

# Molecular energy decompositions

István Mayer<sup>1,a</sup>

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

**Abstract.** Energy decomposition is a valuable tool for understanding the results of a quantum chemical calculation and connecting them with the genuine chemical concepts. We are summarizing here some of our results obtained in the last decade, and a new scheme is proposed which is capable to solve the problems which previously remained open.

## 1 Introduction

The understanding and interpretation of the results obtained in a quantum chemical calculation can be much facilitated by presenting the total energy as a sum of chemically meaningful components: contribution of the individual atoms and of the pairs of atoms. In the era of semiempirical quantum chemistry, this aim could be achieved trivially: the semiempirical model Hamiltonians contained only one- and two-center integrals, so the different terms could be allocated to the atoms or pairs of atoms involved in a quite natural manner [1]. In the *ab initio* theory, however, there are three- and four-center integrals, too, and in the straightforward energy component analysis of Clementi [2] they led to the appearance of significant three- and four-atomic energy contributions, in sharp contradiction with the chemist's way of thinking of molecule as consisting of atoms linked by bonds or exhibiting pairwise non-bonded interatomic interactions.

The solution of this problem is difficult because there is an infinite number of possibilities to present a single number as a sum of components, and we are lacking any strict definition of an atom *within* a molecule. Actually one can perform the analysis either in the Hilbert space of the atomic basis orbitals or in the physical 3-dimensional (3D) space, but neither is unique.

## 2 Energy decomposition in the AIM scheme

In Bader's topological "Atoms in Molecules" (AIM) theory [3] the 3D space is decomposed into *disjunct* atomic domains  $\Omega_A$ . This means that every integral over the space represents a sum of integrals over the individual domains:  $\int f(\mathbf{r})dv \equiv \sum_A \int_{\Omega_A} f(\mathbf{r})dv$ . As a consequence, the SCF energy of the molecule breaks down spontaneously into a sum of one- and two-center energy components [4]:

$$E = \sum_A E_A + \sum_{A<B} E_{AB} \quad (1)$$

The explicit expressions of the energy components are given in the paper quoted.

In our opinion, the existence of this "natural" energy decomposition has a big conceptual importance when the connection of the quantum mechanical and chemical description of a molecule is concerned. From the numerical point of view, this scheme leads to quite reasonable numbers [5], but is so demanding computationally that it is not appropriate for the routine use.<sup>1</sup> An alternative of the AIM

<sup>a</sup> e-mail: mayer.istvan@ttk.mta.hu

<sup>1</sup> Also, the use of the AIM scheme becomes problematic in cases when there appear domains (so called "non-nuclear attractors") which cannot be assigned to either of the atoms.

scheme is the use of “fuzzy atoms” i.e., such divisions of the 3D physical space into atomic regions in which the regions assigned to the individual atoms have no sharp boundaries but exhibit a continuous transition from one to another [6,7]. The energy decompositions performed in terms of fuzzy atoms also gave interesting results [8,9] — they are, however, out of our present scope.

### 3 The CECA scheme

As early as in 1983 the present author introduced the so called “chemical Hamiltonian approach” [10] devoted to connect the Hilbert space description of molecules (the use of atom-centered basis sets) with the genuine chemical concept of atoms exhibiting pairwise interatomic interactions. Already that paper contained the formulae of an energy decomposition scheme; later on it was improved somewhat but these results were left unpublished for about a decade. A discussion with Pierre Valiron, however, changed the situation and he convinced the author that these formulae are worth to be programmed. That was done first by using Pierre’s version of the HONDO *ab initio* program system [11]. The scheme thus obtained [12,13] was named “Chemical Energy Component Analysis” (CECA). In this approach the energy decomposition is exact for diatomic molecules, while the three- and four-center effects are compressed into one- and two-center ones as much as it is possible by performing appropriate projections. Thus one gets an approximate version of the decomposition Eq. (1): the energy of a molecule calculated at the SCF level is expressed *approximately but to a good accuracy* as a sum of atomic and diatomic contributions, the computation of which requires the use of *one- and two-center integrals only*.

The idea of CECA can be illustrated on the example of a three-center electron-nuclear attraction integral  $\langle \chi^A | \frac{Z_B}{r_B} | \chi^C \rangle$ . (The sub- and superscripts indicate the centers.) Function  $\frac{Z_B}{r_B} | \chi^C \rangle$  has a diatomic character, it corresponds to the fragment BC. We *replace* it by its projection on the basis set of that fragment:

$$\frac{Z_B}{r_B} | \chi^C \rangle \Rightarrow \hat{P}^{BC} \frac{Z_B}{r_B} | \chi^C \rangle \quad (2)$$

The projector on some fragment  $X$  for an *overlapping* basis set is

$$\hat{P}^X = \sum_{\mu, \nu \in X} | \chi_\mu \rangle S_{(X)\mu\nu}^{-1} \langle \chi_\nu | \quad (3)$$

where the shorthand  $S_{(X)\mu\nu}^{-1} \equiv (\mathbf{S}_{(X)}^{-1})_{\mu\nu}$  has been introduced for the elements of the *inverse* intrafragment overlap matrix. By substituting the expression of  $\hat{P}^{BC}$ , and defining the projection-type matrix  $A^{BC}$

$$A_{\nu\tau}^{BC} = \sum_{\rho \in BC} S_{\nu\rho} S_{(BC)\rho\tau}^{-1} \quad (4)$$

we get:

$$\langle \chi_\nu^A | \frac{Z_B}{r_B} | \chi_\mu^C \rangle \Rightarrow \sum_{\tau \in BC} A_{\nu\tau}^{BC} \langle \chi_\tau^C | \frac{Z_B}{r_B} | \chi_\mu^C \rangle \quad (5)$$

For the two-electron integrals, function  $\frac{1}{r_{12}} \chi_\rho^C(1) \chi_\tau^D(2)$  is of diatomic (CD) character. Two projectors should be used, one for electron “1”, another for electron “2”. Thus

$$[ \chi_\mu^A \chi_\nu^B | \chi_\rho^C \chi_\tau^D ] \Rightarrow \sum_{\lambda, \eta \in CD} A_{\mu\lambda}^{CD} A_{\nu\eta}^{CD} [ \chi_\lambda^C \chi_\eta^D | \chi_\rho^C \chi_\tau^D ] \quad (6)$$

Similar procedures are used for separating *intraatomic* energy contributions. In this manner one arrives at the *approximate* energy decomposition

$$E \approx \sum_A E_A + \sum_{A < B} E_{AB} \quad (7)$$

which has been found sufficiently accurate even for small basis sets. The explicit expressions of the energy components are given in the paper [12]; they contain only one- and two-center integrals.

It has been found that the approximations applied in the CECA scheme are pretty good, and the results obtained reflects well the different intramolecular interactions; for instance they can be used to identify weak forces like C-H...O interactions. The results show very well the chemical transferability, and thus may have a great interpretative power; moreover in some case they can even used for prediction—*e.g.*, for that of mass spectra [14]. The method is very well suited for routine work.

A characteristic feature of the CECA scheme is that the numerical values are, so to say, “non chemical”: the atomic energies are very positive (as compared with the free atom energies), while the bonding energies are very large negative numbers. The same tendency can be observed for the “straightforward” energy decompositions in the semiempirical methods. We shall return to this problem later.

The formulae of the CECA scheme were subjected to a further analysis, and the diatomic terms have been decomposed into components of different physical nature [15]. Thus one can identify electrostatic terms, which can further be decomposed into terms in point-charge approximation and those describing the deviation from the point charges; for atoms which are at least 2–3 bonds apart the point-charge approximation is usually quite sufficient. Very important is the term reflecting interatomic exchange interactions; essentially it determines whether there is a chemical bonding between the atoms in question. There are two types of terms which should vanish as the basis on the individual atoms approaches completeness. One of them is connected with the overlaps of the orbitals centered on different atoms and for moderate basis sets can be related to the repulsion of closed shells (Pauli repulsion). The other type of terms describes the following effect: the remainder of the approximation of the functions having intraatomic origin by their intraatomic projections can be, in part, expanded in terms of diatomic fragments. These give rise to some “basis extension terms” that in CECA are assigned diatomic character.

## 4 Some further developments

It has been observed that not only the energy components, but even the errors of the projective approximation are pretty transferable and additive (*e.g.*, change by the same value if an additional CH<sub>2</sub> unit is added to an aliphatic chain). That opened the way to further developments.

A very simple “exact” decomposition has been proposed [16] in which each term is simply assigned according to the “ket” (and nucleus). By introducing the integral-approximations of CECA, an approximate version can also be obtained having the same overall error as CECA; it differs from CECA only by assigning the “basis extension” terms to the atoms the atomic Hamiltonians of which are involved in the terms of question. By comparing the results of the exact and approximate versions, the projection errors can be assigned to the respective atoms (pairs of atoms); thus, in turn, an “exact” version of CECA is also obtained.

There is also some freedom in treating kinetic energy: instead of considering it as a component of the intraatomic Hamiltonian, one may treat separately its one- and two-center integrals. This leads to a redistribution between the one- and two-center energy components, and permits to obtain very “chemical” numbers [17] but *only at the equilibrium molecular geometries*: when a bond is stretched, the diatomic energy components become more negative, and not less negative as would expected [16]. This behaviour leads to a dilemma [18]: either the results are “on the chemical scale” for molecules at equilibrium, but behave “counter-intuitive” when the bond-lengths change, or they have a correct distance dependence but have too large absolute values.

## 5 The new scheme

An energy decomposition with wrong distance dependence is hardly acceptable, so in order to solve the dilemma we had to start from the other scheme and clarify why the numbers have too large absolute values of opposite signs. We have found that the clue is the wrong dissociation of the RHF method: it

has too large ionic terms that cause too high energy at the large distances. But the ionic terms are due to forcing the doubly occupied MO-s, so are not atomic properties. The large ionic terms are present at the usual distances, too, and the energy increase they cause should be attributed to the bonds, rather than to the atoms.

Thus the energy of electron-electron repulsion due to the ionic terms *caused by the bonding* (and only that which is caused by the bonding) should be distributed between the individual *pairs* of atoms.

For instance, if we consider the H<sub>2</sub> molecule treated at the *minimal basis* (classical LCAO) level, then from the physical point of view we have two H atoms occupied by a single electron each, like a free H atom, but they are bonded. There is *no room for promotion* owing to the minimal basis set—therefore the atomic energy components must coincide with the free atom's energy, and *all other energy components should be attributed to the bond*. This requirement permits the energy decomposition to become essentially unique.

Thus the new energy partitioning scheme consists of the following steps:

1. Performing the (approximate) CECA energy decomposition;
2. Introducing corrections for the remaining 3- and 4-center effects by comparing a "kets only" exact decomposition and its approximation by the same projected integral expansion scheme as used in CECA ("exact" version of CECA);
3. Regrouping the ionic energy terms *which are due to the bonding* to the respective two-center terms.

One can obtain and redistribute the energy increases due to the ionic terms in the following manner, applicable for polyatomic molecules, too:

- a) One calculates the "effective atomic orbitals"  $\chi_i$  for every atom. These represent generalizations of McWeeny's natural hybrid orbitals and diagonalize the atomic block of the first order density matrix. Actually the algorithm of Mayer [19] is used to obtain the effective AO-s.
- b) One calculates the two-electron integral  $\frac{1}{4}[\chi_i\chi_i|\chi_i\chi_i]$  and multiplies it by the contribution  $V_A^i$  of the given orbital to the valence of the atom. The product is deduced from the one-center energy. (The factor  $\frac{1}{4}$  comes from the weight of the ionic term in the HF wave function.)
- c) Using the relation  $V_A^i = \sum_{B \neq A} B_{AB}^i$  between valence and bond orders, one adds the value  $\frac{1}{4}[\chi_i\chi_i|\chi_i\chi_i]B_{AB}^i$  to each diatomic energy component  $\Delta E_{AB}$ .

## 6 Conclusions

Energy partitioning is useful for:

- Connecting the physical and chemical picture of molecules;
- Identifying and interpreting the primary and secondary interactions, their locality, transferability, *etc.*;
- Can be used for predictions in some cases.

The *different schemes* may lead to much different numbers but not to much different qualitative conclusions.

Some numerical values calculated by the new scheme

Atom(s)	Energy component (6-31G**)		
	$\Delta E_A, E_{AB}, \text{kcal/mol}$		
	Ethane	Ethylene	Acetylene
C	83.91	108.33	155.38
H	54.62	67.74	102.57
C–C	-159.13	-296.62	-481.02
C–H	-167.20	-167.22	-150.97
H–H geminal	5.92	10.05	–
H–H vicinal	2.30 (2x)	2.97	5.03
	1.54 (1x)	1.89	
C–H vicinal	5.92	6.51	-16.80

## References

1. Fischer, H., Kollmar, H. *Theor. Chim. Acta* 16, 163 (1970)
2. Clementi, E. J. *Chem. Phys.* 46, 3842 (1967)
3. Bader, R.F.W. "*Atoms in Molecules: A Quantum Theory*" Oxford University Press (Oxford) (1990)
4. Mayer, I., Hamza, A. *Theor. Chem. Acc.* 105, 360. (2001a)
5. Salvador, P., Duran, M., Mayer, I. *J. Chem. Phys.* (2011)
6. Hirshfeld, F.L. *Theor. Chim. Acta* 44 129 (1977)
7. Mayer, I., Salvador, P. *Chem. Phys. Lett.* 383, 368. (2004)
8. Salvador, P., Mayer, I. *J. Chem. Phys.* 120, 5046 (2004)
9. Salvador, P., Mayer, I. *J. Chem. Phys.* 126, 234113 (2007)
10. Mayer, I. *Internat. J. Quantum Chem.* 23, 341 (1983)
11. Dupuis, M., Farazdel, A. "*HONDO-8 from MOTECC-91*", IBM Corporation, Center for Scientific & Engineering Computations, Kingston, New York.115, 1153 (1991)
12. Mayer, I. *Chem. Phys. Lett.* 332, 381 (2000)
13. Mayer, I., Hamza, A. *Program "APOST"*, <http://occam.chemres.hu>, Budapest (2000)
14. Mayer, I., Hamza, A. *Program "APOST-MS"*, <http://occam.chemres.hu>, Budapest (2011b)
15. Mayer, I., Hamza, A. *Theor. Chem. Acc.* 109, 91. (2003)
16. Mayer, I. *Phys. Chem. Chem. Phys.*, 8, 4630 (2006)
17. Mayer, I. *Chem. Phys. Lett.* 382, 265 (2003)
18. Mayer, I. *Faraday Discussions*, 135, 439 (2007)
19. Mayer, I. *Chem. Phys. Lett.* 498, 366 (2010)