular properties are exceptionally basis sensitive.[5,7] It is expected that this exceptional sensitivity enhances the general trends and correlations between the balance of basis sets and calculated molecular properties.

The balance of the various Gaussian basis sets used in our analysis is characterized by three parameters, $\sigma_S$, $\sigma_O$, and $\sigma_T$. These parameters are the lengths of various projections of the energy gradient vector in the logarithmic orbital-exponent space. Referring to quantity $\sigma$ [eq. (1)], in $\sigma_S$ the summation is restricted to the sulfur atom orbital exponents, in $\sigma_O$ to those of the oxygen atom, whereas in $\sigma_T$ the summation extends over the entire random basis set. These parameters correspond to generalized "forces" acting upon the exponent set of a given basis represented by a point in an abstract space spanned by the $\ln\alpha_i$ coordinate directions. The total length of this force vector is $\sigma_T$, whereas the respective lengths of the projections of this force vector into subspaces corresponding to the sulfur and oxygen AO bases are denoted by $\sigma_S$ and $\sigma_O$, respectively. If these forces, acting upon a given basis, are zero, i.e.

$$\sigma_T = \sigma_S = \sigma_O = 0 \qquad (2)$$

then the basis set is perfectly balanced. Alternatively, the degree of imbalance of a basis can be characterized by the total length $\sigma_T$ of the force vector and the lengths $\sigma_S$ and $\sigma_O$ of its projections. The latter two quantities represent the balance of sulfur and oxygen AO bases, respectively.

A distance parameter, $d_{\ln}^T$, proportional to the usual Euclidean distance in the logarithmic orbital-exponent space,

$$d_{\ln}^T(A,R) = \left(\frac{1}{n}\sum_i \left[\ln\alpha_i^A - \ln\alpha_i^R\right]^2\right)^{1/2} \qquad (3)$$

has also been determined for all 100 basis sets, where $A$ refers to the actual and $R$ to the reference (uniform quality) bases. Distance parameter $d_{\ln}^T(A,R)$ gives a measure of how different the fully variational AO basis $R$ and basis $A$ are. The analogous distance parameters, $d_{\ln}^S$ and $d_{\ln}^O$ within subspaces corresponding to the sulfur and oxygen AO bases, respectively, are defined similarly.

In Table I the pair-correlation coefficients of calculated molecular properties, S=O bond length, $E_{tot}$, $E_{nuc}$, $E_{el}$, $E_{HOMO}$, $E_{LUMO}$ values, and sulfur, oxygen, carbon formal atomic charges, with distance parameters $d_{\ln}^S$, $d_{\ln}^O$, $d_{\ln}^T$, balance parameters $\sigma_S$, $\sigma_O$, $\sigma_T$, and gradient vector components $g_i$ corresponding to both the sulfur and oxygen orbital exponents, are listed. The correlation coefficients between the calculated physical properties have been discussed in Part I and are quoted in Table I only for sake of comparisons.

The distance parameters show only weak correlations with the calculated molecular properties. Apparently, the Euclidean distance in the $\ln\alpha_i$ space is a poor measure of basis-set quality. The largest correlation coefficients are found for the total energy $E_{tot}$, but even these values (0.19 with $d_{\ln}^S$ and 0.13 with $d_{\ln}^T$) are rather small. These numbers are positive, as deviations from the optimum reference basis tend to increase the total energy, $E_{tot}$. As expected, the calculated formal charges also show some, although weak, correlations with the distance parameters (the sulfur charge-$d_{\ln}^S$ correlation coefficient is 0.15, whereas for the oxygen charge-$d_{\ln}^O$ correlation the value is 0.10). The positivity of these numbers indicate that the deviation from optimality in the subspace of a given atomic basis tends to *increase* the calculated charge on the atom.

The balance parameters $\sigma_S$, $\sigma_O$, and $\sigma_T$ show much stronger correlations with the calculated molecular properties. The $\sigma_T - E_{tot}$ correlation is

**Table I.**  Pair-correlation coefficients between distributions of various calculated molecular properties, distance, and balance parameters, as well as abstract force components of random Gaussian basis sets, defined in the logarithmic orbital exponent space.

| | S=0 bond length | $E_{tot}$ | $E_{nuc}$ | $E_{el}$ | $E_{HOMO}$ | $E_{LUMO}$ | Formal charge on | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | | | | | | S | O | C |
| S=0 bond length | 1.000 | -0.107 | -0.993 | 0.820 | -0.610 | -0.530 | -0.646 | 0.315 | 0.641 |
| $E_{tot}$ | -0.107 | 1.000 | 0.119 | 0.471 | 0.100 | 0.042 | 0.142 | -0.097 | -0.152 |
| $E_{nuc}$ | -0.993 | 0.119 | 1.000 | -0.818 | 0.617 | 0.517 | 0.657 | -0.295 | -0.686 |
| $E_{el}$ | 0.820 | 0.471 | -0.818 | 1.000 | -0.490 | -0.434 | -0.501 | 0.205 | 0.521 |
| $E_{HOMO}$ | -0.610 | 0.100 | 0.617 | -0.490 | 1.000 | 0.864 | -0.041 | 0.438 | -0.117 |
| $E_{LUMO}$ | -0.530 | 0.042 | 0.517 | -0.434 | 0.864 | 1.000 | -0.080 | 0.344 | 0.048 |
| Charge on   S | -0.646 | 0.142 | 0.657 | -0.501 | -0.041 | -0.080 | 1.000 | -0.833 | -0.884 |
| O | 0.315 | -0.097 | -0.295 | 0.205 | 0.438 | 0.344 | -0.833 | 1.000 | 0.543 |
| C | 0.641 | -0.152 | -0.686 | 0.521 | -0.117 | 0.048 | -0.884 | 0.543 | 1.000 |
| $d^S_{ln}$ | 0.000 | 0.188 | -0.000 | 0.109 | -0.058 | -0.046 | 0.147 | -0.112 | -0.049 |
| $d^O_{ln}$ | -0.058 | 0.022 | 0.056 | -0.037 | 0.069 | -0.058 | -0.020 | 0.098 | 0.022 |
| $d^T_{ln}$ | -0.031 | 0.129 | 0.029 | 0.048 | -0.005 | -0.055 | 0.098 | -0.019 | -0.024 |
| $\sigma_S$ | -0.052 | 0.899 | 0.065 | 0.461 | 0.050 | 0.024 | 0.097 | -0.061 | -0.099 |
| $\sigma_O$ | -0.163 | 0.773 | 0.171 | 0.295 | 0.062 | -0.037 | 0.167 | -0.172 | -0.181 |
| $\sigma_T$ | -0.070 | 0.927 | 0.082 | 0.462 | 0.050 | 0.013 | 0.112 | -0.082 | -0.117 |
| Force components on S $s_1$ | 0.238 | 0.060 | -0.226 | 0.235 | -0.217 | -0.290 | -0.062 | 0.007 | -0.002 |
| $s_2$ | -0.184 | 0.245 | 0.177 | -0.015 | 0.277 | 0.268 | 0.006 | 0.068 | 0.012 |
| $s_3$ | 0.160 | 0.144 | -0.146 | 0.213 | -0.228 | -0.196 | 0.057 | -0.106 | -0.057 |
| $s_4$ | -0.136 | 0.151 | 0.119 | -0.019 | 0.137 | 0.150 | -0.026 | 0.030 | 0.048 |
| $s_5$ | -0.005 | -0.129 | 0.019 | -0.091 | -0.033 | -0.026 | 0.048 | -0.002 | -0.059 |
| $s_6$ | 0.116 | -0.457 | -0.130 | -0.148 | -0.083 | -0.129 | -0.049 | -0.007 | 0.035 |
| $s_7$ | -0.083 | -0.366 | 0.084 | -0.287 | 0.076 | 0.078 | -0.073 | 0.089 | 0.086 |
| $s_8$ | 0.108 | -0.261 | -0.093 | -0.067 | -0.056 | -0.063 | -0.103 | 0.087 | 0.080 |
| $s_9$ | 0.037 | -0.341 | -0.059 | -0.144 | -0.092 | -0.086 | 0.045 | -0.120 | 0.014 |
| $p_1$ | -0.084 | 0.239 | 0.086 | 0.060 | 0.167 | 0.088 | 0.116 | -0.006 | -0.137 |
| $p_2$ | 0.090 | 0.085 | -0.087 | 0.126 | -0.139 | -0.104 | -0.068 | -0.036 | 0.075 |
| $p_3$ | -0.116 | 0.258 | 0.108 | 0.052 | 0.129 | 0.126 | 0.100 | -0.027 | -0.084 |
| $p_4$ | 0.070 | -0.465 | -0.070 | -0.206 | -0.129 | -0.100 | -0.094 | 0.027 | 0.098 |
| $p_5$ | -0.121 | -0.197 | 0.162 | -0.258 | -0.170 | -0.309 | 0.406 | -0.238 | -0.546 |
| $p_6$ | 0.099 | -0.052 | -0.156 | 0.108 | 0.309 | 0.509 | -0.431 | 0.261 | 0.657 |
| Force components on O $s_1$ | 0.102 | 0.115 | -0.088 | 0.145 | 0.058 | 0.013 | -0.160 | 0.211 | 0.043 |
| $s_2$ | -0.073 | 0.140 | 0.058 | 0.029 | -0.020 | 0.026 | 0.092 | -0.171 | 0.014 |
| $s_3$ | -0.001 | 0.087 | 0.013 | 0.038 | -0.016 | -0.078 | 0.002 | 0.032 | -0.078 |
| $s_4$ | -0.017 | -0.265 | 0.010 | -0.163 | 0.009 | 0.083 | 0.053 | -0.060 | 0.029 |
| $s_5$ | 0.088 | -0.268 | -0.097 | -0.069 | -0.031 | 0.068 | -0.121 | 0.058 | 0.165 |
| $s_6$ | 0.106 | -0.149 | -0.099 | 0.002 | 0.042 | -0.070 | -0.165 | 0.290 | -0.014 |
| $p_1$ | -0.172 | 0.116 | 0.185 | -0.097 | 0.587 | 0.487 | -0.203 | 0.413 | 0.065 |
| $p_2$ | 0.324 | -0.094 | -0.343 | 0.250 | -0.698 | -0.664 | 0.162 | -0.464 | 0.013 |
| $p_3$ | -0.466 | 0.075 | 0.498 | -0.399 | 0.780 | 0.719 | -0.109 | 0.512 | -0.122 |

the largest, 0.93 as expected, since the minimum of $\sigma_T$ is a good approximation to that of $\sigma_T$ (molecular, no LCAO constraints) that in turn corresponds to the minimum of $E_{tot}$. The balance parameters $\sigma_S$ and $\sigma_O$ for the sulfur and oxygen bases both show large correlations with $E_{tot}$ (0.90 and 0.77, respectively). Since it is the electronic energy component $E_{el}$ which is affected directly by basis-set variations, all three balance parameters show much stronger correlations with $E_{el}$ than with $E_{nuc}$. The correlation coefficients with $E_{nuc}$ are close to zero, with the exception of the $E_{nuc}$ −

$\sigma_O$ correlation (0.17). This result is in accord with the atomic charge–balance parameters correlations, where also the correlations involving $\sigma_O$ are the largest in absolute value.

According to the quantum chemical Le Châtelier Principle,[5-8] (i.e. the general trend of partial compensation between variations in $E_{el}$ and $E_{nuc}$, approximately preserving the $E_{tot}$ value), one could expect that $E_{el}$ and $E_{nuc}$ have correlation coefficients of opposite signs with the balance parameters. Although all these correlation coefficients are positive, this result is only an apparent

violation of the above principle. The balance parameters are proportional to the *norms* of formal force vectors in various subspaces of the $\ln \alpha_i$ ($i = 1, \ldots, n$) space. When the directional properties of these formal force vectors are also taken into account (see below), then the calculated trends clearly show the expected compensation between $E_{el}$ and $E_{nuc}$; i.e., the quantum chemical Le Châtelier principle applies.

A greater imbalance of the oxygen basis, i.e. a larger $\sigma_O$ value, correlates with a larger negative charge on the carbon and oxygen atoms and also with a larger positive charge on the sulfur. Poor balance of atomic bases tends to increase the calculated polarization of the molecule.

The correlation coefficients between the balance parameters and the S=O bond length, optimized for each basis, are small negative numbers, indicating that the imbalance of a basis set tends to decrease the optimum bond-length value. Again, the correlation with $\sigma_O$ is the most significant, although even this correlation is rather weak (−0.16).

More detailed information can be obtained when the directional properties of formal force vectors in the logarithmic orbital-exponent space are also taken into account; i.e., when the vector-component–molecular-property correlation coefficients are considered.

The largest correlation coefficients are those involving the outer $p$-type Gaussian functions of sulfur and oxygen. The innermost $s$ functions contributing predominantly to the cores, and functions having approximately equal contributions to two different shells, also appear to have a significant influence on some of the molecular properties.

The calculated optimum S=O bond length shows the largest correlations with the gradient components of the oxygen $p$ exponents. The more stabilizing is the increase of the outermost $p$ exponent; the more likely is a longer optimum S=O bond length. The larger exponents correspond to tighter Gaussian functions, which in general tend to overemphasize electrostatic effects at the expense of delocalized valence interactions. Consequently, the calculated long S=O bonds are related to an overemphasis of electrostatic effects. Moving inward from the outermost $p$ function, the correlation coefficients show a sign alternation; however, the absolute values decrease.

Similar sign alternation is found for the correlation coefficients between the inner four sulfur $s$-exponent force components and the S=O bond

length. Although these latter correlations are not very strong (the largest coefficient is 0.24), nevertheless they are significant. The balance of the sulfur core representation affects the value of the optimum S=O bond length, which result agrees well with earlier conclusions on the balance of core-deficient basis sets.[4]

The calculated total energies show the strongest correlations with those force components, which belong to functions having significant contributions to two shells. Most of these correlation coefficients are negative. A smaller total energy value is likely to be obtained with a basis set where these force components are less negative. This finding indicates that the proper balance of those basis functions which may easily switch over from one shell to another (or from one MO to another) is of particular importance.

Whereas the force components corresponding to the outermost $p$ functins and innermost $s$ functions do not have a large effect on the total energy, these components show dominant correlations with energy components $E_{nuc}$ and $E_{el}$. Although it is the electronic energy $E_{el}$ which is affected the most directly by orbital exponent variations, and the nuclear repulsion component $E_{nuc}$ may readjust only through geometry optimization, nevertheless for the oxygen $p$ orbitals the force components show stronger correlations with $E_{nuc}$ than with $E_{el}$. The dominant correlation is the one between the force component of the outermost oxygen $p$ exponent and the nuclear repulsion energy $E_{nuc}$. The positive sign of the correlation coefficient (0.50) is in agreement with our earlier conclusion on the role of electrostatic effects influencing the optimum S=O bond length. The more stabilizing (or less destabilizing) is the increase of the outermost $p$ exponent, the more likely is a smaller $E_{nuc}$ value; that is, the more likely is a more extended molecule, e.g., a longer S=O bond.

For all force components, if there is any significant correlation at all, the correlation coefficients with $E_{el}$ and $E_{nuc}$ are of opposite signs. Altering the balance of a basis set by changing a force component results in opposing changes in $E_{nuc}$ and $E_{el}$, as predicted by the quantum mechanical Le Châtelier principle. The pair of (small) negative correlations with the force component of the sulfur $s_6$ exponent is exceptional and is likely to reflect the strong involvement of this function in more than one MO.

The HOMO and LUMO orbital energies show the strongest correlations, with the force components

corresponding to the three oxygen $p$ exponents, which is the expected result. It is noteworthy, however, that these correlation coefficients are larger in absolute values than those involving the actual values of orbital exponents (Table III of Part I). Apparently, the calculated $E_{HOMO}$ and $E_{LUMO}$ values are even more sensitive to the force components, i.e. to the balance of the basis, than to the actual choice of orbital exponents. This finding underlines our earlier conclusions[5] on the strong basis-set dependence of results of HOMO–LUMO analyses, often used for a qualitative interpretation of molecular processes.

These correlation coefficients show an alternation of sign, similar to that of correlation coefficients involving the S=O bond length. Apparently, there exists some degree of mutual compensation in the sequence of Gaussian basis functions. Surprisingly, the components of the innermost sulfur $s$ exponents also show small but significant correlations with the $E_{HOMO}$ and $E_{LUMO}$ values. Since these Gaussian functions can have no direct contribution to either the HOMO or the LUMO, their effect must be related to the influence of the lowest-lying eigenvalues on the overall eigenvalue spectrum. The MO with the lowest energy is essentially a sulfur $1s$ core orbital, involving primarily the first few sulfur $s$-basis functions. The influence of this core MO on the entire MO energy spectrum is the explanation of these unexpected correlations, as well as for those correlations involving $E_{el}$, $E_{nuc}$, and the S=O bond length, on the one hand, and the same force components, on the other.

The calculated formal atomic charges show the strongest correlations with the force components of the few outermost $p$-orbital exponents. The calculated charge on oxygen is influenced most by the oxygen $p$-exponent components. Unexpectedly, the outermost sulfur $p$-exponent components of the force vector have the strongest correlations *not* with the charge on sulfur but with the formal charges on the carbon atoms. The corresponding AOs on sulfur, however, are the loosest functions in the entire basis set and have significant contribution to the molecular wavefunction around the carbon atoms. The correlation coefficients show a sign alteration in the sequence of force compo-

nents, similar to those noted for other molecular properties.

## SUMMARY

The balance parameters $\sigma_S$, $\sigma_O$, and $\sigma_T$ are better general measures of basis-set quality than distance parameters $d_{ln}^S$, $d_{ln}^O$, and $d_{ln}^T$, the latter representing deviations between basis sets in an abstract logarithmic orbital-exponent space. Among the individual force components in the above space, those corresponding to the outermost $p$ orbitals, the innermost core orbitals, and functions contributing approximately equally to more than one MO, have the strongest influence on calculated molecular properties. The correlation coefficients show that these components have the opposite effects on energy components $E_{nuc}$ and $E_{el}$, in agreement with the quantum chemical Le Châtelier principle. HOMO and LUMO energies are even more sensitive to the *balance* of the basis than to the actual choice of exponents. The unexpected influence of the balance of sulfur core exponents on various molecular properties is related to the role of the lowest MO eigenvalue in affecting the overall eigenvalue spectrum. The balance of the outermost sulfur $p$ exponents have the strongest influence on the formal charge on carbon atoms and *not* on the calculated formal charge on sulfur.

## REFERENCES

1. Y. Sakai, H. Tatewaki, and S. Huzinaga, *J. Comput. Chem.*, **3**, 6 (1982), and references therein.
2. M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, and W. J. Hehre, *J. Am. Chem. Soc.*, **104**, 2797 (1982), and references therein.
3. F. B. van Duijneveldt, *J. Res. Rev. IBM*, **XX**, 945 (1971).
4. P. G. Mezey and C. V. S. R. Rao, *J. Comput. Chem.*, **1**, 134 (1980), and references therein.
5. P. G. Mezey and E. C. Hass, *J. Chem. Phys.*, **77**, 870 (1982).
6. P. G. Mezey, *Chem. Phys. Lett.*, **47**, 70 (1977).
7. P. G. Mezey and A. Kapur, *Can. J. Chem.*, **58**, 599 (1980).
8. P. G. Mezey, *Int. J. Quant. Chem.*, in press.