

Chapter 2

On Quantum Chemical Topology

Paul L.A. Popelier

Abstract Quantum Chemical Topology (QCT) is a branch of theoretical chemistry that uses the language of dynamical systems (e.g. attractor, basin, homeomorphism, gradient path/phase curve, separatrix, critical points) to partition chemical systems and characterise them via associated quantitative properties. This methodology can be applied to a variety of quantum mechanical functions, the oldest and most documented one being the electron density. We define and discuss the topological atom, and justify the name topology. Then we define the quantum atom without reference to the topological atom. Subsequently, it turns out that each topological atom is a quantum atom, a property that enables the construction of a topologically inspired force field called QCTFF. We briefly discuss the four primary energy contributions governing this force field under development, and how the machine learning method kriging captures the variation in these energies due to geometrical change. Finally, in a more philosophical style, we advocate falsification in the area of chemical interpretation by means of quantum mechanical tools, introducing the concept of a *non-question*.

2.1 Introduction

Recently a chapter on the “Quantum Theory of Atoms in Molecules (QTAIM)” [1] was commissioned by editors Frenking and Shaik for their book on fundamental aspects of chemical bonding. This detailed and lengthy chapter has meanwhile been published [2], and features alongside authoritative chapters on alternative approaches such as EDA, NBO, Valence Bond, conceptual DFT, Block-localised

P.L.A. Popelier (✉)

Manchester Institute of Biotechnology (MIB), 131 Princess Street,
Manchester M1 7DN, UK
e-mail: pla@manchester.ac.uk

P.L.A. Popelier

School of Chemistry, University of Manchester, Oxford Road,
Manchester M13 9PL, UK

wavefunctions, ELF and high-resolution X-ray crystallography. The spirit of that book combined an educational style with an awareness of current scientific boundaries, while avoiding too many equations in the main text, as requested by the editors. That chapter managed to deliver added value in explaining QTAIM again, by means of an alternative angle of exposition, different to that in other sources [3–7]. Moreover, an historic narrative was given there, as well as a discussion of topological energy partitioning. The current chapter selects and re-explains elements from that document, with the new didactic example of HCN, justifies the name of Quantum Chemical Topology (QCT) (which encompasses and supersedes QTAIM), outlines the current state of affairs in a novel QCT-based protein force field, and briefly invites the community to start falsifying interpretative methods (QCT and non-QCT) in case studies where the outcome would make a difference.

The term Quantum Chemical Topology (QCT) was first coined [8] in 2003, and the first dedicated symposium took place in 2013, in Mexico City. Footnote 19 in the paper that coined QCT, gave a detailed justification for this name and it is helpful to quote part of this footnote, with a few modifications: “...The use of the acronym QCT does not downplay the physics behind “Atoms in Molecules” (AIM) by referring to the topology language as the central idea behind it. Instead, the name QCT seeks to capture better what this approach is about. The term (QT) AIM is widely used but 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 AIM) 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. Also, returning to the electron density, one could use the topological analysis to recover molecules inside van der Waals complexes, an important idea in intermolecular forces. Again, as a name, AIM would not describe this result. The name QCT also invites any future developments based on a topological analysis of other 3D or higher-dimensional scalar functions.” This view was elaborated in Sect. 2.2 of a chapter [9] published in 2005, and updated again in the introduction of a paper [10] in 2009, and finally in Box 8.1 in Chap. 8 of the aforementioned book [2] edited by Frenking and Shaik.

The current book should be the right habitat to start thinking more in terms of falsification when interpreting a chemical phenomenon. Unfortunately, not many papers directly and critically compare methods. The papers that do so, however, often terminate with diplomatic and almost vague conclusions. This status quo perhaps adds to the prevailing notion that all methods are equivalent in quality and predictive value. They can all be used at the same time, in spite of their known pitfalls, and even if their results contradict each other. Such view is echoed in Hoffmann’s statement that “any rigorous definition of a chemical bond is bound to be impoverishing” and also in his advice “that one should have fun with the fuzzy richness of the idea”. Such an attitude perpetuates discussions, without prospect of them ever being resolved. Is this really the fate of interpretative theoretical chemistry? Or should one strive for conceptual hygiene? Is chemistry really this

hopelessly complicated universe preventing chemists to ever discover the right tools to give them solid insight in this universe? Interpretative methods urgently need to make predictions that are falsifiable: one method is wrong and the other is right. Or is asking for such binary resolution naïve? Or is calling for this binary clarity a step towards doing better science? This chapter will give a few examples of interpretations disagreeing and thereby setting the scene for falsification.

2.2 The Topological Atom

Surely everyone can agree that there are atoms inside molecules, in the same way that there are living cells inside an animal or a plant. Molecules are not novel aggregates of electrons and nuclei but are rather constructions based on pre-organised entities called atoms. Similarly, an animal is not a totally new form of life but instead built from specialised cells that each represent pre-organised (more elementary) matter such as proteins, lipids, carbohydrates and nucleotides. Several energy production mechanisms inside a unicellular creature are the same as in the cell of an animal. In a loosely similar vein, atoms largely retain their energy in going from an isolated state in the gas phase to an existence inside a molecule. The near-preservation of atoms is exactly what chemistry is about as a science: the study of how atoms change when interacting with other atoms. Therefore it is important that an atom inside a molecule is defined and calculated such that it does a good job in not changing too much while going from the gas phase to the molecule. It is then that one recovers a truly chemical atom rather than a physical atom. The chemist recovers an atom that allows her, or more modernly him, to insulate how the atom interacts with other atoms rather than being distracted by how that atom was built from scratch (i.e. electrons and the nucleus). Only physics is interested in building the atom from its constituents. Chemistry focuses on the small changes an atom undergoes as it interacts with other atoms, small compared to the energy changes involved in building an atom from electrons and a nucleus (all brought together from infinity).

The question is now *how* to define an atom inside a molecular system and this is where opinions differ, perhaps unfortunately. No experiment helps in settling the contentious question of how to define an atom, and even if there was such an experiment then the interpretation of its measured signal would probably be equally contentious. Hence, it appears that an answer to a prime question of chemistry—what is an atom in a molecule?—can only be tackled theoretically. An important guide to value the theoretical proposals on what an atom inside a molecule actually is, is the energetic transferability of that atom. In other words, how much does the energy of a given atom change as it is transferred from one atomic environment to another one? We will come back to energy transferability in Sect. 2.3.

In this section we focus on the molecular electron density and its shape. We seek a theoretical proposal to define an atom inside a molecule, based on the internal differences in the molecular density. In doing so, we avoid introducing a reference

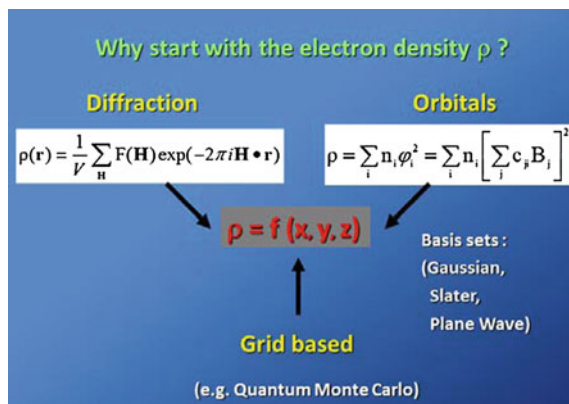


Fig. 2.1 The electron density ρ is a three-dimensional function that can be obtained from three different routes: X-ray diffraction (i.e. crystallography), SCF-LCAO-MO (or “orbitals”) and a method without orbitals where the electron density is known only in given grid points. A method that defines an atom at the level of ρ has the advantage that its definition is independent on how the electron density was obtained

density and this minimality obeys Occam’s razor. This principle of parsimony in assumptions is a quality gauge while assessing theories and models. The fewer parameters a model or theory possesses, while explaining the same number of observed phenomena, the more powerful it is. As a result, the more minimal models are preferable to the more elaborate ones. If experiment does not come to the aid of ranking theoretical proposals by their merit, then Occam’s razor does.

After this philosophical but also strategic interlude we are ready to inspect the electron density of a simple pilot system, the HCN molecule, which is linear. However, we first ask why the electron density is a good starting point to look for an atom in a molecule. Figure 2.1 summarises the argument: the electron density, denoted $\rho(\mathbf{r})$, is independent from the route by which it was generated. In other words, the electron density is an “information platform” describing in detail how electrons distribute themselves in a molecule regardless of the route in which this information was obtained. Figure 2.1 shows three main routes from which the electron density can be acquired.

First, $\rho(\mathbf{r})$ can be obtained from experiment, that is, X-ray crystallography. Routine crystallography only uses local parts of a system’s electron density, namely those at the core of each atom, from which crystallography determines each nuclear position. High-resolution crystallography [11] goes further and collects data on the valence electron density with an eye on measuring chemical features such as bonds and lone pairs. From the 1970s onwards this was done by introducing an artificial reference electron density, which was subtracted from the target electron density, in order to eliminate the huge electron density peaks near the nuclei. This reference density consists of a mere superposition of spherically averaged atomic densities, thereby not allowing any hybridisation to develop and thus missing any chemical

features. As a result, this difference density (technically known as the deformation density) contains all the chemical features. Although a simple and innocent looking approach, the exact form of the reference density is a concern. Different results can be obtained for different choices made in constructing the deformation density. However, there is a more minimal way forward, which avoids such choices in the first place.

Occam's razor proposes to use the molecular electron density as its own reference. Subtracting this density from itself returns a zero density everywhere, which is of course useless but introducing the gradient achieves what is required. The gradient represents an internal difference, via its definition as a derivative, which contains the difference of two function values, each evaluated at two points infinitesimally close to each other. At a given point, the gradient vector contains local information on how the function (in the case the electron density) changes internally. We wonder how the information obtained by the reference-free introspection can be revealed. The key to this goal is simply plotting a succession of gradient vectors, as shown in Fig. 2.2.

The gradient path that results from the primitive construction shown in Fig. 2.2 is all one needs to reveal the internal structure of the electron density. A bundle of gradient paths, called the gradient vector field, naturally exposes two fundamental features a chemist wants to extract from the electron density: the atom and the bond. Figure 2.3 illustrates this for a simple molecule: hydrogen cyanide.

Figure 2.3 clearly shows how a gradient path is everywhere orthogonal to a contour line of constant electron density. This statement is equivalent to the fact that the gradient path traverses the electron density in the direction of maximum ascent. As a result a gradient path also has a direction: its trajectories contain "earlier" points and "later" points in space. The question is now if it has a beginning and an ending. The answer is affirmative to both parts of the question. In fact, the origin and terminus of a gradient path have something in common: they are points where

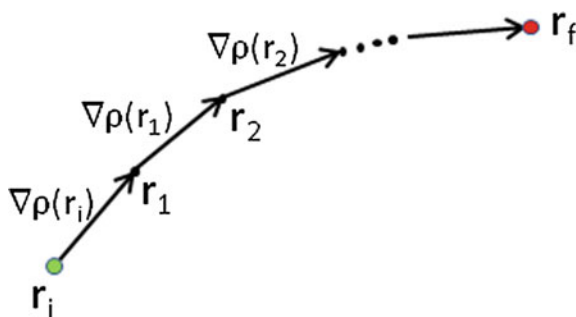


Fig. 2.2 In its most elementary construction a gradient path can be seen as a succession of infinitely short gradient vectors. Starting at \mathbf{r}_i the gradient vector evaluated at this point is followed over a very short stretch, reaching \mathbf{r}_1 , where the gradient is re-evaluated and again followed very briefly. This resulting broken line becomes a gradient vector in the continuous limit, ultimately terminating in point \mathbf{r}_f .

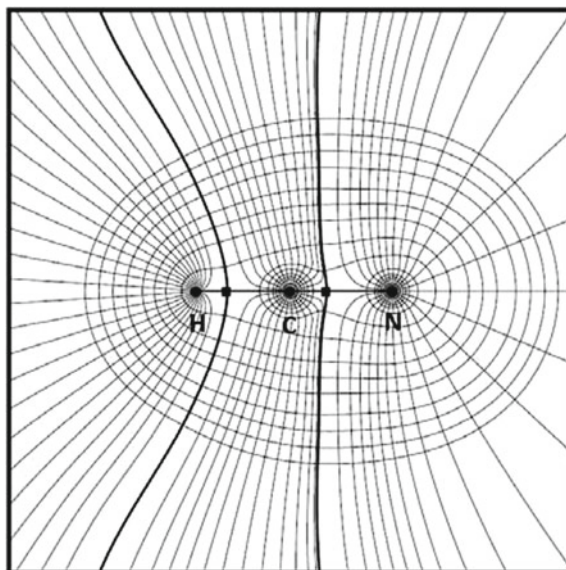


Fig. 2.3 Electron density contour plot of $\text{HC} \equiv \text{N}$ superimposed to its gradient vector field, which consists of an infinite multitude of gradient paths, here represented by a few dozen paths originating at infinity and terminating at the respective nuclei. A special bundle of gradient paths starts at infinity and ends up at the little squares, which are bond critical points. From each bond critical point emerge two gradient paths, each of which is attracted to a different nucleus. This pair of gradient paths is called the atomic interaction line, or in this case of a local energy minimum, the bond path. The carbon is placed at the origin and the bold square box marks the -6 a.u. and $+6$ a.u. horizontal and vertical boundaries of the plot. The electron density values of the contour lines are 1×10^{-n} , 2×10^{-n} , 4×10^{-n} and 8×10^{-n} au where n starts at -3 and increases with unity increments

the gradient vanishes. *Such a point is called a critical point.* Points \mathbf{r}_i and \mathbf{r}_f in Fig. 2.3 are critical points. By deduction all points at infinity are critical points. In three-dimensions there are four types of (non-degenerate) critical points.

Figure 2.3 shows two types of critical points: the maximum and the bond critical point. The maximum is an attractor for an infinite number of gradient paths originating at infinity. *Each such set of gradient paths forms a topological atom.* The bond critical point is a saddle point in that it is a maximum in two directions only (rather than three) and a minimum in the remaining direction. The latter direction is the molecular axis. Indeed, a gradient path originates at the bond critical point and terminates at one of the nuclear maxima. A second gradient path originates at the bond critical point but at the opposite side, and is attracted to the other nuclear maximum at that side. This pair of gradient paths is called an atomic attraction line [12]. When the forces on all nuclei vanish, as is the case for a local energy minimum, then the atomic interaction line becomes a bond path. The set of all bond paths occurring a molecule (or molecular complex) is called a molecular graph. A graph is a mathematical structure that models pairwise relations between objects,

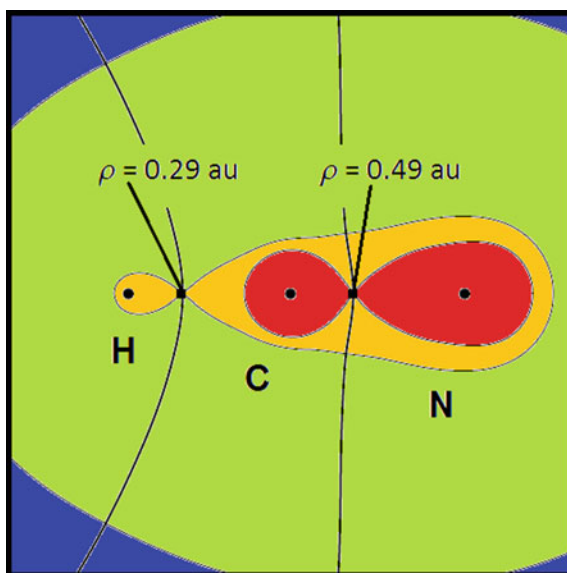
which in this case are atoms. Such a relation is robust under moderate geometric deviations (shrinking and elongation) from the local energy minimum geometry.

It should be pointed out that a gradient path can always be characterised and classified by the *types* of the two critical points that it manifestly connects. This was done exhaustively [13] and for the first time in 2003. This classification focuses on how many gradient paths can originate from a source critical point and how many gradient paths can terminate at the sink critical point.

The attentive reader may ask how a gradient can originate at a critical point if, at that point, the gradient vanishes and hence the gradient path construction of Fig. 2.2 collapses at \mathbf{r}_i . In other words, if there is no gradient there is no directional guidance. This is true but, in practice, the gradient path starts from a position infinitesimally close to the bond critical point, in a direction given by the relevant eigenvector of the Hessian of ρ evaluated at the bond critical point. At that new point the gradient does not vanish. The meaning of the bond critical point is still a matter of debate but an explanation of it, not given in the original literature by what Bader et al., will be given just below.

Let us look again at $\text{HC} \equiv \text{N}$ and fix the value of the electron density at 0.001 a.u. Figure 2.4 show the contour associated with this value, which can be taken as the practical edge of the molecule when in the gas phase. Note that when a molecule is part of a condensed phase then there is no need for such a practical edge; the whole molecule will then be topologically bounded, by a surface that is parameter-free, as explained below. The $\rho = 0.001$ a.u. contour is the boundary between the blue region ($\rho < 0.001$ a.u.) and the green region ($0.001 \text{ a.u.} < \rho < 0.29 \text{ a.u.}$). If the electron density is increased beyond 0.29 a.u. then the hydrogen atom becomes disconnected from the rest of the molecule. In other words, this hydrogen, while still being inside

Fig. 2.4 The atomic disconnection process in $\text{HC} \equiv \text{N}$. Each bond critical point (*little black square*) marks the contact point between two adjacent atoms



the molecule, is now completely enclosed by its own contour lines (not shown). The value of 0.29 a.u. is special because it is that at the bond critical point between *H* and *C*. While increasing the electron density starting from 0.001 a.u., the value of 0.29 a.u. is the highest electron density for which the hydrogen is still attached to the rest of the molecule. For any higher value the contours encompassing the whole molecule become disconnected. The same disconnection process occurs when ρ increases above 0.49 a.u., which of course is the electron density at the second bond critical point. It is then that *C* and *N* also become disconnected. Now, all three atoms in $\text{HC}\equiv\text{N}$ are fully encircled by their own contours. Overall, this process shows that bond critical points are “contact points” between certain atoms. *A bond critical point between two given atoms represents the transition point of them being connected or disconnected.* When connected, they are encompassed by the same contours. When disconnected, the respective atoms have their own “atomic” contours.

It is clear that topological atoms are non-overlapping. This is an important property that has attractive consequences in the area of intermolecular forces, where the thinking is dominated by overlapping molecules. The second important feature of topological atoms is that there are no gaps between the atoms. As a consequence, every point in space belongs to a topological atom; there is no “empty” (i.e. unallocated) space. The absence of the void has consequences for how one thinks about pockets in enzymes, including active sites and allosteric sites. The familiar ball-and-stick, or even “helix/turn/sheet ribbon” representation of the protein modelling world, gives the impression that there is empty space. A molecular view according to topological atoms challenges [14] this impression. Instead, if a ligand enters an enzymatic pocket, it will have to deform a host of topological atoms, each of which has an energy cost. Steric hindrance then becomes a more gradual and continuous concept as opposed to the simple on-off picture that van der Waals radii give. In other words, whereas traditional atoms act as billiard balls, topological atoms behave like sponges.

Figure 2.5 gives a three-dimensional view of the topological partitioning of the pilot molecule. The vertical solid lines appearing in Figs. 2.3 and 2.4 now show

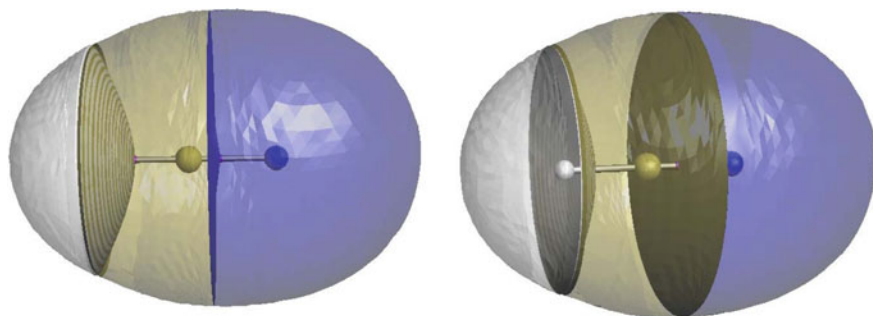


Fig. 2.5 Two views of the same three-dimensional representation of the three topological atoms (grey *H*, gold *C*, blue *N*) in $\text{HC}\equiv\text{N}$. The interatomic surfaces are bundles of gradient paths originating at infinity and terminating at a bond critical point (little bright purple sphere)

their full 3D extent as so-called *interatomic surfaces*. These surfaces are the sharp boundaries between atoms inside a molecule. Because a molecule is simply the union of its topological atoms the boundary between molecules is also sharp. Hence, a molecule in condensed matter is fully bounded by interatomic surfaces, which are parameter free. The use of molecular contour surfaces of constant electron density (e.g. $\rho = 0.001$ a.u.) is artificial and exists for practical (i.e. visualisation) purposes only. Note that a molecule in the gas phase, alone in the Universe, is a fiction: sooner or later one will find another molecule far away that still shares a topological boundary with the original “isolated” molecule.

The hydrogen cyanide molecule, $\text{H}-\text{C}\equiv\text{N}$, can be isomerised to hydrogen isocyanide, $\text{H}-\text{N}^+\equiv\text{C}^-$, by tilting the hydrogen over the carbon and gradually rotating it to the right. Eventually, this hydrogen ends up at the right hand side of a new linear arrangement, which can also be written as $\text{C}^- \equiv \text{N}^+-\text{H}$ for easy comparison with $\text{H}-\text{C}\equiv\text{N}$. At some sharp transition point during the rotation of hydrogen, the atomic interaction line flips over: where it originally connected *H* and *C*, it then connects *H* and *N*. This means that the connectivity of the atoms suddenly changes, which is a topological feature. This is an example of a so-called conflict mechanism.

We now ask in which way(s) the name topology is appropriate for what we have discovered. Topology is the mathematical study of shapes and topological spaces. It is an area of mathematics concerned with the properties of space that are preserved under continuous deformations including stretching and bending. Avoiding precise mathematical terms, one can define topology as the study of qualitative properties of certain objects that are invariant under continuous transformations. For example, Euler’s work on the Königsberg bridge problem was one of the earliest topological studies. He showed it was impossible to find a route through the city of Königsberg that crosses each of the seven bridges exactly once. This solution only depended on which bridges are *connected* to which islands and riverbanks. In other words, only connectivity mattered, not the length of the bridges or the distances between them. This work also marked the beginning of graph theory and thereby establishes a link between topology and graph theory. In general, the motivation behind topology is that some geometric problems do not depend on the exact shape of the objects involved, but rather on the way they are put together. How does this way of thinking apply to the analysis of the electron density discussed above?

The keyword “topology” was first used [15] in the expression “quantum topology” by Bader et al. in 1979. Unfortunately, this name is already taken by a branch of mathematics that connects quantum mechanics with low-dimensional topology, which has little to do with Bader et al. developed and which culminated in QCT. In any event, the name “quantum topology” was inspired by a paper [16] by Collard and Hall published in 1977. These authors were the first to use the Poincaré-Hopf theorem, which links the respective numbers of each of the four possible types of critical points to the so-called Euler characteristic. The latter is a purely topological concept and hence justifies the name topology. The Euler characteristic is an application of *algebraic topology*, one of the four branches of topology, which uses tools from abstract algebra to study topological spaces. Collard and Hall also pointed out that the analysis of the discontinuous change in

the topological characteristics of a molecular charge distribution resulting from the continuous change in its nuclear coordinates is given by *catastrophe theory* [17]. This theory is a branch of *bifurcation theory* in the study of *dynamical systems*. In turn, bifurcation theory is the mathematical study of changes in the qualitative or topological structure of a family of vector fields. A second branch of topology called *differential topology*, which is the field dealing with differentiable functions on differentiable manifolds and which is closely related to differential geometry, also applies to QCT. In fact, as early as 1996, a differential geometry study [18] appeared on the Gaussian curvature of interatomic surfaces. Differential topology also makes use of the Poincaré-Hopf theorem, and hence the two branches of topology overlap. Finally, we mention that there are two remaining branches in topology, called *geometric topology* and *general topology*. The first branch, which includes knot theory, is the study of manifolds and maps between them, particularly embeddings of one manifold into another. This branch does not appear to have been applied as of yet in QCT. The second branch (general topology, also known as point-set topology) establishes the foundational aspects of topology (point-set topology, compactness and connectedness etc.). It deals with the basic set-theoretic definitions and constructions used in topology, and thereby underpins the three other branches (differential, geometric and algebraic topology).

Figure 2.6 shows an example of a 2D dynamical system that has nothing to do with quantum mechanics but shows key features of QCT. The equations state how the time derivative (dot signifies d/dt) varies as a function of the position in (x, y) space, as a non-linear function of x and y. A particle at position (x, y) will travel

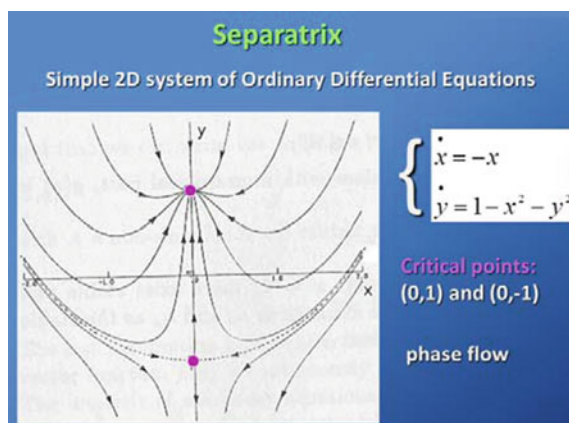


Fig. 2.6 Simple system of two ordinary differential equations, which shows a separatrix (*dashed line*) and two critical points (*pink*). This topological object separates the basin dominated by the attractor critical point (0, 1) (*top*). The second critical point shown is a saddle-type critical point, at which the separatrix trajectories (*dashed line*) terminate. The collection of trajectories (*phase flow*) can be seen as the paths followed by imagined particles travelling in time. The superscripted dots in the equations signify differentiation with respect to time

with a direction and magnitude given by the vector $(-x, 1-x^2-y^2)$. The resulting trajectories form the phase flow in Fig. 2.6, which is reminiscent of the gradient vector field of the electron density of a diatomic molecule with one nucleus shown (at the attractor critical point $(0, 1)$). The second critical point shown lies at point $(0, -1)$ and clearly has a qualitatively different flow pattern locally. This critical point is both a maximum and a minimum, depending on the direction of approach to it, and is reminiscent of a bond critical point. The two dashed trajectories lines that terminate at this critical point form a separatrix, which is reminiscent of an inter-atomic surface.

A non-exhaustive list of quantum mechanical functions that have hitherto been partitioned includes the electron density $\rho(\mathbf{r})$ (the analysis of which started with Ref. [19]), its Laplacian $\nabla^2\rho(\mathbf{r})$ (started off with Refs. [20, 21] and studied for the first time in terms of the full topology in Refs. [13, 22, 23]), the nuclear potential $V_{\text{nuc}}(\mathbf{r})$ (studied [24] already in 1980 but the first elaborate and self-contained study [10] appeared only 30 years later), the electron localization function (ELF) [25] (started with Ref. [26] and reviewed in Ref. [27]), the electrostatic potential [28] (started with thorough but stubbornly named “topographic” instead of topological studies [29, 30] and continued with more modern work [31–34]), the virial field [35], the magnetically induced molecular current distributions (started with [36]), the intracule density (started with Ref. [37]), the Ehrenfest force field (topology first investigated [38] in 2012 and then improved [39] in 2015), and finally the topological energy partitioning (Coulomb potential energy partitioning started with [40] and culminated into the theory of Interacting Quantum Atoms (IQA) [41] (see below), leading to energetic underpinning for the topological expression of chemical bonding [42]) By bundling all these QCT studies under the umbrella of the topology, the combined method is strengthened and can start competing with the more traditional interpretative method of quantum chemistry [43–49]. This competition should be seen in the light of falsification.

At the end of Sect. 2.2 it is useful to pause and muse about the character of the topology as an instrument to study Nature. The language of dynamical systems, which is rooted in topology, is at the heart of QCT. A hallmark of QCT’s partitioning is its binary character: a point in space belongs to a QCT subspace or not. Whereas non-QCT approaches allow for more gradual transitions from one subspace (e.g. an atom) to another, QCT works with step functions. The 3D step function defines a finite-volume subspace that remains well defined under (possibly large) geometrical deformations. Topology does allow for large deformations in the geometry of the objects it defines while still characterising them by the same invariant measures. However, once beyond a certain degree of deformation, the topological object changes. The suddenness of this change makes some researchers uncomfortable. The comfort of a gradual change may look appealing but then one can ask if this is a false comfort. Can a world view with only gradual change make any clear decision on what is *A* and what is *B*? Or should one not care about being able to make this decision? Or can one make the decision at the price of introducing a parameter? But is then the problem of state allocation (i.e. making the aforementioned decision) not simply deferred to fixing a parameter value?

The topology, even with its abrupt decisions, focuses on the essence of a system in the same way that a phase diagram does. A molecular system can be in the liquid state or the solid state, and the transition between these two states is abrupt (in the limit of an infinitely large system). A phase diagram shares the characteristics of a topologically partitioned space: it has sharp boundaries and it ignores the geometrical details of the system. Indeed, a phase diagram looks beyond the exact positions of the atoms in the molecules that make up a system; the atoms can vibrate while the molecules can translate and rotate. In the same way, QCT looks beyond the exact trajectories of the gradient paths but focuses on their connectivities, which are robust over large deformations of the gradient paths themselves. A topological atom is then analogous to a phase. This is an example of how Nature itself apparently imposes binary structures onto reality: it makes sense to say that a piece of matter is either a liquid or a solid and the boundary between the two is sharp.

There are more examples of sharp compartmentalisation in Nature. One of the deepest examples is the architecture of thermodynamics, which discerns the system and the surroundings. It is essential to the theoretical and practical functioning of thermodynamics that a point in space either belongs to the system or to the surroundings. Any fuzzy partitioning or delay in decision would paralyse any thermodynamic calculations or predictions. Secondly, Life itself, this most complex of structures, has organised and evolved under the very existence of sharp boundaries. Due to its small size, a cell membrane is a relatively sharp boundary between the cytoplasm and the extracellular space. Of course, the boundaries are open (under the control of specialised proteins in the cell's lipid membrane). The boundaries of a topological atom are also open in that electrons can swirl through them. A topological atom is a pattern, comparable to the shape of water as it rapidly cascades over a rock in a river. From a distance, the water appears standing still in a barely fluctuating shape but of course the water itself streams through the pattern. Thirdly, at a higher level, human societies have also carved up the Earth's space in non-overlapping subspaces with sharp boundaries, called countries. When a territory is not allocated to a single clear "attractor" such as China, Pakistan or India, as in the case of Kashmir, then a dispute arises, proving the inherent human nature of partitioning land into non-overlapping sections. Further examples of binary statuses are found in the legal atmosphere where one is either alive or dead, married or not, or guilty or innocent. The question then remains why Chemistry is not the right locale to propose non-overlapping partitioning. What is so intrinsically fuzzy about atoms and electron densities that would prevent sharp boundaries? Is life or human society perhaps less fuzzy?

At the very end of this section on the topological atom, and on the wider topological approach with its fundamental characteristics and consequences, we put the topology to rest and look at energy instead. Energy is a quantum mechanical observable and the main question is how it can be partitioned. This is the topic of the next section, where we forget about the gradient vector field of the electron density, at least at the start.

2.3 The Quantum Atom

Energy is as important as the electron density. The Schrödinger equation presents the energy and the wave function as prime quantities, joined at the hip and being of equal status. For each eigenvalue (energy), there is an eigenfunction (wave function) and they both come as an inseparable pair. Because the electron density immediately derives from the wave function, the molecular electron density and the molecule's energy are also twinned. Hence, because this electron density is of prime importance due to the first Hohenberg-Kohn theorem, energy shares this importance. Indeed, energy is in charge of the way a molecular system behaves and understanding it is therefore crucial. Phenomena, such as steric hindrance, ultimately reduce to energy considerations, even if sterics appear irreducible intuitively (based on daily life experience). The natural way to understand something (at least in the Western tradition of doing science) is to study its parts. Such an approach calls for the spatial partitioning of energy.

The key question is how to define a molecular *fragment* that has a well-defined kinetic energy. This question is attacked by starting with *local* kinetic energy, which is the kinetic energy at a particular point *per unit volume*. This quantity is thus a kinetic energy *density*, which when integrated over a volume, gives the kinetic energy of the electrons in that volume. The kinetic energy of a molecular fragment is then obtained from a 3D integral of the kinetic energy density over the volume of that fragment. However, there is a practical problem in that there is no “the” kinetic energy density; at best, there is “a” kinetic energy density. We write “at best” because if one starts from the quasiprobability distribution function the quantum mechanical treatment of kinetic energy, partitioned or not, is actually problematic. Local kinetic energy is then ambiguous. However, within the paradigm that Anderson et al. [50] call the “Laplacian family of local kinetic energies”, the deduction below is valid [51].

Although there are an infinite number [52] of expressions for the kinetic energy density, it is sufficient to choose only two possible expressions to develop this deduction [53], as given by Eqs. 2.1 and 2.2,

$$K(\mathbf{r}) = -\frac{1}{4}N \int d\tau' [\psi^* \nabla^2 \psi + \psi \nabla^2 \psi^*] \quad (2.1)$$

$$G(\mathbf{r}) = \frac{1}{2}N \int d\tau' \nabla \psi^* \cdot \nabla \psi \quad (2.2)$$

where N is the total number of electrons in the system, ψ the system's N -electron wave function, and $\int d\tau'$ signifies integration over all electrons except one. Note that the electron spin is not considered here. It is easy to show that the two kinetic energy densities, $K(\mathbf{r})$ and $G(\mathbf{r})$, are linked via the Laplacian of the electron density, $\nabla^2 \rho$, or

$$\mathbf{K}(\mathbf{r}) = \mathbf{G}(\mathbf{r}) - \frac{1}{4} \nabla^2 \rho(\mathbf{r}) \quad (2.3)$$

The Laplacian of the electron density vanishes when integrated over whole space or

$$\int_{\text{whole space}} dV \nabla^2 \rho(\mathbf{r}) = 0 \quad (2.4)$$

Integrating both sides of Eq. 2.3 over whole space then gives a unique value of the molecule's kinetic energy,

$$\mathbf{K}(\text{molecule}) = \mathbf{G}(\text{molecule}) = \mathbf{T}(\text{molecule}) \quad (2.5)$$

where T expresses *the* kinetic energy regardless of whether it was calculated from $\mathbf{K}(\mathbf{r})$ or $\mathbf{G}(\mathbf{r})$. Because a single molecule in the gas phase occupies whole space, one indeed recovers the kinetic energy of the molecule by integration over whole space. This energy is well-defined because it is unique: indeed, both $\mathbf{K}(\mathbf{r})$ and $\mathbf{G}(\mathbf{r})$ give the same answer.

The main question is now if this same unique result can also be obtained for a molecular *fragment*. Let us consider the subspace of an *arbitrary* fragment, denoted \oplus . For such an arbitrary subspace in 3D space we find that

$$\int_{\oplus} dV \nabla^2 \rho(\mathbf{r}) \neq 0 \quad (2.6)$$

From this equation and integration over both sides of Eq. 2.3, one deduces that

$$\mathbf{K}(\oplus) \neq \mathbf{G}(\oplus) \quad (2.7)$$

Hence, we do not obtain a unique kinetic energy for an arbitrary subspace. However, if we can find a *special* subspace Ω such that

$$\int_{\Omega} dV \nabla^2 \rho(\mathbf{r}) = 0 \quad (2.8)$$

then it makes sense to speak of a unique and hence well-defined kinetic energy $\mathbf{T}(\Omega)$ associated with such a special subspace,

$$\mathbf{K}(\Omega) = \mathbf{G}(\Omega) = \mathbf{T}(\Omega) \quad (2.9)$$

An atom that occupies such a special subspace Ω , and thereby obeys Eq. 2.9, is called a *quantum atom*. At this moment we do not worry about what this quantum atom looks like nor about how many possible such atoms there are. The only matter

we are interested in is whether the topological atom is also a quantum atom. A way to find out is to reformulate Eq. 2.9 using Gauss’s divergence theorem, which yields

$$\int_{\Omega} dV \nabla^2 \rho(\mathbf{r}) = \int_{\partial\Omega} dS \nabla \rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \quad (2.10)$$

where $\partial\Omega$ is the boundary of Ω . Equation (2.10) shows how a volume integral over Ω is equal to a surface integral over $\partial\Omega$. Now we focus on the integrand of the surface integral and also look at the gradient vector field in Fig. 2.3. The interatomic surface $\partial\Omega$, separating H and C for example, is a surface that consists of gradient paths. Hence the normal to this surface, denoted $\mathbf{n}(\mathbf{r})$, is orthogonal to a gradient path at any point belonging to this surface including the bond critical point, or

$$\nabla \rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \quad \forall \mathbf{r} \in \partial\Omega \quad (2.11)$$

If Eq. 2.11 is true then Eq. 2.10 is also true. *Thus a topological atom is a quantum atom.* Note that, unlike Bader et al. do we claim the reverse, which is that each quantum atom is also a topological atom. In fact, we now know that this statement is not correct. So, in summary, all topological atoms are quantum atoms but not all possible quantum atoms are topological atoms [54]. Therefore, any criticism [50] against the orthodox version of QTAIM which is the one propagated by Bader, does not apply to the approach presented here. In other words, we do not insist that the topological atoms are the *only* quantum atoms. We have deliberately introduced and justified topological atoms on their own merit, independently from quantum mechanics. They are indeed remarkable and attractive objects, and one can ask why not more scientific disciplines use the elegant idea of partitioning by gradient vector field subspace (called *basin* in short).

It is important to properly appreciate the result obtained above (Eq. 2.9) as a “gateway” to a fully quantum-mechanically based force field. Traditional force fields ignore kinetic energy, or more precisely, they do not explicitly account for it. However, kinetic energy is a physical quantity and cannot be switched off; it does influence the behaviour of atoms in a system and hence must somehow be incorporated in a force field or what one could call a “rapid energy predictor”. A traditional force field only mimics the effect of kinetic energy, and only indirectly, by including it in a Morse-like potential, for example. Such a methodology does not isolate the kinetic energy in an atomic way. Instead, it lumps the behaviour of the kinetic energy of two interacting atoms into *bond-based* parameters. QCT offers a completely different route, one where the parameterisation is atom-based. Moreover, this novel parameterisation recognises the explicit existence of kinetic energy, at atomic level. That a topological atom offers this route, by virtue of being a quantum atom (with a well-defined kinetic energy) is enticing. In the next section we give a very brief outline of the QCT force field strategy.

2.4 Towards a QCT Protein Force Field

2.4.1 Topological Energy Partitioning

An early and important result in the development of QTAIM was that an atom in a molecule has its own (atomic) virial theorem. This means that, for any given topological atom, there is a (simple) relation between the kinetic energy of this atom and its potential energy. This in turn means that the potential energy of an atom (i.e. interacting with itself and all remaining atoms) can be trivially calculated from the atom's kinetic energy (which we already know to be well defined). As a further consequence, the total energy of an atom (which is the sum of kinetic and potential energy) can be calculated from the kinetic energy alone. The sum of all total atomic energies forming a molecule then yields the total energy of that molecule. However, all of this is only true if the forces on the atomic nuclei vanish. If not, one is left with a residual virial term consisting of nuclear position vectors dotted into non-vanishing forces on the nuclei. Partitioning the latter (molecular) quantity over the respective atoms has always been a problem, until in 2001 the potential energy of an atom was calculated [40] independently from the kinetic energy.

The calculation of the interatomic electrostatic potential energy V_{elec} involves a six-dimensional integral, over the volume of each of the two topological atoms A and B , or

$$V_{\text{elec}}^{AB} = \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \frac{\rho_{\text{tot}}(\mathbf{r}_1)\rho_{\text{tot}}(\mathbf{r}_2)}{r_{12}} \quad (2.12)$$

where the *total* charge density, $\rho_{\text{tot}}(\mathbf{r})$, is the sum of the nuclear charge density and minus the electron density $-\rho(\mathbf{r})$ (i.e. electronic and hence corrected by a minus sign catering for the negative electronic charge), while r_{12} is the distance between two infinitesimal pieces of charge density [40]. This work was further developed with the calculation of non-Coulomb interaction energies [55, 56].

The use of Eq. 2.12 implies that the condition of vanishing forces no longer restricts the topological partitioning of the molecular energy into intra- and inter-atomic contributions. This advance led to the development of Interacting Quantum Atoms (IQA) [41], which since its implementation in AIMALL [57] has become an increasingly popular tool in the armoury of interpretative quantum chemical tools. A second and parallel development from the advance in the aforementioned 2001 paper [40] is that of a quantum mechanical force field based on the energies associated with topological atoms (at any nuclear configuration and including non-stationary points on the potential energy surface). This is indeed what our lab started doing, initially much focusing on multipolar electrostatics, under the acronym QCTFF (Quantum Chemical Topology Force Field) [9, 58–60]. There is sustained and consistent evidence [61] that multipole moments are more accurate and realistic than point charges. In spite of the latter's inherent and well

documented limitations some researchers are still looking (e.g. Ref. [62]) for the “magic point charge” that accurately reproduces the molecular electrostatic potential, even if that point charge is then just a mathematical number without any connection to the physical process of charge transfer. We believe that the atomic monopole is primarily a measure of charge transfer; at long range this monopole becomes increasingly representative of the electrostatic potential that this atom generates.

Applying the Laplace multipole expansion leads to

$$V_{elec}^{AB} = \sum_{l_A l_B m_A m_B} T_{l_A l_B m_A m_B} Q_{l_A m_A} Q_{l_B m_B} \quad (2.13)$$

where $Q_{\ell m}$ represents the m -th component of a rank ℓ atomic multipole moment, while T is a purely geometrical interaction tensor. The convergence properties of this series expansion have been thoroughly studied [63–67] by our lab. There are three conceptual and technical advantages associated with QCT multipole moments. They are more compact than Cartesian multipole moments, avoiding redundancies, they demonstrate good convergence at short-range, and they escape penetration effects (and hence damping functions) due to their non-overlapping nature.

Note that V_{elec}^{AB} consists of 4 contributions, exhausting the purely electronic and nuclear contribution on both A and B (i.e. $4 = 2 \times 2$) that is, the electron-electron Coulomb energy $V_{ee,coul}^{AB}$, the electron-nucleus attraction (potential) energy, denoted V_{en}^{AB} , its dual V_{en}^{BA} , and the nucleus-nucleus repulsion, V_{nn}^{AB} . When added, these terms lead to the full electrostatic interaction between two atoms A and B , V_{elec}^{AB} , or

$$V_{elec}^{AB} = V_{ee,coul}^{AB} + V_{en}^{AB} + V_{en}^{BA} + V_{nn}^{AB} \quad (2.14)$$

The electron-nucleus attraction energy is calculated as a three-dimensional integral,

$$V_{en}^{AB} = -Z_B \int_{\Omega_A} d\mathbf{r} \frac{\rho(\mathbf{r})}{r_{1B}} \quad (2.15)$$

where r_{1B} is the distance between an electron inside the volume of atom A and the nucleus of atom B . This calculation can also be performed if $A = B$, which features in the intra-atomic energy discussed below.

The energy $V_{ee,coul}^{AB}$ can be related to the second-order reduced matrix, $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$. To understand how exactly, one needs to know the fine structure of $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$, or

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \rho_2^{coul} + \rho_2^{exch} + \rho_2^{corr} = \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) - \rho_1(\mathbf{r}_1, \mathbf{r}_2)\rho_1(\mathbf{r}_2, \mathbf{r}_1) + \rho_2^{corr}(\mathbf{r}_1, \mathbf{r}_2) \quad (2.16)$$

where the first term refers to the quantum-mechanically uncorrelated Coulomb-like pair density, the second term to the Fock-Dirac exchange (which is dominated by and associated with the Fermi hole), while the third term is at least an order of magnitude smaller [41, 68] than the second term, and connected with the Coulomb hole. The energy quantity V_{ee}^{AB} is associated with the whole of $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$, collecting the three types of interactions that electrons experience when interacting with each other. Each term in Eq. 2.16 is associated with a type of potential energy, so that the corresponding fine-structure of V_{ee}^{AB} automatically follows,

$$\begin{aligned}
 V_{ee}^{AB} &= \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \frac{\rho_2(\mathbf{r}_1, \mathbf{r}_2)}{r_{12}} \\
 &= \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} - \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \frac{\rho_1(\mathbf{r}_1, \mathbf{r}_2)\rho_1(\mathbf{r}_2, \mathbf{r}_1)}{r_{12}} + \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \frac{\rho_2^{corr}(\mathbf{r}_1, \mathbf{r}_2)}{r_{12}} \\
 &= V_{ee,coul}^{AB} + V_{ee,exch}^{AB} + V_{ee,corr}^{AB}
 \end{aligned} \tag{2.17}$$

The second term in Eq. 2.17 represents the exchange delocalisation energy, $V_{ee,exch}^{AB}$, which is (already) present at Hartree-Fock level. This term teases out the interaction that keeps bonded atoms together. The degree to which atoms are bonded can be estimated by a non-energy measure, which is typically a quantum-mechanical bond order. QCT offers such a measure [69]. However, it was shown by our lab [56] that this bond order is only the first term of the multipolar expansion of $V_{ee,exch}^{AB}$. Hence, the latter quantity contains more information than a bond order. However, in the construction of QCTFF, the route of expanding $V_{ee,exch}^{AB}$ as so-called *exchange moments* was abandoned because they have an imprint of the molecular orbitals they are derived from. This imprint hampers transferability. The energy quantity $V_{ee,exch}^{AB}$ can remain unexpanded because it drops off so quickly with distance [70] in saturated systems, which proteins largely are. However, multipole moments are essential in the representation of electrostatics because this type of interaction drops off more slowly than $V_{ee,exch}^{AB}$. Therefore the number of non-negligible V_{elec}^{AB} values is much larger than the number of $V_{ee,exch}^{AB}$ values. The trouble with this observation is the rapidly increase in the number of possible distances between *A* and *B*. In other words, atoms that are further apart can appear in more possible configurations than atoms that are closer to each other. This is why it is undesirable to calculate all possible $1, n$ ($n > 4$) V_{elec}^{AB} interactions. A multipole series succeeds in avoiding the calculation of all these V_{elec}^{AB} interactions. The series separates a geometrically entangled (since r_{12} involves simultaneously \mathbf{r}_1 and \mathbf{r}_2) energy quantity into single atom quantities, i.e. multipole moments. This separation enables the calculation of the interatomic interaction to be free of large geometric variations. Conversely, short-range interactions (1, 2; 1, 3 and 1, 4) are geometrically much more constrained and hence would not benefit that much from multipole moments. This is why it is alright to *not* expand the electrostatic energy V_{elec}^{AB} at short range.

The third term in Eq. 2.17 completes the discussion on the three types of electron-electron energy contributions. The quantity $V_{ee,corr}^{AB}$ covers the effect of dynamic correlation and hence dispersion. It is absent at Hartree-Fock level or $V_{ee,corr}^{AB} = 0$.

$$V_{ee,corr}^{AB} = \sum_{j=1}^{n_G} \sum_{k=1}^j \sum_{l=1}^{n_G} \sum_{m=1}^l d_{jklm} \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \frac{1}{r_{12}} G_j(\mathbf{r}_1 - \mathbf{R}_j) G_k(\mathbf{r}_1 - \mathbf{R}_k) G_l(\mathbf{r}_2 - \mathbf{R}_l) G_m(\mathbf{r}_2 - \mathbf{R}_m) \quad (2.18)$$

where d_{jklm} are 4-index coefficients that we have extracted from the computer program GAUSSIAN, G_p is the p -th Gaussian primitive centered on \mathbf{R}_p and n_G is the number of primitives. The number of d-coefficients rapidly increases with the number of primitives, in particular as $\frac{1}{4}[n_G(n_G + 1)]^2$. Hence truncation schemes must be devised and I/O optimised.

The energy contribution $V_{ee,corr}^{AB}$ was calculated for the first time [71] as late as 2015, for the four simple case studies of H_2 , N_2 , H_2O and CO , operating on CCSD/cc-pVDZ wave functions obtained by the program MOLPRO. The effect of dynamic correlation is dual: an increase in the magnitude of the nucleus-electron attraction energy, and a decrease in the electronic repulsion. Representing dispersion accurately and consistently within the QCT framework (rather than by a bolt-on [72]) is important for future-proof success in the modelling of the conjugated residues (imidazole, phenol, indole and benzyl) of the four aromatic amino acids [73]. This streamlined approach will avoid penetration effects, which the non-overlapping topological atoms naturally preclude. Hence, there is no need for damping functions in QCTFF. It appears that satisfactory expressions for damping functions are problematic in view of the complexity of atom typing [74]. The dynamic correlation part of QCTFF is currently under investigation in lab (in connection with the program GAUSSIAN). Figure 2.7 summarises the three types of interatomic energy contributions of QCTFF.

The remaining energy contribution is intra-atomic in nature, denoted E_{intra} , and measures the intrinsic stability of an atom. It cannot be written as “V” because this symbol is reserved for potential energy only and the atomic “self-energy” [40] also contains kinetic energy, which is well-defined for a topological atom, as clearly argued above. Broadly speaking, E_{intra} features in (and indeed may control) rotation barriers, steric hindrance, the anomeric effect, the gauche effect or other stereo-electronic effects. We note that in typical potentials, such as the Lennard-Jones potential, repulsion is formulated as an *inter*-atomic effect, whereas within QCT, steric “interaction” is a *mono*-atomic property. The full consequence of this philosophical difference still needs to be worked out because it already appears to have an impact on the way we should think about “steric clashes”. Some support against the traditional view that steric effects are due to precise one-to-one interaction, and hence in favour of the QCT view, comes from a non-QCT angle.

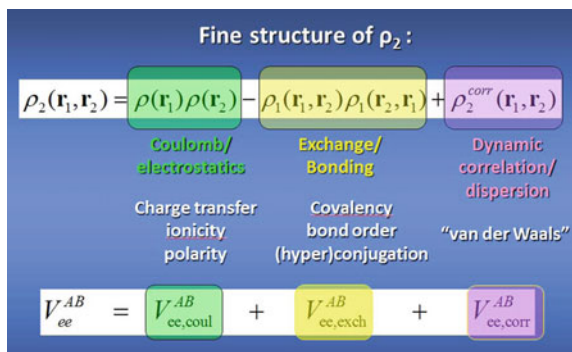


Fig. 2.7 Overview of the three types of inter-atomic energy contributions: Coulomb, exchange and correlation, each with the specific chemical insight they offer

Weinhold, who pioneered the natural bond orbital (NBO) method, writes in a paper [75] dating from 1997: “A persistent theme of this work is that steric exchange repulsion is not simply a sum of pairwise interactions between two electron pairs, but rather a complex function of the entire N-electron distribution. The pattern of orbital energy changes due to exchange repulsion is more complex than a simple “atom–atom repulsion” picture would suggest.” The contribution E_{intra} has indeed an imprint of the whole molecule, although it has a practical cut-off, which we call the atomic horizon (see next paragraph and energetic transferability between tri- and penta-peptides).

Setting $A = B$ in Eqs. 2.15 and 2.17 allows one to write the intra-atomic energy of topological atom A as

$$E_{\text{intra}}^A = T^A + V_{ee}^{AA} + V_{en}^{AA} \quad (2.19)$$

where T^A is its kinetic energy. The intra-atomic energy E_{intra}^A is the energy that a single atom possesses inside a system, regardless of whether this system is a single molecule or a cluster of molecules (including even ions). Work from our lab (to be published in Molecular Physics 2016) shows that an oxygen, nitrogen or carbon has the same energy, within maximum 2.3 kJmol^{-1} , when appearing in a tri-peptide (three amino acids) compared to appearing in a penta-peptide, with these peptides’ common nuclear skeleton in the same configuration. This energetic transferability was observed in seven test cases, i.e. the homo-oligopeptides of Ala, Ser, Thr, Gly, Val, Leu and Ile. This high degree of energetic transferability is an asset to QCT. Transferability has also been detected [76] in terms of atomic charges by those who develop alternative partitioning schemes (such as the Hirshfeld partitioning [77] scheme and all its variants).

2.4.2 Machine Learning

The final strand of QCTFF to discuss briefly is the way atomic properties (both intra- and inter-) are predicted from the coordinates of the nuclei surrounding the atom of interest. In general, this mapping is so complex that it needs a machine learning method, and one that can handle high-dimensional spaces, given the large number of coordinates that influence the atom of interest. Kriging [78] is such a method. Originating in geostatistics, Kriging is a powerful interpolation technique that can capture the behaviour of an output as a function of many inputs, using a relatively small amount of data points. In its infancy [79] it succeeded in predicting where the best location for a mine would be in a two-dimensional landscape, based on measurements of a precious material (originally gold but could be diamond, oil, uranium or any ore) at various locations in this landscape. The basic idea of Kriging is to predict the value of a function at a given point by computing a weighted average of the known values of the function in the neighborhood of the point. An accessible account of the details of Kriging as used within the QCTFF context has been given elsewhere [80].

Here we highlight one key idea, namely that of maximising the likelihood L , which has not been clarified in that previous account [80]. To fix thoughts, let us start with a simple example: a coin is being tossed thrice. If the coin is fair, then the probability to observe head up, (denoted p_H), is one half, that is $p_H = 0.5$. Equally, the probability of observing tail denoted p_T is one half, or $p_T = 0.5$. The probability to observe head up twice and then tail (HHT) is $p_{HHT} = p_H p_H p_T = p_H^2 (1 - p_H) = 0.125$. An equivalent way of saying this is to reverse this statement: the likelihood L that the coin was fair (i.e. $p_H = 0.5$), given the observation of two heads being up (HHT), is one eighth, i.e. $L = 0.125$. This is formally written as follows:

$$L(p_H = 0.5 | HHT) = 0.125 \quad (2.20)$$

In summary, the likelihood L is a function returning the probability of observed outcomes (e.g. HHT), given a parameter value (i.e. p_H). We now ask ourselves how the likelihood $L = p_H^2 (1 - p_H)$ can be maximised. Mathematically this is easy: calculus tells us that $dL/dp_H = d/dp_H [p_H^2 (1 - p_H)] = 2p_H - 3p_H^2$, which vanishes when $p_H = 2/3$. A plot, or a quick calculation of the second derivative, tells us that $p_H = 2/3$ is indeed a maximum, at which point $L = 4/27$. The result that $p_{H, \max L} = 2/3$ can be intuitively understood by stating that the coin is biased towards heads up, by a factor 2 over tail up. Indeed, with such a bias, the probability of the observed outcomes HHT, given $p_H = 2/3$, is maximal. How does all this help understanding a key aspect behind Kriging?

Kriging uses the same strategy of maximising the likelihood: it finds the parameters θ_h and p_h ($h = 1, 2, \dots, d$) in the so-called Gram matrix \mathbf{R} ,

$$R_{ij} = \exp \left[- \sum_{h=1}^d \theta_h |x_h^i - x_h^j|^{p_h} \right] \quad \theta_h > 0, \quad 0 < p_h \leq 2 \quad (2.21)$$

such that the likelihood L of the Kriging model, given the observed data points y^i ,

$$L(\boldsymbol{\theta}, \mathbf{p}, \mu, \sigma | y^i; i = 1, 2, \dots, N) \sim \exp \left[- \frac{(\mathbf{y} - \mathbf{1}\mu)^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{1}\mu)}{2\sigma^2} \right] \quad (2.22)$$

is maximal, where σ^2 is the process variance, $\mathbf{1}$ is a column vector of ones, and μ models the global trend of the observable \mathbf{y} . The observations are fixed (i.e. atomic properties = output and coordinates of surrounding nuclei = input) but the parameters are being varied such that what we observed becomes maximally likely. How this maximisation is achieved is beyond the scope of this intentionally non-technical text but this is a very important active research topic in our lab.

We refer the interested reader to the literature [51, 73, 80–86] for the use of Kriging in the construction of QCTFF. Here we can only afford three general remarks. Firstly, we know that three of the four types of energy contributions described above can be Kriged successfully for all 20 amino acids, cholesterol, small carbohydrates and small water clusters (also in the presence of a cation) and a few pilot systems (NMA, ethanol, water, etc.). Proof-of-concept of successful Kriging of the dynamic correlation energy contribution still needs to be obtained but we do not expect any fundamental problems.

Secondly, the term “successful” needs to be qualified. The performance of a Kriging model is validated by an external test set of molecular configurations. This is where we display the full performance of a Kriging model over the *whole* test set, by means of a so-called S-curve. From the latter one can read off which percentage of test configurations scores an energy prediction error up to any desired value. For example, if this value is set to 4 kJmol^{−1} (referring the old-fashioned and arbitrary unit of kcalmol^{−1}) then 70 % of test configurations containing all local energy minima found in the Ramachandran map of the doubly-capped amino acid isoleucine, return an error of less than 4 kJmol^{−1}. While the mean error over all 200 test configurations is 3.3 kJmol^{−1}, there is a small percentage (~ 2 %) of test configurations that have errors just over 10 kJmol^{−1}. While this behavior is typical, matters are worse for cysteine where only 50 % return an error of less than 4 kJmol^{−1}, while the average is 5.3 kJmol^{−1} and just over 10 % have an error within the interval 10–20 kJmol^{−1}. The reported errors are all purely electrostatic and involve all interactions of the type 1, 4 and higher. This is a rather severe test because it involves short-range interactions, switching on all multipole moments up to the hexadecapole moment. The average of mean errors over all 20 amino acids is 4.2 kJmol^{−1}, while the worse values are for cysteine, alanine and arginine, all at 5.3 kJmol^{−1}. The best mean error is 2.8 kJmol^{−1} for tyrosine.

Thirdly, the kriging method covers all polarisation effects, but without introducing polarisabilities. The QCTFF method focuses on the end result of the

polarisation process, not the process itself. As a result, when used in a molecular dynamics simulation, QCTFF renders immediately the energies of all atoms in response to a given nuclear configuration. There is no need to iteratively converge towards a self-consistent field at each simulation step.

2.5 An Invitation to Falsification

As announced in the Introduction there is a need for more falsification in the area of chemical interpretation by means of quantum mechanical tools. In a first stage contradictions need to be spotted: when are two methods providing (semi)-quantitatively or qualitatively different interpretations? The second stage is more challenging: how can an experiment judge one interpretation to be right and the other wrong? A valiant but strongly disputed [87] example of this kind of scientific activity was published [88] in 2009 where it was claimed that experiment could disprove QTAIM's interpretation of an attractive interaction between the two hydrogens in the bay region of phenanthrene.

With regards to the second stage, one could broaden the decision process, not through experiment, but by teasing out a clash with a theoretical principle or another theoretical interpretation that is more firmly established. For example, a number of electronegativity scales all agree that boron is a very electropositive element. One may then ask how it is possible that a population analysis allocates a negative net charge to boron. Yet this happens. For example, in 1995 Siegbahn allocated a net charge $-0.26e$ to boron in $(\text{BH}_3\text{NH}_3)_2$ using the Mulliken population analysis. Of course the QTAIM charge of boron is emphatically positive. Another candidate for a falsifiable case study is that of 1, 2-difluoroethene, which was discussed in a 2009 publication [89] comparing the IQA method (i.e. QCT) with the non-QCT method EDA and NBO, in connection with interpreting stereo-electronic effects. IQA rules that there is significant FF' delocalisation in the *cis* isomer, which is "not easily found in NBO" according to the article. This is a fine example of one method spotting an effect and the other not. The challenge is to exploit this difference, either via an experiment that can confirm one or the other method, or demonstrate that guidance (in synthesis for example) is more reliable by one method than by the other.

A final example is that of diborane. A pivotal question is: can QTAIM and hence QCT extract a Lewis diagram from a given molecular wave function? A very recent study, published [70] in 2013, set out to answer precisely this question, and the answer is yes. It is possible by inspecting motives in calculated V_{exch}^{AB} values, and this 2013 work systematically investigated V_{exch}^{AB} values, for all atom-atom interactions in 31 small covalent molecules (including ions) and 3 van der Waals complexes. For the first time, clear clusters were revealed in the values of V_{exch}^{AB} , clusters separated by almost an order of magnitude in energy, starting with hundreds of kilojoules per mole, and decreasing in a stepwise manner to less than

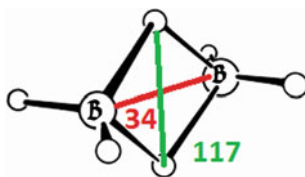


Fig. 2.8 A ball-and-stick diagram of diborane, B_2H_6 , endowed with numerical values of $-V_{xc}^{AB}$ in kJmol^{-1} . The unlabelled disks represent hydrogen atoms

0.1 kJmol^{-1} . This quantitative information reveals where to draw the lines in a Lewis diagram.

A useful example illustrating this success is that of diborane, B_2H_6 , which at one time was controversial in terms of its Lewis structure. Figure 2.8 shows a ball-and-stick diagram of B_2H_6 , endowed with $| -V_{xc}^{AB} |$ values (in kJmol^{-1} and for HF/6-311G(d,p) wave functions). The two largest (absolute) values are 385 and 222 kJmol^{-1} , corresponding to the covalent bonds BH_{term} and BH_{bridge} , respectively. They provide the “sticks” of the molecular graph, which was controversial until 1951. We note that Pauling got the structure wrong while Longuet-Higgins got it right. The next strongest interaction is that between the bridging hydrogens (117 kJmol^{-1}) (green in Fig. 2.8), which is three times larger than the value between the two borons (34 kJmol^{-1}) (red in Fig. 2.8).

In a private communication, Roald Hoffmann spontaneously pointed out that the HH interaction is something new to him and that there is some BB bonding in B_2H_6 is easier to understand. The latter assertion followed from his explanation of a MO diagram. Here we have a clear example of two theories (QCT and MO) stating qualitatively different things. Again, the challenge is to find an experiment that could settle this contradiction.

In closing this section, it is helpful to philosophise a little more, in order to contemplate if the suggested falsification is useful or even possible. Let us start with a simple and clear but ridiculous example (plucked from reality). A reality TV celebrity said that she thought that the sun and the moon were the same object. Sure: when the sun has gone down then it is time for the moon to appear a bit later on. A casual observer may not have a problem with this view. In fact, contemporary political correctness and encouragement of diversity in personal views may even support this assertion. Her “theory” can be adhered to for a long time until one evening, a less casual and alert observer (hopefully herself) sees a red sun at dusk *at the same time* as the moon, and both objects and *clearly separated* in the sky, almost at opposite ends in fact. Such special evenings exist and this one “experiment” kills off the old theory. In summary, at the edge of our knowledge, competing theories (e.g. sun = moon and sun \neq moon) coexist. However, one crucial observation can banish one theory to the history books while promoting the other to build on and use for asking the next exciting question. It is important that a decision is obtained as to which theory is right because otherwise, any planning on how to put a person on the moon is clearly futile, just as an example. To recap briefly with a

now serious example, Alan Guth's inflation theory is really at the edge of our collective human knowledge. This theory predicts the existence of gravitational waves but they have never been detected with certainty (in spite of a premature announcement in 2014). A momentous multi-year, multi-team data collection project set up on the South Pole, can still not decide if gravitational waves exist or not. The signal could be due only to galactic dust. It remains important to settle this question for once or for all. Falsification is needed to make true progress.

The main question is now if the process of falsification (whether smoothly controlled or erratic is not important) is relevant for the interpretation of chemical phenomena. Is the above analysis of the two examples discussed too simple? Does the mechanism of falsification apply to theoretical interpretative chemistry? I believe it does but we should be aware of one more notion. Let us discuss what I call the *non-question*, again starting with an example.

How can we come to grips with the particle-wave duality? Is there an experiment that can decide, for once and for all, if a quantum object is either a particle or a wave? The textbook answer is no: the quantum object is both. But then we ask why this question cannot be settled. In fact, there are enough experiments that decide in favour of a wave, while others decide in favour of a particle. How can this be? A way out is suggesting that the wave-particle question is a *non-question*. *A non-question is a question that cannot be settled one way or the other because it makes a fundamentally wrong assumption.* The problem here, however, is that it is not clear which assumption exactly. But there are several clear examples of non-questions. One is "what is north of the North Pole?", another is "What came first: the chicken or the egg?" The latter cannot be solved because it ignores what really happened in evolution: a pre-chicken "lays" a pre-egg and so on. As one moves back in time the distinction between the two becomes problematic and the question actually dissolves. This chicken-egg "question" is a non-question because it wrongly assumes that one can project a binary end point of evolution onto the very beginning of this evolution. Such a false projection is clearer in a typical child-like question such as "why does a tree not weep if it is chopped down?" Well, one needs a pretty highly developed nervous system in order to weep, a system that the tree clearly lacks. Again, this is a simple example, perhaps ludicrous to adults, but unfortunately spilling over to the world of some adults who believe plants feel pain.

Of course, one should be open to new initially mysterious phenomena, such as X-rays killing living cells. Clearly, X-rays do not emerge from the wonderfully mature edifice of classical mechanics. But one should beware of the non-question. One can spend a life time thinking about the question of the beginning and end of a thing until someone shows that this thing is actually circular. This then means that the wrong image has been projected onto an object, giving rise to a natural and innocent question, which turns out to be a vicious non-question. Imagine if this thing is the Universe. One will then have wasted a life time thinking about this question because it actually turned out to be a non-question. The earlier one spots a non-question, the earlier one can ask the real question. This transition is a major

advance: perhaps from childhood into adulthood, from deluded or blinkered adult to informed adult, from scientist in a desert to scientist in an oasis.

As a community, we should increase the effort in spotting non-questions. Maybe “how aromatic is this compound?” is a non-question. In order to save time and focus on the real question I advocated [90] to carry out bottom-up research and focus on the emergence of patterns of primary quantities (closely linked to the Schrödinger equation). If aromaticity is captured some way along this bottom-up approach then this concept is lucky, as it were, and will survive. Those that spotted this property, more than a century ago, at high and intuitive level, without knowing about quantum chemistry, will then be vindicated. But if aromaticity falls apart into two or more new concepts, or even worse, evaporates altogether, then that is progress. We have then turned the non-question into a question and we can then re-explore the complex world of chemical phenomena armoured with more powerful insight.

2.6 Conclusion

Quantum Chemical Topology has a long and rich history of about four decades. It started with an innocuous paper [19] in 1972, which however marked the birth of a completely novel way of thinking about how to partition and characterise a quantum mechanical system. That the topological atom, which can exist in its own right, is also a quantum atom, makes it possible to build a force field using these atoms. Perhaps this force field is better “called a rapid energy predictor” because it overhauls the architecture of traditional force fields and probes deeper into the quantum mechanics that underpin them. Finally, we point out the need for falsification of theoretical interpretative tools and theories. Experimentalists need more reliable and predictive guidance from theoretical interpretation. If methods contradict each other there is an opportunity to establish one method as the way forward. However, one should beware of the non-question.

Acknowledgments I thank the EPSRC for the award of an Established Career Fellowship.

References

1. Bader RFW (1991) A quantum-theory of molecular-structure and its applications. *Chem Rev* 91:893–928
2. Popelier PLA (2014) The quantum theory of atoms in molecules. In: Frenking G, Shaik S (eds) *The nature of the chemical bond revisited*, Chapter 8. Wiley, pp 271–308
3. Bader RFW (1985) Atoms in molecules. *Acc Chem Res* 18:9–15
4. Bader RFW (1990) *Atoms in molecules. A quantum theory*. Oxford Univ. Press, Oxford, Great Britain
5. Popelier PLA (2000) *Atoms in molecules. An introduction*. Pearson Education, London

6. Matta CF, Boyd RJ (2007) The quantum theory of atoms in molecules. From solid state to DNA and drug design. Wiley, Weinheim
7. Popelier PLA (2012) Quantum chemical topology: on descriptors, potentials and fragments. In: Banting L, Clark T (eds) Drug design strategies: computational techniques and applications, vol 20, Chapter 6. Roy Soc Chem, pp 120–163
8. Popelier PLA, Aicken FM (2003) Atomic properties of amino acids: computed atom types as a guide for future force field design. *Chem Phys Chem* 4:824–829
9. Popelier PLA (2005) Quantum chemical topology: on bonds and potentials. In: Wales DJ (ed) Structure and bonding. intermolecular forces and clusters, vol 115. Springer, Heidelberg, pp 1–56
10. Popelier PLA, Brémond ÉAG (2009) Geometrically faithful homeomorphisms between the electron density and the bare nuclear potential. *Int J Quant Chem* 109:2542–2553
11. Koritsanszky TS, Coppens P (2001) Chemical applications of X-ray charge-density analysis. *Chem Rev* 1583–1627
12. Bader RFW, Essen H (1984) The characterization of atomic interactions. *J Chem Phys* 80:1943–1960
13. Malcolm NOJ, Popelier PLA (2003) An improved algorithm to locate critical points in a 3D scalar field as implemented in the program MORPHY. *J Comp Chem* 24:437–442
14. Popelier PLA (2012) New insights in atom-atom interactions for future drug design. *Curr Top Med Chem* 12:1924–1934
15. Bader RFW, Anderson SG, Duke AJ (1979) Quantum topology of molecular charge distributions. 1. *J Am Chem Soc* 101:1389–1395
16. Collard K, Hall GG (1977) Orthogonal trajectories of the electron density. *Int J Quant Chem* 12:623–637
17. Thom R (1975) Structural stability and morphogenesis; english ed. Benjamin: Reading, MA
18. Popelier PLA (1996) On the differential geometry of interatomic surfaces. *Can J Chem* 74:829–838
19. Bader RFW, Beddall PM (1972) Virial field relationship for molecular charge distributions and the spatial partitioning of molecular properties. *J. Chem. Phys.* 56:3320–3329
20. Bader RFW, MacDougall PJ, Lau CDH (1984) Bonded and nonbonded charge concentrations and their relation to molecular-geometry and reactivity. *J Am Chem Soc* 106:1594–1605
21. Bader RFW, Gillespie RJ, MacDougall PJ (1988) A physical basis for the VSEPR model of molecular geometry. *J Am Chem Soc* 110:7329–7336
22. Popelier PLA (2000) On the full topology of the Laplacian of the electron density. *Coord Chem Rev* 197:169–189
23. Malcolm NOJ, Popelier PLA (2003) The full topology of the Laplacian of the electron density: scrutinising a physical basis for the VSEPR model. *Faraday Discuss* 124:353–363
24. Tal Y, Bader RFW, Erkkü J (1980) Structural homeomorphism between the electron density and the nuclear potential of a molecular system. *Phys Rev A* 21:1–11
25. Becke AD, Edgecombe KE (1990) A simple measure of electron localization in atomic and molecular systems. *J Chem Phys* 92:5397–5403
26. Silvi B, Savin A (1994) Classification of chemical bonds based on topological analysis of electron localization functions. *Nature(London)* 371:683–686
27. Polo V, Andres J, Berski S, Domingo LR, Silvi B (2008) Understanding reaction mechanisms in organic chemistry from catastrophe theory applied to the electron localization function topology. *J Phys Chem A* 112:7128–7136
28. Naray-Szabo G, Ferenczy GG (1995) Molecular electrostatics. *Chem Rev* 95:829–847
29. Gadre SR, Kulkarni SA, Shrivastava IH (1992) Molecular electrostatic potentials: a topographical study. *J Chem Phys* 96:5253–5261
30. Gadre SR, Shrivastava IH (1991) Shapes and sizes of molecular anions via topographical analysis of electrostatic potential. *J Chem Phys* 94:4384–4391
31. Balanarayan P, Gadre SR (2003) Topography of molecular scalar fields. I. algorithm and poincare-hopf relation. *J Chem Phys* 119:5037–5043

32. Balanarayan P, Kavathekar R, Gadre SR (2007) electrostatic potential topography for exploring electronic reorganizations in 1,3 dipolar cycloadditions. *J Phys Chem A* 111:2733–2738
33. Aray Y, Rodriguez J, Coll S, Rodriguez-Arias EN, Vega D (2005) Nature of the lewis acid sites on molybdenum and ruthenium sulfides: an electrostatic potential study. *J Phys Chem B* 109:23564–23570
34. Tsirelson VG, Avilov AS, Lepeshov GG, Kulygin AK, Stahn J, Pietsch U, Spence JCH (2001) Quantitative analysis of the electrostatic potential in rock-salt crystals using accurate electron diffraction data. *J Phys Chem B* 105:5068–5074
35. Keith TA, Bader RFW, Aray Y (1996) Structural homeomorphism between the electron density and the virial field. *Int J Quant Chem* 57:183–198
36. Keith TA, Bader RFW (1993) Topological analysis of magnetically induced molecular current distributions. *J Chem Phys* 99:3669–3682
37. Cioslowski J, Liu GH (1999) Topology of electron-electron interactions in atoms and molecules. II. The correlation cage. *J Chem Phys* 110:1882–1887
38. Pendas AM, Hernandez-Trujillo J (2012) The Ehrenfest force field: topology and consequences for the definition of an atom in a molecule. *J Chem Phys* 137:134101
39. Dillen J (2015) The Topology of the Ehrenfest force density revisited. a different perspective based on slater-type orbitals jan. *J Comput Chem*. doi:[10.1002/jcc.23869](https://doi.org/10.1002/jcc.23869)
40. Popelier PLA, Kosov DS (2001) Atom-atom partitioning of intramolecular and intermolecular Coulomb energy. *J Chem Phys* 114:6539–6547
41. Blanco MA, Pendas AM, Francisco E (2005) Interacting quantum atoms: a correlated energy decomposition scheme based on the quantum theory of atoms in molecules. *J Chem Theor Comput* 1:1096–1109
42. Pendas AM, Francisco E, Blanco MA, Gatti C (2007) Bond paths as privileged exchange channels. *Chem Eur J* 13:9362–9371
43. Webster B (1990) Chemical bonding theory. Blackwell, Oxford
44. Reed AE, Curtiss LA, Weinhold F (1988) Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem Rev* 88:899–926
45. Stone AJ (1981) Distributed multipole analysis, or how to describe a molecular charge-distribution. *Chem Phys Lett* 83:233–239
46. Mulliken RS (1955) Electronic population analysis on LCAO-MO molecular wave functions. *I J Chem Phys* 23:1833–1840
47. Pearson RG (2007) Applying the concepts of density functional theory to simple systems. *Int J Quant Chem* 108:821–826
48. Kovacs A, Esterhuysen C, Frenking G (2005) The nature of the chemical bond revisited: an energy-partitioning analysis of nonpolar bonds. *Chem Eur J* 11:1813–1825
49. McWeeny R (1992) Methods of molecular quantum mechanics, 2nd edn. Academic Press, SanDiego
50. Anderson JSM, Ayers PW, Hernandez JIR (2010) How ambiguous is the local kinetic energy? *J Phys Chem A* 114:8884–8895
51. Fletcher TL, Kandathil SM, Popelier PLA (2014) The prediction of atomic kinetic energies from coordinates of surrounding atoms using kriging machine learning. *Theor Chem Acc* 133 (1499):1–10
52. Cohen L (1978) Local kinetic energy in quantum mechanics. *J Chem Phys* 70:788–799
53. Bader RFW, Preston HJT (1969) The kinetic energy of molecular charge distributions and molecular stability. *Int J Quant Chem* 3:327–347
54. Nasertayoob P, Shahbazian S (2008) Revisiting the foundations of quantum theory of atoms in molecules (QTAIM): the variational procedure and the zero-flux conditions. *Int J Quant Chem* 108:1477–1484
55. Rafat M, Popelier PLA (2007) Atom-atom partitioning of total (super)molecular energy: the hidden terms of classical force fields. *J Comput Chem* 28:292–301

56. Rafat M, Popelier PLA (2007) Topological atom-atom partitioning of molecular exchange energy and its multipolar convergence. In: Matta CF, Boyd RJ (eds) *Quantum theory of atoms in molecules*, vol 5. Wiley, Weinheim, pp 121–140
57. AIMall: Todd A, Keith TK (2014) Gristmill software, overland Park KS, USA, (aim.tkgristmill.com)
58. Popelier PLA (2012) Quantum chemical topology: knowledgeable atoms in peptides. *AIP Conf Proc* 1456:261–268
59. Popelier P, Rafat M, Devereux M, Liem SY, Leslie M (2005) Towards a force field via quantum chemical topology. *Lect Series Comput Comput Sci* 4:1251–1255
60. Popelier PLA (2012) A generic force field based on quantum chemical topology. In: Gatti C, Macchi P (eds) *Modern charge-density analysis*, vol 14. Springer, Germany, pp 505–526
61. Cardamone S, Hughes TJ, Popelier PLA (2014) Multipolar electrostatics. *Phys Chem Chem Phys* 16:10367–10387
62. Ivanov MV, Talipov MR, Timerghazin QD (2015) Genetic algorithm optimization of point charges in force field development: challenges and insights. *J Phys Chem A* 119:1422–1434
63. Rafat M, Popelier PLA (2006) A convergent multipole expansion for 1,3 and 1,4 Coulomb interactions. *J Chem Phys* 124(144102):1–7
64. Rafat M, Popelier PLA (2007) Long range behaviour of high-rank topological multipole moments. *J Comput Chem* 28:832–838
65. Popelier PLA, Joubert L, Kosov DS (2001) Convergence of the electrostatic interaction based on topological atoms. *J Phys Chem A* 105:8254–8261
66. Yuan Y, Mills MJL, Popelier PLA (2014) Multipolar electrostatics for proteins: atom-atom electrostatic energies in crambin. *J Comp Chem* 35:343–359
67. Solano CJF, Pendás AM, Francisco E, Blanco MA, Popelier PLA (2010) Convergence of the multipole expansion for 1,2 Coulomb interactions: the modified multipole shifting algorithm. *J Chem Phys* 132:194110
68. Pendas AM, Francisco E, Blanco MA (2006) Binding energies of first row diatomics in the light of the interacting quantum atoms approach. *J Phys Chem A* 110:12864–12869
69. Angyan JG, Loos M, Mayer I (1994) Covalent bond orders and atomic valence indices in the topological theory of atoms in molecules. *J Phys Chem* 98:5244–5248
70. Garcia-Revilla M, Francisco E, Popelier PLA, Martín-Pendas AM (2013) Domain-averaged exchange correlation energies as a physical underpinning for chemical graphs. *Chem Phys Chem* 14:1211–1218
71. Chávez-Calvillo R, García-Revilla M, Francisco E, Martín Pendás A, Rocha-Rinza T (2015) Dynamical correlation within the Interacting quantum atoms method through coupled cluster theory. *Comput Theor Chem* 1053:90–95
72. Grimme S (2004) Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J Comput Chem* 25:1463–1473
73. Fletcher T, Davie SJ, Popelier PLA (2014) Prediction of intramolecular polarization of aromatic amino acids using kriging machine learning. *J Chem Theory Comput* 10:3708–3719
74. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J Chem Phys* 132:154104–154122
75. Badenhoop JK, Weinhold F (1997) Natural bond orbital analysis of steric interactions. *J Chem Phys* 107:5406–5421
76. Verstraelen T, Ayers PW, Van Speybroeck V, Waroquier M (2013) Hirshfeld-E partitioning: AIM charges with an improved trade-off between robustness and accurate electrostatics. *J Chem Theory Comput* 9:2221–2225
77. Hirshfeld FL (1977) Bonded-atom fragments for describing molecular charge densities. *Theor Chem Acta* 44:129–138
78. Rasmussen CE, Williams CKI (2006) *Gaussian processes for machine learning*. The MIT Press, Cambridge
79. Krige DG (1951) A statistical approach to some basic mine valuation problems on the Witwatersrand. *J Chem Metall Mining Soc South Africa* 52:119–139

80. Kandathil SM, Fletcher TL, Yuan Y, Knowles J, Popelier PLA (2013) Accuracy and tractability of a Kriging model of intramolecular polarizable multipolar electrostatics and its application to histidine. *J Comput Chem* 34:1850–1861
81. Mills MJL, Popelier PLA (2011) Intramolecular polarisable multipolar electrostatics from the machine learning method Kriging. *Comput Theor Chem* 975:42–51
82. Mills MJL, Popelier PLA (2012) Polarizable multipolar electrostatics from the machine learning method Kriging: an application to alanine. *Theor Chem Acc* 131:1137–1153
83. Mills MJL, Hawe GI, Handley CM, Popelier PLA (2013) Unified approach to multipolar polarisation and charge transfer for ions: microhydrated Na⁺. *Phys Chem Chem Phys* 15:18249–18261
84. Handley CM, Hawe GI, Kell DB, Popelier PLA (2009) Optimal construction of a fast and accurate polarisable water potential based on multipole moments trained by machine learning. *Phys Chem Chem Phys* 11:6365–6376
85. Yuan Y, Mills MJL, Popelier PLA (2014) Multipolar electrostatics based on the Kriging machine learning method: an application to serine. *J Mol Model* 20:2172–2186
86. Hughes TJ, Kandathil SM, Popelier PLA (2015) Accurate prediction of polarised high order electrostatic interactions for hydrogen bonded complexes using the machine learning method kriging. *Spectrochim Acta A* 136:32–41
87. Bader RFW (2009) Bond paths are not chemical bonds. *J Phys Chem A* 113:10391–10396
88. Grimme S, Mück-Lichtenfeld C, Erker G, Kehr G, Wang H, Beckers H, Willner H (2009) When do interacting atoms form a chemical bond? spectroscopic measurements and theoretical analyses of dideuteriophenanthrene. *Angew Chem Int Ed* 48:2592–2595
89. Pendas AM, Blanco MA, Francisco E (2009) Steric repulsions, rotation barriers, and stereoelectronic effects: a real space perspective. *J Comput Chem* 30:98–109
90. Ayers PL, Boyd RJ, Bultinck P, Caffarel M, Carbó-Dorca R, Causá M, Cioslowski J, Contreras-García J, Cooper DL, Coppens P, Gatti C, Grabowsky S, Lazzeretti P, Macchi P, Pendás AM, Popelier PLA, Ruedenberg K, Rzepa H, Savin A, Sax A, Schwarz WEH, Shabbazian S, Silvi S, Solà M, Tsirelson V (2015) Six questions on topology in theoretical chemistry. *Comput Theor Chem* 1053:2–16

Applications of Topological Methods in Molecular
Chemistry

Chauvin, R.; Lepetit, C.; Silvi, B.; Alikhani, E. (Eds.)

2016, IX, 586 p. 248 illus., 81 illus. in color., Hardcover

ISBN: 978-3-319-29020-1