# Energy partitioning schemes: a dilemma

## I. Mayer

Two closely related energy partitioning schemes, in which the total energy is presented as a sum of atomic and diatomic contributions by using the "atomic decomposition of identity", are compared on the example of *N,N*-dimethylformamide, a simple but chemically rich molecule. Both schemes account for different intramolecular interactions, for instance they identify the weak C–H···O intramolecular interactions, but give completely different numbers. (The energy decomposition scheme based on the virial theorem is also considered.) The comparison of the two schemes resulted in a dilemma which is especially striking when these schemes are applied for molecules distorted from their equilibrium structures: one either gets numbers which are "on the chemical scale" and have quite appealing values *at the equilibrium molecular geometries*, but exhibiting a counter-intuitive distance dependence (the two-center energy components increase in absolute value with the increase of the interatomic distances)—or numbers with too large absolute values but "correct" distance behaviour. The problem is connected with the quick decay of the diatomic kinetic energy components.

## I.   Introduction and theoretical background

For chemists molecules consist of atoms held together by bonds; quantum chemists treat them as systems of electrons and nuclei. It is not trivial to bridge these images—both of which are legitimate in their own context—because there is no strict quantum mechanical definition of an atom *within* a molecule. At the same time, it is desirable to introduce such a definition—even if it is not unique. Thus, one can decompose different physical quantities, among them the total energy, into the sum of contributions corresponding to the individual atoms and pairs of atoms (possibly, also larger groups). This may help to explain and exploit the widespread effects of near additivity and transferability of different effects characteristic of chemistry. As stressed by Hall,[1] such an analysis can be performed either in the 3-dimensional (3D) physical space, assigning to an atom, a part of the space near the nucleus, or in the Hilbert space of the basis orbitals assigned to the individual atoms.

Of course, the total energy of a molecule is a single number which can be decomposed into components in an infinite number of different ways. There are, however, occasions when there is some "natural" decomposition. In the 3D analysis, this may be due to the disjunct character of the atomic domains, as is the case[2,3] *e.g.*, in Bader's AIM theory.[4] Similarly, in the semiempirical theories one applies only one- and two-center integrals, and therefore the energy breaks down spontaneously to one-center and two-center energy components. However, in *ab initio* theories

*Chemical Research Center, Hungarian Academy of Sciences, P.O. Box 17, H-1525 Budapest, Hungary*

there are three- and four-center integrals, too, and the "natural" energy analysis of Clementi[5] based on such a systematization of the integrals lead to non-chemically large three- and four-center energy components. We shall not enter here into a discussion of other energy component analyses (*e.g.*, ref. 6–25), we wish only to concentrate on the recent method[26] of "atomic decomposition of identity" permitting most different decomposition tasks to be treated by using a common general formalism.

In this scheme one writes down the identity operator as

$$\hat{I} = \sum_A \hat{\rho}_A \qquad (1)$$

where operators $\hat{\rho}_A$ are *in some sense* assigned to the individual atoms A. The individual operators $\hat{\rho}_A$ need not be Hermitian, as far as they sum up to the identity operator. The actual form of the operators $\hat{\rho}_A$ varies for different types of analysis. In the 3D analysis it is expressed *via* a weight-function,[26] in the Hilbert-space analysis it may be defined as

$$\hat{\rho}_A = \sum_{\mu \in A} |\chi_\mu\rangle\langle\tilde{\chi}_\mu| \qquad (2)$$

where $\tilde{\chi}_\mu$ is the *biorthogonal* counterpart of the basis function $\chi_\mu$, having the property $\langle\tilde{\chi}_\mu|\chi_\nu\rangle = \delta_{\mu\nu}$. Then acting with operator $\hat{\rho}_A$ on an molecular orbital $\varphi_i = \sum_\rho c_\rho^i \chi_\rho$ one gets its part truncated to the basis orbitals centered on the atom considered: $\hat{\rho}_A \varphi_i = \sum_{\mu \in A} c_\mu^i \chi_\mu$.

Now, we consider the total RHF energy† of the molecule by using the [12|12] convention for the two-electron integrals

$$E = 2\sum_i^{\text{occ.}} \langle\varphi_i|\hat{h}|\varphi_i\rangle + \sum_{i,j}^{\text{occ.}} (2[\varphi_i\varphi_j|\varphi_i\varphi_j] - [\varphi_i\varphi_j]\varphi_j\varphi_i) \qquad (3)$$

and introduce one "atomic resolution of identity" to each "ket"—one for each electron in the case of the two-electron integrals. Using eqn (2), taking into account that the one-electron Hamiltonian contains interactions with all the nuclei and recollecting the terms we get trivially the energy decomposition[26]

$$E = \sum_A E_A + \sum_{A<B} E_{AB} \qquad (4)$$

where the individual energy components are

$$E_A^1 = \sum_{\mu \in A}\sum_\nu D_{\mu\nu}h_{\nu\mu}^A + \frac{1}{2}\sum_{\mu,\varrho \in A}\sum_{\nu,\tau}\left(D_{\mu\nu}D_{\varrho\tau} - \frac{1}{2}D_{\varrho\nu}D_{\mu\tau}\right)[\chi_\nu\chi_\tau|\chi_\mu\chi_\varrho] \qquad (5)$$

$$E_{AB}^1 = \frac{Z_A Z_B}{R_{AB}} - \sum_{\mu \in A}\sum_\nu D_{\mu\nu}\left\langle\chi_\nu\left|\frac{Z_B}{r_B}\right|\chi_\mu\right\rangle - \sum_{\mu \in B}\sum_\nu D_{\mu\nu}\left\langle\chi_\nu\left|\frac{Z_A}{r_A}\right|\chi_\mu\right\rangle$$
$$+ \sum_{\mu \in A}\sum_{\varrho \in A}\sum_{\nu,\tau}\left(D_{\mu\nu}D_{\varrho\tau} - \frac{1}{2}D_{\varrho\nu}D_{\mu\tau}\right)[\chi_\nu\chi_\tau|\chi_\mu\chi_\rho] \qquad (6)$$

Here $h^A = \hat{T} - Z_A/r_A$ is the atomic one-electron Hamiltonian, $\boldsymbol{D}$ is the usual "density matrix" and we have introduced the superscripts "1" to the energy components owing to the fact that a second energy decomposition will also be

---

† Generalization to the UHF case is trivial, while the use of electron correlated methods (or DFT theory) in the analysis would require modifications which are too large to be considered here.

considered below. This energy decomposition scheme has not been considered previously.‡

From the physical point of view, the above formulae represent a compression of the three- and four-center interactions into one- and two-center ones in a manner which is closely related to the basic idea of Mulliken's population analysis. When one has an interaction of a diatomic overlap population coming from basis orbitals centered on atoms A and B, respectively, with a nucleus of atom C, then it is equally divided into two terms, one assigned to the pair of atoms AC and another to the pair of atoms BC. Similarly, the interaction of two overlap densities coming from A and B from one side and C and D from the other, should be equally divided between four two-center ones, AC, BC, AD and BD.[27] Owing to the occurrence of one- and two-electron integrals which have equal values and differ only in the order of indices, these assignments are automatically performed by eqn (5) and (6).

As we have used one "atomic resolution of identity" for every one-electron integral and two for every two-electron one, this scheme may be considered very straightforward from a formal mathematical point of view. At the same time, it has the peculiarity that the kinetic energy operator $\hat{T}$ enters only the one-center (atomic) energy components. One may consider this to be in contradiction with the complex role the kinetic energy plays in chemical bond formation.[28] Therefore, it may be reasonable to introduce another resolution of identity into the expectation value of the kinetic energy—to the "bra"-s of the integrals, applying there $\hat{I}$ presented as a sum of the adjoints of operators $\hat{\rho}_A$. We get the modified energy components (superscripts "2") as

$$
\begin{aligned}
E_A^2 = \sum_{\mu,\nu \in A} D_{\mu\nu} T_{\nu\mu} - \sum_{\mu \in A} \sum_{\nu} D_{\mu\nu} \left\langle \chi_\nu \left| \frac{Z_A}{r_A} \right| \chi_\mu \right\rangle \\
+ \frac{1}{2} \sum_{\mu,\varrho \in A} \sum_{\nu,\tau} \left( D_{\mu\nu} D_{\varrho\tau} - \frac{1}{2} D_{\varrho\nu} D_{\mu\tau} \right) [\chi_\nu \chi_\tau | \chi_\mu \chi_\varrho]
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
E_{AB}^2 = 2 \sum_{\mu \in A} \sum_{\nu \in B} D_{\mu\nu} T_{\nu\mu} - \sum_{\mu \in A} \sum_{\nu} D_{\mu\nu} \left\langle \chi_\nu \left| \frac{Z_B}{r_B} \right| \chi_\mu \right\rangle - \sum_{\mu \in B} \sum_{\nu} D_{\mu\nu} \left\langle \chi_\nu \left| \frac{Z_A}{r_A} \right| \chi_\mu \right\rangle \\
+ \sum_{\mu \in A} \sum_{\varrho \in B} \sum_{\nu,\tau} \left( D_{\mu\nu} D_{\varrho\tau} - \frac{1}{2} D_{\varrho\nu} D_{\mu\tau} \right) [\chi_\nu \chi_\tau | \chi_\mu \chi_\varrho] + \frac{Z_A Z_B}{R_{AB}}
\end{aligned} \tag{8}
$$

where $T_{\mu\nu} = \langle \chi_\nu | -\frac{1}{2} \Delta | \chi_\mu \rangle$ is a matrix element of the kinetic energy.

These expressions coincide with those, introduced on intuitive grounds in ref. 27, if written down for the RHF case.§ The test-calculations on a few simple molecules gave chemically appealing numbers and showed moderate basis sensitivity.[27]

In the framework of the Born–Oppenheimer separation the virial theorem holds only for the *equilibrium geometries* (and other stationary points of the potential energy surfaces) and sufficiently good wavefunctions. In that case we have for the expectation values of the potential and kinetic energies the ratio $\langle \hat{U} \rangle / \langle \hat{T} \rangle = -2$, or the relationship $E = -\langle \hat{T} \rangle$ for the energy $E$. As the kinetic energy has only one- and two-center integrals, the virial theorem provides us with an extremely simple energy decomposition scheme.[13,29]

---

‡ The assumption in ref. 26 according to which APEX4 program realizes this scheme is in error: it corresponds to another, somewhat different partitioning scheme (with some minor terms regrouped between one- and two-center energy components) and, likewise, the CECA method ref. 22 represents an approximation not to this scheme but to the latter.

§ Papers 17 and 21 propose energy decomposition formulae which are similar but not identical to either of our schemes or to each other. However, both contain exchange energy components which are decomposed in a manner not symmetric with respect to the two electrons, therefore seem to me unacceptable.

$$E \cong -\langle \hat{T} \rangle = -\sum_A \sum_{\mu,\nu \in A} D_{\mu\nu} T_{\nu\mu} - 2 \sum_{A<B} \sum_{\mu \in A} \sum_{\nu \in B} D_{\mu\nu} T_{\nu\mu} \qquad (9)$$

which, however, has not been widely used in practice.

It appears to me that the above two energy decomposition schemes "1" and "2" (and, possibly, that which is based on the virial theorem) are the most straightforward ones as far as one can see for one- and two-center components of the Hartree–Fock energy in the framework of a Hilbert-space analysis. The aim of the present paper is to compare these schemes on the example of a simple, but chemically non-trivial molecule, and to discuss a dilemma which one encounters in connection with the distance dependence of the energy components provided by these two schemes.

## II.    Dimethylformamide—a case study

$N,N$-dimethylformamide $(CH_3)_2N$–CHO is a very simple molecule but nevertheless exhibits a rich chemistry. According to gas-phase electron diffraction data,[30] it is (at least nearly) planar, indicating a "resonance" between the carbonyl group and the nitrogen lone pair—essentially the formation of a small π-electron system to which the methyl groups are also connected *via* hyperconjugation. From the other side, the C–N–C angle which is *syn* to the carbonyl group was found somewhat smaller than the other one, indicating that there must be some attractive C–H···O interaction, despite the relatively large $(2.40 \pm 0.03$ Å) H···O distance. Thus, this molecule is one of the simplest in which the existence of a C–H···O interaction has been postulated on the basis of experimental results.[30]

In the SCF calculations the molecule has been found completely planar for all basis sets considered (from 6-31G to cc-pVTZ), except for the minimal (STO-3G, 6G) basis level. (No piramidality has been observed at the MP2 or DFT levels either.)

Fig. 1 shows the numbering of atoms applied. Table 1 summarizes some of the geometrical data which seem to be of interest, calculated for several basis sets, Table 2 shows selected bond orders.[31] The results are in good agreement with the notion of a delocalized π-electron system mentioned above: *e.g.*, the N–C(O) distance is significantly shorter than the N–C(Me) distances, as the respective bond order significantly exceeds one. [The N–C(Me) bonds are essentially single ones.]

The C–N–C angle in the *syn* position is consistently smaller than that in the *anti* position, in perfect agreement with the experiment. The agreement with the experiment is also remarkable concerning the small but consistent difference between the N–C(Me) bond lengths indicating that the bond in the *syn* position is slightly
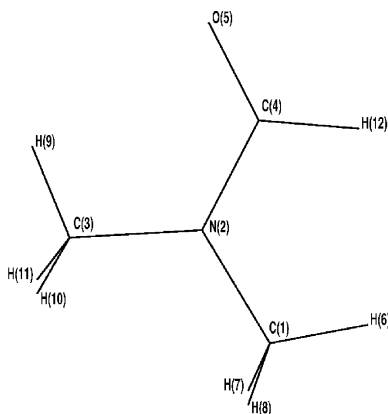


**Fig. 1**   Numbering of the atoms in the *N,N*-dimethylformamide molecule.

**Table 1** Selected geometrical data of the *N,N*-dimethylformamide molecule (Å and °)

| Basis | R(H$_9\cdots$O) | ∠C$_3$H$_9$O | ∠C$_3$NC$_4$ | ∠C$_1$NC$_4$ | R(NC$_3$) | R(NC$_1$) |
|---|---|---|---|---|---|---|
| 6-31G | 2.365 | 103.2 | 120.7 | 121.9 | 1.454 | 1.450 |
| 6-31G** | 2.353 | 102.8 | 120.5 | 121.9 | 1.446 | 1.442 |
| 6-311G** | 2.360 | 102.6 | 120.7 | 121.8 | 1.447 | 1.443 |
| 6-311G (2d,2p) | 2.355 | 102.6 | 120.7 | 121.7 | 1.444 | 1.439 |
| 6-311 + + G(2d,2p) | 2.366 | 102.3 | 120.9 | 121.7 | 1.445 | 1.441 |
| cc-TVZ | 2.366 | 102.4 | 120.9 | 121.7 | 1.444 | 1.439 |
| Exp.[a] | 2.40 ± 0.03 | | 120.8 ± 0.3 | 122.3 ± 0.4 | 1.454 ± 0.002 | 1.449 ± 0.002 |

| Basis | R(NC4) | R(CO) | R(C$_1$H$_6$) | R(C$_1$H$_7$) | R(C$_3$H$_9$) | R(C$_3$H$_{10}$) |
|---|---|---|---|---|---|---|
| 6-31G | 1.349 | 1.225 | 1.081 | 1.084 | 1.077 | 1.084 |
| 6-31G** | 1.349 | 1.197 | 1.082 | 1.086 | 1.078 | 1.087 |
| 6-311G** | 1.348 | 1.191 | 1.083 | 1.087 | 1.078 | 1.087 |
| 6-311G(2d,2p) | 1.344 | 1.191 | 1.080 | 1.084 | 1.075 | 1.084 |
| 6-311 + + G(2d,2p) | 1.343 | 1.193 | 1.080 | 1.084 | 1.075 | 1.084 |
| cc-TVZ | 1.344 | 1.192 | 1.080 | 1.085 | 1.076 | 1.085 |
| Exp.[a] | 1.391 ± 0.007 | 1.224 ± 0.003 | Average R(CH): 1.112 ± 0.003 | | | |

[a] Ref. 30.

stretched, probably to provide better conditions for the C–H$\cdots$O interaction. The closest H$\cdots$O distance has been found to be 2.35–2.36 Å, and is practically independent of the basis set applied. This distance is slightly, but not significantly, shorter than the experimental one. Thus, we may claim that these results confirm the existence of some attractive C–H$\cdots$O interaction, as far as that can be deduced from the geometrical data. The C–H bond lengths exhibit some characteristic differences which we shall consider later. (These geometrical features are present also at the DFT and MP2 levels; the H$\cdots$O distance even reduces a little.)

Tables 3 and 4 display some two-center energy components obtained according to schemes "1" and "2" *i.e.*, by using eqn (6) and eqn (8), respectively. Table 5 contains the analogous results obtained by the virial decomposition, eqn. (9).

**Table 2** Selected bond orders of the *N,N*-dimethylformamide molecule

| Basis | C$_4$–O | N–C$_4$ | N$\cdots$O | N–C$_3$ | N–C$_1$ | O–H$_9$ |
|---|---|---|---|---|---|---|
| 6-31G | 1.723 | 0.977 | 0.044 | 0.769 | 0.793 | 0.018 |
| 6-31G** | 1.798 | 1.091 | 0.048 | 0.868 | 0.894 | 0.020 |
| 6-311G** | 1.889 | 1.166 | 0.029 | 0.900 | 0.931 | 0.033 |
| 6-311G(2d,2p) | 1.898 | 1.213 | 0.031 | 0.968 | 1.011 | 0.039 |
| 6-311 + + G(2d,2p) | 1.636 | 1.183 | 0.141 | 0.909 | 0.943 | 0.035 |
| cc-TVZ | 2.020 | 1.216 | −0.003 | 0.937 | 0.983 | 0.034 |

| Basis | O$\cdots$H$_6$ | O$\cdots$H$_{12}$ | C$_1$–H$_6$ | C$_1$–H$_7$ | C$_3$–H$_9$ | C$_3$–H$_{10}$ |
|---|---|---|---|---|---|---|
| 6-31G | −0.0004 | 0.024 | 0.960 | 0.949 | 0.938 | 0.955 |
| 6-31G** | 0.0006 | 0.018 | 0.975 | 0.966 | 0.959 | 0.972 |
| 6-311G** | 0.0013 | 0.020 | 0.985 | 0.980 | 0.974 | 0.986 |
| 6-311G(2d,2p) | 0.0028 | 0.023 | 0.991 | 0.984 | 0.982 | 0.990 |
| 6-311 + + G(2d,2p) | 0.0026 | −0.004 | 1.014 | 1.002 | 1.009 | 1.005 |
| cc-TVZ | 0.0017 | 0.014 | 0.982 | 0.968 | 0.968 | 0.977 |

**Table 3** Selected diatomic energy components of the *N*,*N*-dimethylformamide molecule calculated by using scheme "1" [eqn (6)] (atomic units)

| Basis | $C_4$–O | N–$C_4$ | N$\cdots$O | N–$C_3$ | N–$C_1$ | O–$H_9$ |
|---|---|---|---|---|---|---|
| 6-31G | −1.148 | −0.842 | 0.150 | −0.510 | −0.525 | −0.039 |
| 6-31G** | −1.331 | −0.902 | 0.126 | −0.590 | −0.609 | −0.033 |
| 6-311G** | −1.350 | −0.890 | 0.101 | −0.613 | −0.639 | −0.031 |
| 6-311G(2d,2p) | −1.423 | −0.899 | 0.094 | −0.672 | −0.709 | −0.027 |
| 6-311++G(2d,2p) | −1.176 | −0.871 | −0.025 | −0.644 | −0.662 | −0.031 |

| Basis | O$\cdots$$H_6$ | O$\cdots$$H_{12}$ | $C_1$–$H_6$ | $C_1$–$H_7$ | $C_3$–$H_9$ | $C_3$–$H_{10}$ |
|---|---|---|---|---|---|---|
| 6-31G | −0.013 | 0.007 | −0.546 | −0.543 | −0.556 | −0.544 |
| 6-31G** | −0.010 | 0.015 | −0.565 | −0.562 | −0.573 | −0.562 |
| 6-311G** | −0.005 | 0.018 | −0.573 | −0.573 | −0.581 | −0.574 |
| 6-311G(2d,2p) | −0.002 | 0.022 | −0.579 | −0.578 | −0.581 | −0.579 |
| 6-311++G(2d,2p) | −0.003 | 0.031 | −0.590 | −0.584 | −0.597 | −0.583 |

Of course, when doing Hilbert-space analyses, one has to keep in mind that the results do not have, in general, any definite basis set limit and, therefore, only results obtained by using strictly the same basis set should be compared. Also, it is obvious that only basis sets of sufficiently "atomic character" must be considered. Thus, diffuse basis functions lacking true atomic nature may spoil the analysis. (In our case there is only one such basis set—the 6-31++G(2d,2p) one—and it indeed shows a deviation from the smooth change of results observed for the other basis sets. Also, in some instances—in particular for the nonbonding N$\cdots$O interaction—it seems to exhibit a non-physical behaviour, so we shall not discuss in detail the results obtained by this basis set in the following.)

By inspecting the numbers, one may see that schemes "1" and "2" both account for the different chemical effects in the molecule, of both the bonding and non-bonding type. Besides the covalent bonds, they permit identifying an attractive (negative) energy component reflecting the secondary C–H$\cdots$O interaction described above on the basis of geometry data. (We shall return later to its discussion.) The repulsive character of the *geminal* O$\cdots$H interaction due to overlap can also be established. The significant repulsive O$\cdots$N energy contribution is in agreement with

**Table 4** Selected diatomic energy components of the *N*,*N*-dimethylformamide molecule calculated by using scheme "2" [eqn. (8)] (atomic units)

| Basis | $C_4$–O | N–$C_4$ | N$\cdots$O | N–$C_3$ | N–$C_1$ | O–$H_9$ |
|---|---|---|---|---|---|---|
| 6-31G | −0.349 | −0.593 | 0.126 | −0.270 | −0.265 | −0.038 |
| 6-31G** | −0.290 | −0.432 | 0.108 | −0.241 | −0.237 | −0.032 |
| 6-311G** | −0.258 | −0.336 | 0.081 | −0.228 | −0.223 | −0.026 |
| 6-311G(2d,2p) | −0.317 | −0.292 | 0.071 | −0.223 | −0.221 | −0.023 |
| 6-311++G(2d,2p) | −0.214 | −0.235 | −0.022 | −0.210 | −0.192 | −0.028 |

| Basis | O$\cdots$$H_6$ | O$\cdots$$H_{12}$ | $C_1$–$H_6$ | $C_1$–$H_7$ | $C_3$–$H_9$ | $C_3$–$H_{10}$ |
|---|---|---|---|---|---|---|
| 6-31G | −0.013 | −0.006 | −0.188 | −0.188 | −0.199 | −0.188 |
| 6-31G** | −0.010 | 0.002 | −0.169 | −0.170 | −0.175 | −0.169 |
| 6-311G** | −0.005 | 0.010 | −0.156 | −0.160 | −0.160 | −0.160 |
| 6-311G(2d,2p) | −0.004 | 0.017 | −0.158 | −0.165 | −0.156 | −0.163 |
| 6-311++G(2d,2p) | −0.003 | 0.016 | −0.164 | −0.164 | −0.163 | −0.164 |

**Table 5** Selected diatomic energy components of the *N,N*-dimethylformamide molecule calculated by using the decomposition scheme based on the virial theorem [eqn (9)] (atomic units)

| Basis | $C_4$–O | N–$C_4$ | N$\cdots$O | N–$C_3$ | N–$C_1$ | O–$H_9$ |
|---|---|---|---|---|---|---|
| 6-31G | −0.799 | −0.249 | 0.024 | −0.240 | −0.260 | −0.0015 |
| 6-31G** | −1.041 | −0.471 | 0.017 | −0.348 | −0.372 | −0.0016 |
| 6-311G** | −1.092 | −0.554 | 0.020 | −0.385 | −0.416 | −0.0045 |
| 6-311G(2d,2p) | −1.106 | −0.606 | 0.023 | −0.449 | −0.487 | −0.0049 |
| 6-311 + +G(2d,2p) | −0.963 | −0.636 | −0.002 | −0.434 | −0.471 | −0.0035 |
| cc-pVTZ | −1.200 | −0.647 | 0.032 | −0.470 | −0.507 | −0.0039 |

| Basis | O$\cdots H_6$ | O$\cdots H_{12}$ | $C_1$–$H_6$ | $C_1$–$H_7$ | $C_3$–$H_9$ | $C_3$–$H_{10}$ |
|---|---|---|---|---|---|---|
| 6-31G | 0.0001 | 0.013 | −0.358 | −0.355 | −0.357 | −0.356 |
| 6-31G** | 0.0001 | 0.013 | −0.396 | −0.392 | −0.399 | −0.393 |
| 6-311G** | 0.0001 | 0.008 | −0.418 | −0.413 | −0.421 | −0.414 |
| 6-311G(2d,2p) | 0.0001 | 0.005 | −0.421 | −0.413 | −0.425 | −0.416 |
| 6-311 + +G(2d,2p) | 0.0002 | 0.015 | −0.426 | −0.420 | −0.434 | −0.419 |
| cc-pVTZ | 0.0001 | 0.009 | −0.428 | −0.419 | −0.429 | −0.425 |

| Basis | Virial ratio |
|---|---|
| 6-31G | 1.999 370 |
| 6-31G** | 2.001 586 |
| 6-311G** | 1.999 997 |
| 6-311G(2d,2p) | 1.999 785 |
| 6-311 + +G(2d,2p) | 2.000 020 |
| cc-pVTZ | 2.000 385 |

the negative overlap population between these atoms, and may be in the final outcome understood as the topological effect of four π-electrons on three centers. (Simple Hückel theory gives a negative 1–3 Coulson's bond order in such a situation.) At the same time, the O$\cdots$N bond order is a small but non-negligible positive quantity, indicating the advantage of using energy partitioning as compared with the bond order indices that cannot always distinguish between bonding and antibonding situations.

The striking difference between the results given by the two schemes is in the *absolute values* of the numbers given by the two methods: scheme "1" assigns to the formally single C–H and C–N bonds negative energy contributions as large as 0.55 to 0.65 au (350–400 kcal mol$^{-1}$), and even numbers about 1.2–1.3 au (750–800 kcal mol$^{-1}$) occur for the C=O double bond. At the same time, scheme "2" assigns values about 0.16–0.24 au (100–150 kcal mol$^{-1}$) to the single bonds and about 0.35 au (220 kcal mol$^{-1}$) to the double one. These latter values are "on the chemical scale", while the numbers given by scheme "1" are of unusual magnitude for practical chemists.

Of course, as both schemes give results which exactly sum up to the total energy, similar large differences are observed in the results obtained for the one-center (monoatomic) energy components: while scheme "2" corresponds to moderate atomic promotion energies, scheme "1" predicts for them quite large values. (Thus, for instance, the one-center energies corresponding to hydrogens are about −0.26 to −0.29 au according to scheme "1", but between −0.46 and −0.49 for scheme "2".) Interestingly enough, there are no such dramatic differences between the results of

the schemes "1" and "2", as far as non-bonded interactions are concerned, which is obviously due to the fast decay of the kinetic energy integrals with the distance.

The results provided by the virial decomposition are somewhat different. For the case of covalent chemical bonds it gives numbers which again have rather large absolute values (though somewhat less than those given by scheme "1"), while for non-bonded interactions the numbers are systematically smaller than in either schemes "1" or "2". This may again be connected with the fast decay of the kinetic energy integrals. Nonetheless, one may consider that the performance of the virial energy decomposition is surprisingly good. In fact, while one can directly link the different energy terms occurring in eqn (5)–(8) to some well-defined physical interactions (kinetic energy, electron–nuclear attraction and electron–electron repulsion), eqn (9) holds only by virtue of a rather subtle interrelation between the overall kinetic and potential energies which is valid *only in stationary points of the potential surface*. It was by far not so obvious that this complex relationship should lead to such a "well cut" interpretation of the individual atoms and bonds in polyatomic molecules. (The one-center terms are quite reasonable, too.) We should recall here a characteristic peculiarity of the virial theorem in the framework of the Born–Oppenheimer separation: the energy on the left-hand site of the expansion (5) is the total energy including the nuclear–nuclear repulsion, while the right-hand site contains only the kinetic energy of the electrons.[32] In addition, one should realize that the basis sets applied here are not very large and, accordingly, the virial theorem is not particularly well fulfilled: the virial ratios do not approach particularly closely the ideal value of 2.

It is a well-known general tendency that longer bonds are weaker (and exhibit lower stretching frequencies). As noted above, the N–C(Me) bond length $R(C_3N)$ in the *syn* position to the carbonyl is somewhat longer than $R(C_4N)$ in the *anti* position. In full accord with this, the respective bond orders are somewhat smaller as well as the absolute values of the respective energy components given by scheme "1" and by the virial decomposition. At the same time, the more "chemical" scheme "2" leads to a more negative two-center energy component for the longer $R(C_3N)$ in the *syn* position. This counterintuitive behaviour is a manifestation of the dilemma which will be discussed in the next section.

The situation is more complex as far as one considers the C–H bonds. The bond length $R(C_3H_9)$ of the hydrogen entering the C–H$\cdots$O interaction is significantly shorter than all the other ones (which is a feature similar to the "anomalous" or "blue shift" C–H$\cdots$O bonds[33]) but the bond orders are smaller—obviously because some bond order is built up with the oxygen atom. The values of the C–H energy components given by scheme "1" and the virial decomposition are in accord with the bond lengths: the shorter the bond, the larger is the absolute value of the energy component. Scheme "2" does not exhibit any clear tendency in this case and the relative magnitudes of the C–H energy components depend on the basis used—which is a disturbing behaviour.

The data in Tables 3 and 4 clearly show the existence of some attractive interaction between the hydrogen and oxygen atoms. The question arises whether this is simply the electrostatic interaction between a negative oxygen and a positive hydrogen or if there is some bonding, too. The answer is "yes", as indicated by the appearance of a non-negligible bond order component. No such bond order can be observed for the other methyl hydrogens—as it is also absent in the acetone molecule. It may also be noted that both the bond orders and the H$\cdots$O energy component are about half of the analogous values calculated for the water dimer by using the same basis sets.¶ Thus we may conclude that energy component analysis also confirms the assumption[30] about the presence of a non-trivial C–H$\cdots$O interaction in the *N,N*-

---

¶ Obviously, the water dimer calculations performed in these small basis sets require correction for BSSE. No doubt, analogous "effects" can arise within a given molecule, too, although this problem is not usually considered explicitly.

dimethylformamide molecule. (However, it may be an exaggeration to call this interaction a "hydrogen bond": not only is the H···O distance longer than that for a standard H-bond,[30] but the C–H···O angle is only about 102–103°.)

When considering the numbers obtained for the C–H···O interaction, which are scattered around 0.03 au (20 kcal mol$^{-1}$), one should keep in mind that *it is not the net effect of the interaction*—one has to subtract the energy, necessary to stretch other bonds and angles; one-center energy components can also change. The resulting net effect can be order(s) of magnitude less than the primary C–H···O energy component; the latter, however, points to the driving force behind all these changes.

## III.  The dilemma

For all systems studied to date the energy decomposition scheme "2" gave very "chemical" energy components for molecules at their equilibrium geometries. Therefore, at least at first sight, one might indeed consider scheme "2" (and its 3D counterpart[23]) as *the* appropriate energy partitioning scheme which may be recommended for general use. This scheme should also be able to please the practical chemists, some of whom were reluctant to deal with the "too large" numbers given by our previous CECA scheme[23] that gives results comparable with (though not identical to) those of scheme "1" above. Indeed, the results provided by scheme "2" are pretty reasonable from the chemist's point of view: the two-center energy components are comparable with the bond dissociation energies, and some atomic promotion energies are also predicted.[27] For that reason this energy partitioning method might appear to be the most attractive among the known Hilbert-space schemes.

However, we have already seen in the example of *N,N*-dimethylformamide that the situation is not that simple by far: there are cases in which scheme "2" assigns the more negative value to the diatomic energy component of the longer of two similar bonds and not to that of the shorter, as one would expect (and as scheme "1" does). Although, obviously, the diatomic energy component is a static parameter which is not directly related to the dissociation energy of the given bond, one is inclined to consider the diatomic energy component as some measure of the bond strength and such examples are, therefore, disturbing. It appears that this problem is not related to some peculiar properties of that or another molecule, but it is a manifestation of a general counterintuitive behaviour of scheme "2": if a bond is stretched with respect to its equilibrium length then this leads to the increase (in absolute value) of the respective diatomic energy component and not to its decrease as one would expect. At the same time, the energy components given by scheme "1" are too large in absolute value, but exhibit a "correct" behaviour with the distance.

Fig. 2 compares the distance dependence of the C–C energy components given by the two schemes "1" and "2" for the ethane molecule. It shows that the problem is quite acute even for such a simple system, and should be considered an inherent property of these decomposition schemes. Thus, one is faced with a *dilemma*: one either gets numbers which are "on the chemical scale" and have quite appealing values *at the equilibrium molecular geometries*, but exhibit a counterintuitive distance dependence (the two-center energy components increase in absolute value with the increase of the interatomic distances)—or numbers which have too large absolute values but "correct" distance behaviour.‖ This dilemma is obviously due to the fast decay of the kinetic energy integrals with the distance, as compared with the other terms, in particular, to those of an electrostatic nature. (We recall that the two schemes differ only in regrouping some kinetic energy terms.)

---

‖ One does not expect the two-atomic energy contribution to reflect directly the minimum on the potential energy surface, as the one-center energy components are also changing with the bond distances.
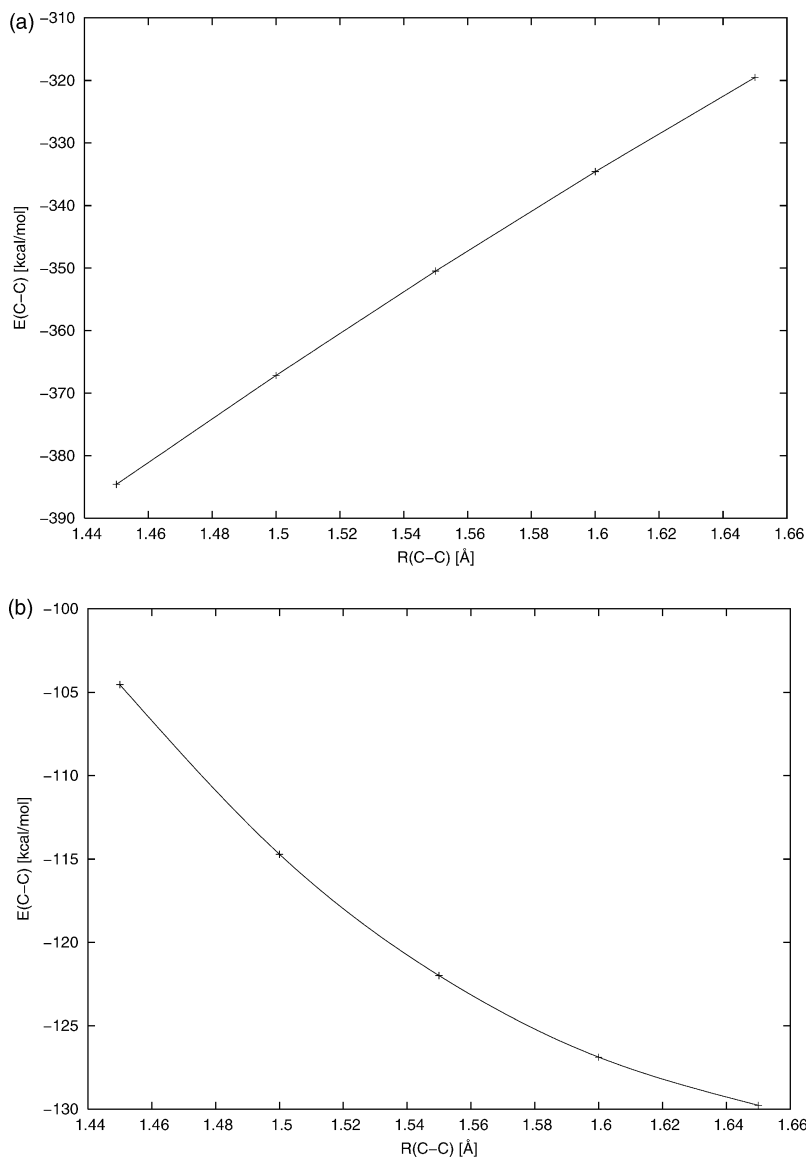
Fig. 2 Distance dependence of the C–C energy components of the ethane molecule (basis 6-31G\*\*) calculated (a) by scheme "1" and (b) by scheme "2", respectively.

When considering this dilemma, one has to ponder that there is no actual proof telling us that scheme "2" should *always* give "chemical" numbers; although one cannot exclude that this behaviour may be in some manner connected with the fact that the usual form of the virial theorem is satisfied only at the stationary points in the Born–Oppenheimer framework, but no such explicit connection has as yet been discovered.

The dilemma is not limited to the Hilbert-space analyses because in the 3D analysis performed in terms of "fuzzy atoms" (divisions of the three dimensional physical space into atomic regions without sharp boundaries) we have also obtained a similar pair of energy decomposition schemes and again the one involving two-

center kinetic energy components—*i.e.*, the analogue of scheme "2"—was that which gave chemically appealing numbers.[23]

One may collect some arguments (besides the "correct" distance dependence) which may help one to be reconciled to the use of the numbers out of the "chemical scale", as are predicted by scheme "1". For instance, in the 3D analysis, if the physical space is decomposed into disjunct atomic domains, as is the case in the AIM theory, then the kinetic energy contributions are strictly one-center ones, similar to our scheme "1". As in this case the 3D energy decomposition is a quite "natural" one (each one-electron integral spontaneously breaks down into atomic contributions), this comparison may motivate us to stick to the atomic (and not diatomic) decomposition of the kinetic energy. Also, the approximate CECA scheme[23] which gives similar (though not identical) numbers as does scheme "1", had been found applicable not only for interpretations but also for predicting mass spectra;[34] most probably scheme "1" would be equally good for that purpose. Thus these unfamiliar "too large" numbers may appear useful. One may also mention in this context that the energy components in the semiempirical schemes, like MNDO, are also obtained in a natural fashion, and have similarly large "non-chemical" values as those of our scheme "1". Nonetheless they are very often able to reproduce well the heats of formation and other molecular properties. So the large absolute values of the energy component do not necessarily contradict the chemistry of the problem.

Of course, one would prefer to have a scheme which gives both appealing numbers and exhibits a correct distance-dependence. It is unclear whether this is possible at all, definitely not without extending the theoretical toolbox which is presently at our disposal.

## Acknowledgements

## References

1   G. G. Hall, *Chairman's remarks*, 5th International Congress on Quantum Chemistry, Montreal, 1985.
2   I. Mayer and A. Hamza, *Theor. Chem. Acc.*, 2001, **105**, 360.
3   P. Salvador, M. Duran and I. Mayer, *J. Chem. Phys.*, 2001, **115**, 1153.
4   R. F. W. Bader, *Atoms in Molecules. A Quantum Theory*, Oxford University Press, Oxford, UK, 1990.
5   E. Clementi, *J. Chem. Phys.*, 1967, **46**, 3842.
6   J. B. Moffat and H. E. Popkie, *Int. J. Quantum Chem.*, 1968, **2**, 565.
7   H. Fischer and H. Kollmar, *Theor. Chim. Acta*, 1970, **16**, 163.
8   K. Kitaura and M. Morokuma, *Int. J. Quantum Chem.*, 1976, **10**, 325.
9   F. Driessler and W. Kutzelnigg, *Theor. Chim. Acta*, 1976, **43**, 1.
10   T. Ziegler and A. Rauk, *Theor. Chim. Acta*, 1977, **46**, 1.
11   T. Ziegler and A. Rauk, *Inorg. Chem.*, 1979, **18**, 1558, 1755.
12   H. Kollman, *Theor. Chim. Acta*, 1978, **50**, 235.
13   E. Kapuy, C. Kozmutza, R. Daudel and M. E. Stephens, *Theor. Chim. Acta*, 1979, **53**, 147.
14   A. E. Reed, L. A. Curtiss and F. Weinhold, *Chem. Rev.*, 1988, **88**, 899.
15   E. D. Glendening and A. J. Streitwieser, *J. Chem. Phys.*, 1994, **100**, 2900.
16   A. Sierraalta and G. Frenking, *Theor. Chim. Acta*, 1997, **95**, 1.
17   H. Ichikawa and A. Yoshida, *Int. J. Quantum Chem.*, 1999, **71**, 35.
18   Y. Mo, J. Gao and S. D. Peyerimhoff, *J. Chem. Phys.*, 2000, **112**, 5530.
19   P. L. A. Popelier and D. S. Kosov, *J. Chem. Phys.*, 2001, **114**, 6539.

20  G. Frenking, K. Wichmann, N. Fröhlich, C. Loschen, M. Lein, J. Frunzke and V. M. Rayón, *Coord. Chem. Rev.*, 2003, **238–239**, 55.
21  H. Nakai and Y. Kikuchi, *J. Theor. Comput. Chem.*, 2005, **4**, 317.
22  I. Mayer, *Chem. Phys. Lett.*, 2000, **332**, 381.
23  P. Salvador and I. Mayer, *J. Chem. Phys.*, 2004, **120**, 5046.
24  B. Jeziorski, R. Moszynski and K. Szalewicz, *Chem. Rev.*, 1994, **94**, 1887.
25  M. A. B. Blanco, A. Martin Pendás and E. Francisco, *J. Chem. Theor. Comput.*, 2005, **1**, 1096.
26  I. Mayer and A. Hamza, *Int. J. Quantum Chem.*, 2005, **103**, 798.
27  I. Mayer, *Chem. Phys. Lett.*, 2003, **382**, 265.
28  K. Ruedenberg, *Rev. Mod. Phys.*, 1962, **34**, 326.
29  I. Mayer, Program ''BO-VIR'', Budapest, 2005. May be downloaded from the web-site http://occam.chemres.hu.
30  G. Schultz and I. Hargittai, *J. Phys. Chem.*, 1993, **97**, 4966.
31  I. Mayer, *Chem. Phys. Lett.*, 1983, **97**, 270 (addendum: 1985, **117**, 396).
32  I. Mayer, in *Simple Theorems, Proofs, and Derivations in Quantum Chemistry*, Kluwer Academic/Plenum Publishers, New York, 2003.
33  Y. Gu, T. Kar and S. Scheiner, *J. Am. Chem. Soc.*, 1999, **121**, 9411.
34  I. Mayer and A. Hamza, Program ''APOST-MS'', Budapest, 2001. May be downloaded from the web-site http://occam.chemres.hu (*cf.* also ref. 35).
35  I. Mayer and Á. Gömöry, *Chem. Phys. Lett.*, 2001, **344**, 553.