Atoms in Molecules and Population Analysis

2 authors:

Paul L A Popelier
The University of Manchester
273 PUBLICATIONS **10,927** CITATIONS

Patrick Bultinck
Ghent University
237 PUBLICATIONS **5,782** CITATIONS

Some of the authors of this publication are also working on these related projects:

Electronegativity Equalization Method View project

Raman optical activity of proteins and IDPs View project

15 Atoms in Molecules and Population Analysis

Patrick Bultinck and Paul Popelier

CONTENTS

15.1   INTRODUCTION

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known ...". This assertion stated in 1929 by P. A. M. Dirac [1] sets a clear goal for theoretical chemists. Put simply, every chemist, and the theoretical chemists probably first, should strive to frame chemistry within the context of quantum mechanics. Unfortunately, after nearly 80 years, we are far from having achieved this aspiration. Chemistry is still described by a wealth of concepts that are neither rooted in quantum mechanics nor directly derived from it [2]. Two reasons immediately come to mind. Firstly, chemistry has a long history whereas quantum mechanics is a relatively young discipline. Recasting an older science with its own traditional views within the context of a newer one is hard and often meets opposition. Secondly, in the distant past, chemists had already developed their own parlance, which continues to suit their needs very well till now. A good example is the chemical bond. The chemical bond was already mentioned by Frankland in 1866 [3], long before the Heitler–London wave function for H_2 [4] was derived, and thus much before one could put it in a quantum mechanical framework.

It would have been easier to put vague and intuitive concepts on a firmer quantum mechanical footing if they had not been so successful. Indeed, one can only admire the amount of interpretation and prediction accomplished by chemists using their toolbox of concepts. Some of these "toolbox users" may be distrustful of theoretical chemists trying to fix something that, according to these users, is not

broken [5]. On the other hand, chemistry does need the input of quantum mechanics for a different reason. Chemists want to fundamentally understand why a chemical bond between the two same atoms is so similar in different molecules and why molecules in a homologous series behave in such a similar way, for example.

Current computational resources enable highly accurate predictions of molecular structures and properties. As a result, the theoretical chemist is appreciated within a broader chemical community as being able to contribute to the ever growing field of chemistry. This mutual understanding is partially due to all involved speaking the same (often conceptual) chemical language. Suddenly abandoning this language may quickly turn the theoretical chemist into the outsider he once was in the chemical community. But every theoretician should remember Dirac's remark mentioned above and speak the language of quantum mechanics.

The debate on several chemical concepts between those following a more ''intuitive'' path and those following a ''physically rigorous'' path remains lively. The present chapter deals exactly with such a concept: the atom in the molecule (AIM). Some consider it a product of the mind, a noumenon [6], others accept only a strict quantum mechanical definition. The dust does not seem to have settled yet as far as this argument is concerned. In order to give the reader an idea of the discussions arising from confronting different AIM methods, emphasis is put on describing different AIM techniques. Some thoughts on the deeper roots of the AIM will be shared at the end. This chapter will also introduce some key ideas on population analysis. Nevertheless, we limit the coverage of population analyses because the concept of an AIM is wider than a mere atomic charge.

15.2 ATOMS IN MOLECULES VERSUS POPULATION ANALYSIS

Although the title of this chapter reads Atoms in Molecules and Population Analysis, it should be clear from the beginning that the two topics in the title need to be differentiated. A population analysis is a computational technique to obtain atomic charges. An intuitively plausible population analysis should quantitatively recover the qualitative consequences of electronegativity, where more electronegative atoms ''draw'' more electrons to themselves than less electronegative ones. Any sound approach that yields a definition of an AIM should allow obtaining an AIM population and thus an AIM charge. The reverse is not always possible. To show this we discuss the ubiquitous electrostatic potential (ESP)-derived charges [7–10].

The electron density $\rho(\mathbf{r})$ and the molecular M-nuclear framework $\{Z_A, \mathbf{R}_A\}$ together generate an EPS $V(\mathbf{r}_i)$ in all points of space $\mathbf{r}_i$:

$$V(\mathbf{r}_i) = \sum_A^M \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} - \int \frac{\rho(\mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|} \, d\mathbf{r} \qquad (15.1)$$

There is much interest in the ESP because it plays a major role in molecular recognition. Repeated evaluation of the ESP is very costly, especially for large molecules during simulations. Therefore finding an alternative way to compute ab initio quality ESPs at much lower cost is important. Methods that simplify

Equation 15.1 have been derived. The key idea is to replace the continuous density function and the nuclear framework with the following much simpler expression (for a review, see Ref. [9]):

$$V'(\mathbf{r}_i) = \sum_A^M \frac{Q_A}{|\mathbf{r}_i - \mathbf{R}_A|} \tag{15.2}$$

The introduction of point charges in Equation 15.2, usually placed at nuclear positions, is a popular simplification of Equation 15.1. Naturally these charges need to be fitted in order to attain the best possible agreement between $V(\mathbf{r}_i)$ and $V'(\mathbf{r}_i)$. The atomic charges Q_A are optimized under the constraint of adding up to the molecular charge, and often also under the constraint of reproducing some molecular multipole moments. The input for the optimization consists usually of a (very) large set of data points $\{\mathbf{r}_i, V(\mathbf{r}_i)\}$, where the $V(\mathbf{r}_i)$ values were computed from Equation 15.1. The points $\{\mathbf{r}_i\}$ are chosen on some predetermined molecular surface so that the procedure effectively reproduces the ESP in regions around the molecule where interactions are assumed to take place. This simple description of ESP-derived charges shows how a population analysis is performed without any notion of an AIM. In fact, one can easily think of adding more sites to position charges, for instance, in the middle of each bond. Nevertheless, point charges continue to dominate applications both in biochemistry and material science in spite of their well-documented deficiencies. The literature on determination of point charges is relatively large and highlights a surprising lack of mathematical understanding of the fitting procedures used to determine them. A thorough and informed analysis of this problem [11] demonstrated that the fitting procedure is underdetermined. In other words, there are too many point charges (i.e., degrees of freedom) to reproduce the exact ESP V as best as possible. This leads to point charge values being erratically assigned and further reduction of their chemical meaning. Overall, it is more desirable to model intermolecular interactions by means of multipole moments [12–14].

15.3 BASIC REQUIREMENTS OF AN AIM METHOD

If population analysis is not synonymous with the concept of an AIM, it becomes necessary to introduce a proper set of requirements before one can speak of an AIM. An AIM is a quantum object and as such has an electron density of its own. This atomic electron density must obviously be positive definite and the sum of these atomic densities must equal the molecular density. Each atomic density $\rho_A(\mathbf{r})$ can be obtained from the molecular density $\rho(\mathbf{r})$ in the following way:

$$\rho_A(\mathbf{r}) = \hat{w}_A \rho(\mathbf{r}) \tag{15.3}$$

where $\hat{w}_A$ is a positive definite operator. As the atomic densities sum to the molecular density,

$$\sum_A \rho_A(\mathbf{r}) = \rho(\mathbf{r}) \tag{15.4}$$

it follows that

$$\sum_A \hat{w}_A = \mathbf{1} \tag{15.5}$$

All the different AIM methods that will be discussed below basically use this same approach but quite different in the nature of $\hat{w}_A$. Chronologically, we will discuss the Mulliken AIM, the Hirshfeld AIM, and the Bader AIM. This last approach will henceforth be called quantum chemical topology (QCT)*. There are more AIM methods, but most of them can be easily understood by the three selected emblematic approaches.

15.4 MULLIKEN APPROACH

Before introducing the form of the Mulliken operator, $\hat{w}_A^{\text{Mull}}$, it is appropriate to return to the concepts of early days of quantum mechanics. Heitler and London wrote down the singlet wave function for H_2 in terms of the hydrogen 1s atomic orbitals on both hydrogen atoms A and B [4]:

$$\Psi = N(1s_A(1)1s_B(2) + 1s_B(1)1s_A(2))(\alpha(1)\beta(2) - \alpha(2)\beta(1)) \tag{15.6}$$

From this wave function, one sees how even in the early beginning of molecular quantum mechanics, atomic orbitals were used to construct molecular wave functions. This explains why one of the first AIM definitions relied on atomic orbitals. Nowadays, molecular ab initio calculations are usually carried out using basis sets consisting of basis functions that mimic atomic orbitals. Expanding the electron density in the set of natural orbitals and introducing the basis function expansion leads to [15]

$$\rho(\mathbf{r}) = \sum_{\nu\mu} D_{\nu\mu} |\nu\rangle\langle\mu| \tag{15.7}$$

---

* In the literature of the 1980s and 1990s, the acronym AIM was uniquely used to refer to the ''quantum theory of atoms in molecules'' pioneered by the Bader group. To differentiate this approach from others the acronym, QTAIM was introduced later. Extensive work by the Bader group provided QTAIM with a rigorous quantum mechanical foundation by proving that the topological condition of zero-flux serves as the boundary condition for the application of Schwinger's principle of stationary action in the definition of an open system. One of us has proposed the name quantum chemical topology (QCT) to better capture the essential and unique features of QTAIM. Secondly, QCT facilitates future developments, generalizations, and applications. The acronym AIM (or QTAIM) is actually too narrow because, strictly speaking, it only makes sense as a term if one analyses the electron density topologically. Only then does one recover an ''atom in a molecule.'' A topological analysis of the Laplacian of the electron density (which is part of QTAIM) or the topology of the electron localization function (ELF), for example, does not yield ''atoms in molecules.'' However, they can both be put under the umbrella of QCT since they share the central topological idea of partitioning space by means of a gradient vector field. Also, returning to the electron density, one could use the topological analysis to recover molecules in van der Waals complexes or condensed matter. Again, as a name, AIM does not describe this result.

Atoms in Molecules and Population Analysis   **219**

where we use Dirac's bra-ket notation for the basis functions $|\nu\rangle$ and $\langle\mu|$. The symbol $D_{\nu\mu}$ denotes an element of the charge and bond order matrix [15]. Greek letters refer to the basis functions used in the calculation. In order to establish the Mulliken AIM, the required operator can be written as [16]

$$\hat{w}_A^{\text{Mull}} = \sum_{\sigma \in A} \sum_{\lambda} S_{\sigma\lambda}^{(-1)} |\sigma\rangle\langle\lambda| \tag{15.8}$$

where $S_{\sigma\lambda}^{(-1)}$ is the element $\sigma\lambda$ of the inverse of the overlap matrix. Letting this operator act on Equation 15.7 one finds that

$$\begin{aligned}
\rho_A^{\text{Mull}}(\mathbf{r}) &= \hat{w}_A^{\text{Mull}}\rho(\mathbf{r}) \\
&= \sum_{\nu\mu} \sum_{\sigma \in A} \sum_{\lambda} D_{\nu\mu} S_{\sigma\lambda}^{(-1)} |\sigma\rangle\langle\lambda|\nu\rangle\langle\mu| \\
&= \sum_{\nu\mu} \sum_{\sigma \in A} \sum_{\lambda} D_{\nu\mu} S_{\sigma\lambda}^{(-1)} S_{\lambda\nu} |\sigma\rangle\langle\mu| \\
&= \sum_{\nu\mu} \sum_{\sigma \in A} D_{\nu\mu} \delta_{\sigma\nu} |\sigma\rangle\langle\mu| \\
&= \sum_{\sigma \in A} \sum_{\mu} D_{\sigma\mu} |\sigma\rangle\langle\mu|
\end{aligned} \tag{15.9}$$

The well-known expression embodying the Mulliken population analysis [17–20] then follows after integration over all space,

$$\int \rho_A^{\text{Mull}}(\mathbf{r}) d\mathbf{r} = \sum_{\sigma \in A} (\mathbf{DS})_{\sigma\sigma} \tag{15.10}$$

The operator $\hat{w}_A^{\text{Mull}}$ has an interesting characteristic, namely that

$$\begin{aligned}
\hat{w}_A^{\text{Mull}} \hat{w}_B^{\text{Mull}} &= \sum_{\sigma \in A} \sum_{\lambda} \sum_{\nu \in B} \sum_{\mu} S_{\sigma\lambda}^{(-1)} S_{\lambda\nu} S_{\nu\mu}^{(-1)} |\sigma\rangle\langle\mu| \\
&= \sum_{\sigma \in A} \sum_{\nu \in B} \sum_{\mu} \delta_{\sigma\nu} S_{\nu\mu}^{(-1)} |\sigma\rangle\langle\mu| \\
&= 0 \qquad \text{if } A \neq B \\
&= \hat{w}_A^{\text{Mull}} \quad \text{if } A = B
\end{aligned} \tag{15.11}$$

This means that the operators are mutually exclusive and that the operator is idempotent. Nevertheless, in three-dimensional Cartesian space the atoms do overlap, often even to a large extent. So they have no boundaries.

It is clear that the Mulliken operator works in the Hilbert space of the basis functions, which has repercussions on the way the electron density is assigned to the nuclei. The basis functions are allocated to the atomic nuclei they are centered on; the decision as to which portion of the electron density belongs to which nucleus rests on

the centering itself. The decision is encapsulated in expressions such as "$\sigma \in A$" appearing in Equations 15.8 through 15.11. This allocation is a direct consequence of an approach that invokes atom-centered basis functions. Such an allocation scheme will fail when basis functions centered on other locations were included in the basis set, or if the basis functions have no center, as in plane waves. The Mulliken approach is thus intertwined with linear combination of atomic orbitals (LCAO) theory. Secondly, if diffuse basis functions are included in the basis set, the Mulliken population analysis may become "unstable." Large fluctuations in the atomic charges may appear upon changing the number of basis functions in the basis set. Diffuse functions decline slowly while moving away from the nucleus. Hence they contribute to a relatively large part of the electron density that is remote from this nucleus. However, the contribution of the diffuse functions to the electron density is still allocated to the nucleus. This leads to spurious results [21].

15.5   HIRSHFELD APPROACH

This approach, also often called the stockholder scheme, was introduced in 1977 by Hirshfeld [22]. The central idea of the Hirshfeld method originates in x-ray crystallography. It proposes to divide the electron density among the atoms in a molecule, guided by a promolecular density. More precisely, once a molecular geometry is known, a promolecular density $\rho^0(\mathbf{r})$ is composed by simply summing the density of each atom A (denoted $\rho_A^0(\mathbf{r})$) in an isolated state:

$$\rho^0(\mathbf{r}) = \sum_A \rho_A^0(\mathbf{r}) \tag{15.12}$$

where the sum runs over all constituent atoms. The idea of using a promolecule to distinguish AIM was not new in 1977 since it had been proposed earlier by Daudel and coworkers [23–25].

At each point in space, the share of the atom is calculated as

$$w_A^H(\mathbf{r}) = \frac{\rho_A^0(\mathbf{r})}{\sum_A \rho_A^0(\mathbf{r})} \tag{15.13}$$

This share is used as the Hirshfeld operator. The assumption behind the Hirshfeld AIM is that this same weight operator can be used to divide the electron density of the molecule via

$$\rho_A(\mathbf{r}) = w_A^H(\mathbf{r})\rho(\mathbf{r}) = \frac{\rho_A^0(\mathbf{r})}{\sum_A \rho_A^0(\mathbf{r})}\rho(\mathbf{r}) \tag{15.14}$$

This so-called Hirshfeld scheme is particularly popular within the so-called conceptual density functional theory (DFT) [26,27]. The weighting function, which identifies the AIM as one that is most similar to the isolated atom [28], has been shown to be directly derivable from information entropy [6,29–33]. Here again, the atoms do not

Atoms in Molecules and Population Analysis   **221**

have sharp boundaries and extend to infinity although a relatively sharp drop in the weight of an atom may be expected in the neighborhood of other atoms. As such, it is also reminiscent of other techniques such as the fuzzy atoms of Mayer and Salvador [34] or Becke's scheme [35].

Hirshfeld atomic populations were found to be mildly dependent on the basis set [36,37], thus remedying one of the important problems of the Mulliken technique. However, new problems are introduced. First, a promolecular density is a nonphysical concept since it does not comply with the Pauli principle. Several other questionable aspects of the Hirshfeld technique have been revealed as well. Firstly, the promolecule is not defined uniquely. Davidson and Chakravorty [38], Bader and Matta [39], Matta and Bader [40], and Bultinck et al. [41] have questioned the apparently conventional choice of neutral ground state isolated atom densities as building blocks for the promolecular density. A different promolecular density results in a different set of atomic charges. Davidson and Chakravorty [38] showed this for the case of N_2, where the atomic charges differed significantly whether one took it as a promolecule N^+ and N^- or as two neutral nitrogen atoms. This is clearly undesirable and Bader and Matta [39] and Matta and Bader [40] raised the question as to what would be the proper Li density to use in the promolecule for LiF: Li^+ or neutral Li? Bultinck and coworkers [37,41,42] investigated this matter and found that using $Li^0 + F^0$ as promolecule resulted in charges of $+0.57$ and -0.57, respectively. Using the combination, $Li^+ + F^-$ as promolecule yielded a charge of ± 0.98, and a charge of ± 0.30 was found when using $Li^- + F^+$ as promolecule. The only constant is that Li is assigned the positive charge. The reported charge fluctuations are dramatic, which corroborates the criticism by Bader and Matta [39] and Matta and Bader [40]. A second concern is that, for covalently bonded systems the Hirshfeld atomic charges are virtually zero. For instance, in the case of H_2O the atomic charges are 0.16 (H) and -0.32 (O). Compared with any other type of population analysis, these values are rather small. Even for cases where one expects significant contributions from charge transfer between atoms, the charges remain very small. Relying on the proof by Ayers that the Hirshfeld, AIM is the one that keeps the AIM as similar as possible to the reference isolated atom [28], this may indicate that indeed the AIM also wants to keep its charge as close as possible to that of the reference atom, i.e., zero. This problem becomes even larger for charged systems.

A third concern follows from the information theory background provided by Parr, Nalewajski, and coworkers [6,29–32]. In order to use the Kullback–Liebler formula for missing information for the AIM, an essential criterion is that the AIM density and the density of the isolated atom used in the promolecule must normalize to the same number [6]. This is almost never the case when using neutral atoms as a reference. All these concerns led Bultinck and coworkers [37,41,42] to develop the iterative Hirshfeld method, denoted Hirshfeld-I. This method, which actually coincides with the suggestion of Davidson and Chakravorty [38] to find a self-consistent Hirshfeld method, proceeds in the following way. First, a regular Hirshfeld population analysis is carried out, resulting in a first set of atomic populations. Then a new promolecule is constructed using atomic densities that normalize to the populations produced in the previous step. Using that promolecule, the Hirshfeld analysis is again carried out. This procedure is repeated until eventually the populations that

result from the ith iteration are the same as those that were used in the promolecule in the same iteration. Once self-consistency is obtained, one can proceed with the extraction of Hirshfeld-I AIMs and use them for further analysis.

The results [41,42] of implementing this scheme reveal that the atomic charges tend to grow much bigger, roughly by a factor of 3. Even charges below -1 or above $+1$ do appear now. In many studies, Bader's AIM charges (detailed explanation in the following sections) were criticized for being too large. However, the charges of the self-consistent Hirshfeld scheme seem to grow quite large as well. Secondly, the Hirshfeld-I scheme also makes charged molecules tractable without using promolecular densities that normalize differently from the molecular density. Finally, the Kullback–Liebler formula is strictly valid in assessing information loss [43]. Also, the basis set dependence has been shown to be very small [37]. Returning to LiF, the Hirshfeld-I procedure yields AIM charges of ± 0.932, which agree well with those of an ionic bond. LiH is another molecule for which Hirshfeld and similar approaches are problematic according to Bader and coworkers [39,40]. The regular Hirshfeld method yields charges of ± 0.43, which are indeed very low for this kind of species. On the other hand, using the Hirshfeld-I scheme, the charges become ± 0.930. As is the case for the Bader method [39], the Hirshfeld-I charges of LiH and LiF have become very similar.

Introducing self-consistency in the Hirshfeld-I scheme removes most of the arbitrariness in choosing a promolecule. It was also proven that the corresponding AIM populations [37] are independent from the starting point of the iterative process. Still, one arbitrary decision remains. The states of the isolated atoms used in constructing the promolecule are chosen arbitrarily as the ground state. In principle, this problem can also be solved in the spirit of information theory. One could initiate the Hirshfeld-I procedure for every possible combination of states of every atom, carry it out until convergence, and then compute the information loss. Routine application of this procedure is of course impossible. On the other hand, Rousseau et al. [45] showed that, in the regular Hirshfeld scheme, changing the states of the neutral reference atoms has little impact on the final AIM populations.

15.6 BADER APPROACH

The original ''atoms in molecules'' method [45–47] as developed by the Bader group is based on the topology of the electron density ρ and that of the Laplacian of ρ. Instead of an in-depth discussion of the topological features of the electron density we just introduce a few essential characteristics here. A key concept is the gradient vector field, which is a collection of gradient paths. Here, a gradient path is a trajectory of steepest ascent in the electron density. The topology of the electron density is best revealed through its gradient vector field. At the so-called critical points, the gradient of the electron density vanishes. The critical points are best classified in terms of their rank and signature. At a critical point, the Hessian of ρ is computed and its eigenvalues are obtained. The rank is the number of nonzero eigenvalues and the signature is the sum of the signs of the Hessian eigenvalues. For example, a maximum in ρ has three negative eigenvalues, and hence its signature is $-3 = (-1) + (-1) + (-1)$. The rank is usually 3, which gives rise to four possible signatures, namely -3, -1, $+1$, and $+3$.

Atoms in Molecules and Population Analysis   **223**

The nature of a critical point is denoted as (rank, signature). For example, a minimum in ρ is designated by $(3, +3)$ and a maximum by $(3, -3)$. The two remaining types of critical points, $(3, -1)$ and $(3, +1)$ are saddle points, called bond critical point and ring critical point, respectively.

Gradient paths can be classified by means of the type of critical points (i.e., their signature) that they connect. This was achieved exhaustively for the first time in 2003 [48]. In the gradient vector field of ρ, the vast majority of gradient paths terminate at $(3, -3)$ critical points, which (approximately) coincide with nuclear positions. A collection of gradient paths that terminates at a given nucleus is called an atomic basin. These basins are mutually exclusive and thus partition three-dimensional space into disjoint (i.e., nonoverlapping) domains. The basin together with the nucleus inside it is then defined as the AIM within the context of QCT. Figure 15.1 shows a few examples of QCT atoms, using a new algorithm based on finite elements [49]. It is clear that a water dimer, which is a van der Waals complex, can also be partitioned into QCT atoms, in the same manner as a single molecule. The electron density's topology does not distinguish intramolecular interactions from intermolecular interactions. In principle, a QCT atom can be completely bounded by topological surfaces called interatomic surfaces provided there are enough neighboring atoms. The hydrogen atom in the middle of the water dimer, for example, is bounded by an interatomic surface on the left and one on the right. Only a very small edge of a nontopological surface bounds it at the top and at the bottom. This nontopological surface is an envelope of constant electron density, typically set at $\rho = 0.001$ a.u. For most atoms in the systems shown, such an envelope is vital to bound an atom visually. Secondly, an atom also needs to be bounded in order to have a finite volume for a numerical integration [50–52] to be possible over the atomic basin. The picture showing cyclopropane marks the bond critical points as purple points. These carbon–carbon bond critical points are at the center of the interatomic surface that separates the two carbons.

Till now, the AIM in QCT comes from an entirely topological origin. The single most important step forward in the theory was the realization that a quantum mechanical AIM coincides exactly with this topological atom [53]. The definition of an interatomic surface is given by

$$\nabla\rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \quad \text{for all } \mathbf{r} \in S \tag{15.15}$$
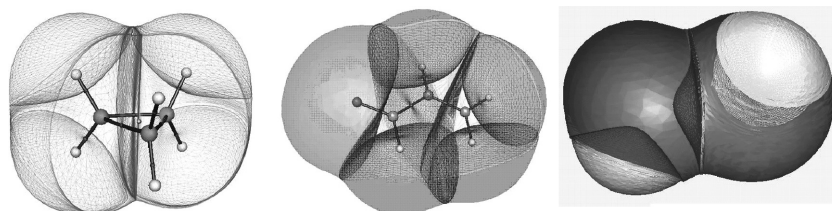
FIGURE 15.1 (See color insert following page 302.)   Examples of the QCT partitioning of the electron density. (left) All atoms in cyclopropane (except for the front methylene group); (middle) acrolein; and (right) a water dimer (global minimum).

This equation means that the normal to the surface S, $\mathbf{n}(\mathbf{r})$, is orthogonal to the gradient of the electron density. In other words, the surface is parallel to $\nabla\rho$, or rephrased again, the surface consists of gradient paths. The interatomic surface is a bundle of gradient paths that terminate at the bond critical point at the center of the surface.

It can be shown that Equation 15.15 means no less than the QCT AIM itself is a quantum mechanical object within the global quantum object. A common misunderstanding is that the AIM in this case becomes a closed system. This is incorrect. The QCT AIM should be seen as an open system [54], free to exchange electronic charge, for instance.

In order to frame the QCT method within the general expression (3), we need to give an expression for the weight operator $\hat{w}_A^{QCT}$. From the above discussion, it is clear that this operator will depend on $\mathbf{r}$ and that it is a binary operator, so its value is either 0 or 1. For a given atom A, the operator vanishes at every point in space, except within the basin of A, where it is equal to one. This way all atomic basins are indeed mutually exclusive.

Figure 15.1 also shows molecular graphs, which is a collection of special gradient paths that embody chemical bonds. Two such gradient paths originate from a given bond critical point and when traced in opposite directions, each terminate at a nucleus. This pair of gradient paths is called an atomic interaction line. In an equilibrium geometry, they are known as bond paths. A debate has emerged about whether bond paths can always be identified with a chemical bond. Unfortunately, this debate remained circular until the careful work of the Oviedo group and Gatti in 2007. They recently published an important paper [55], in which the exchange-correlation energy between atoms, $V_{xc}(A,B)$, is invoked to cut the vicious circle that fueled the bond path debate. The authors show that bond paths are ''privileged exchange channels'' and highlight a remarkable set of observations for classical test systems. It is clear that the competition between various $V_{xc}(A,B)$ terms is expressed by the presence or absence of bond paths. One should note that QCT not only proposes an AIM but also recovers bonding.

15.7 AIM PROPERTIES AND COMPARISON OF AIM METHODS

The properties of a quantum mechanical system such as an AIM are readily calculated from any method as long as they involve an operator acting on the electron density, e.g., for the case of the dipole moment. The problem would seem to become harder for other properties, although the introduction of property densities allows us to generally introduce AIM expectation values [45]. The expectation value of a property A for atom α in the Hirshfeld and QCT methods can be written as

$$A_\alpha = \int w_\alpha(\mathbf{r})\mathrm{d}\mathbf{r} \int \mathrm{d}\tau' \frac{N}{2} \left[\Psi^*\hat{A}\Psi + \left(\hat{A}\Psi\right)^*\Psi \right] \qquad (15.16)$$

where $\mathrm{d}\tau'$ denotes integration over all coordinates of all electrons in the system, except one. For QCT, this means that computing the average value of an atomic property warrants integration over the atomic basin only. In the case of both the Hirshfeld schemes (original and I), one integrates over all space but weights the

Atoms in Molecules and Population Analysis   225

result with the appropriate $w_A^H(\mathbf{r})$. In the case of the Mulliken method, the expressions are analogous, except that manipulations are carried out in the Hilbert space of the basis functions.

Many studies have focused on the comparison of AIM methods. Most concentrate on comparing the results from the AIM populations. Therefore, no such comparison will be repeated here. The issue to be addressed here is of a more fundamental nature. Although all the three methods addressed here have been and continue to be used widely, none of them is without its own problems. These may be computational in nature, such as the basis set dependence of the Mulliken method, the high CPU cost of QCT, or the arbitrariness of the promolecule in the Hirshfeld method. The most fundamental criticism is that in the Mulliken and Hirshfeld methods, the AIM is introduced whereas the AIM can be deduced from the QCT analysis of the electron density. In particular, in the Hirshfeld method, one composes a promolecule based on the atoms that constitute the molecule, thus from the beginning the concept of an atom is introduced. In the Mulliken approach, one needs to rely on the use of atom-centered basis functions that mimic the atomic orbitals, thereby introducing the atoms a priori. In QCT, the atoms can be obtained solely from the electron density, although one must then accept finding atoms without nucleus. These are the so-called nonnuclear attractors or pseudoatoms that appear occasionally. The appearance of such nonnuclear attractors can be an artifact from the use of small basis sets, although some can be genuine in that they survive at high level of theory [56]. These nonnuclear attractors do not fit well within our usual view of the molecular graph or molecular structure, but accepting QCT, one needs to accept the existence of pseudoatoms as well, and thus the concept of an atom without a nucleus because the atom has been defined as the union of the nucleus with its basin.

Finally, it is worth turning back to the question whether atoms in molecules can be defined uniquely. According to Parr et al., the AIM cannot be defined uniquely [6]. According to these authors, an AIM can neither be observed directly from experiment, nor could one gather enough properties of the AIM to define it unambiguously. They cannot conceive any experimental measurement that would show that one definition is uniquely correct, and so they accept the existence of several different approaches, all possibly useful in specific contexts. The AIM is therefore considered a noumenon: an object or pure thought not connected with sense perception. This point of view was scrutinized by Bader and Matta, claiming QCT to be the only correct and coherent way to define the AIM [40]. According to the latter authors, the QCT AIM is also confirmed by experiment.

Since Dalton, chemists have thought of molecules in terms of a collection of atoms. The entire discussion between the two different points of view, namely the noumenon view and Bader's view, is not about the usefulness of the AIM concept. There seems to be agreement on the usefulness of the concept. Rather, it is about pinpointing the AIM through a definition that excludes every other possibility. In any event, given the philosophy of Dirac's statement at the beginning of the chapter, QCT does have the advantage of allowing AIM to be derived from quantum mechanics. An atom was neither directly nor indirectly introduced at any stage. Yet, domains that do conform to the concept of an atom in a molecule resulted.

REFERENCES

1. Dirac, P.A.M. *Proc. Roy. Soc. (London)*, 1929, *123*, 714–733.
2. Tapia, O. *J. Math. Chem.*, 2006, *39*, 637–669.
3. Frankland, E.W. *J. Chem. Soc.*, 1866, *19*, 372–395.
4. Heitler, W. and London, F. *Z. Phys.*, 1927, *44*, 455–472.
5. Popelier, P.L.A. *Faraday Discuss*, 2007, *135*, 3–5.
6. Parr, R.G., Ayers, P.W., and Nalewajski, R.F. *J. Phys. Chem. A*, 2005, *109*, 3957–3959.
7. Francl, M.M., Carey, C., Chilian, L.E., and Gange, D.M. *J. Comput. Chem.*, 1996, *17*, 367–383.
8. Breneman, C.M. and Wiberg, K.B. *J. Comput. Chem.*, 1990, *11*, 361–373.
9. Francl, M.M. and Chirlian, L.E. *Rev. Comput. Chem.*, 2000, *14*, 1–31.
10. Singh, U.C. and Kollman, P.A. *J. Comput. Chem.*, 1984, *5*, 129–145.
11. Francl, M.M., Carey, C., Chirlian, L.E., and Gange, D.M. *J. Comput. Chem.*, 1996, *17*, 367–383.
12. Stone, A.J. and Price, S.L. *J. Phys. Chem.*, 1988, *92*, 3325–3335.
13. Popelier, P.L.A., Rafat, M., Devereux, M., Liem, S.Y., and Leslie, M. *Lecture Series on Computer and Computational Sciences*, 2005, *4*, 1251–1255.
14. Popelier, P.L.A. *Structure and Bonding. Intermolecular Forces and Clusters,* D.J.Wales (Ed.), vol. 115, pp. 1–56, Springer, Heidelberg, 2005.
15. Szabo, A. and Ostlund, N.S. *Modern Quantum Chemistry—Introduction to Advanced Electronic Structure Theory*, Dover Publications, New York, 1996.
16. Carbó-Dorca, R. and Bultinck, P. *J. Math. Chem.*, 2004, *36*, 201–210.
17. Mulliken, R.S. *J. Chem. Phys.*, 1955, *23*, 1833–1840.
18. Mulliken, R.S. *J. Chem. Phys.*, 1955, *23*, 1841–1846.
19. Mulliken, R.S. *J. Chem. Phys.*, 1955, *23*, 2338–2342.
20. Mulliken, R.S. *J. Chem. Phys.*, 1955, *23*, 2343–2346.
21. Jensen, F. *Introduction to Computational Chemistry*, John Wiley & Sons, Chichester, 1999.
22. Hirshfeld, F.L. *Theor. Chim. Acta*, 1977, *44*, 129–138.
23. Daudel, R. *Compt. Rend. Acad. Sci.*, 1952, *235*, 886–888.
24. Roux, M. and Daudel, R. *Compt. Rend. Acad. Sci.*, 1955, *240*, 90–92.
25. Roux, M., Besnainou, S., and Daudel, R. *J. Chim. Phys.*, 1956, *54*, 218–221.
26. Parr, R.G. and Yang, W. *Density Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
27. Geerlings, P., De Proft, F., and Langenaeker, W. *Chem. Rev.*, 2003, *103*, 1793–1873.
28. Ayers, P.W. *J. Chem. Phys.*, 2000, *113*, 10886–10898.
29. Nalewajski, R.F., Parr, R.G. *Proc. Natl. Acad. Sci. U.S.A.* 2000, *97*, 8879–8882.
30. Nalewajski, R.F., Switka, E., and Michalak, A. *Int. J. Quantum Chem.* 2002, *87*, 198–213.
31. Nalewajski, R.F. *Chem. Phys. Lett.*, 2003, *372*, 28–34.
32. Nalewajski, R.F. and Broniatowska, E. *Int. J. Quantum Chem.*, 2005, *101*, 349–362.
33. Ayers, P.W. *Theor. Chem. Acc.*, 2006, *115*, 370–378.
34. Mayer, I. and Salvador, P. *Chem. Phys. Lett.*, 2004, *383*, 368–375.
35. Becke, A.D. *J. Chem. Phys.*, 1988, *88*, 2547–2553.
36. Fonseca Guerra, C., Handgraaf, J.-W., Baerends, E.J., and Bickelhaupt, F.M. *J. Comput. Chem.* 2003, *25*, 189–210.
37. Bultinck, P., Ayers, P.W., Fias, S., Tiels, K., and Van Alsenoy, C. *Chem. Phys. Lett.*, 2007, *444*, 205–208.
38. Davidson, E.R. and Chakravorty, S. *Theor. Chim. Acta*, 1992, *83*, 319–330.
39. Bader, R.F.W. and Matta, C.F. *J. Phys. Chem. A*, 2004, *108*, 8385–8394.
40. Matta, C.F. and Bader, R.F.W. *J. Phys. Chem. A*, 2006, *110*, 6365–6371.

Atoms in Molecules and Population Analysis   **227**

41. Bultinck, P., Van Alsenoy, C., Ayers, P.W., and Carbó-Dorca, R. *J. Chem. Phys.*, 2007, *126*, 144111.
42. Bultinck, P. *Faraday Discussions*, 2007, *135*, 244–246.
43. Pierce, J.R. *An Introduction to Information Theory, Symbols, Signals and Noise,* Dover Publications, New York, 1980.
44. Rousseau, B., Peeters, A., and Van Alsenoy, C. *Chem. Phys. Lett*. 2000, *324*, 189–194.
45. Bader, R.F.W. *Atoms in Molecules: A Quantum Theory*, Clarendon, Oxford 1990.
46. Bader, R.F.W. *Chem. Rev*. 1991, *91*, 893–928.
47. Popelier, P. *Atoms in Molecules: An Introduction*, Prentice-Hall, Harlow, 2000.
48. Malcolm, N.O.J. and Popelier, P.L.A. *J. Comput. Chem*., 2003, *24*, 437–442.
49. Rafat, M. and Popelier, P.L.A. *J. Comput. Chem*., 2007, *28*, 2602–2617.
50. Popelier, P.L.A. *Comp. Phys. Commun*., 1998, *108*, 180–190.
51. Popelier, P.L.A. *Theor. Chem. Acc*., 2001, *105*, 393–399.
52. Popelier, P.L.A. *Molec. Phys*., 1996, *87*, 1169–1187.
53. Bader, R.F.W. *Phys. Rev. B* 1994, *49*, 13348–13356.
54. Bader, R.F.W. *J. Phys. Chem., A* 2007, *111,* 7966–7972.
55. Pendás, A.M., Francisco, E., Blanco, M.A., and Gatti, C. *Chem. Eur. J*., 2007, *13*, 9362–9371.
56. Gatti, C., Fantucci, P., and Pacchioni, G. *Theoret. Chim. Acta*, 1987, *72*, 433–458.