1
An Introduction to the Quantum Theory of Atoms in Molecules

Chérif F. Matta and Russell J. Boyd

1.1
Introduction

The observation that some properties attributed to atoms and functional groups are transferable from one molecule to another has played a key role in the development of chemistry. This observation provides a basis for group additivity schemes and is exemplified by the constancy of group contributions to thermodynamic and spectroscopic properties. But what is the electronic basis of this empirical transferability? The quantum theory of atoms in molecules (QTAIM) [1], developed by Professor Richard F. W. Bader and his coworkers, relies on quantum observables such as the electron density $\rho(\mathbf{r})$ and energy densities to answer such a question. Other important (related) questions addressed by QTAIM include:

- What is an atom in a molecule or a crystal?
- How can an atom or a group of atoms be transferable sometimes in *very* different external potentials?
- Can one define bonding in molecules unambiguously especially in borderline cases?

This chapter contains a summary of some of the main concepts of QTAIM. A more comprehensive and mathematically elegant treatment can be found in Bader's book [1].

(Often in this chapter, the word "molecule" includes extended systems such as polymers, weakly bonded molecular complexes, and molecular and ionic crystals, in addition to its more traditional meaning of a single, finite, isolated chemically bonded group of atoms. It will be clear from the context when this term is used in its traditional or in its larger sense.)

1.2
The Topology of the Electron Density

The topology of the electron density is dominated by the attractive forces of the nuclei imparting it with its principal topological feature – a substantial local max-

The Quantum Theory of Atoms in Molecules. Edited by Chérif F. Matta and Russell J. Boyd

ISBN: 978-3-527-30748-7

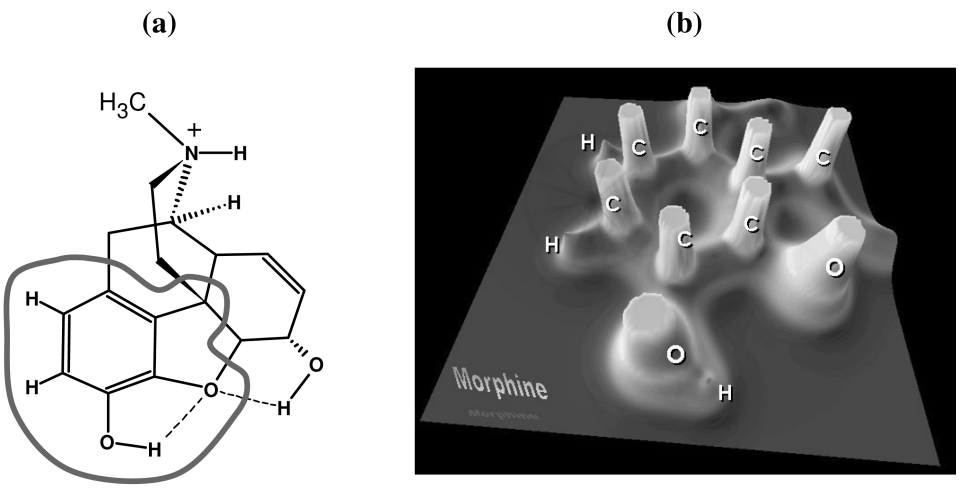

Fig. 1.1 (a) The molecular structure of the morphine molecule with an indication of the region shown in the relief map in (b). (b) A relief map representation of the electron density in the plane of the aromatic ring showing marked maxima at the positions of the carbon and oxygen nuclei (truncated at $\rho(\mathbf{r}) = 1.0$ au) and much smaller peaks at the position of the hydrogen nuclei.

imum at the position of each nucleus. A consequence of the dominance of nuclear maxima in the electron density distribution is the association of an atom with a region of space the boundaries of which are determined by the balance in the forces the neighboring nuclei exert on the electrons. Figure 1.1b is a relief map of the electron density of the phenolic region of the morphine molecule, in the plane of the aromatic ring, showing the maxima at the C, O, and H nuclei.

A "critical point" (CP) in the electron density is a point in space at which the first derivatives of the density vanish, i.e.:

$$\nabla\rho = \mathbf{i}\frac{d\rho}{dx} + \mathbf{j}\frac{d\rho}{dy} + \mathbf{k}\frac{d\rho}{dz} \rightarrow \begin{cases} = \vec{\mathbf{0}} & \text{(At critical points and at } \infty\text{)} \\ \text{Generally} \neq \vec{\mathbf{0}} & \text{(At all other points)} \end{cases} \quad (1)$$

where the zero vector signifies that each individual derivative in the gradient operator, ∇, is zero and not just their sum. The gradient of a scalar function such as $\rho(\mathbf{r})$ (Eq. 1) at a point in space is a vector pointing in the direction in which $\rho(\mathbf{r})$ undergoes the greatest rate of increase and having a magnitude equal to the rate of increase in that direction. The maximum at the position of a nucleus constitutes one type of CP, namely, a nuclear critical point (NCP). (The neglect of the finite size of atomic nuclei in quantum chemical calculations, an exceptionally good approximation, results in cusps in the potential and in the electron density $\rho(\mathrm{r})$ at the position of the nuclei. Because of this cusp, the derivatives of the electron

density at the position of a nucleus are not defined and so, in a formal mathematical sense, this position is not a true critical point. The nuclear maxima behave *topologically* as critical points, however.)

One can discriminate between a local minimum, a local maximum, or a saddle point by considering the second derivatives, the elements of the tensor $\nabla\nabla\rho$. There are nine second derivatives of $\rho(\mathbf{r})$ that can be arranged in the so-called "Hessian matrix", which when evaluated at a CP located at $\mathbf{r}_c$ is written:

$$\mathbf{A}(\mathbf{r}_c) = \begin{pmatrix} \frac{\partial^2\rho}{\partial x^2} & \frac{\partial^2\rho}{\partial x\partial y} & \frac{\partial^2\rho}{\partial x\partial z} \\ \frac{\partial^2\rho}{\partial y\partial x} & \frac{\partial^2\rho}{\partial y^2} & \frac{\partial^2\rho}{\partial y\partial z} \\ \frac{\partial^2\rho}{\partial z\partial x} & \frac{\partial^2\rho}{\partial z\partial y} & \frac{\partial^2\rho}{\partial z^2} \end{pmatrix}_{\mathbf{r}=\mathbf{r}_c} . \tag{2}$$

The Hessian matrix can be diagonalized because it is real and symmetric. The diagonalization of $\mathbf{A}(\mathbf{r}_c)$ is equivalent to a rotation of the coordinate system $\mathbf{r}(x, y, z) \rightarrow \mathbf{r}(x', y', z')$ superimposing the new axes x', y', z' with the principal curvature axes of the critical point. The rotation of the coordinate system is accomplished via a unitary transformation, $\mathbf{r}' = \mathbf{rU}$, where $\mathbf{U}$ is a unitary matrix constructed from a set of three eigenvalue equations $\mathbf{Au}_i = \lambda_i\mathbf{u}_i$ $(i = 1, 2, 3)$ in which $\mathbf{u}_i$ is the ith column vector (eigenvector) in $\mathbf{U}$. A similarity transformation $\mathbf{U}^{-1}\mathbf{AU} = \mathbf{\Lambda}$ transforms the Hessian into its diagonal form, which is written explicitly as:

$$\mathbf{\Lambda} = \begin{pmatrix} \frac{\partial^2\rho}{\partial x'^2} & 0 & 0 \\ 0 & \frac{\partial^2\rho}{\partial y'^2} & 0 \\ 0 & 0 & \frac{\partial^2\rho}{\partial z'^2} \end{pmatrix}_{\mathbf{r}'=\mathbf{r}_c} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \tag{3}$$

in which λ_1, λ_2, and λ_3 are the curvatures of the density with respect to the three principal axes x', y', z'.

An important property of the Hessian is that its trace is invariant to rotations of the coordinate system. The trace of the Hessian of the density is known as the Laplacian of the density $[\nabla^2\rho(\mathbf{r})]$ and, when $x = x'$, $y = y'$, and $z = z'$, is given by:

$$\nabla^2\rho(\mathbf{r}) = \nabla\cdot\nabla\rho(\mathbf{r}) = \underbrace{\frac{\partial^2\rho(\mathbf{r})}{\partial x^2}}_{\lambda 1} + \underbrace{\frac{\partial^2\rho(\mathbf{r})}{\partial y^2}}_{\lambda 2} + \underbrace{\frac{\partial^2\rho(\mathbf{r})}{\partial z^2}}_{\lambda 3} \tag{4}$$

where we have dropped the primes of the principal axes.

Critical points are classified according to their *rank* (ω) and *signature* (σ) and are symbolized by (ω, σ). The rank is the number of non-zero curvatures of ρ at the critical point. A critical point that has $\omega < 3$ is mathematically unstable and will vanish or bifurcate under small perturbations of the density caused by nuclear motion. The presence of such a CP (with a rank less than three) indicates a change in the topology of the density and, hence, a change in the molecular structure. For this reason, critical points with $\omega < 3$ are generally not found in equilibrium charge distributions and one nearly always finds $\omega = 3$. The signature is the algebraic sum of the signs of the curvatures, i.e. each of the three curvatures contributes ± 1 depending on whether it is a positive or negative curvature.

There are four types of stable critical points having three non-zero eigenvalues:

- $(3, -3)$ Three negative curvatures: ρ is a local maximum.
- $(3, -1)$ Two negative curvatures: ρ is a maximum in the plane defined by the corresponding eigenvectors but is a minimum along the third axis which is perpendicular to this plane.
- $(3, +1)$ Two positive curvatures: ρ is a minimum in the plane defined by the corresponding eigenvectors and a maximum along the third axis which is perpendicular to this plane.
- $(3, +3)$ Three curvatures are positive: ρ is a local minimum.

Each type of critical point described above is identified with an element of chemical structure: $(3, -3)$ *nuclear critical point* (NCP); $(3, -1)$ *bond critical point* (BCP); $(3, +1)$ *ring critical point* (RCP); and $(3, +3)$ *cage critical point* (CCP).

The number and type of critical points that can coexist in a molecule or crystal follow a strict topological relationship which states that:

$$n_{\mathrm{NCP}} - n_{\mathrm{BCP}} + n_{\mathrm{RCP}} - n_{\mathrm{CCP}} = \begin{cases} 1 & \text{(Isolated molecules)} \\ 0 & \text{(Infinite crystals)} \end{cases} \tag{5}$$

where n denotes the number of the subscripted type of CP. The first equality is known as the Poincaré–Hopf relationship (PH) [1] and applies for isolated finite systems such as a molecule, the second equality is known as the Morse equation and applies in cases of infinite periodic lattices [2]. The set $\{n_{\mathrm{NCP}}, n_{\mathrm{BCP}}, n_{\mathrm{RCP}}, n_{\mathrm{CCP}}\}$ for a given system is known as the "characteristic set".

Violation of Eq. (5) implies an inconsistent characteristic set, that a critical point has been missed, and that a further search for the missing critical point(s) is necessary. On the other hand, the satisfaction of this equation does not *prove* its completeness. For example, if we miss both a BCP and an RCP for a molecule, Eq. (5) becomes $n_{\mathrm{NCP}} - (n_{\mathrm{BCP}} - 1) + (n_{\mathrm{RCP}} - 1) - n_{\mathrm{CCP}} = 1$ which is clearly still valid [3]. The likelihood of missing both a BCP and a RCP is small, however, and, in practice, satisfaction of Eq. (5) is taken as a proof of the consistency *and* completeness of the characteristic set.

A ring critical point will always be found in the interior of a ring of chemically bonded atoms. When several rings are connected in a manner which encloses an

interstitial space, a cage critical point arises in the enclosed space. Figure 1.2 shows the molecular graph (the set of bond paths and critical points) of two molecules: (a) cubane, and (b) 4-methyl-1,12-difluoro[4]helicene. The bond path is a single line of maximum electron density linking the nuclei of two chemically-bonded atoms. (The bond path is discussed in more detail later in this chapter.) In cubane, the bond paths are arranged between the vertices of a cube forming six rings with the consequent appearance of one-ring critical point at the centre of each face of the cube. These six ring surfaces completely enclose the volume of the cube and, as a result, a cage critical point forms in the center of the cube. In Fig. 1.2a, the reader may also note the marked curvature of the bond paths in cubane, indicative of a significant ring strain in this unstable molecule.

All cage critical points reported in the literature until 2005 were found to be enclosed by at least three ring surfaces, as stated by Bader in 1990 [1]: *''While it is mathematically possible for a cage to be bounded by only two ring surfaces, the minimum number found in an actual molecule so far is three, as in bicyclo [1.1.1] pentane, for example''*, a statement reiterated in 2000 [3]. In Fig. 1.2b there is nothing unusual about the aromatic system, but the nuclei of the two fluorine atoms in the ''Fjord region'' are linked by a bond path [4] closing a seven-membered ring which has quite an unusual topology – it gives rise to *two ring* critical points and *a cage* critical point [5]. We have, thus, recently reported the first example of an actual molecular system in which a cage is bounded by only *two* ring surfaces [5]. Such a CCP (enclosed by two ring surfaces) arises in all the studied derivatives of 1,12-difluoro[4]helicenes [5]. In these molecules, the seven-membered ring in the Fjord region is so distorted out of planarity that its ring surface splits into two, giving rise to this CCP [5]. In all cases, the Poincaré–Hopf relationship is satisfied [5].

1.3 The Topology of the Electron Density Dictates the Form of Atoms in Molecules

The pronounced maxima in the electron density at the positions of the nuclei give rise to a rich topology. This topology embodies a natural partitioning of the molecular space into separate mononuclear regions, Ω, identified as atoms in molecules. The surface bounding an atom in a molecule is one of zero flux in the gradient vector field of the electron density, i.e. it is not crossed by any of the gradient vectors $[\nabla\rho(\mathbf{r})]$ at any point, a statement which is equivalent to satisfying the condition:

$$\nabla\rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0, \quad \text{for all } \mathbf{r} \text{ belonging to the surface } S(\Omega) \tag{6}$$

where $\mathbf{r}$ is the position vector and $\mathbf{n}(\mathbf{r})$ the unit vector normal to the surface $S(\Omega)$.

The plot in Fig. 1.3a represents the electron density and its gradient vector field in the molecular plane of BF_3. The figure contrasts the zero-flux surfaces which partition the molecular space into separate mononuclear ''atomic basins'' and an arbitrary surface cutting though the density. The left side of Fig. 1.3a is a contour

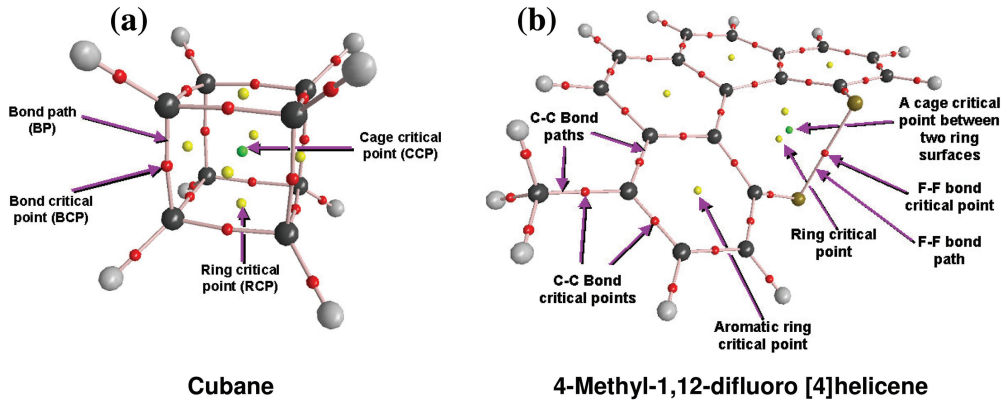

Fig. 1.2 The molecular graph of (a) cubane and (b) 4-methyl-1,12-difluoro[4]helicene showing the bond paths (lines) and the different critical points: nuclear (color-coded by element: C = black, H = grey, F = golden), bond (small red dots), ring (yellow dots), and cage (green dots) critical points.

plot of $\rho(\mathbf{r})$, the contours decreasing in value from the nuclei outward. Instead of plotting $\rho(\mathbf{r})$ in the right half of Fig. 1.3a (which is a mirror image of the left side by virtue of the molecular symmetry), we have depicted, instead, the corresponding gradient vector field $\nabla\rho(\mathbf{r})$. The gradient vector field lines partition the molecular space naturally into three fluorine basins and a central boron basin (Fig. 1.3a).

Gradient vector field lines belonging to an atomic basin all converge to *one* nucleus which acts as an attractor to these gradient vector field lines. In doing so, these gradient vector field lines sweep a portion of physical space associated with one nucleus and which is identified as the basin of an atom in a molecule (AIM). Three-dimensional volume renderings of the atoms and groups of atoms within the BF_3 molecule are shown in Fig. 1.3b. An atom in a molecule is defined as the union of a nucleus and its associated basin. Each basin is bounded by one (or by the union of a number of) zero-flux surface(s) one of which may occur at infinity. An atom in a molecule may be defined, alternatively and equivalently, as a region of space bounded by one or more zero-flux surface(s).

Occasionally, local maxima in the electron density can occur at positions other than those of atomic nuclei, especially in metals [6, 7] and semiconductors [8, 9], but also in systems such as the solvated electron [10] and at the positions of defects in crystals and color F-centers [11]. The non-nuclear maxima, also known as non-nuclear attractors (NNA), are topologically indistinguishable from the nuclear maxima. Just like a nucleus, an NNA is associated with a basin swept by gradient vector field lines and is bounded by a zero-flux surface. Consequently, NNA

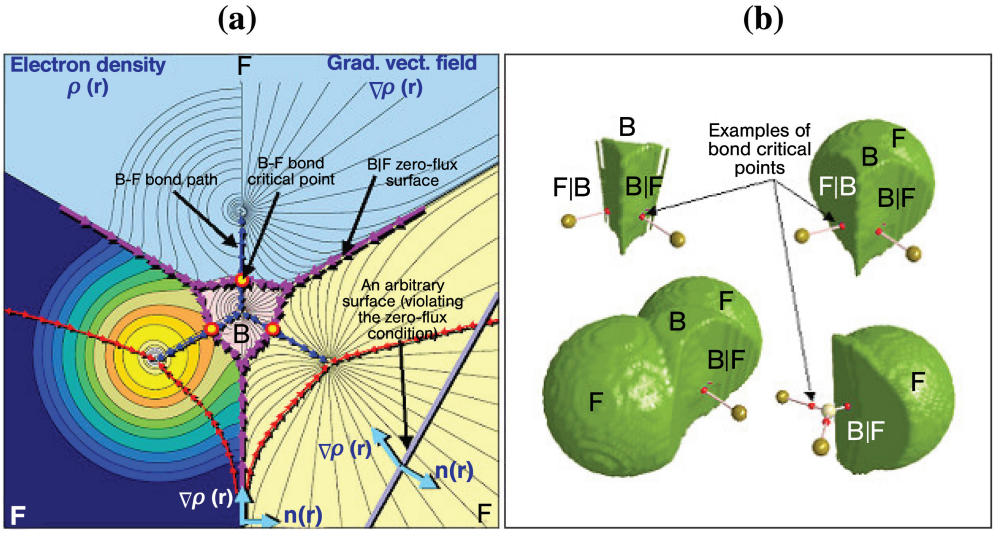

Fig. 1.3 (a) The electron density (left) and the gradient vector field (right) of the density in the molecular plane of BF_3. The blue arrows connecting the nuclei trace the bond paths. The magenta arrows delimiting atomic basins trace the intersections of the zero-flux surfaces with the plane. The contours increase from the outermost 0.001 au contour followed by 2×10^n, 4×10^n, and 8×10^n au with n starting at -3 and increasing in steps of unity. The small circles drawn on the three bond paths are the B–F bond critical points (BCP). The intersection of an arbitrary surface with the plane of the figure, the straight line on the lower right part of (a), is shown to be crossed by gradient vectors and is contrasted with a zero-flux surface. (b) Four three-dimensional renderings of the density of atoms and groupings of atoms in BF_3. The outer surface is the 0.002 au isodensity envelope. The zero-flux surfaces are denoted by the vertical bars between the atomic symbols. Large spheres represent the nuclei of the fluorine atoms (golden) and of the boron atom (blue–gray). The lines linking the nuclei represent the bond paths. The BCPs are denoted by the small red dots. A BCP always lies on the zero-flux surface shared by the two bonded atoms.

basins constitute proper open quantum systems and are therefore termed ''pseudo-atoms''. Pseudo-atoms can be bonded (i.e. share a common interatomic zero-flux surface, a bond critical point, and a bond path) to atoms and other pseudo-atoms in a molecule. Non-nuclear attractors and their basins are of great importance in characterizing metallic bonding and are of substantial theoretical interest. A detailed discussion of NNA can be found in Chapter 7 of this book.

There is a unique set of gradient vectors lines which originate at infinity and terminate at a point *between* two bonded atoms, the lines of this set fall by definition on the zero-flux surface because they satisfy Eq. (6) locally. It should be noted that the three zero-flux surfaces depicted in Fig. 1.3 are between the boron and fluorine atoms, the boron atom being bounded three zero-flux surfaces which merge in pairs at infinity between fluorine basins. There are no zero-flux surfaces

between any pair of fluorine atoms in BF_3, these surfaces only exist between bonded atoms and are characteristic of bonding interactions.

The topological definition of an atom follows from the boundary condition expressed in Eq. (6) and which applies to every point on the surface. This real space partitioning of the electron density has been shown to be rooted in quantum mechanics bringing into coincidence the topological definition of an atom in a molecule with that of a proper open quantum system (see Chapter 2 and also the detailed derivation of the quantum mechanics of proper open systems [12] from Schwinger's principle of stationary action [13]).

1.4 The Bond and Virial Paths, and the Molecular and Virial Graphs

The presence of an interatomic zero-flux surface between any two bonded atoms in a molecule is always accompanied by another key topological feature – there is, *in real space*, a single line of *locally* maximum density, termed the "bond path" (BP), linking their nuclei. The bond path is a universal indicator of chemical bonding of all kinds; weak, strong, closed-shell, and open-shell interactions [14]. The point on the bond path with the lowest value of the electron density (minimum along the path) is the bond critical point (BCP) and it is at that point where the bond path intersects the zero-flux surface separating the two bonded atoms.

The collection of bond paths linking the nuclei of bonded atoms in an equilibrium geometry, with the associated critical points, is known as the *molecular graph*. (In a non-equilibrium geometry, lines of maximum electron density linking the nuclei are known as "atomic interaction lines", because these may or may not persist when the geometry is energy-minimized, i.e. optimized.) The molecular graph provides an unambiguous definition of the "molecular structure" and can thus be used to locate changes in structure along a reaction path.

Mirroring every molecular graph is a "shadow" graph, again in real space, but this time the graph is defined by a set of lines of maximally negative *potential energy density*. In other words, there is a single line of maximally negative potential energy density linking the same attractors which share a bond path [15]. This line of "maximum stability" in real space is termed a "*virial path*". The collection of virial paths and the associated critical points constitute the *virial graph*. The virial graph defines the same molecular structure as the molecular graph, the virial field and the electron density being homeomorphic [15].

Figure 1.4 shows the chemical structure and the molecular and virial graphs of the phenanthrene molecule. This polycyclic aromatic hydrocarbon molecule has a bond path between the two hydrogen atoms in the bay region, a mode of closed-shell bonding which has been recently characterized in detail and termed hydrogen–hydrogen bonding (to be contrasted with dihydrogen bonding) [16, 17]. The virial graph is shown to faithfully map each bond path with a corresponding virial path including the bond path of a weak closed-shell bonding interaction such as the hydrogen–hydrogen bonding interaction (Fig. 1.4).

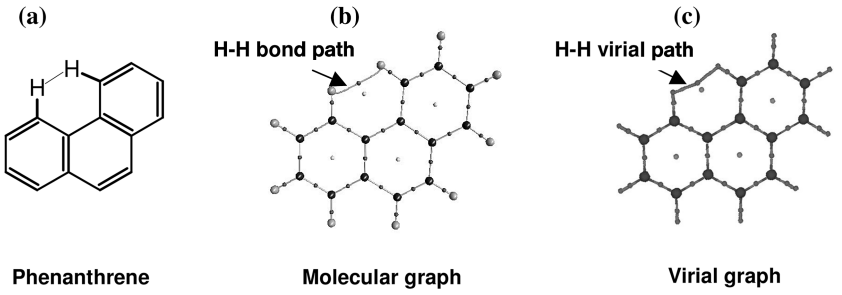

Fig. 1.4 (a) The chemical structure of phenanthrene. (b) The molecular graph of phenathrene showing the collection of bond paths and associated critical points. (c) The corresponding virial graph.

We conclude this section by stating that *atoms that are chemically bonded have their nuclei linked by a (single) bond path and by an accompanying virial path and they share a bond critical point and a common interatomic zero-flux surface.*

1.5 The Atomic Partitioning of Molecular Properties

The quantum theory of atoms in molecules is a generalization of quantum mechanics to open quantum systems. Bader has shown that the topological partitioning of the molecules into atomic basins is essential for development of the quantum mechanics of open systems [12]. The zero-flux condition, Eq. (6), is the necessary constraint for the application of Schwinger's principle of stationary action [13] to part of a quantum system [12].

The partitioning of the molecular space into atomic basins enables the partitioning of electronic properties into atomic contributions in one consistent theoretical framework. Among the properties often discussed are the atomic charges and higher multipolar electric polarizations, atomic volumes, atomic total energies (and the different contributions to the atomic energies), and the electron localization within one basin or delocalization between two basins [1, 18].

The expectation value of an operator averaged over all space is given by the sum of the expectation values of this operator averaged over all the atoms in the molecule or the crystal, in atomic units:

$$\langle \hat{O} \rangle_{molecule} = \sum_{i}^{\substack{all\ atoms\\ in\ the\\ molecule}} \left(N \int_{\Omega_i} \left\{ \int \frac{1}{2} [\Psi^* \hat{O} \Psi + (\hat{O}\Psi)^* \Psi] \, d\tau' \right\} d\mathbf{r} \right) \tag{7a}$$

$$= \sum_{i}^{\substack{all\ atoms\\ in\ the\\ molecule}} \left(\int_{\Omega_i} \rho_O \, d\mathbf{r} \right) = \sum_{i}^{\substack{all\ atoms\\ in\ the\\ molecule}} O(\Omega_i) \tag{7b}$$

where $\langle \hat{O} \rangle_{molecule}$ is the molecular expectation value of the operator $\hat{O}$, $O(\Omega_i)$ is the average of this operator over an atom Ω_i, and where the sum runs over all the atoms in the molecule or crystal. Integration over the coordinates of all electrons but one and summation over all spins is symbolized by $\int d\tau'$. Equation (7b) implies that any molecular property O which can be expressed in terms of a corresponding property density in space $\rho_O(\mathbf{r})$ can be written as a sum of atomic contributions obtained by averaging the appropriate operator over the volume of the atom, i.e. it exhibits atomic additivity.

1.6
The Nodal Surface in the Laplacian as the Reactive Surface of a Molecule

Because the Laplacian is essentially a second derivative, its sign indicates regions of local electronic charge concentration or depletion with respect to the immediate neighborhood. Thus, where $\nabla^2\rho(\mathbf{r}) > 0$ the density is locally depleted and expanded relative to its average distribution; where $\nabla^2\rho(\mathbf{r}) < 0$ the density is locally concentrated, tightly bound, and compressed relative to its average distribution. A local charge concentration behaves as a Lewis base (electron donor) whereas a local charge depletion acts as a Lewis acid (electron acceptor).

The Laplacian reproduces the spherical shell structure of isolated atoms in terms of alternating shells of charge concentration followed by shells of charge depletion [19, 20]. The spherical nodes in the Laplacian are envelopes bounding regions of density depletion or concentration. The outer shell of charge concentration, which is followed by a shell of charge depletion extending to infinity, is called the valence shell charge concentration (VSCC). When an atom is involved in bonding the spherical symmetry of the VSCC is broken. A chemical reaction corresponds to the combination of a ''lump'' in the VSCC of the base with a ''hole'' in the VSCC of the acid.

Covalently bonded atoms have bonding charge concentrated in the region between their nuclei. In addition to bonding charge concentrations, lone pairs are associated with non-bonding charge concentrations. These observations reflect an underlying mapping between the Laplacian of the electron density and the Laplacian of the conditional pair density when electrons tend to be localized [21].

The Laplacian of the density is characterized by a rich topology which provides a basis for the VSEPR model [22–24] of molecular geometry [1, 25–27]. More details on this topic are available elsewhere [1, 3, 25–29] and Chapter 19 explores the use of the reactive surface in drug design and drug–receptor molecular complementarity.

1.7
Bond Properties

A zero-flux surface is defined by a particular set of $\nabla\rho(\mathbf{r})$ trajectories all the members of which terminate at a single point, the bond critical point, where

$\nabla\rho(\mathbf{r}) = 0$. There is one BCP between each pair of atoms that are bonded, i.e., two atoms linked by a bond path and sharing a common interatomic zero-flux surface. In addition to the set of trajectories which terminate at the BCP and define an interatomic surface, a pair of trajectories originates at the BCP with each member of the pair terminating at one of the nuclei of the chemically bonded atoms. This latter pair of trajectories defines the bond path [14]. Chemical bonding interactions are characterized and classified according to the properties of the electron and energy densities at the BCP, collectively know as ''bond properties''.

1.7.1
The Electron Density at the BCP (ρ_b)

The strength of a chemical bond, its bond order (BO), is reflected in the electron density at the BCP (ρ_b) [1]:

$$\mathrm{BO} = \exp[A(\rho_b - B)] \tag{8}$$

where A and B are constants which depend on the nature of the bonded atoms. In general, ρ_b is greater than 0.20 au in shared (covalent) bonding and less than 0.10 au in a closed-shell interaction (for example ionic, van der Waals, hydrogen, dihydrogen, H–H bonding, etc.). ρ_b has been shown to be strongly correlated with the binding energy for several types of bonding interaction [30–36] and with the bond length of S–S bonding interactions [37]. Proposals to generalize Eq. (8) by including more than two elements in the same fitting have recently appeared in the literature [38, 39].

1.7.2
The Bonded Radius of an Atom (r_b), and the Bond Path Length

The distance of a BCP from nucleus A determines the ''bonded radius'' of atom A relative to the interaction defined by the BCP, and is denoted r_b(A). If the bond path is coincident with the internuclear axis, then the sum of the two associated bond radii, termed the bond path length, equals the bond length. If, however, the bond path is curved, or strained chemically, the bond path length will exceed the bond length. Examples of this latter behavior are found for hydrogen-bonded interactions and for bonding within strained cyclic molecules (e.g. the curved C–C bond paths in the cubane molecule, Fig. 1.2a).

1.7.3
The Laplacian of the Electron Density at the BCP ($\nabla^2\rho_b$)

The Laplacian at the BCP is the sum of the three curvatures of the density at the critical point (Eq. 4), the two perpendicular to the bond path, λ_1 and λ_2, being negative (by convention, $|\lambda_1| > |\lambda_2|$) whereas the third, λ_3, lying along the bond

path, is positive. The negative curvatures measure the extent to which the density is concentrated along the bond path and the positive curvature measures the extent to which it is depleted in the region of the interatomic surface and concentrated in the individual atomic basins.

In covalent bonding the two negative curvatures are dominant and $\nabla^2\rho_b < 0$, for example, $\nabla^2\rho_b = -1.1$ au for a typical C–H bond. In contrast, in closed-shell bonding, for example ionic, hydrogen-bonding or van der Walls interactions, the interaction is characterized by a depletion of density in the region of contact of the two atoms and $\nabla^2\rho_b > 0$. An N–(H···O)=C hydrogen bond, for instance, is characterized by $\nabla^2\rho_b = +0.03$ au. In strongly polar bonding, (e.g. C–X, where $X = O, N, F$), there is a significant accumulation of electron density between the nuclei, as in all shared interactions, but the Laplacian in this type of bonding can be of either sign.

1.7.4
The Bond Ellipticity (ε)

The ellipticity measures the extent to which density is preferentially accumulated in a given plane containing the bond path. The ellipticity is defined as:

$$\varepsilon = \frac{\lambda_1}{\lambda_2} - 1 \quad (\text{where } |\lambda_1| \geq |\lambda_2|) \tag{9}$$

If $\lambda_1 = \lambda_2$, then $\varepsilon = 0$, and the bond is cylindrically symmetrical; examples are the C–C single bond in ethane or the triple bond in acetylene. Thus, ε is a measure of the π-character of the bonding up to the limit of the "double bond" for which the ellipticity reaches a maximum. On going from a double to a triple bond, the trend is reversed and the ellipticity decreases with increasing bond order, because at the limit of BO = 3 the bonding regains its cylindrical symmetry (two π-bonding interactions in two orthogonal planes in addition to a cylindrically symmetric σ-bonding interaction). The ellipticity of an aromatic bond is ca. 0.23 in benzene and that of a formal double bond is ca. 0.45 in ethylene.

1.7.5
Energy Densities at the BCP

Energy densities require information contained in the one-electron density matrix (and not just the density, its diagonal elements). The energy densities (potential, kinetic, and total) are used to summarize the mechanics of a bonding interaction.

The potential energy density, $\mathcal{V}(\mathbf{r})$, also known as the virial field, is the average effective potential field experienced by a single electron at point $\mathbf{r}$ in a many-particle system. The virial field evaluated at any point in space is always negative and its integral over all space yields the total potential energy of the molecule. The local statement of the virial theorem expresses the relationship between the

virial field, the kinetic energy density, and the Laplacian, which when written for a stationary state is [1, 12, 40]:

$$\left(\frac{\hbar^2}{4m}\right)\nabla^2\rho(\mathbf{r}) = 2G(\mathbf{r}) + \mathscr{V}(\mathbf{r}) \tag{10}$$

where

$$G(\mathbf{r}) = \frac{\hbar^2}{2m} N \int d\tau' \nabla\Psi^* \cdot \nabla\Psi \tag{11}$$

and where $G(\mathbf{r})$ is the gradient kinetic energy density and Ψ is an antisymmetric many-electron wavefunction.

Because we always have $G(\mathbf{r}) > 0$ and $\mathscr{V}(\mathbf{r}) < 0$, the local virial theorem when applied at a BCP implies that interactions for which $\nabla^2\rho_b < 0$ are dominated by a local reduction of the potential energy. Conversely, interactions for which $\nabla^2\rho_b > 0$ are dominated by a local excess in the kinetic energy.

To compare the kinetic and potential energy densities on an equal footing (instead of the 2:1 virial ratio) Cremer and Kraka [41] proposed evaluating the total electronic energy density $[H(\mathbf{r}) = G(\mathbf{r}) + \mathscr{V}(\mathbf{r})]$ at the BCP:

$$H_b = G_b + \mathscr{V}_b \tag{12}$$

The total energy density yields the total electronic energy when integrated over all space. H_b is negative for interactions with significant sharing of electrons, its magnitude reflecting the "covalence" of the interaction [41].

1.7.6
Electron Delocalization between Bonded Atoms: A Direct Measure of Bond Order

The number of electron pairs *shared* between two bonded atoms is often called the *bond order*. QTAIM provides a bookkeeping of the number of pairs shared between two atoms by integrating the exchange density once over each of the two atomic basins. This property may as well be classified under "atomic properties" because it involves the double integration of the exchange density over the basins of two atoms, but, because it "counts" the number of electron pairs shared between two atoms, when reported for *bonded* atoms, it can be regarded as a bond property.

The magnitude of the exchange of the electrons in the basin of atom A with those in the basin of atom B is termed the delocalization index between them, $\delta(A, B)$, and is defined for a closed-shell system as [42]:

$$\delta(A, B) = 2|F^{\alpha}(A, B)| + 2|F^{\beta}(A, B)| \tag{13}$$

where the Fermi correlation is defined as:

$$F^{\sigma}(A,B) = -\sum_{i}\sum_{j}\int_{A} d\mathbf{r}_1 \int_{B} d\mathbf{r}_2 \{\phi_i^*(\mathbf{r}_1)\phi_j(\mathbf{r}_1)\phi_j^*(\mathbf{r}_2)\phi_i(\mathbf{r}_2)\}$$
$$= -\sum_{i}\sum_{j} S_{ij}(A)S_{ji}(B) \tag{14}$$

where $S_{ij}(\Omega) = S_{ji}(\Omega)$ is the overlap integral of two spin orbitals over a region Ω and σ represents spin (α or β).

The second-order density matrix obtained from a configuration interaction (CI) calculation can also be expressed in terms of products of basis functions multiplied by the appropriate coefficients enabling one to express the integrated pair density in terms of overlap contributions. Thus, terms similar to those in Eq. (14) multiplied by the appropriate coefficients appear in the CI expression for $F^{\sigma}(A,B)$ and electron delocalization is still described in terms of the exchange of electrons between molecular orbitals, but this time in a wavefunction incorporating Coulomb in addition to Fermi correlation [42].

If the double integration in Eq. (14) is performed over only one atomic basin, say atom A, this would yield the total Fermi correlation for the electrons in region A [43]:

$$F^{\sigma}(A,A) = \int_{A} d\mathbf{r}_1 \int_{A} d\mathbf{r}_2 \rho^{\sigma}(\mathbf{r}_1)h^{\sigma}(\mathbf{r}_1,\mathbf{r}_2) \tag{15}$$

where its limiting value is $-N^{\sigma}(A)$, the negative of the σ-spin population of atom A, i.e. the number of σ electrons in A being totally localized within this atom because all remaining σ-spin density would then be excluded from A. In other words, if this limiting value is reached, it implies that the electrons in A do not exchange with electrons outside A. Thus a *localization index* $[\lambda(A)]$ is defined as:

$$\lambda(A,A) = |F^{\alpha}(A,A)| + |F^{\beta}(A,A)| \tag{16}$$

The limit of total localization, while approached quite closely ($\geq 95\%$) in ionic systems, cannot usually be achieved and one finds that $|F^{\sigma}(A,A)| < N^{\sigma}(A)$, indicating that the electrons in region A always exchange, to some extent, with electrons outside the boundaries of A, i.e., they are delocalized.

Because the Fermi correlation counts all electrons, the sum of the localization indices and half of all the delocalization indices is N, the total number of electrons in the molecule. This, in turn, provides a measure of how these electrons are localized within the individual atomic basins and delocalized between them, in effect resulting in bookkeeping of electrons in the molecule:

$$N(A) = \lambda(A) + \frac{1}{2}\sum_{B\neq A}\delta(A,B) \tag{17}$$

How closely the sum of the localization and the delocalization indices (Eq. 17) recovers the total molecular electron population is a global measure of the quality of the atomic integrations.

The localization and delocalization indices can be calculated from the atomic overlap matrices using readily available software such as AIMDELOC [44] or LI-DICALC [45, 46].

It is important to realize that a delocalization index can be calculated between *any* pair of atoms whether bonded or not. When $\delta(A, B)$ is calculated between *bonded* atoms it yields a measure of the bond order between them if the electron pairs are equally shared (i.e. there is no appreciable charge transfer) [42, 47].

Because ρ_b and the bond order are strongly correlated (Eq. 8), Matta and Hernández-Trujillo [48] suggested calibrating this correlation using the delocalization index rather than arbitrarily assigned bond orders:

$$\delta(A, B) = \exp[A(\rho_b - B)] \tag{18}$$

Equation (18) enables calibration of experimental ρ_b with delocalization indices obtained by calculation. The fitted equation can then be used to obtain experimental estimates for information on electron sharing contained in a full density matrix, information which is not accessible in a conventional X-ray diffraction experiment, from experimentally determined ρ_b [48]. Data for the 21 carbon–carbon bonds in the estrone hormone could be fitted to the following equation [49]:

$$\delta(\mathrm{C}, \mathrm{C}') = \exp\{4.7427 \times [\rho_b(\text{in a.u.}) - 0.2538]\} \tag{19}$$

with $r^2 = 0.939$, a variance of 0.002, and a root mean square deviation of 0.010, and in which $\delta(\mathrm{C}, \mathrm{C}')$ were calculated at the B3LYP/6-311++$G(d, p)$ level and ρ_b are the experimentally determined electron density values at the C–C BCPs.

1.8 Atomic Properties

The average of a property O over an atomic basin Ω, $O(\Omega)$, is calculated from:

$$O(\Omega) = \langle \hat{O} \rangle_\Omega = \frac{N}{2} \int_\Omega d\mathbf{r} \int d\tau' [\Psi^* \hat{O} \Psi + (\hat{O}\Psi)^* \Psi] \tag{20}$$

where $\hat{O}$ is a one-electron operator or a sum of one-electron operators. Some examples of commonly computed atomic properties are discussed in the subsections below.

1.8.1
Atomic Electron Population [$N(\Omega)$] and Charge [$q(\Omega)$]

The total electron population of an atom in a molecule is obtained by setting $\hat{O} = \hat{1}$ in Eq. (20). This yields:

$$N(\Omega) = \int_{\Omega} \rho(\mathbf{r})\, d\mathbf{r} \tag{21}$$

which can also be expressed explicitly in terms of the separate spin populations as the expectation value of the number operator, an integral operator, averaged over a proper open quantum subsystem:

$$N(\Omega) = \sum_{i} [\langle \psi_i(\mathbf{r}) | \psi_i(\mathbf{r}) \rangle_{\Omega}^{\alpha} + \langle \psi_i(\mathbf{r}) | \psi_i(\mathbf{r}) \rangle_{\Omega}^{\beta}] \tag{22}$$

in which the separate spin populations are given by:

$$\langle \psi_i(\mathbf{r}) | \psi_i(\mathbf{r}) \rangle_{\Omega}^{\sigma} = \int_{\Omega} \psi_i^{\sigma *}(\mathbf{r}) \psi_i^{\sigma}(\mathbf{r})\, d\mathbf{r} \equiv S_{ii}^{\sigma}(\Omega) \tag{23}$$

where σ refers to either α-spin or β-spin, and $S_{ii}^{\sigma}(\Omega)$ is the ith diagonal element of the atomic overlap matrix.

The atomic charge is obtained by subtracting $N(\Omega)$ from the nuclear charge Z_{Ω}:

$$q(\Omega) = Z_{\Omega} - N(\Omega) \tag{24}$$

Because of the manner by which atomic populations are defined, Eqs (22) and (23), QTAIM populations and charges are true quantum expectation values. That is, they are "observables" in the quantum mechanical sense [18, 50]. Observables are not necessarily measurable in practice, but any measurable quantity is an observable or can be expressed in terms of one or more observables. Indirect experimental evidence lends strong support to the physical nature of QTAIM atomic populations and charges [51] (see also Section 1.9.2).

The deviation of the sum of the atomic populations (or charges) from the corresponding molecular value is an indicator of the quality of the numerical integrations. Deviations of less than ca. 0.001–0.002 electrons are regarded as acceptable for molecules of medium size (up to ~100 first to third row atoms).

1.8.2
Atomic Volume [Vol.(Ω)]

The atomic volume is defined as the space bounded by the intersection of the zero-flux surface(s) bounding the atom from the molecular interior and a chosen

outer isodensity envelope (if a side of this atom's basin extends to infinity). While a molecule extends in principle to infinity, an outer isodensity of $\rho(\mathbf{r}) = 0.001$ au is usually chosen as its outer bounding surface for two reasons:

1. this isosurface closely recovers the experimental van der Waals volumes in the gas phase, and
2. it usually encloses more than 99% of the electron population of the molecule [1].

The van der Waals surface in condensed phases is closer to the 0.002 au isodensity envelope [1].

1.8.3 Kinetic Energy [$T(\Omega)$]

There are at least two forms of the kinetic energy operator [52] with two corresponding expressions for the atomic average of the kinetic energy, the Schrödinger kinetic energy:

$$K(\Omega) = -\frac{\hbar^2}{4m} N \int_{\Omega} d\mathbf{r} \int d\tau' [\Psi \nabla^2 \Psi^* + \Psi^* \nabla^2 \Psi] \tag{25}$$

and the gradient kinetic energy:

$$G(\Omega) = \frac{\hbar^2}{2m} N \int_{\Omega} d\mathbf{r} \int d\tau' \nabla_i \Psi^* \cdot \nabla_i \Psi \tag{26}$$

For the total system and for a proper open quantum system, Eqs (25) and (26) must yield an identical value for the kinetic energy, of course, i.e. $K(\Omega) = G(\Omega) = T(\Omega)$. Because the difference between $K(\Omega)$ and $G(\Omega)$ should vanish for an atom in a molecule, the (small) departure from zero of this difference as gauged by the Laplacian (Section 1.8.4) is a measure of the numerical accuracy of the atomic integrations.

1.8.4 Laplacian [$L(\Omega)$]

The Laplacian function has the dimensions of electrons $\times$ (length)$^{-5}$. Because of the zero-flux boundary condition, Eq. (6), the Laplacian of the electron density, vanishes when integrated over an atomic basin, as can be seen from:

$$\begin{aligned} L(\Omega) &= K(\Omega) - G(\Omega) \\ &= -\frac{\hbar^2}{4m} \int_{\Omega} d\mathbf{r} [\nabla^2 \rho(\mathbf{r})] \\ &= -\frac{\hbar^2}{4m} \int dS(\Omega, \mathbf{r}) \nabla \rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \end{aligned} \tag{27}$$

the last equality is valid only for the total system or if the integration is performed over a proper open quantum system bounded by zero-flux surfaces.

How close the integrated Laplacian approaches zero is often used as an indicator of the numerical accuracy of atomic integrations. Deviations from zero are a measure of integration error. $L(\Omega) \leq$ ca. 1.0×10^{-3} au for second and third-row atoms and $L(\Omega) \leq$ ca. 1.0×10^{-4} au for hydrogen atoms are regarded as acceptable and are usually paralleled by atomic energies which add up to within a kcal mol^{-1} of the directly calculated molecular total energy for a medium size molecule ($\sim$100 atoms or fewer). The smaller $L(\Omega)$ the better the quality of an atomic integration.

1.8.5 Total Atomic Energy [$E_e(\Omega)$]

The partitioning of the total molecular energy into a set of additive atomic energies is a non-trivial problem that was solved by Bader [1]. To see the difficulties in partitioning the total energy, one may ask, for instance, how can the nuclear–nuclear repulsion contribution to the total molecular energy be partitioned on an atom-by-atom basis?

The kinetic energy density can be expressed:

$$K(\mathbf{r}) = -\frac{\hbar^2}{4m} N \int d\tau' [\Psi \nabla^2 \Psi^* + \Psi^* \nabla^2 \Psi] \tag{28}$$

which when compared with Eq. (11) yields:

$$K(\mathbf{r}) = G(\mathbf{r}) - \frac{\hbar^2}{4m} \nabla^2 \rho(\mathbf{r}) \tag{29}$$

It is clear from Eq. (29) that the integral of the kinetic energy densities $K(\mathbf{r})$ and $G(\mathbf{r})$ over a volume ω would usually yield different values because the integral of the Laplacian does not usually vanish when integrated over an arbitrary volume, in which case the kinetic energy is *not* well defined. The kinetic energy is well defined if, and only if, the integral of the Laplacian term vanishes, i.e. when this integral is performed over the total system or over an atomic basin bounded by a zero-flux surface. Integrating Eq. (29) over ω, one obtains:

$$K(\omega) = G(\omega) - \frac{\hbar^2}{4m} N \int_\omega d\mathbf{r}\, \nabla \cdot \nabla \rho \tag{30}$$

Using Gauss's theorem, the volume integral in Eq. (30) can be transformed into a surface integral:

$$K(\omega) = G(\omega) - \frac{\hbar^2}{4m} N \int dS(\omega, \mathbf{r}) \nabla \rho \cdot \mathbf{n}(\mathbf{r}) \tag{31}$$

Now it is clear that the second term in the R.H.S. will vanish only for systems bounded by a zero-flux surface satisfying Eq. (6) (or for the whole system, because the Laplacian integrated over the entire space also vanishes). Thus only the total system and proper open sub-systems will have a definite kinetic energy. A proper open system (one bounded by a zero-flux surface and/or infinity) will be referred to as Ω to distinguish it from an arbitrary bounded region of space ω. For such a proper open system one has:

$$K(\Omega) = G(\Omega) = T(\Omega) \tag{32}$$

and, because the integral of the Laplacian vanishes over Ω, the integral of the local statement of the virial theorem (Eq. 10) over Ω yields the *atomic virial theorem*:

$$-2T(\Omega) = \mathscr{V}(\Omega) \tag{33}$$

where the $\mathscr{V}(\Omega)$ is the total atomic virial.

The atomic electronic energy $E_e(\Omega)$ is given by:

$$E_e(\Omega) = T(\Omega) + \mathscr{V}(\Omega) \tag{34}$$

For systems in equilibrium there are no Hellmann–Feynman forces acting on the nuclei and the virial equals the average potential energy of the molecule, i.e. $\mathscr{V} = V$. Under this condition Eq. (33) becomes:

$$-2T(\Omega) = V(\Omega) \tag{35}$$

where $V(\Omega)$ is the potential energy of atom Ω, and Eq. (34) becomes:

$$E(\Omega) = E_e(\Omega) = T(\Omega) + V(\Omega) = -T(\Omega) = \frac{1}{2}V(\Omega) \tag{36}$$

where $E(\Omega)$ is the total energy of atom Ω.

Thus, the energy of an atom in a molecule at its equilibrium geometry is obtained from the atomic statement of the virial theorem, and $E(\Omega) = -T(\Omega)$. The sum of atomic energies yields, naturally, the total energy of the molecule (obtained directly from the electronic structure calculation) to within a small numerical integration error. This additivity of the atomic energies is expressed as:

$$E_{total} = \sum_{\Omega} E(\Omega) \tag{37}$$

The result shown in Eq. (37) is remarkable. The equation expresses the partitioning of the total molecular energy into atomic contributions, a partitioning which includes, for example, the nuclear–nuclear repulsion contribution to the

molecular energy. Such partitioning of the total energy is indispensable if one is to understand the atomic origins of the energy difference between two isomers [4, 16, 53], for example, or the atomic origins of potential energy barriers [17, 54].

The deviation of the sum of the atomic energies from the directly calculated total molecular energy is another global measure of the quality of atomic integrations. A deviation of no more than ca. 1 kcal mol^{-1} is usually regarded as an indicator of accurate integrations.

The discussion above is based upon the assumption that the calculated molecular wavefunction satisfies the virial theorem exactly, i.e. the molecular virial ratio $-V/T = 2$ to infinite accuracy. In practice, the calculated virial ratio deviates slightly from this ideal value of 2 because of the truncation of the basis set, residual forces on the nuclei, and the finite nature of the convergence thresholds in a typical calculation. The manner by which AIMPAC corrects for this deviation is described in the Appendix.

1.8.6
Atomic Dipolar Polarization [μ(Ω)]

Also known as the *first atomic electrostatic moment*, atomic dipolar polarization is the atomic space average of the electronic position vector. It is a three-dimensional vector with components and magnitude defined in Eqs (38) and (39), respectively:

$$\boldsymbol{\mu}(\Omega) = \begin{pmatrix} \mu_x \\ \mu_y \\ \mu_z \end{pmatrix} = \begin{pmatrix} -e\int_\Omega x\rho(\mathbf{r})\,d\mathbf{r} \\ -e\int_\Omega y\rho(\mathbf{r})\,d\mathbf{r} \\ -e\int_\Omega z\rho(\mathbf{r})\,d\mathbf{r} \end{pmatrix} \equiv -e\int_\Omega \mathbf{r}_\Omega\rho(\mathbf{r})\,d\mathbf{r} \tag{38}$$

$$|\boldsymbol{\mu}(\Omega)| = \sqrt{{\mu_x}^2 + {\mu_y}^2 + {\mu_z}^2} \tag{39}$$

with the origin for the vector $\mathbf{r}_\Omega$ at the nucleus of atom Ω, i.e. $\mathbf{r}_\Omega = \mathbf{r} - \mathbf{R}_\Omega$, $\mathbf{r}$ being the electronic coordinates and $\mathbf{R}_\Omega$ the nuclear coordinates of atom Ω. The first moment measures the polarization of the charge density, that is to say the departure from sphericity of the electron density.

The dipolar polarization can be used to understand the origins of permanent and induced molecular dipole moments and dielectric polarization in materials [55, 56]. Carbon monoxide is an instructive example [57]. The dipole moment of the CO molecule has the (unexpected) polarity $^{\delta-}C{=}O^{\delta+}$, opposite to intuition based on the relative electronegativities of carbon and oxygen. This observation is readily explained when one considers both atomic charges and atomic dipoles. Calculated atomic charges are indeed in accordance with the expected relative electronegativities of these two atoms (an electronegativity of 2.5 for carbon and

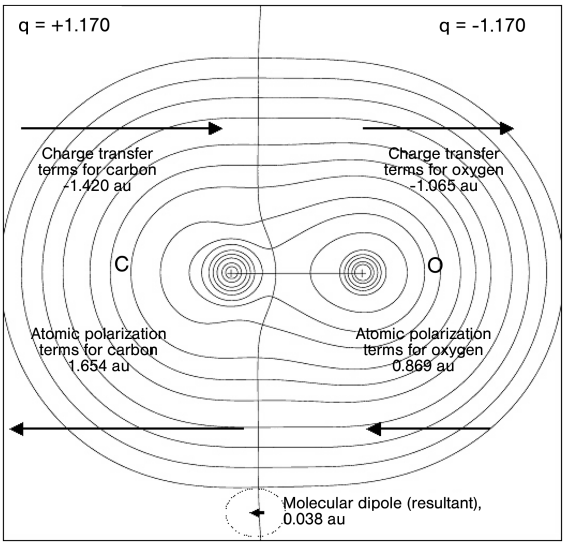

Fig. 1.5 Contour plot of the electron density of CO, showing the magnitudes and directions of atomic and charge-transfer dipoles (arrow lengths are proportional to the dipoles magnitudes). The head of an arrow points to the negative end of a dipole. The molecular dipole moment is given by the vector sum of the charge-transfer terms ($\boldsymbol{\mu}_{CT}$) and the atomic polarizations ($\boldsymbol{\mu}_{AP}$). The directly calculated SCF molecular dipole is 0.096 debyes (D) at the B3LYP/6-311+G(3df) and the corresponding dipole obtained from group contributions is 0.096 D, (experimental: 0.110 D). (Reproduced from Ref. [57] with the permission from the *American Chemical Society*).

3.5 for oxygen) [58] with the carbon bearing a positive charge [$q(\mathrm{C}) = +1.17$ au] and the oxygen a negative charge, resulting in a *charge-transfer dipole* with the direction $^{+1.17}\overrightarrow{\mathrm{C{=}O}}^{-1.17}$. The electron density of each of the two atomic basins responds to this charge-transfer dipole with an opposing dipolar polarization $\overleftarrow{\mathrm{C{=}O}}$ with a magnitude which is not only sufficient to cancel the charge-transfer dipole but to slightly exceed it. The net result is a small dipole in complete accordance with the unexpected experimental result (Fig. 1.5). Thus, it is necessary to take the vectorial sum of both the charge transfer and the atomic polarization dipoles in defining atomic or group contributions to the molecular dipole moment [55, 56].

A program, FRAGDIP [59], is available for calculation of additive atomic and group contributions to the molecular dipole moment. As an illustration, Table 1.1 lists the group contributions to the dipole moments of several naturally occurring amino acids with their vector sum and compares this sum with the dipole moments calculated directly from an SCF calculation (second line). Each amino acid in its neutral form, with general formula R–CH(NH_2)COOH, was regarded as consisting of two groups – the side-chain (R–) and the "main chain" (–CH(NH_2)COOH). The reader can see how closely the group contributions sum to the molecular dipole.

Table 1.1 Comparison of the molecular dipole moments of some amino acids obtained from group contributions with those obtained directly from SCF calculations.[a,b] The table also compares the second letter of the genetic code of each amino acid with side-chain dipole. (Adapted from Refs. [56, 60]).

Amino acid	Second base in the mRNA genetic codon[c]	Nature of the 2nd base[d] in the codon	Side-chain dipole[a,b] (au) R–				Main-chain dipole[a,b] (au) –CH(NH_2)COOH				Total molecular dipole[a,b] (au) R–CH(NH_2)COOH			
			μ_x	μ_y	μ_z	$\lvert\mu\rvert$	μ_x	μ_y	μ_z	$\lvert\mu\rvert$	μ_x	μ_y	μ_z	$\lvert\mu\rvert$
Gly	G	I	−0.006	0.068	−0.098	0.119	0.415	−0.376	0.095	0.568	0.409	−0.307	−0.002	0.512
											0.402	−0.317	−0.001	0.512
Ile	U	II	0.130	0.012	−0.074	0.150	−0.363	0.521	−0.267	0.689	−0.233	0.533	−0.341	0.674
											−0.207	0.537	−0.269	0.635
Ala	C	II	0.034	−0.108	−0.120	0.165	0.217	0.590	−0.073	0.632	0.251	0.481	−0.193	0.576
											0.252	0.480	−0.193	0.576
Val	U	II	−0.103	0.149	−0.062	0.191	0.097	−0.621	−0.266	0.683	−0.006	−0.473	−0.328	0.575
											−0.013	−0.473	−0.319	0.571
Leu	U	II	0.069	0.182	0.081	0.211	−0.195	−0.607	0.163	0.658	−0.127	−0.424	0.244	0.506
											−0.119	−0.412	0.264	0.503
Phe	U	II	0.217	−0.142	−0.118	0.285	−0.099	0.642	−0.057	0.652	0.118	0.501	−0.175	0.543
											0.077	0.495	−0.187	0.535
Tyr	A	I	−0.292	−0.401	−0.201	0.536	0.138	−0.622	−0.065	0.641	−0.154	−1.024	−0.266	1.069
											−0.137	−1.026	−0.272	1.071
Thr	C	II	0.399	−0.259	−0.548	0.726	0.114	−0.645	0.065	0.658	0.514	−0.904	−0.483	1.147
											0.515	−0.904	−0.479	1.146
Met	U	II	−0.377	−0.636	−0.327	0.809	−0.288	−0.005	−0.382	0.479	−0.665	−0.642	−0.709	1.165
											−0.661	−0.647	−0.695	1.157

Trp	G	I	0.354	0.727	−0.140	0.820	−0.084	0.498	−0.173	0.534	0.270	1.225	−0.313	1.293
											0.245	1.233	−0.315	1.296
Cys	G	I	−0.692	0.090	0.450	0.830	−0.366	0.017	−0.479	0.603	−1.058	0.107	−0.029	1.064
											−1.047	0.118	−0.033	1.055
Ser	G/C[e]	I/II[e]	−0.823	−0.091	0.174	0.847	0.390	−0.428	−0.047	0.581	−0.433	−0.519	0.127	0.688
											−0.441	−0.528	0.130	0.700
Gln	A	I	−0.206	−1.296	−1.030	1.668	−0.056	0.436	0.695	0.823	−0.261	−0.860	−0.335	0.959
											−0.255	−0.894	−0.339	0.989
His	A	I	−1.510	−0.722	0.782	1.848	−0.380	0.097	−0.487	0.625	−1.890	−0.625	0.295	2.013
											−1.860	−0.637	0.303	1.989
Asn	A	I	−0.963	1.631	−0.197	1.904	−0.019	−0.651	−0.249	0.697	−0.982	0.980	−0.446	1.457
											−0.983	0.988	−0.458	1.467

[a] The total molecular dipole of an amino acid symbolized by R–CH(NH_2)COOH is given in the last four columns. For each amino acid, the top line in the last four columns lists the dipole moment obtained from the "side-chain" (R–) and the "main chain" (–CH(NH_2)COOH) group contributions using the FRAGDIP program [55, 56, 59], and the second line lists the corresponding dipole moment calculated directly from the SCF using the Gaussian program [61].

[b] Electronic structure calculations were performed at the HF/6-311++G(d,p)//HF/6-31+G(d) level of theory.

[c] The second base in the RNA triplet genetic code of the amino acid: A = adenine, C = cytosine, G = guanine, U = uracil.

[d] The chemical nature of the second base in the triplet code: I = purine (adenine and guanine), and II = pyrimidine (cytosine, uracil).

[e] Serine is the only amino acid which has a degenerate genetic code in the second position, all other amino acids have a unique base in the second position in all of their synonymous codons.

Further, the listings in Table 1.1 have been sorted in terms of the magnitude of the side-chain dipole magnitude, a sorting that reveals a striking regularity in the genetic code. Most amino acids listed in the upper part of the table with side-chain dipole magnitude less than 0.81 au (i.e. with non-polar side-chains) are encoded by a genetic triplet code having a pyrimidine base as the middle letter in the mRNA codon (except glycine, which lacks a side-chain, and tyrosine). On the other hand, most polar amino acids (having side-chain dipole magnitudes greater than 0.81 au) are encoded by a purine base, serine being the only "degenerate" amino acid, having codons of both types [56, 60]. Whereas this regularity in the genetic code has been well known for a long time, it is given a quantitative basis derived directly from the electron density distributions of the amino acids for the first time [60].

1.8.7
Atomic Quadrupolar Polarization [Q(Ω)]

The atomic quadrupolar polarization tensor is also known as the *second atomic electrostatic moment*. It is a symmetric traceless tensor defined as:

$$\mathbf{Q}(\Omega) = \begin{pmatrix} Q_{xx} & Q_{xy} & Q_{xz} \\ Q_{yx} & Q_{yy} & Q_{yz} \\ Q_{zx} & Q_{zy} & Q_{zz} \end{pmatrix}$$

$$\equiv -\frac{e}{2} \begin{pmatrix} \int_\Omega (3x_\Omega^2 - r_\Omega)\rho(\mathbf{r})\, d\mathbf{r} & 3\int_\Omega x_\Omega y_\Omega \rho(\mathbf{r})\, d\mathbf{r} & 3\int_\Omega x_\Omega z_\Omega \rho(\mathbf{r})\, d\mathbf{r} \\ 3\int_\Omega y_\Omega x_\Omega \rho(\mathbf{r})\, d\mathbf{r} & \int_\Omega (3y_\Omega^2 - r_\Omega)\rho(\mathbf{r})\, d\mathbf{r} & \int_\Omega y_\Omega z_\Omega \rho(\mathbf{r})\, d\mathbf{r} \\ 3\int_\Omega z_\Omega x_\Omega \rho(\mathbf{r})\, d\mathbf{r} & \int_\Omega z_\Omega y_\Omega \rho(\mathbf{r})\, d\mathbf{r} & \int_\Omega (3z_\Omega^2 - r_\Omega)\rho(\mathbf{r})\, d\mathbf{r} \end{pmatrix} \tag{40}$$

where, as for the first moment, the origin is placed at the nucleus. If the atomic electron density has spherical symmetry, then $\int_\Omega x_\Omega^2 \rho(\mathbf{r})\, d\mathbf{r} = \int_\Omega y_\Omega^2 \rho(\mathbf{r})\, d\mathbf{r} = \int_\Omega z_\Omega^2 \rho(\mathbf{r})\, d\mathbf{r} = \frac{1}{3}\int_\Omega \mathbf{r}_\Omega^2 \rho(\mathbf{r})\, d\mathbf{r}$, and $Q_{xx} = Q_{yy} = Q_{zz} = 0$. Thus, the quadrupole moment is another measure of the deviation of the atomic electron density from spherical symmetry. For example, if a diagonal component of $\mathbf{Q}(\Omega)$ is <0, the electron density is concentrated along that axis, and vice versa. It is always possible to find a coordinate system such that the original tensor in Eq. (40) [$\mathbf{Q}(\Omega)$] is diagonalized [$\mathscr{Q}(\Omega)$]. The diagonalization of $\mathbf{Q}(\Omega)$ corresponds to a rotation of the original coordinate system. The diagonalized quadrupole tensor corresponding to Eq. (40) is written:

$$\mathscr{Q}(\Omega) = \begin{pmatrix} \mathscr{Q}_{x'x'} & 0 & 0 \\ 0 & \mathscr{Q}_{y'y'} & 0 \\ 0 & 0 & \mathscr{Q}_{z'z'} \end{pmatrix} \tag{41}$$

where $\mathcal{Q}_{x'x'}$, $\mathcal{Q}_{y'y'}$, and $\mathcal{Q}_{z'z'}$ are the principal values of the quadrupole moment with regard to the principal (rotated) axes, the x', y', and z' axes, which correspond to axes of symmetry if they exist in the electron density distribution (the primes will be dropped for simplicity).

The traceless property of the tensor defined in Eq. (40) (or in its diagonalized form, Eq. 41) is a consequence of the equality:

$$\mathbf{r}_\Omega^2 = x_\Omega^2 + y_\Omega^2 + z_\Omega^2 \tag{42}$$

which is always true in any coordinate system. Therefore:

$$(Q_{xx} + Q_{yy} + Q_{zz}) = (\mathcal{Q}_{xx} + \mathcal{Q}_{yy} + \mathcal{Q}_{zz}) = 0 \tag{43}$$

and only five independent components completely specify $\mathbf{Q}(\Omega)$ in the original coordinate system and only two are sufficient to specify its diagonalized form $\mathcal{Q}(\Omega)$.

Finally, the magnitude of the quadrupolar polarization moment is defined as [62]:

$$|\mathbf{Q}| = \sqrt{\frac{2}{3}(Q_{xx}^2 + Q_{yy}^2 + Q_{zz}^2)} = \sqrt{\frac{2}{3}(\mathcal{Q}_{xx}^2 + \mathcal{Q}_{yy}^2 + \mathcal{Q}_{zz}^2)} \tag{44}$$

1.9
"Practical" Uses and Utility of QTAIM Bond and Atomic Properties

1.9.1
The Use of QTAIM Bond Critical Point Properties

Several QTAIM bond properties have been shown to be correlated with experimental molecular properties. For example, the electron density at the BCP, ρ_b, has been shown on several occasions to be strongly correlated with the bond energies, and hence provide a measure of bond order (Eq. 8) [1, 30]; the potential energy density at the BCP has been shown to be highly correlated with hydrogen bond energies [32]; full interaction potentials in hydrogen bonds were recovered from the potential energy density at the BCP [63]; π–π stacking interactions in benzene dimers and in stacked DNA bases and base-pairs have been found to be highly correlated to BCP and cage critical point data between π-stacked monomers [64–66].

The use of BCP properties in drug design is a field pioneered by Popelier and coworkers. These authors proposed the construction of a vector space from bond properties evaluated at the bond critical points, i.e. a point in this space is specified by a set of bond properties [67–70]. This space was used as a basis for comparing related molecules, the smaller the distance between two molecules in this

space the more they are similar. Quantification of molecular similarity in this manner has several advantages over other similarity measures (for example Carbo's similarity index [71]):

1. it is much faster because it involves no spatial integration (the density of each molecule is only sampled at the positions of the BCPs);
2. it is not dominated by nuclear maxima but rather emphasizes the more interesting chemical bonding regions of the molecule; and
3. it is not plagued with the alignment problem, in which one must often choose how to align the molecules to be compared before the integration.

The new method has been successful in accurately predicting a number of properties of several series of molecules [67–70].

1.9.2 The Use of QTAIM Atomic Properties

The review in this section follows closely Table 1 of Ref. [51].

Atomic properties have been used to recover and *directly* predict several experimentally additive atomic and group contributions to molecular properties, including, for example, heats of formation [72], magnetic susceptibility (Refs [73–77] and Chapter 3 in this book), molecular volumes [78], electric moments (Chapter 3) and polarizability [79–81], Raman intensities [79, 81–84] (see also Chapter 4), IR intensities [85–88] (see also Chapter 4), spectroscopic transition probabilities [89], dielectric polarization in crystals and molecular dipole and quadrupole moments [55, 90, 91], Wigner–Seitz cells in crystals [92], group additivity in silanes [93], and Pascal's aromatic exaltations [1]. They have also been used to provide an atomic basis for electron localization and delocalization [42, 43, 47, 94, 95].

Atomic properties have also been used *empirically* to predict several experimental properties including for example, the pK_a of weak acids from the atomic energy of the acidic hydrogen [96], a wide array of biological and physicochemical properties of the amino acids, including the genetic code itself, and the effects of mutation on protein stability [60], protein retention times [97], HPLC column capacity factors of high-energy materials [98], NMR spin–spin coupling constants from the electron delocalization indices [99, 100], simultaneous consistent prediction of five bulk properties of liquid HF in MD simulation [101], classification of atom types in proteins with future potential applications in force-field design [60, 102–104], reconstructing large molecules from transferable fragments or atoms in molecules [60, 105–119] (see also Chapters 11 and 12), atomic partitioning of the molecular electrostatic potential [120–122], prediction of hydrogen-bond donor capacity [123] and basicity [124], and to provide an atomic basis for curvature-induced polarization in carbon nanotubes and nanoshells [125].

1.10 Steps of a Typical QTAIM Calculation

It should be clear from the outset that QTAIM applies equally well to experimental [2, 126] and calculated electron densities [1, 3]; in this tutorial, however, we will discuss calculated densities.

The starting point for the application of the QTAIM theory is the electron density. The density can be calculated from the many-electron single-determinant or many-determinant wavefunction (or Slater-like determinants built from Kohn–Sham orbitals in density functional theory [127]) obtained by a variety of methods and software. The electron density necessary for meaningful analysis by means of QTAIM must be obtained with a basis set flexible enough for an accurate representation of the bonding regions, in other words it must include polarization functions. In the case of anions, excited states, and weak bonding interactions between atoms separated by relatively large distances one must augment the basis set with diffuse functions. When heavy atoms are present in the molecule, which usually necessitates the use of effective core potentials (ECP), it is necessary to treat the valence shell and at least one sub-valence shell explicitly to obtain meaningful results from the integrations. It is important to note that bond paths cannot be traced to the nuclei of atoms described by ECPs. Alternatively, often the geometry is optimized with a basis set including the ECP on the heavy atoms followed by a single point calculation at the optimized geometry using a full basis set on all atoms.

The first step in a molecular QTAIM calculation is, thus, the generation of a wavefunction (or wavefunction-like single determinant in a DFT calculation [127]) from an electronic structure calculation with software such as Gaussian [61] or GAMESS [128].

The electron density derived from the wavefunction is then subjected to a point-by-point topological analysis to locate the bond critical points and the bond paths, by use of software such as EXTREME [129–131]. The space of the molecule is then partitioned by the zero-flux surfaces and atomic integrations are performed to obtain the atomic contributions to the molecular properties using software such as PROAIM and its variants [129–131]. EXTREME and PROAIM are both part of the AIMPAC suite of programs developed in Professor Bader's laboratory (*McMaster University*) [129–131].

Several other software packages derived from AIMPAC are available for analysis of the electron density according to QTAIM. Among the widely used programs are MORPHY [132], developed by Dr Paul Popelier's group (*University of Manchester*), and AIM2000 [133–135], developed by Professor Friederich Biegler–König (*University of Bielfeld*), both of which apply to molecular calculations. The program TOPOND [136] (described in Chapter 7) was developed by Dr Carlo Gatti (*National Research Council of Italy*) for the analysis of periodic densities obtained from CRYSTAL [137].

After all atomic integrations have been performed to the desired accuracy (as measured by the value of the integrated Laplacian) one typically uses shell scripts

and/or simple UNIX/Linux commands such as "grep" to extract the relevant information from the electronic integration files. In this manner the summarized results can be further imported to a spreadsheet or a plotting program. Integration files – which contain the atomic overlap matrices – can be subsequently analyzed by software such as AIMDELOC [44] or LI-DICALC [45, 46] to obtain the localization and delocalization indices.

Further, the wavefunction files can be used as input to plotting routines. GRDVEC can be used to generate two-dimensional plots of the gradient vector field and/or the interatomic surfaces and bond paths projected on a plane selected by the user (right half of Fig. 1.3a). Contour diagrams of the density (such as those in the left half of Fig. 1.3a, and Fig. 1.5), the Laplacian, or energy densities can be generated by first calculating the corresponding grid by the use of GRIDV software followed by the generation of the graphics file from the grid

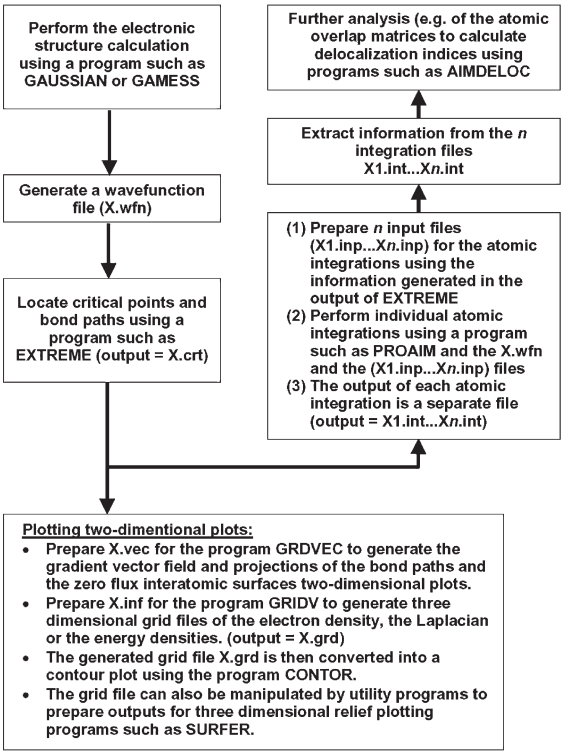

Fig. 1.6 The main steps in a simple QTAIM calculation. The software cited in the figure is part of the AIMPAC suite of programs [129–131]. Other programs are available that can perform most or all of these steps, including, for example, AIM2000 [133–135], MORPHY [132], and AIMALL97 [139].

by using a program such as CONTOR. (GRDVEC, GRIDV, and CONTOR are components of AIMPAC [129–131]).

The grid generated by GRIDV can be manipulated by utility programs such as GridV_REFORMATTER (available from the authors) to generate inputs for programs such as Surfer [138] which produce three-dimensional relief maps of the field represented by the calculated grid (Fig. 1.1b is an example).

The main steps of a typical QTAIM calculation are summarized in Fig. 1.6.

Appendix: The Inexact Satisfaction of the Molecular Virial Theorem in Electronic Structure Calculations

For a molecule in an equilibrium geometry (with vanishing forces on the nuclei), the molecular virial theorem is expressed as:

$$\gamma = -\frac{V}{T} = 2 \tag{45}$$

Because of the propagation of numerical errors, small (but non-vanishing) thresholds of convergence of both the SCF and the geometry optimization steps, and the use of incomplete basis sets, electronic structure calculations do not usually satisfy the virial theorem exactly and the virial ratio (γ) can deviate by perhaps as much as 0.01 from the ideal value of 2. As a result of this deviation, atomic energies will not sum to yield the molecular energy with acceptable accuracy.

Atomic integration software such as PROAIM [129–131] correct for this error numerically. Thus, instead of simply multiplying each atomic kinetic energy $T(\Omega)$ by (-1) to obtain the total atomic energy $E(\Omega)$, the latter is obtained by multiplying $T(\Omega)$ by $(1-\gamma)$. These corrected atomic energies do satisfy Eq. (37), and their sum equals the total molecular energy to within a small numerical *integration* error. The virial corrections usually scale linearly with regard to $T(\Omega)$ which, fortunately, leaves the relative stabilities of the atoms unchanged.

The integration software obtains the virial ratio from the wavefunction files generated by Gaussian [61] or GAMESS [128]. The virial is printed in the last line in the wavefunction file. For Hartree–Fock or density functional calculations, Gaussian prints the correct virial in the wavefunction file and the integrations proceed without problems. For wavefunction files calculated at a post Hartree–Fock level, for example those obtained using Møller–Plesset perturbation theory (MP*n*) or configuration interaction methods (CI), the virial printed in the wavefunction file generated using the Gaussian 98 or 03 [61] programs (which are available at the time of writing) is *the Hartree–Fock virial and not that of the current post-Hartree–Fock method* [even if the key word "DENSITY = CURRENT" is invoked and despite the fact that the correct (current) wavefunction is printed]. If such a wavefunction file is fed directly to an integration program, the calculated atomic energies will be rectified using the Hartree–Fock γ (instead of the post Hartree–Fock γ), resulting in atomic energies which do not add up to the molecular value.

In these circumstances the user must calculate the virial of the current method ''by hand'' from information contained in the Gaussian ''log'' or ''out'' output file [140], by dividing, for example, the MP2 (or other correlated total energy) by the kinetic energy listed just after the final electrical multipoles in the Gaussian output. The wavefunction files must then be edited to reflect this new ''correct'' virial before submitting it to the integration software [140].

In highly accurate calculations it is sometimes necessary to perform atomic integrations of energy densities obtained from systems which satisfy the molecular virial theorem exactly [16, 141]. The author of Chapter 3 of this book, Dr Todd A. Keith, has written a link [142] for Gaussian [61] implementing Löwdin's self-consistent virial scaling (SCVS) [143, 144] which produces final wavefunctions satisfying the virial theorem to a very high accuracy.

Acknowledgments

We thank Professor George Heard (University of North Carolina at Asheville) and Dr Katherine N. Robertson (Dalhousie University) for their helpful comments on this chapter, and Dr Todd Keith (Semichem, Inc.) and Dr Jamie Platts (Cardiff University) for discussions on the virial correction. The American Chemical Society is acknowledged for its permission to reproduce Fig. 1.5.

References

1 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press: Oxford, U.K., 1990.
2 P. Coppens, *X-ray Charge Densities and Chemical Bonding*, Oxford University Press, Inc.: New York, 1997.
3 P. L. A. Popelier, *Atoms in Molecules: An Introduction*, Prentice Hall: London, 2000.
4 C. F. Matta, N. Castillo, R. J. Boyd, *J. Phys. Chem. A* **2005**, *109*, 3669–3681.
5 N. Castillo, C. F. Matta, R. J. Boyd, *Chem. Phys. Lett.* **2005**, *409*, 265–269.
6 C. Gatti, P. Fantucci, G. Pacchioni, *Theor. Chem. Acc. (Formerly, Theoret. Chim. Acta)* **1987**, *72*, 433–458.
7 W. L. Cao, C. Gatti, P. J. MacDougall, R. F. W. Bader, *Chem. Phys. Lett.* **1987**, *141*, 380–385.
8 M. Sakata, *Acta Cryst. A* **1990**, *46*, 263–270.
9 R. Y. de Vries, W. J. Briels, D. Feil, G. te Velde, E. J. Baerends, *Can. J. Chem.* **1996**, *74*, 1054–1058.
10 A. Taylor, C. F. Matta, R. J. Boyd, submitted for publication **2006**.
11 R. F. W. Bader, J. A. Platts, *J. Chem. Phys.* **1997**, *107*, 8545–8553.
12 R. F. W. Bader, *Phys. Rev. B* **1994**, *49*, 13348–13356.
13 J. Schwinger, *Phys. Rev.* **1951**, *82*, 914–927.
14 R. F. W. Bader, *J. Phys. Chem. A* **1998**, *102*, 7314–7323.
15 T. A. Keith, R. F. W. Bader, Y. Aray, *Int. J. Quantum Chem.* **1996**, *57*, 183–198.
16 C. F. Matta, J. Hernández-Trujillo, T. H. Tang, R. F. W. Bader, *Chem. Eur. J.* **2003**, *9*, 1940–1951.
17 C. F. Matta, Chapter 9 in: *Hydrogen Bonding – New Insight, (S. J. Grabowski, Ed.)*, Springer: 2006, pp 337–376.
18 R. F. W. Bader, C. F. Matta, *J. Phys. Chem. A* **2004**, *108*, 8385–8394.
19 R. P. Sagar, A. C. T. Ku, V. H. Jr. Smith, A. M. Simas, *J. Chem. Phys.* **1988**, *88*, 4367–4374.

20 Z. Shi, R. J. Boyd, *J. Chem. Phys.* **1988**, *88*, 4375–4377.

21 R. W. F. Bader, G. L. Heard, *J. Chem. Phys.* **1999**, *111*, 8789–8797.

22 R. J. Gillespie, R. S. Nyholm, *Quart. Rev. Chem. Soc.* **1957**, *11*, 339.

23 R. J. Gillespie, R. S. Nyholm, in: *Progress in Stereochemistry* (W. Klyne, P. B. D. de la Mare, Eds.), Butterworths: London, 1958.

24 R. J. Gillespie, I. Hargittai, *The VSEPR Model of Molecular Geometry*, Allyn and Bacon: Boston, 1991.

25 R. F. W. Bader, P. J. MacDougall, C. D. H. Lau, *J. Am. Chem. Soc.* **1984**, *106*, 1594–1605.

26 R. F. W. Bader, R. J. Gillespie, P. J. MacDougall, *J. Am. Chem. Soc.* **1988**, *110*, 7329–7336.

27 R. J. Gillespie, I. Bytheway, T.-H. Tang, R. F. W. Bader, *Inorg. Chem.* **1996**, *35*, 3954–3963.

28 M. T. Carroll, C. Chang, R. F. W. Bader, *Mol. Phys.* **1988**, *63*, 387–405.

29 M. T. Carroll, J. R. Cheeseman, R. Osman, H. Weinstein, *J. Phys. Chem.* **1989**, *93*, 5120–5123.

30 R. J. Boyd, S. C. Choi, *Chem. Phys. Lett.* **1986**, *129*, 62–65.

31 M. T. Carroll, R. F. W. Bader, *Mol. Phys.* **1988**, *65*, 695–722.

32 E. Espinosa, E. Molins, C. Lecomte, *Chem. Phys. Lett.* **1998**, *285*, 170–173.

33 S. J. Grabowski, *J. Phys. Chem. A* **2001**, *105*, 10739–10746.

34 M. Domagala, S. Grabowski, K. Urbaniak, G. Mloston, *J. Phys. Chem. A* **2003**, *107*, 2730–2736.

35 S. Grabowski, W. A. Sokalski, J. Leszczynski, *J. Phys. Chem. A* **2005**, *109*, 4331–4341.

36 M. Domagala, S. Grabowski, *J. Phys. Chem. A* **2005**, *109*, 5683–5688.

37 O. Knop, R. J. Boyd, S. C. Choi, *J. Am. Chem. Soc.* **1988**, *110*, 7299–7301.

38 J. L. Jules, J. R. Lombardi, *J. Mol. Struct. (Theochem)* **664–665**, *2003*, 255–271.

39 S. T. Howard, O. Lamarche, *J. Phys. Org. Chem.* **2003**, *16*, 133–141.

40 R. F. W. Bader, T. T. Nguyen-Dang, *Adv. Quantum Chem.* **1981**, *14*, 63–124.

41 D. Cremer, E. Kraka, *Angew. Chem. Int. Ed. Engl.* **1984**, *23*, 627–628.

42 X. Fradera, M. A. Austen, R. F. W. Bader, *J. Phys. Chem. A* **1999**, *103*, 304–314.

43 R. F. W. Bader, M. E. Stephens, *J. Am. Chem. Soc.* **1975**, *97*, 7391–7399.

44 C. F. Matta, *AIMDELOC (QCPE 0802)* Quantum Chemistry Program Exchange, Indiana University, 2001. (http://qcpe.chem.indiana.edu/).

45 Y.-G. Wang, C. F. Matta, N. H. Werstiuk, *J. Comput. Chem.* **2003**, *24*, 1720–1729.

46 Y.-G. Wang, N. H. Werstiuk, *J. Comput. Chem.* **2003**, *24*, 379–385.

47 M. A. Austen, *A New Procedure for Determining Bond Orders in Polar Molecules, with Applications to Phosphorus and Nitrogen Containing Systems, Ph.D. Thesis*, McMaster University: Hamilton, Canada, 2003.

48 C. F. Matta, J. Hernández-Trujillo, *J. Phys. Chem. A* **2003**, *107*, 7496–7504 (Correction: *J. Phys. Chem. A* **2005**, *109*, 10798).

49 E. A. Zhurova, C. F. Matta, N. Wu, V. V. Zhurov, A. A. Pinkerton, *J. Am. Chem. Soc.* **2006**, *128*, 8849–8861.

50 R. F. W. Bader, P. F. Zou, *Chem. Phys. Lett.* **1992**, *191*, 54–58.

51 C. F. Matta, R. F. W. Bader, *J. Phys. Chem. A* **2006**, *110*, 6365–6371.

52 (a) L. Cohen, *J. Chem. Phys.* **1979**, *70*, 788–789; (b) L. Cohen, *J. Chem. Phys.* **1984**, *80*, 4277–4279.

53 F. Cortés-Guzmán, J. Hernández-Trujillo, G. Cuevas, *J. Phys. Chem. A* **2003**, *107*, 9253–9256.

54 R. F. W. Bader, J. R. Cheeseman, K. E. Laidig, K. B. Wiberg, C. Breneman, *J. Am. Chem. Soc.* **1990**, *112*, 6530–6536.

55 R. F. W. Bader, C. F. Matta, *Int. J. Quantum Chem.* **2001**, *85*, 592–607.

56 C. F. Matta, *Applications of the Quantum Theory of Atoms in Molecules to Chemical and Biochemical Problems, Ph.D. Thesis*, McMaster University: Hamilton, Canada, 2002. (Available on line, http://chem.utoronto.ca/~cmatta/).

57 C. F. Matta, R. J. Gillespie, *J. Chem. Educ.* **2002**, *79*, 1141–1152.

58 L. Pauling, *The Nature of the Chemical Bond, (Third Ed.)*, Cornell University Press: Ithaca, N.Y., 1960.

59 C. F. Matta, *FRAGDIP (QCPE 0801)* Quantum Chemistry Program Exchange, Indiana University, 2001. (http://qcpe.chem.indiana.edu/).

60 C. F. Matta, R. F. W. Bader, *Proteins: Struct. Funct. Genet.* **2003**, *52*, 360–399.

61 M. J. Frisch, *et al.*, Gaussian 03, Gaussian Inc.: Pittsburgh PA, 2003.

62 C. G. Gray, K. E. Gubbins, *Theory of Molecular Fluids, (Vol. 1)*, Clarendon Press: Oxford, 1984.

63 E. Espinosa, E. Molins, *J. Chem. Phys.* **2000**, *113*, 5686–5694.

64 O. A. Zhikol, O. Shishkin, K. A. Lyssenko, J. Leszczynski, *J. Chem. Phys.* **2005**, *122*, 144104-1–144104-8.

65 M. P. Waller , A. Robertazzi, J. A. Platts, D. E. Hibbs, P. A. Williams, *J. Comput. Chem.* **2006**, *27*, 491–504.

66 C. F. Matta, N. Castillo, R. J. Boyd, *J. Phys. Chem. B* **2006**, *110*, 563–578.

67 S. E. O'Brien, P. L. A. Popelier, *Can. J. Chem.* **1999**, *77*, 28–36.

68 P. L. A. Popelier, *J. Phys. Chem. A* **1999**, *103*, 2883–2890.

69 S. E. O'Brien, P. L. A. Popelier, *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 764–775.

70 U. A. Chaudry, P. L. A. Popelier, *J. Org. Chem.* **2004**, *69*, 233–241.

71 R. Carbó, L. Leyda, M. Arnau, *Int. J. Quantum Chem.* **1980**, *17*, 1185–1189.

72 K. B. Wiberg, R. F. W. Bader, C. D. H. Lau, *J. Am. Chem. Soc.* **1987**, *109*, 1001–1012.

73 T. A. Keith, R. F. W. Bader, *Chem. Phys. Lett.* **1992**, *194*, 1–8.

74 T. A. Keith, R. F. W. Bader, *Int. J. Quantum Chem.* **1996**, *60*, 373–379.

75 R. F. W. Bader, T. A. Keith, *J. Chem. Phys.* **1993**, *99*, 3683–3693.

76 T. A. Keith, R. F. W. Bader, *Chem. Phys. Lett.* **1993**, *210*, 223–231.

77 T. A. Keith, R. F. W. Bader, *J. Chem. Phys.* **1993**, *99*, 3669–3682.

78 R. F. W. Bader, M. T. Carroll, J. R. Cheeseman, C. Chang, *J. Am. Chem. Soc.* **1987**, *109*, 7968–7979.

79 R. F. W. Bader, T. A. Keith, K. M. Gough, K. E. Laidig, *Mol. Phys.* **1992**, *75*, 1167–1189.

80 K. E. Laidig, *Can. J. Chem.* **1996**, *74*, 1131–1138.

81 K. M. Gough, M. M. Yacowar, R. H. Cleve, J. R. Dwyer, *Can. J. Chem.* **1996**, *74*, 1139–1144.

82 K. M. Gough, H. K. Srivastava, K. Belohorcová, *J. Chem. Phys.* **1993**, *98*, 9669–9677.

83 K. M. Gough, H. K. Srivastava, K. Belohorcová, *J. Phys. Chem.* **1994**, *98*, 771–776.

84 K. M. Gough, H. K. Srivastava, *J. Phys. Chem.* **1996**, *100*, 5210–5216.

85 R. L. A. Haiduke, A. E. de Oliveira, R. E. Bruns, *J. Phys. Chem. A* **2004**, *108*, 6788–6796.

86 J. V. da Silva, R. L. A. Haiduke, R. E. Bruns, *J. Phys. Chem. A* **2006**, *110*, 4839–4845.

87 P. H. César, S. H. D. M. Faria, J. V. da Silva Jr., R. L. A. Haiduke, R. E. Bruns, *Chem. Phys.* **2005**, *317*, 35–42 .

88 R. L. A. Haiduke, R. E. Bruns, *J. Phys. Chem. A* **2005**, *109*, 2680–2688.

89 R. F. W. Bader, D. Bayles, G. L. Heard, *J. Chem. Phys.* **2000**, *112*, 10095–10105.

90 K. E. Laidig, *Chem. Phys. Lett.* **1991**, *185*, 483–489.

91 R. F. W. Bader, *Mol. Phys.* **2002**, *100*, 3333–3344.

92 P. F. Zou, R. F. W. Bader, *Acta Cryst. A* **1994**, *50*, 714–725.

93 R. F. W. Bader, D. Bayles, *J. Phys. Chem. A* **2000**, *104*, 5579–5589.

94 R. F. W. Bader, A. Streitwieser, A. Neuhaus, K. E. Laidig, P. Speers, *J. Am. Chem. Soc.* **1996**, *118*, 4959–4965.

95 R. F. W. Bader, S. Johnson, T.-H. Tang, P. L. A. Popelier, *J. Phys. Chem.* **1996**, *100*, 15398–15415.

96 K. R. Adam, *J. Phys. Chem. A* **2002**, *106*, 11963–11972.

97 M. Song, C. M. Breneman, J. Bi, N. Sukumar, K. P. Bennett, S. Cramer, N. Tugcu, *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1347–1357.

98 C. M. Breneman, M. Rhem, *J. Comput. Chem.* **1997**, *18*, 182–197.

99 C. F. Matta, J. Hernández-Trujillo, R. F. W. Bader, *J. Phys. Chem. A* **2002**, *106*, 7369–7375.

100 N. Castillo, C. F. Matta, R. J. Boyd, *J. Chem. Inf. Mod.* **2005**, *45*, 354–359.

101 S. Y. Liem, P. L. A. Popelier, *J. Chem. Phys.* **2003**, *119*, 4560–4566.

102 H. J. Bohórquez, M. Obregón, C. Cárdenas, E. Llanos, C. Suárez, J. L. Villaveces, M. E. Patarroyo, *J. Phys. Chem. A.* **2003**, *107*, 10090–10097.

103 P. L. A. Popelier, F. M. Aicken, *ChemPysChem* **2003**, *4*, 824–829.

104 P. L. A. Popelier, F. M. Aicken, *J. Am. Chem. Soc.* **2003**, *125*, 1284–1292.

105 C. Chang, R. F. W. Bader, *J. Phys. Chem.* **1992**, *96*, 1654–1662.

106 C. M. Breneman, T. R. Thompson, M. Rhem, M. Dung, *Comput. Chem.* **1995**, *19*, 161–179.

107 C. M. Breneman, L. W. Weber, in: *The application of charge density research to chemistry and drug design (NATO ASI Series)*, (G. A. Jeffrey, J. F. E. Piniella, Eds.) Plenum Press, New York, 1991, pp 357–358.

108 C. F. Matta, R. F. W. Bader, *Proteins: Struct. Funct. Genet.* **2000**, *40*, 310–329.

109 C. F. Matta, R. F. W. Bader, *Proteins: Struct. Funct. Genet.* **2002**, *48*, 519–538.

110 R. F. W. Bader, F. J. Martín, *Can. J. Chem.* **1998**, *76*, 284–291.

111 F. J. Martín, *Theoretical Synthesis of Macromolecules from Transferable Functional Groups, Ph.D. Thesis*, McMaster University: Hamilton, 2001.

112 R. F. W. Bader, C. F. Matta, F. J. Martín, Chapter 7 in: *Medicinal Quantum Chemistry*, (F. Alber, P. Carloni, Eds.) Wiley–VCH: Weinheim, 2003, pp 201–231.

113 V. Pichon-Pesme, C. Lecomte, R. Wiest, M. Bénard, *J. Am. Chem. Soc.* **1992**, *114*, 2713–2715.

114 R. Wiest, V. Pichon-Pesme, M. Bénard, C. Lecomte *J. Phys. Chem.* **1994**, *98*, 1351–1362.

115 V. Pichon-Pesme, C. Lecomte, H. Lachekar, *J. Phys. Chem.* **1995**, *99*, 6242–6250.

116 C. Jelsch, V. Pichon-Pesme, C. Lecomte, A. Aubry, *Acta Cryst. D* **1998**, *54*, 1306–1318.

117 C. F. Matta, *J. Phys. Chem. A* **2001**, *105*, 11088–11101.

118 B. Dittrich, T. Koritsánszky, M. Grosche, W. Scherer, R. Flaig, A. Wagner, H. G. Krane, H. Kessler, C. Riemer, A. M. M. Schreurs, P. Luger, *Acta Cryst. B* **2002**, *58*, 721–727.

119 S. Scheins, M. Messerschmidt, P. Luger, *Acta Cryst. B* **2005**, *61*, 443–448.

120 D. S. Kosov, P. L. A. Popelier, *J. Chem. Phys.* **2000**, *113*, 3969–3974.

121 D. S. Kosov, P. L. A. Popelier, *J. Phys. Chem. A* **2000**, *104*, 7339–7345.

122 P. L. A. Popelier, L. Joubert, D. S. Kosov, *J. Phys. Chem. A* **2001**, *105*, 8254–8261.

123 J. A. Platts, *Phys. Chem. Chem. Phys.* **2000**, *2*, 973–980.

124 J. A. Platts, *Phys. Chem. Chem. Phys.* **2000**, *2*, 3115–3120.

125 T. Dumitrica, C. M. Landis, B. I. Yakobson, *Chem. Phys. Lett.* **2002**, *360*, 182–188.

126 T. S. Koritsanszky, P. Coppens, *Chem. Rev.* **2001**, *101*, 1583–1628.

127 W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory, (Second Edition)*, Wiley–VCH: New York, 2001.

128 M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem.* **1993**, *14*, 1347–1363.

129 Bader, R. F. W., http://www.chemistry.mcmaster.ca/aimpac/.

130 F. W. Biegler-König, T. T. Nguyen-Dang, Y. Tal, R. F. W. Bader, A. J. Duke *J. Phys. B: At. Mol. Phys.* **1981**, *14*, 2739–2751.

131 F. W. Biegler-König, R. F. W. Bader, T.-H. Tang, *J. Comput. Chem.* **1982**, *13*, 317–328.

132 Popelier, P. L. A., MORPHY, UMIST, England, EU, **1998**.

133 F. W. Biegler-König, J. Schönbohm, D. Bayles, *J. Comput. Chem.* **2001**, *22*, 545–559.

134 F. W. Biegler-König, *J. Comput. Chem.* **2000**, *21*, 1040–1048.

135 F. W. Biegler-König, J. Schönbohm, D. Bayles, AIM2000 Website: http://gauss.fh-bielefeld.de/aim2000.
136 C. Gatti, TOPOND, CNR-CSRSRC: Milano, 1998.
137 V. R. Saunders, R. Dovesi, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, N. M. Harrison, K. Doll, B. Civalleri, I. J. Bush, Ph. D'Arco, M. Llunell, *CRYSTAL* 2003.
138 Surfer 08, Golden Software, Inc., Golden, Colorado, USA, 2002.
139 T. A. Keith, *AIMALL97 for Windows*, 1997.
140 J. A. Platts, CCL: MP2 virial value (8 July 2005), http://www.ccl.net/cgi-bin/ccl/message.cgi?2005+07+08+004.
141 F. Cortés-Guzmán, R. F. W. Bader, *Chem. Phys. Lett.* **2003**, *379*, 183–192.
142 T. A. Keith, Link for self-consistent virial scaling (SCVS) in Gaussian 94/98, **1998**.
143 P.-O. Löwdin, *J. Mol. Spectr.* **1959**, *3*, 46–66.
144 D. E. Magnoli, J. R. Murdoch, *Int. J. Quantum Chem.* **1982**, *22*, 1249–1262.