Information theory, atoms in molecules, and molecular similarity

Roman F. Nalewajski*† and Robert G. Parr†‡

*K. Guminski, Department of Theoretical Chemistry, Jagiellonian University, R. Ingardena 3, 30-060 Cracow, Poland; and †Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599-3290

Contributed by Robert G. Parr, May 15, 2000

Using information theory, it is argued that from among possible definitions of what an atom is when it is in a molecule, a particular one merits special attention. Namely, it is the atom defined by the ''stockholders partitioning'' of a molecule invented by Hirshfeld [(1977) *Theor. Chim. Acta* 44, 129]. The theoretical tool used is the minimum entropy deficiency principle (minimum missing information principle) of Kullback and Liebler [(1951) *Ann. Math. Stat.* 22, 79]. A corresponding analysis is given of the problem of assessing similarity between molecules or pieces of molecules.

Fundamental to chemistry is the understanding of molecules as combinations of atoms. It is not surprising, then, that the concept of atoms in molecules (AIM) has been much discussed in the literature (1). Chemistry mainly involves small changes among atoms and molecular fragments, with reasonably well-understood molecular invariants, e.g. AIM, functional groups, molecular subsystems, etc., which tend to maintain their identity. Most molecular systems may be thought of as consisting of slightly perturbed atoms (or atomic ions), possibly deformed by the presence of molecular remainders and exhibiting modified net charges arising from charge transfers and/or the formation of chemical bonds. These chemical atoms therefore are open subsystems.

The imposing variety of published theoretical methods for partitioning a molecular density into AIM contributions, e.g. refs. 1–15, testifies to the importance of this theme for chemistry. The different methods are based on different principles, some to a degree arbitrary (5) or heuristic (15), which can produce conflicting trends in the associated atomic net charges (effective oxidation states). Methods differ in the theoretical technique used, e.g., topological analysis of the density, wave-function description, or density-functional description. They also differ in the physical/heuristic principles invoked, e.g., electronegativity equalization, zero flux, and minimum-promotion energy rules. They can have specific disadvantages, e.g., basis set dependence. The well-known and appealing quantum-topological approach (1–4) suffers from the fact that its defined atomic densities are not ''v-representable.'' [An atomic density is v-representable when there exists an external potential that has this density as a ground-state density.]

Appropriate isolated atom densities constitute good references for defining properties of chemical atoms in terms of fast convergent Taylor series in the external potential and charge transfer displacements relative to the potentials and charges of the free atoms. The sum of the isolated atom/ion densities, with nuclear cusps at the actual positions of the nuclei, defines the density distribution of the promolecule (15). The promolecule is a key ingredient in the density difference analysis of the chemical bond done by Hirshfeld (15), in which the assumption is made that in forming a molecule each atom partakes of a local gain or loss in proportion to its local contribution to the promolecule density. In the present paper, we will recover this Hirshfeld ''stockholder partitioning'' of a molecular density into atomic components.

One would hope to find that a chemical atom, like its free analog, would possess a single cusp in its electron density, linked to the effective atomic number of the nucleus (16). AIM densities should be related to both promolecule and molecular densities, as representing the atomic fragments in a particular molecular system. One would want some degree of overlap between the densities of these chemical atoms, to reflect the presence of chemical bonds (1, 17, 18).

That atomic ground states, and/or small perturbations thereof, are uniquely appropriate reference states for a detailed description of AIM, was well understood by the pioneers, for instance Pauling and Mulliken, in their use of the concepts of hybridization, promotion, polarization, and ionic character. Ideally, an AIM definition will preserve as much information as possible about the separated atoms. [Here and later, the word atoms usually stands for atoms or ions.]

In defining AIM, how can one preserve, to the extent possible, the information content of ground-state atoms? It is natural to use some information—theoretic principle (19–26) for this purpose. And here is where density-functional theory (DFT) (27, 28) helps. For, DFT states that the electron density itself carries all of the information about a ground state. So, we may define AIM in a way that makes the atomic densities resemble as much as possible the isolated atom densities, and thereby achieve the ''best'' atoms we can have in a molecule in an information theoretic sense.

Kullback–Leibler Entropy Deficiency Functional

Suppose that $P_0(\vec{r})$ is a given (reference) well-behaved probability distribution and that $P(\vec{r})$ is some trial probability distribution, the information content of which we want to make as close as possible to that of P_0, subject to one or more constraints. Define (25) the entropy deficiency (missing information) functional by

$$\Delta S[P/P_0] = \int P(\vec{r}) 1n\left[\frac{P(\vec{r})}{P_0(\vec{r})}\right] d\vec{r}, \qquad [1]$$

and let there be constraints of the form

$$F_k[P] = F_k^0 \qquad k = 1, 2, \ldots, n. \qquad [2]$$

Then the P which is the ''best'' approximation to P_0 that satisfies the constraints is obtained by solving the minimum entropy deficiency (missing information) principle

$$\delta\left\{\Delta S[P/P_0] + \sum_k \lambda_k F_k[P]\right\} = 0\,, \qquad [3]$$

Abbreviation: AIM, atoms in molecules.

‡To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

where $\{\lambda_k\}$ are Lagrange multipliers. ΔS and the λ_k are found by solving Eqs. **2** and **3** simultaneously. A given Lagrange multiplier measures how sensitive the entropy deficiency is to the corresponding constraint. If there are several solutions for ΔS, one takes the minimum one.

This constrained entropy extremization (24) constitutes the most unbiased manner to assimilate (as completely as possible) the information included in the reference distribution and the auxiliary conditions. For specific P and P_0, Eq. **1** gives a non-negative number, which defines the information distance between the two distributions, or the entropy deficiency of P relative to P_0.

Example: To begin to approach the AIM problem, suppose that an accurate AB electron density is known, ρ, and that we want to represent it as the sum of two densities ρ_A and ρ_B, with electron numbers N_A^0 and N_B^0, both fixed, and that we want to keep ρ_A as close to ρ_A^0 and ρ_B as close to ρ_B^0 as possible. We take

$$\Delta S[\rho_A,\rho_B|\rho_A^0,\rho_B^0] = \int \rho_A 1n\left[\frac{\rho_A}{\rho_A^0}\right]d\vec{r} + \int \rho_B 1n\left[\frac{\rho_B}{\rho_B^0}\right]d\vec{r}, \quad [4]$$

and the constraints

$$N_A^0 = \int \rho_A d\vec{r} \equiv N[\rho_A]\,,\ N_B^0 = \int \rho_B d\vec{r} \equiv N[\rho_B]\,, \quad [5]$$

with the Lagrange multipliers λ_A and λ_B, respectively. Eqs. **2** and **3** then give

$$1n\left(\frac{\rho_A}{\rho_A^0}\right) = \text{constant} = 1;\ \rho_A = \rho_A^0\,,$$
$$1n\left(\frac{\rho_B}{\rho_B^0}\right) = \text{constant} = 1;\ \rho_B = \rho_B^0\,. \quad [6]$$

That is, if the "atoms" A and B are closed (not allowed to transfer electrons), entropy does not drive any density change. One will need to implicate electron transfer to define the atom in a molecule.

Derivation of the Hirshfeld AIM Stockholder Partitioning

An elementary modification of the example just given produces what we want, open atoms A and B, with electron numbers N_A and N_B not necessarily equal to N_A^0 and N_B^0 but still conforming with $N_A + N_B = N_A^0 + N_B^0 = N^0$. We can require exhaustive allocation of the whole electronic distribution in AB to either A or B,

$$\rho_A(\vec{r}) + \rho_B(\vec{r}) = \rho(\vec{r})\,. \quad [7]$$

This constraint requires using a Lagrange multiplier $\lambda(\vec{r})$ that is a function of $\vec{r}$. Here ρ denotes the accurate (presumed known) density of the AB system, which is normalized to N^0 electrons. Also normalized to N^0 is the isoelectronic promolecule density $\rho^0 = \rho_A^0 + \rho_B^0$. Implementing the minimum entropy deficiency principle of Eq. **3**, with the constraint of Eq. **7**, gives the equation for the AIM densities ρ_A and ρ_B,

$$\sum_{\propto = A,B}\left\{1n\left[\frac{\rho_\propto(\vec{r})}{\rho_\propto^0(\vec{r})}\right] - 1nD(\vec{r})\right\}\delta\rho_\propto(\vec{r}) = 0, \quad [8]$$

where $1n\ D(\vec{r}) = \lambda\ (\vec{r}) - 1$. Hence

$$\rho_\propto(\vec{r}) = \rho_\propto^0(\vec{r})D(\vec{r})\,, \qquad \propto = A,B. \quad [9]$$

The proportionality factor $D(\vec{r})$ is determined from the constraint of Eq. **7**:

$$D(\vec{r}) = \rho(\vec{r})/\rho^0(\vec{r})\,. \quad [10]$$

We therefore find

$$\rho_\propto(\vec{r}) = w_\propto(\vec{r})\rho(\vec{r}),\ w_\propto(\vec{r}) = \rho_\propto^0(\vec{r})/\rho^0(\vec{r}),\ \propto = A,B. \quad [11]$$

This is the "stockholder partition" of the electron density proposed long ago by Hirshfeld (7).

The sharing factor $w_\propto(\vec{r})$ determines the relative share of atom $\propto$ in the promolecule density $\rho^0(\vec{r})$. The result is trivially generalized to any number of constituent atoms. The "atoms" could be neutral atoms, ions, or functional groups.

The Kullback–Liebler entropy deficiency minimization of Eq. **8** is seen to provide a solid theoretical basis for the Hirshfeld prescription, which is known to yield fairly transferable charge distributions and moments (15,30), which can be used e.g. in calculations of the electrostatic potential and the interaction energy. The AIM fragments defined by Eq. **11** have continuous but well-localized densities. This partitioning procedure is basis set independent and it can be used for either theoretical or experimental densities.

The AIM densities of Eq. **11** possess correct behavior at the separated (isolated) atoms limit, where all internuclear distances $R_{\propto\beta} \to \infty$, because $D(\vec{r}) \to 1$ when $\rho(\vec{r}) \to \rho^0\ (\vec{r})$. The AIM densities fulfill, quite well through not perfectly, the proper nuclear cusp and long-range decay conditions. These densities are most probably v-representable or at least ensemble v-representable. [Recent unpublished work by Paul Ayers on the v-representability problem implies that Hirshfeld atomic densities are either v-representable or infinitely close to v-representable.]

The bonding factor $D(\vec{r})$ of Eq. **10** indicates how the free atom density has been modified in the chemical atom. In the usual case where there is density accumulation in the molecular bonding region and depletion in the nonbonding regions of atoms bonded in a molecule, relative to ρ^0, bonded atoms are polarized toward their bonding partners.

The chemical potentials of the bonded atoms each must be equal to the molecular chemical potential μ (8, 9, 31). The argument follows. For the fixed molecular external potential v we have:

$$E = E(N) = E(N_A + N_B),\ N = N_A + N_B, \quad [12]$$

so that

$$dE = \frac{\partial E}{\partial N}dN \equiv \mu dN \quad [13]$$

$$= \left(\frac{\partial E}{\partial N_A}\right)_{N_B} dN_A + \left(\frac{\partial E}{\partial N_B}\right)_{N_A} dN_B \quad [14]$$

$$\mu dN = \mu dN_A + \mu dN_B \quad [15]$$

Subtracting Eq. **15** from Eq. **14** gives:

$$0 = \left[\mu - \left(\frac{\partial E}{\partial N_A}\right)_{N_B}\right]dN_A + \left[\mu - \left(\frac{\partial E}{\partial N_B}\right)_{N_A}\right]dN_B \quad [16]$$

Therefore, if we define

$$\mu_A = (\partial E/\partial N_A)_{N_B} \text{ and } \mu_B = (\partial E/\partial N_B)_{N_A}\,, \quad [17]$$

we find:

$$\mu = \mu_A = \mu_B\,. \quad [18]$$

AIM from a Modified Entropy Functional

To illustrate the many possibilities for extensions and variations of this analysis, we examine another entropy deficiency functional, in which in addition we include the information distance between $\rho(\vec{r})$ and the resultant molecular density $\bar{\rho}_A(\vec{r}) + \bar{\rho}_B(\vec{r}) \equiv \bar{\rho}(\vec{r})$, which is not fixed but is to be determined. The modified Eq. **4** becomes

$$\Delta S[\bar{\rho}_A,\bar{\rho}_B|\rho_A^0,\rho_B^0,\rho] = \Delta S[\bar{\rho}_A,\bar{\rho}_B|\rho_A^0,\rho_B^0] + \int \bar{\rho}(\vec{r}) 1n\left[\frac{\bar{\rho}(\vec{r})}{\rho(\vec{r})}\right] d\vec{r}. \quad [19]$$

We minimize this subject to the constraint on the fixed overall number of electrons,

$N[\bar{\rho}] = N^0 = N_A^0 + N_B^0$:

$$\delta\left\{\Delta S[\bar{\rho}_A\bar{\rho}_B|\rho_A^0,\rho_B^0,\rho] - \lambda \int [\bar{\rho}_A + \bar{\rho}_B] d\vec{r}\right\} = 0, \quad [20]$$

where λ is a global Lagrange multiplier. Setting $1nc = \lambda - 2$ we obtain the optimum AIM densities,

$$\bar{\rho}_\propto = \sqrt{c}\rho_\propto^0(\rho/\rho^0)^{1/2} \equiv \sqrt{c}\rho_\propto^0\bar{D};\ \propto = A,B, \quad [21]$$

and the optimum overall density

$$\bar{\rho} = \sqrt{c}[\rho\rho^0]^{1/2}. \quad [22]$$

The proportionality constant can be determined from the global constraint in Eq. **20**:

$$\sqrt{c} = N^0/\int [\rho\rho^0]^{1/2} d\vec{r} \simeq 1. \quad [23]$$

Thus, the variational principle of Eq. **20** gives the optimum molecular density $\bar{\rho}(\vec{r})$, as the geometric mean of the promolecule and the true ground-state density of the molecule.

Comparison of Eq. **21** with Eqs. **9** and **10** indicates that this modified entropy deficiency minimization also gives chemical atoms, which are properly polarized toward the bonding region of a covalent bond. However, this effect is weaker in Eq. **21** because $\bar{D}(\vec{r}) = [D(\vec{r})]^{1/2}$, where $D(\vec{r})$ (Eq. **10**) stands for the bonding factor of the Hirshfeld stockholder atoms. Nevertheless, it is reassuring that the last "similarity" term in Eq. **19**, which drives $\bar{\rho}$ toward ρ, is seen to modify the promolecule density ρ^0 in the qualitatively correct direction, toward ρ, thus generating a partial bonding in a molecule. The density $\bar{\rho}$ is a transition density between ρ^0 and ρ.

Both the Hirshfeld atoms of Eq. **11** and the modified chemical atoms of Eq. **21** represent open atomic fragments in a molecule, with the effective charges

$$q_\propto = Z_\propto - N[\rho_\propto] \text{ or } \bar{q}_\propto = Z_\propto - N[\bar{\rho}_\propto] \quad [24]$$

generally different from the free atom (ion) charge, $q_\propto^0 = Z_\propto - N[\rho_\propto^0]$, caused by the charge transfer component of the chemical bond. Equal chemical potentials do not generally obtain in this case, however.

Information-Theoretic Measure of Molecular Similarity

A molecular similarity concept sometimes is invoked to characterize structural resemblance of different molecules or their fragments. Measures that have been used have included overlap integrals for electronic density, electrostatic potential, and Fukui function (32). Here, we assume that similarity of electronic structure implies a closeness of the information content of electronic density distributions. This again calls for a measure of the information distance as provided by the entropy deficiency functional of Eq. **1**.

In a typical screening search through candidate species with densities $\{\rho_i\}$, which are being tested for similar chemical activity to the reference system with density ρ_0, one would select those with the minimum value of the entropy deficiency $\Delta S[\rho_i|\rho_0]$. When reactivity of a specific "active site" is required, the densities of the relevant molecular fragments should be chosen as the input information carriers.

The simplest variational criterion for the minimum entropy deficiency between the test molecular/fragment density ρ_i and the reference density ρ_0 normalized to $N[\rho_0] = N_0$, subject to the normalization constraint $N[\rho_i] = N_i^0$, would be

$$\delta\{\Delta S[\rho_i|\rho_0] - \lambda N[\rho_i]\} = 0, \quad [25]$$

giving

$$P_i(\vec{r}) \equiv \frac{\rho_i(\vec{r})}{N_i^0} = P_0(\vec{r}) \equiv \frac{\rho_0(\vec{r})}{N_0}. \quad [26]$$

The most favorable matching therefore is achieved when the shape functions $P_i(\vec{r})$ and $P_0(\vec{r})$ are as close as possible. This result is reminiscent of the recent extension of the Hohenberg–Kohn theorems (26) by Ayers (33), demonstrating that the shape function of the bound states of a molecule determines all properties of the system under consideration. Requiring the maximum amount of common information in ρ_i and ρ_0 indeed calls for the same shape functions for the two compared densities.

As a final example, consider the optimum matching criterion of a trial electron density ρ, with two reference densities $\rho_1^0(N[\rho_1^0] = N_1^0)$ and $\rho_2^0(N[\rho_2^0] = N_2^0)$, subject to the usual normalization constraint, $N[\rho] = N^0$:

$$\delta\{\Delta S[\rho|\rho_1^0] + \Delta S[\rho|\rho_2^0] - \lambda N[\rho]\} = 0. \quad [27]$$

Solution of this equation gives

$$\rho = C(\rho_1^0\rho_2^0)^{1/2},\ C = N^0/N. \quad [28]$$

The optimum match is the normalized geometric mean.

One encounters such a problem in heterogeneous catalysis, when matching the adsorbate density (ρ) with the two surface active sites (ρ_1^0, ρ_2^0) involved in bonding the adsorbate. Another example is a molecule binding to two sites of another molecule. The maximum value of $\rho(\vec{r})$ is expected in the overlapping region between ρ_1^0 and ρ_2^0, because the least biased information distance criterion of Eq. **27** calls for ρ to resemble (appreciably overlap with) the overlap density between the two reference densities.

Summary and Concluding Remarks

In this paper we have used the information theory to decompose a molecular electron density into its component atomic densities. More particularly, we have selected the entropy deficiency (informative distance) approach of Kullback and Leibler (25), using the free atom densities as the reference distributions, to obtain AIM densities. Such a reference is both natural and unique in chemistry, because by the Hohenberg–Kohn theorem the isolated atom densities carry all the information of the periodic table, which is the fountainhead of chemical thought.

We have demonstrated that the minimum entropy deficiency (missing information) principle subject to the constraint of the exhaustive partitioning of the molecular density recovers the "stockholder" definition of bonded atoms done by Hirshfeld (15), thus providing it with a fundamental theoretical derivation.

These Hirshfeld chemical atoms are unbiased pieces of the molecular density, which are the least distant in their information content from their isolated atom analogs. The chemical atoms have equalized chemical potentials at the global, molecular level, and they are probably v-representable. Their overlap in a molecule accords with the familiar classical interpretation of the origin of the chemical bond (17). The quantum-topological atoms, in contrast, have zero overlap (1).

In an extension, we have demonstrated how the entropy deficiency functional can be used to generate the bonding character of chemical atoms. The resulting optimum density of a molecule then is given by the geometric mean of the true ground-state density and that of the promolecule (having free atom densities centered at the actual positions of nuclei in a molecule). Both the intermediate polarization (promoted) and final charge transfer stages of the atomic density reorganization in a molecule have been discussed. We also have illustrated the use of the entropy deficiency minimization principle in molecular similarity problems.

Finally, we note that the same ideas apply to excited electronic states. In this case, however, the reference, ground-state densities of free atoms might not be the ones that give the minimum value of the informative distance. Using the excited-state densities of separated atoms as reference densities may give more realistic excited-state (promoted) AIM densities, while still linking the resulting chemical atoms to the isolated atoms (periodic table) information. The information-theoretic approach should be helpful in quantitative studies of charge transfer and effective oxidation states.

We have had helpful discussions with several people, most importantly Paul Ayers, Richard Bader, Peter Politzer, and Vedene Smith. This research has been supported by a grant from the Committee for Scientific Research in Poland and by a grant from the Petroleum Research Fund of the American Chemical Society.

1. Bader, R. F. W. (1970) *An Introduction to the Electronic Structure of Atoms and Molecules* (Clarke, Toronto).
2. Bader, R. F. W. (1994) *Atoms in Molecules* (Oxford, New York).
3. Bader, R. F. & Nguen-Dang, T. F. (1981) *Adv. Quant. Chem.* **14,** 63–124.
4. Bader, R. F. W. & Becker, P. (1998) *Chem. Phys. Lett.* **148,** 452–458.
5. Moffitt, W. (1951) *Proc. R. Soc. London Ser. A* **210,** 245–268.
6. Mulliken, R. S. (1935) *J. Chem. Phys.* **3,** 573–585.
7. Mulliken, R. S. (1955) *J. Chem. Phys.* **23,** 1833–1840.
8. Parr, R. G., Donnelly, R. A., Levy, M. & Palke, W. E. (1978) *J. Chem. Phys.* **68,** 3801–3807.
9. Parr, R. G. (1984) *Int. J. Quantum Chem.* **26,** 687–692.
10. Rychlewski, J. & Parr, R. G. (1986) *J. Chem. Phys.* **84,** 1696–1703.
11. Li, L. & Parr, R. G. (1986) *J. Chem. Phys.* **84,** 1704–1711.
12. Bader, R. F. W. (1986) *J. Chem. Phys.* **85,** 3133.
13. Parr, R. G. (1986) *J. Chem. Phys.* **85,** 3135.
14. Cedillo, A., Chattaraj, P. K. & Parr, R. G. (2000) *Int. J. Quantum Chem.* **71,** 403–407.
15. Hirshfeld, F. L. (1977) *Theo. Chim. Acta* **44,** 129–138.
16. Kato, T. (1957) *Commun. Pure Appl. Math.* **10,** 151–171.
17. Ruedenberg, K. (1962) *Rev. Mod. Phys.* **34,** 326–376.
18. Feinberg, M. J. & Ruedenberg, K. (1971) *J. Chem. Phys.* **54,** 1495–1511.
19. Shannon, C. F. (1948) *Bell System Tech. J.* **27,** 379–493, 623–656.
20. Kullback, S. (1959) *Information Theory and Statistics* (Wiley, New York).
21. Mathai, A. M. & Rathie, P. N. (1975) *Basic Concepts in Information Theory and Statistics* (Wiley, New York).
22. Sears, S. B. (1980) Ph.D. thesis (University of North Carolina, Chapel Hill).
23. Sears, S. B., Parr, R. G. & Dinur, U. (1980) *Isr. J. Chem.* **19,** 165–173.
24. Jaynes, E. T. (1957) *Phys. Rev.* **106,** 620–630.
25. Kullback, K. & Leibler, R. A. (1951) *Ann. Math. Stat.* **22,** 79–86.
26. Esquivel, R. O., Rodriquez, A. L., Sagar, R. P., Hō, M. & Smith, V. H., Jr. (1996) *Phys. Rev. A* **54,** 259–265.
27. Hohenberg, P. & W. Kohn, W. (1959) *Phys. Rev.* **136,** B864–B871.
28. Parr, R. G. & Yang, W. (1989) *Density Functional Theory of Atoms and Molecules* (Oxford, New York).
29. Hō, M., Sagar, R. B., Schmider, H., Weaver, D. F. & Smith, V. H., Jr. (1995) *Int. J. Quantum Chem.* **53,** 627–633.
30. Brook, C. P., Dunitz, J. P. & Hirshfeld, F. L. (1991) *Acta Crystallogr. B* **47,** 789–797.
31. Sanderson, R. T. (1951) *Science* **114,** 670–672.
32. Boon, G., De Proft, F., Langenaeker, W. & Geerliings, P. (1998) *Chem. Phys. Lett.* **295,** 122–128.
33. Ayers, P. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 1959–1964.