One- and two-center energy components in the atoms in molecules theory

P. Salvador,[a] M. Duran, and I. Mayer[b]
Department of Chemistry and Institute of Computational Chemistry, University of Girona,
17071 Girona, Spain

The energy decomposition scheme proposed in a recent paper has been realized by performing numerical integrations. The sample calculations carried out for some simple molecules show excellent agreement with the chemical picture of molecules, indicating that such an energy decomposition analysis can be useful from the point of view of connecting quantum mechanics with the genuine chemical concepts. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1381407]

I. INTRODUCTION

Most recently, one of us demonstrated[1] that the Hartree–Fock (HF) energy of a molecule can be presented *exactly* as a sum of atomic and diatomic contributions in the framework of the Atoms in Molecules (AIM) theory,[2] introduced by Bader, or in any other scheme in which the three-dimensional physical space is decomposed into disjunct atomic domains (basins). This result follows from the simple facts that

(a) the integrals over the whole space are equal to a sum of integrals over the individual domains,

(b) the nuclei and the atomic basins can usually be put into one-to-one correspondence with each other.

The HF energy contains both one- and two-electron terms, described by single and two-fold integrals over the 3D space, respectively. The kinetic energy integrals and the electrostatic interaction of the electronic charge within a given basin with its own nucleus contribute to the atomic (one-center) energy component, while the electrostatic interaction of the electronic charge in one domain with a nucleus in another one represents a diatomic effect. Each two-electron (Coulomb or exchange) integral over the molecular orbitals (MO's) will contain both monoatomic and diatomic components, depending on whether the two integration variables are actually in the same or different basins. The total molecular energy contains also the intermolecular repulsions, which are obviously of diatomic nature.

The authors of Ref. 1 have also demonstrated the close conceptual correspondence of such a decomposition in the AIM theory with the recent chemical energy decomposition analysis (CECA) performed in the linear combination of atomic molecules (LCAO) framework,[3,4] which seems to be a promising simple tool for the *a posteriori* chemical analysis of the results obtained in the *ab initio* calculations. The equations of these two schemes are connected with a mathematical mapping which, however, does not mean that the numerical values should coincide; moreover, the energy decomposition in the AIM case is, in principle, exact while in CECA it is only a (good) approximation. We consider the possibility of presenting the molecular energy as a sum of atomic and diatomic contributions to be a result of upmost conceptual importance from the point of view of connecting quantum mechanics with the genuine chemical concepts. In fact, such an energy decomposition can help us to obtain a deeper insight into the physical content which is behind the chemical structural formulas. Of course, the abstract mathematical results of Ref. 1 will appear truly useful from the practical chemical point of view only if the decomposition leads to numerical results which correlate well with the chemical picture of the molecules. The aim of the present work is to demonstrate that this is really the case. As the integration of molecular orbitals over the atomic basins cannot be performed—at least at time being—analytically, we have had to recur to numerical integrations.

It is worthwhile to point out that this decomposition is not restricted to the HF level of theory, as the expectation value of the energy can always be written in the form of sums in which the one- and two-electron integrals over the molecular orbitals are multiplied by the respective first or second order density matrix elements.

It should also be mentioned that the AIM theory itself contains an energy decomposition scheme which is based on the atomic virial theorem.[2] This scheme presents the molecular energy as a sum of atomic energies, $\Sigma_A E(A)$, integrated for each atomic basin. That type of decomposition is extremely important from a physical point of view, but it is lacking a direct connection with the chemical notion of atoms interacting with each other. This aspect of the original AIM energy decomposition motivated Sierraalta and Frenking to assign part of the atomic energies to the diatomic interactions, by using some overlap integrals over the atomic basins as proportionality factors.[5] Their theory, however, always gives negative diatomic energy components, thus it is unable to distinguish between bonding and non-bonding—or more generally, between attractive and repulsive—interactions between the individual pairs of atoms.

[a]Author to whom correspondence should be addressed; Fax: +34 972 418356; electronic mail: perico@iqc.udg.es
[b]Also at: Institute of Chemistry, Chemical Research Center, Hungarian Academy of Sciences, H-1525 Budapest, P.O. Box 17, Hungary.

II. THEORY

As discussed in Ref. 1, the self-consistent field (SCF) energy

$$E = 2\sum_{i=1}^{N/2} \langle \varphi_i | \hat{h} | \varphi_i \rangle$$

$$+ \sum_{i,j=1}^{N/2} (2[\varphi_i\varphi_j|\varphi_i\varphi_j] - [\varphi_i\varphi_j|\varphi_j\varphi_i]) \quad (1)$$

can be strictly decomposed into one-center and two-center components:

$$E \equiv \sum_A E_A + \sum_{A<B} E_{AB}, \quad (2)$$

$$E_A = 2\sum_{i=1}^{N/2} \langle \varphi_i | -\frac{1}{2}\nabla^2 - \frac{Z_A}{R_A} | \varphi_i \rangle_A$$

$$+ \sum_{i,j=1}^{N/2} (2[\varphi_i\varphi_j|\varphi_i\varphi_j]_{A,A} - [\varphi_i\varphi_j|\varphi_j\varphi_i]_{A,A}), \quad (3)$$

$$E_{AB} = -2\sum_{i=1}^{N/2} \left(\langle \varphi_i | \frac{Z_B}{R_B} | \varphi_i \rangle_A + \langle \varphi_i | \frac{Z_A}{R_A} | \varphi_i \rangle_B \right) + 2\sum_{i,j=1}^{N/2}$$

$$\times (2[\varphi_i\varphi_j|\varphi_i\varphi_j]_{A,B} - [\varphi_i\varphi_j|\varphi_j\varphi_i]_{A,B}) + \frac{Z_A Z_B}{R_{AB}}, \quad (4)$$

where $N/2$ is the number of doubly filled MO's and the $[12|12]$ convention is used for the two-electron integrals. In the one-electron integrals, the subscript "A" denotes that the integration is performed over the Ath atomic basin

$$\langle \varphi_i | \hat{g} | \varphi_i \rangle_A = \int_{\Omega_A} \varphi_i^*(\vec{r}) \hat{g} \varphi_i(\vec{r}) d\nu, \quad (5)$$

with $\hat{g}$ representing operators $-\frac{1}{2}\nabla^2$ or Z_C/R_C and $C=A$ or B.

For the two-electron integrals, the subscript "A,B" indicates that the integration for the electrons 1 and 2 are carried out over the atomic domains Ω_A and Ω_B, respectively:

$$[\varphi_i\varphi_j|\varphi_k\varphi_l]_{A,B}$$

$$= \int_{\Omega_A} d\nu_1 \int_{\Omega_B} d\nu_2 \varphi_i^*(\vec{r}_1) \varphi_j^*(\vec{r}_2) \frac{1}{r_{12}} \varphi_k(\vec{r}_1) \varphi_l(\vec{r}_2). \quad (6)$$

III. COMPUTATIONAL ASPECTS

The integrals defined over the atomic basins must be computed numerically. The identification of the individual AIM basins and the numerical calculation of the one-electron (kinetic energy and nuclear attraction) integrals over them have been performed by using a slightly modified version of the PROAIMV[6] program. The same program was utilized to generate some arrays of data, which were saved on disk and have been used to calculate the two-electron integrals as follows:

The PROAIMV program performs the numerical integration over the atomic domains by generating a grid of points with the radius vectors $\vec{r}_\mu$ and assigning them respective weight factors w_μ, so that the integral of some scalar function $f(\vec{r})$ over the basin Ω_A is approximated as

$$\int_{\Omega_A} f(\vec{r}) d\nu \cong \sum_{\mu \in A} f(\vec{r}_\mu) w_\mu. \quad (7)$$

Here and further on, the notation $\mu \in A$ indicates the set of points belonging to the basin of the atom A.

We have used such numerical integrations to calculate the kinetic energy and nuclear attraction integrals for each nucleus C and each occupied MO φ_i, by defining $f(\vec{r}) = -\frac{1}{2}\varphi_i(\vec{r})\nabla^2\varphi_i(\vec{r})$ and $f(\vec{r}) = -(Z_C/R_C)|\varphi_i(\vec{r})|^2$ for each atom C, respectively. Only real orbitals have actually been considered in the calculations.

The numerical calculation of the two-electron integrals can be performed in a straightforward manner by a repeated use of the integration rule given in Eq. (7). As a preparatory step to such calculations, we have generated by the properly modified program PROAIMV the set of data defined as

$$G_{ij}^A(\mu) = \varphi_i(\vec{r}_\mu)\varphi_j(\vec{r}_\mu)w_\mu \quad (8)$$

for each atomic basin A, and each pair of occupied MO's φ_i and φ_j (including the case $i=j$) in every point $\vec{r}_\mu$, $\mu \in A$.

Using these quantities, the repeated use of Eq. (7) for calculating the integrals of exchange type leads simply to

$$[\varphi_i\varphi_j|\varphi_j\varphi_i]_{A,B}$$

$$= \int_{\Omega_A} d\nu_1 \int_{\Omega_B} d\nu_2 \varphi_i^*(\vec{r}_1)\varphi_j^*(\vec{r}_2) \frac{1}{r_{12}} \varphi_j(\vec{r}_1)\varphi_i(\vec{r}_2)$$

$$\cong \sum_{\mu \in A} \sum_{\nu \in B} G_{ij}^A(\mu) G_{ij}^B(\nu) |\vec{r}_\mu - \vec{r}_\nu|^{-1}. \quad (9)$$

A similar formula holds for the individual Coulomb integrals, too. However, one can compute the sum of the Coulomb type contributions at once, because

$$2\sum_{i,j=1}^{N/2} [\varphi_i\varphi_j|\varphi_i\varphi_j]_{A,B} = \frac{1}{2}\int_{\Omega_A} d\nu_1 \int_{\Omega_B} d\nu_2 \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{r_{12}}, \quad (10)$$

$\rho(\vec{r}_1)$ being the charge density

$$\rho(\vec{r}) = 2\sum_i^{N/2} |\varphi_i(\vec{r}_1)|^2. \quad (11)$$

For this case, the functions G_{ij}^A with $i=j$ have been utilized. By introducing the auxiliary quantities J^A

$$J^A(\mu) = \sum_i^{N/2} G_{ii}^A(\mu) \quad (12)$$

the sum of the integrals of Coulomb-type over the pairs of atomic basins can be written as

TABLE I. Integration parameters and energies (a.u.) computed for the N_2 molecule using the $6-31G(d,p)$ basis set. Total energy stands for the sum of one and two-center energy components and AIM energy corresponds to the sum of the AIM atomic energies $[\Sigma_A E(A)]$. Exact one-electron energy: $-194.963\ 3$; exact two-electron energy: $86.019\ 3$; exact total energy: -108.9439.

Phi planes	Theta planes	Points per Arc	Number of points	One-electron energy	Two-electron energy	Total energy	AIM Energy
16	24	24	28 192	−194.6835	85.9899	−108.6936	−109.2366
16	24	32	32 224	−194.9090	86.0020	−108.9070	−109.0078
32	24	24	44 352	−194.8616	86.1287	−108.7329	−109.2579
32	24	32	48 384	−195.0871	86.1409	−108.9462	−109.0291
64	48	48	153 344	−194.9633	85.9905	−108.9727	−108.9441
64	48	96	177 536	−194.9635	86.0296	−108.9339	−108.9439

$$2\sum_{i,j}^{N/2} \left[\varphi_i\varphi_j|\varphi_i\varphi_j\right]_{A,B}$$

$$=2\sum_{i,j=1}^{N/2} \int_{\Omega_A} d\nu_1 \int_{\Omega_B} d\nu_2\, \varphi_i^*(\vec{r}_1)\varphi_j^*(\vec{r}_2)\frac{1}{r_{12}}\varphi_i(\vec{r}_1)\varphi_j(\vec{r}_2)$$

$$\cong \sum_{\mu\in A}\sum_{\nu\in B} J^A(\mu)J^B(\nu)|\vec{r}_\mu-\vec{r}_\nu|^{-1}. \qquad (13)$$

All the above formulas also hold in the case when basins Ω_A and Ω_B coincide.

The numerical calculation of the two-electron integrals over the pairs of atomic basins according to Eqs. (8)–(13) has been performed by a small FORTRAN program written by us. It also handles the symmetry of the molecule in order to avoid repetitive calculation of identical quantities. This program has been used on different Linux machines in a single node regime and on a SGI Origin 2000 in a parallel way, with a very good scaling with the number of nodes. Moreover, since we are only using the set of data generated by the PROAIMV program, the timing of the two-electron energy contribution depends only on the size of the grid, but is independent of the basis set used.

As given by the PROAIMV program, the total number of points is generated from three input parameters, the number of planes φ, planes θ and the number of radial points. A calculation using some standard integration parameters (64,48,96) involves more than 140.000 points per atomic basin for the H_2 molecule. That means that to compute the two-electron integrals ca. $2\times140.000^2\approx4\cdot10^{10}$ points would appear in the numerical integration, which makes the two-electron integrals extremely costly from a computational point of view. Therefore, we are forced to use the smallest grid of points for each atomic basin which still gives an acceptable accuracy.

It has been found that, in general, about 40.000 points are needed to get a good accuracy. Table I gathers the results of the integration for the model N_2 molecule with respect to the number of points used for each atomic basin. It can be seen that the one-electron part converges faster to the exact value than the two-electron one. Our integration shows, in general, similar accuracy to the sum of the AIM atomic energies. However, when a relatively small grid is used, a better accuracy is obtained. This may be due to the fact that AIM atomic energies are computed assuming the fulfilment of the atomic virial theorem within the atomic domains, so that a more accurate numerical integration is desirable. The energy partitioning we propose is only based on the assumption that the whole space is partitioned in domains, and each domain must contain one nucleus.

IV. ILLUSTRATIVE CALCULATIONS

We have performed test calculations for several diatomic, i.e., H_2, N_2, BH, and HF, and more complex molecular systems such as C_2H_2, C_2H_4, C_2H_6, and B_2H_6. We have used the $6-31G(d,p)$ basis set whenever it was possible. It is to be mentioned that the AIM analysis sometimes yields basins with so-called non-nuclear attractors (NNA). These cases can be included in the above frame by assigning to the NNA a dummy nucleus with zero nuclear charge. However, it is not chemically very appealing that then some atomic and diatomic (bonding) energy will also be assigned to the dummy atoms. In such cases it may be worthwhile

TABLE II. One and two-center energy (a.u.) components for H_2, N_2, BH, and HF molecules using the $6-31G(d,p)$ basis set.

Molecule	ϵ_a	ϵ_{ab}	Total energy	AIM energy	Exact energy
H_2	−0.4565	−0.2178	−1.1308	−1.1306	−1.1313
N_2	−54.1203	−0.7056	−108.9462	−109.029	−108.9439
BH	−24.1501 −0.2803	−0.7079	−25.1406	−25.1200	−25.1182
HF	−0.2020 −99.3280	−0.4931	−100.0230	−100.0128	−100.0117

TABLE III. One and two-center energy (a.u.) components for ethane, ethene and acetylene.

Atom/s	Energy		
	Ethane[a]	Ethylene[a]	Acethylene[b]
C	−37.3567	−37.3767	−37.4304
H	−0.4330	−0.4339	−0.4298
C–C	−0.2625	−0.4567	−0.6061
C–H	−0.2608	−0.2586	−0.2346
H–H geminal	−0.0051	−0.0041	
H–H vicinal	−0.0003(×3)	−0.0004(×2)	0.0048
	0.0007(×6)	−0.0008(×2)	
C–H vicinal	−0.0088	−0.0073	−0.0057
Total energy	−79.2087	−77.9911	−76.8026
AIM energy	−79.2875	−78.0890	−76.8190
Exact energy	−79.2382	−78.0388	−76.7927

[a]$6-31G(d,p)$
[b]$6-31G$

TABLE V. Energy (a.u.) components for acetylene with non-nuclear attractor (X) in the middle of the C–C bond using the $6-31G(d,p)$ basis set. The $(C–C)_{eff}$ value is computed as $E(C–C)+2E(C–X)+E(X)$.

Atom/s	Energy
C	−37.1811
H	−0.4272
X	1.6988
C–C	0.5105
$(C–C)_{eff}$	−1.1468
C–H	−0.2027
C–X	−1.6778
H–H vicinal	0.0031
C–H vicinal	0.0314
H–X	−0.0786
Total energy	−76.8546
AIM energy	−76.8072
Exact energy	−76.8218

either to introduce another decomposition of the space or to check whether the appearance of the non-nuclear attractor is not an artifact of the basis set applied. For instance, for the acetylene, the $6-31G(d,p)$ basis set and many others like $6-311G(d,p)$ or even the $6-311G(2df,2p)$ and $6-311++G(3df,2pd)$ exhibited the undesirable non-nuclear attractors in the middle of the C–C bond. Hence, we have been forced to use the $6-31G$ for this specific case. Nevertheless, the results of the integration for the $6-31G(d,p)$ basis set using a dummy nucleus will be discussed in more detail later.

Also, the molecule of diborane presented integration problems with the $6-31G(d,p)$ basis set, due to a bad location of the bond critical points in the molecular plane. For this reason we have used the $6-31G(d)$ basis set instead. The number of planes φ, planes θ, and the number of radial points of the atomic grid has been set to 32, 24, and 32, respectively for all the calculations described in Tables II–V.

Tables II, III, and IV collect the results obtained for several molecules. The accuracy of the integration is very good and comparable to the AIM integration. The dominating energy components are the large negative atomic (one-center)

TABLE IV. One and two-center energy (a.u.) components for diborane using the $6-31G(d)$ basis set. Subscripts b and t in H atoms hold for bridge and terminal, respectively.

Atom/s	Energy
B	−23.2686
H_b	−0.2044
H_t	−0.3173
B–B	1.2891
B–H_b	−0.7461
B–H_t	−0.8289
H_b–H_b	0.1418
H_b–H_t	0.1191
H_t–H_t geminal	0.1165
cis H_t–H_t	0.0876
trans H_t–H_t	0.0735
Total energy	−52.8098
AIM energy	−52.8109
Exact energy	−52.8124

ones, while the diatomic (two-center) interaction energy components are smaller in absolute value—except for the H-atom energies in polar molecules such as HF or BH. The diatomic energy components are negative for chemically bonded atoms and can be of either sign for the nonbonded interactions. It is to be stressed that these energy components are *static parameters* corresponding to the given geometry and wave function, so the diatomic energy components are not directly related to the dissociation energies. The one-center components are somewhat higher than the free atomic energies, reflecting the promotion of atoms during the bond formation. (This energy is then regained in the form of bond energy.) Accordingly, the static diatomic energy components are more negative than the respective dissociation energies, as the latter give the *net* energy effect with respect to the sum of free atomic energies.

The changes in the two-center energy components in the hydrocarbon series are in good agreement with the chemical intuition. The C–C energy monotonically increases (in absolute value) from ethane to ethylene and acetylene, and the opposite trend is observed, in turn, for the C–H energy.

The accuracy of the energy expansion which can be practically achieved by performing the numerical integrations in the present—conceptually exact—method is approximately the same as that which one gets in the recent approximate LCAO energy decomposition scheme[3,4] (CECA). The results of the two schemes agree qualitatively but not quantitatively: the AIM bonding energy components are usually less negative, and the one-center components are not as negative as observed for the CECA method. This behavior seems to be closer to the intuitive chemical picture than that of CECA, as CECA gives numbers which may be considered somewhat exaggerated. (At the same time, CECA is a computationally very cheap method which can be routinely applied to large systems, too.)

The only surprising result corresponds to the large repulsive interaction between the two boron atoms in the diborane molecule. A previous result obtained with the CECA method[3] produced attractive interactions between the boron atoms (−0.227 a.u.) and also for the B–H interactions (−0.279 and −0.517 a.u. for the B–H bridge and B–H terminal, respec-

tively). In the present method, the B–H energies are much more negative (-0.746 and -0.829, respectively) and compensate for the repulsive B–B and H–H energy components.

Finally, the results for the acethylene molecule exhibiting a non-nuclear attractor in the middle of the C–C bond are collected in Table V. The energy of the dummy atom X corresponding to the NNA is positive. This is due to the fact that the kinetic energy is not compensated by any electron-nuclear interaction, because there is no nucleus to be assigned to the NNA. The C–C interaction is also repulsive whereas the interaction between the carbon atom and the dummy atom at the NNA (C–X), is strongly attractive due to the proximity of both basins. The one-center carbon energy components are less negative than the values obtained with the 6–31G basis set as shown in Table III. An effective C–C interaction could be computed by summing up the two-center C–C and C–X and the one-center X contributions. The computed $(C–C)_{eff}$ energy contribution is still more negative than the corresponding C–C value in Table III by more than 0.5 a.u. This difference can be assigned to the decrease of the C energy component mentioned above.

V. SUMMARY

The partitioning of the SCF energy in terms of one- and two-center interactions in the AIM framework has been carried out. The one-electron contributions have been calculated by using a slightly modified version of the PROAIMV program. The two-electron integration over the atomic basins has been computed with a small program written by us. These large scale numerical integrations have proved to be extremely costly from a computational point of view and large supercomputation resources have been necessary. The results obtained in the test calculations are in excellent agree-

ment with the chemical notion of molecules consisting of interacting atoms. Also, a generally good qualitative agreement has been observed with the recently introduced approximate LCAO energy decomposition scheme (CECA). Nevertheless, important differences have also been found for some molecular systems such as diborane. The presence in some cases of non-nuclear attractors destroys the chemical picture of the molecule and is, therefore, undesirable.

ACKNOWLEDGMENTS

[1] I. Mayer and A. Hamza, Theor. Chem. Acc. **105**, 374 (2001).
[2] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory* (Oxford University Press, Oxford, United Kingdom, 1990).
[3] I. Mayer, Chem. Phys. Lett. **332**, 381 (2000).
[4] I. Mayer and A. Hamza, Program APOST, Version 1.0, Budapest, April 2000 (available via the Internet at http://occam.chemres.hu/programs or by anonymous ftp at ftp://kvk.chemres.hu).
[5] A. Sierraalta and G. Frenking, Theor. Chim. Acta **95**, 1 (1997).
[6] F. W. Biegler-König, R. F. W. Bader, and T.-H. Tang, J. Comput. Chem. **3**, 317 (1982).