The full topology of the Laplacian of the electron density: Scrutinising a physical basis for the VSEPR model

ARTICLE *in* FARADAY DISCUSSIONS · FEBRUARY 2003

Impact Factor: 4.61 · DOI: 10.1039/B211650M · Source: PubMed

2 AUTHORS:

Noj Malcolm
Schrodinger
21 PUBLICATIONS **413** CITATIONS
SEE PROFILE

Paul L A Popelier
The University of Manchester
187 PUBLICATIONS **6,992** CITATIONS
SEE PROFILE

The full topology of the Laplacian of the electron density: scrutinising a physical basis for the VSEPR model

Nathaniel O. J. Malcolm† and Paul L. A. Popelier*

Dept. of Chemistry, UMIST, Manchester, UK M60 1QD

Received 22nd November 2002, Accepted 16th January 2003
First published as an Advance Article on the web 14th May 2003

Within the framework of quantum chemical topology (QCT) the function $L(\boldsymbol{r})$, which equals the negative of the Laplacian of the electron density, has been proposed before as a physical basis for the valence shell electron pair repulsion (VSEPR) model. The availability of a new algorithm to integrate property densities over the basins of $L(\boldsymbol{r})$ enabled a re-evaluation of this physical basis. We optimised a set of nine molecules at B3LYP/6-311+G(2d,p) level and partitioned the corresponding $L(\boldsymbol{r})$ function for each molecule into basins. For the first time we visualise these basins in $L(\boldsymbol{r})$, by directly showing their boundaries. We identify the basins in $L(\boldsymbol{r})$ with the domains of the VSEPR model. Observations drawn from the populations and volumes of L-basins are contrasted with the three subsidiary VSEPR postulates. We find unexpectedly small populations, nearer to one than to two, for non-hydrogen cores and bonding domains, and populations much larger than two for non-bonding domains. We conclude that non-bonding or lone pairs have larger domains than bonding pairs in the same valence shell, in accordance with VSEPR. We also confirm that double and triple bond domains are larger than single-bond domains. However we cannot substantiate the effect of the electronegativity of central atom or ligand on the volume of bonding domains. In summary, the full topology of $L(\boldsymbol{r})$ supports two out of three subsidiary VSEPR postulates.

1. Introduction

Chemistry benefits from consistency between its concepts and the properties of molecular wave functions obtained by modern computational schemes. In other words, the rules, models and explanations that chemists have introduced over the years ultimately need to find a physical basis. Quantum chemical topology (QCT)[1–3] is an attractive way to accomplish this goal. QCT captures chemical information emerging from contemporary mathematical expressions of molecular wave functions. In this paper we focus on finding a physical basis for the valence shell electron pair repulsion (VSEPR) model by examining the *full* topological characteristics of the function $L(\boldsymbol{r}) = -\nabla^2\rho(\boldsymbol{r})$, which is the negative of the Laplacian of the electron density.

The VSEPR model,[4,5] also known as the Gillespie–Nyholm rules,[6] has for many years provided a rationale for molecular geometry in the spirit of Lewis structures. The predecessor of the VSEPR model, already formulated in 1940 by Sidgwick and Powell,[7] stated, on the basis of structures of singly bonded AX_n molecules known at the time, that two pairs of electrons in a valence shell have

† Current address: Tripos UK Ltd., Sunningdale House, Caldecotte Lake Drive, Milton Keynes, UK MK7 8LF.

a linear arrangement, three pairs a triangular arrangement, four pairs a tetrahedral arrangement, five pairs a trigonal bipyramidal arrangement and six pairs an octahedral arrangement. Surveying a substantially larger structural data set Gillespie and Nyholm were able to demonstrate that the electron pairs around a central atom are arranged such that they maximise their mutual distances. Their model, which we review in relevant detail in the next section, is valid for the vast majority of molecules of the main group elements, particularly those of the nonmetals. This success explains why the VSEPR model has found its way into many undergraduate textbooks. It features as a prime example of a successful empirical rationalisation of chemical structure, and is hence in need of a physical basis. Rather than deducing this physical basis[8] from simple orbital models[9] it was recognised early on that the spatial correlation of electrons in molecules was predominantly governed by the Pauli principle[10,11] and thus independent from orbital models. Curiously the Pauli principle does not figure as much as it should in the average chemist's arsenal of explanations. Although much more important in determining the shape of molecules and appearance of materials, electrostatics still seems to play a disproportionately large role.

In its original formulation the VSEPR model was built on the concept that valence shell electron pairs behave as if they repel each other and thus keep as far apart as possible. However, in recent years more emphasis has been placed on the space occupied by a valence shell electron pair, called the domain of the electron pair.[12] Bader, Gillespie and McDougall were the first to propose[13] $L(\boldsymbol{r})$ as a physical basis for the VSEPR model. Their proposition was founded on the observation[14] that local maxima of $L(\boldsymbol{r})$ faithfully duplicate in number, location, and size the spatially localised electron pairs of the VSEPR model. This work surveyed a sizeable set of molecules, including compounds with three maxima (SO_2), four maxima (CH_4, SiH_4, NH_3, PH_3, OH_2, SH_2, NF_3, PF_3, ClF_2^+, $ClCl_2^+$), five maxima (ClF_3, SF_4, SF_4O), and six maxima (ClF_5). The analysis of Bader *et al.*[14] was purely based on the location of so-called critical points in $L(\boldsymbol{r})$ (in this case maxima).

Although very valuable, critical point information is local and makes it hard to gauge the size (*i.e.* volume) of a domain. Thanks to a newly developed algorithm,[15] referred to as the "octree" algorithm and briefly explained below, we can now visualise $L(\boldsymbol{r})$-basins and compute their volumes and electronic populations. This algorithm enables the integration of properties over $L(\boldsymbol{r})$-basins and hence provides global rather than local (point) properties. Consequently we unveil a more complete picture of the relationship between VSEPR and $L(\boldsymbol{r})$, and are put in a position to scrutinise previously made statements. Moreover we can now back up (with rigorous data) the shift of emphasis in the VSEPR model towards an understanding of the space occupied by a valence shell electron pair. Since this is a preliminary study, and in view of the computational expense of the octree algorithm, we only revisit a small selection of molecules in this article. They are CH_4, NH_3, H_2O, NF_3, PF_3, ClF_3, C_2H_6, C_2H_4 and $CH_3C{\equiv}CH$.

2. Theoretical background

2.1 The VSEPR model

This model is a rationale based on a large number of observations, which continues to predict geometrical features of new recently prepared molecules. The value of this model is further underlined by its success as a pedagogical tool. It is perhaps one of the finest examples of *emergence* in chemistry, *i.e.* the appearance of simple patterns and explanations from the realm of complexity that chemistry is recognised to be.

The VSEPR model is sometimes mistaken to be a consequence of classical electrostatic theory, a prime instrument of physical rationalisation of an average bench chemist. Instead, the VSEPR model is an expression of the Pauli exclusion principle, the effect of which is much more powerful in determining the shape of molecules than that of electrostatics. Already in the fifties this fact was emphasised[11] by Lennard-Jones and others. One of the reasons that Hartree–Fock theory is so successful in predicting molecular geometry, in spite of lacking Coulomb correlation, is that it incorporates the Pauli exclusion principle, which dominates Coulomb effects.

Typical treatments of bonding and molecular geometry are almost always based on hydrogen-like atomic orbitals. A bond in this model is regarded as an "*overlapping* of the orbitals of the bonded atoms to create a charge cloud that spreads over and between the nuclei and binds them

together".[12] In contrast the VSEPR model views molecules as built up from non-overlapping regions of space in which there is a high probability of finding an electron pair. These regions negotiate volumes in 3D space, according to the postulates of the VSEPR model, and thereby determine molecular geometry. Strikingly QCT shares with the VSEPR model the important attribute of *non-overlapping* regions. Consequently QCT is a suitable and natural way to link the VSEPR model with modern wave functions.

The first basic assumption of the VSEPR model is that the valence electron density is spatially localised into electron pairs. One can show that this follows from the Pauli exclusion principle (for a vivid illustration in the case of water, see p. 87 in ref. 16). The region of space (in the valence shell) in which there is a high probability of finding an electron pair, or in which a large fraction of an electron pair charge cloud is found is called a *domain*.[16] In this picture an electron pair surrounds the point at which the electron cloud is a maximum probability of finding a pair of electrons, that is, at which the charge cloud is most concentrated.

The second basic assumption of the VSEPR model is that a given number of (valence shell) electron pairs, *considered as points*, adopt that arrangement that keeps them as far apart as possible. In its original form the VSEPR model regarded the domains as spheres that pack as closely as possible around a point representing the core of the central atom.

An important refinement of the VSEPR model is embodied in its three subsidiary postulates. It is here that the importance of the shape and volume of the domains becomes clear since not all the valence shell electron pairs (or domains) are equivalent. The postulates are:

1. Non-bonding (lone pair) domains are larger than bonding domains in the same valence shell.
2. (a) Bonding domains decrease in size with increasing electronegativity of the ligand, and (b) increase in size with increasing electronegativity of the central atom.
3. Double and triple-bond domains are larger than single-bond domains.

Below we go over how and to what extent the challenge of linking the VSEPR model to the information contained in modern *ab initio* wave functions has been met.

2.2 Electron pair localisation

The study of the localisation of electron pairs in a molecule should ultimately refer to properties of the pair density[17] or the second-order reduced density matrix.[18] This is a six-dimensional function, directly obtained from the wave function, which is hard to visualise. A way forward is to fix three dimensions by positioning a so-called reference electron in space and to plot remaining 3D function. This procedure is followed for the same-spin conditional pair density or Lennard-Jones function.[19] This function measures the amount of same-spin density allowed (*i.e.* not excluded) at a point in space by the spreading out of the density of a reference electron at another position. Gillespie *et al.*[19] showed that the Lennard-Jones function is very successful in recovering the geometrical models associated with differing numbers of electron pairs, as suggested by the work of Lennard-Jones.[10,11] In their work[19] Gillespie *et al.* adopted the Lennard-Jones function as a physical basis for the VSEPR model.

Much earlier, in 1988, the Laplacian of the electron density, or more precisely $L(\boldsymbol{r}) = -\nabla^2\rho(\boldsymbol{r})$ had been proposed[13] as a physical basis for the VSEPR model. The observational nature of this physical basis or mapping has been highlighted on several occasions, for example in Bader's second book:[2] "There is no suggestion in this demonstration of the correspondence between the number and relative positions of the maxima in $L(\boldsymbol{r})$ and electron pairs assumed in the Lewis model that the former proves the existence of the latter." This acknowledgement stems from a concern about consistency with earlier results. In 1975 it was demonstrated[20] that valence electrons in a molecule are, in general, not spatially localised into pairs. However, very recently new computational work supported the previously observed link between $L(\boldsymbol{r})$ and the VSEPR model. Bader and Heard[21] observed that most maxima in $L(\boldsymbol{r})$ can be associated with maxima in the Laplacian of the Lennard-Jones function. This discovery strengthens the role of $L(\boldsymbol{r})$ as a physical basis for the VSEPR model, and warrants a more complete investigation, as the current study aims to offer.

2.3 The function $L(\boldsymbol{r})$

Other than being widely used[12,22–69] $L(\boldsymbol{r})$ is simple and orbital independent. It measures the extent to which ρ is locally concentrated ($L(\boldsymbol{r}) > 0$) or depleted ($L(\boldsymbol{r}) < 0$). In which sense does $L(\boldsymbol{r})$

capture the "presence of an electron pair"? Bader and Heard[21] introduced the term *partial pair condensation*, which refers to regions of space exhibiting a pronounced reduction in the number of contributing electron pairs compared to a completely random distribution. For example, it can be rigorously shown[20] that the average pair population of the non-bonding pair on N in ammonia is well below the number expected from a distribution with a random pairing of electrons throughout the molecule.

The function $L(\boldsymbol{r})$ reveals the shell structure of a free atom. Zones where $L(\boldsymbol{r}) > 0$ are shells of charge concentration (CC), and where $L(\boldsymbol{r}) < 0$ we have charge depletion (CD) zones. These zones or shells survive after the formation of a molecule by adding ligands to a central atom. For example, in water we encounter[70] a core shell concentration (CSCC), a core shell charge depletion (CSCD), a valence shell charge concentration (VSCC) and a valence shell charge depletion (VSCD) zone. For every zone of charge concentration there is zone of charge depletion. In the link between $L(\boldsymbol{r})$ and the VSEPR model established before[13] the VSCC is identified with the sphere of maximum charge concentration on which the electron pairs are originally represented as points that avoid each other as much as possible. This picture will be abandoned in favour a more direct connection between domains and so-called basins in $L(\boldsymbol{r})$.

2.4 Quantum chemical topology (QCT)

The quantum chemical topological approach materialised from the development of subspace quantum mechanics,[1,71–73] which in turn derived from earlier work on atomic virial fragments.[74–76] A key concept in QCT is the *gradient path*, which is a curve in real space that follows the gradient of the 3D property density at hand, at any point on its track. A gradient path pursues the direction of steepest ascent until it terminates at a *critical point*, which is a point where the gradient vanishes. Details on the four possible types of critical point, (3,−3), (3,−1), (3,+1) and (3,+3) can be found in ref. 3. If the critical point is a maximum then it acts as an attractor to a multitude of gradient paths that collectively form a *basin*. Because a basin represents a region of space dominated by an attractor, it *is natural to assign any property density integrated over the basin to its attractor*. For example, ρ integrated over an atomic basin yields the atom's population. We adhere to this principle of assignment of integrated properties to basin attractors in the case of $L(\boldsymbol{r})$. Gradient paths that terminate at a (3,−1) saddle type of critical point form a two dimensional topological manifold called a *separatrix*, which is in effect the boundary between two basins.

2.5 Water: a case study

Fig. 1 shows how the critical points, their connecting gradient paths and the separatrices bounding the basins interrelate in one of the simplest possible non-trivial molecules: water. In Fig. 1 we show a wealth of topological information, which we discuss in turn.

In Fig. 1a six maxima are shown (green spheres). Two maxima practically coincide with the nuclear positions of the hydrogens. A so-called bonded maximum lies approximately on a line connecting the oxygen nucleus and a hydrogen nucleus. Another such maximum appears between the oxygen and the other hydrogen nucleus. Finally, two non-bonded maxima can be seen, at either side of the molecular plane. The latter two maxima can be related to the two familiar lone pairs of oxygen in water. Similarly the former (bonded) maxima are to be associated with the electron pairs of the OH bonds. It is in this sense that, almost two decades ago, it was observed that the maxima of $L(\boldsymbol{r})$ can faithfully represent the Lewis picture.

Fig. 1a shows a number of special gradient paths, marked in grey, each of which terminates at one of the six maxima shown. These gradient paths originate at (3,−1) critical points (purple spheres). As a network of gradient paths, the graph in grey should be associated with the VSCC. The other set of gradient paths, marked in blue, should be associated with the VCSD, a name originally coined in the first study[70] on the full topology of $L(\boldsymbol{r})$ in terms of critical points. It consists of paths that terminate at (3,+1) critical points (yellow) and typically originate from (3,+3) critical points. This set of gradient paths also spans the whole molecule. We refer to previous work[77,78] for more details on critical points and graphs in $L(\boldsymbol{r})$. Finally it should be mentioned that for any molecule there is a sea of maxima in $L(\boldsymbol{r})$ at infinity (not shown in the figure), which is sometimes described as a global attractor.

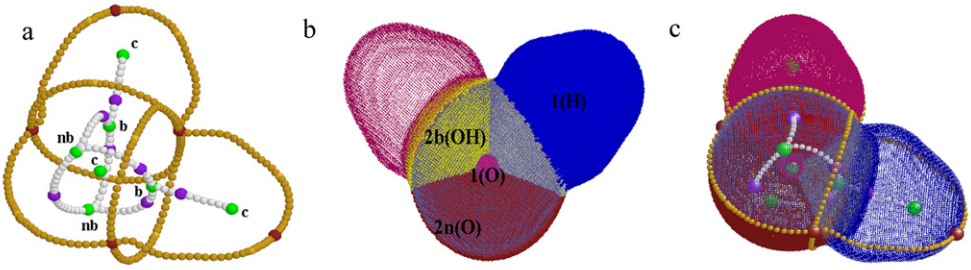

Fig. 1 The topology of $L(\boldsymbol{r})$ in water. (a) Network of gradient paths connecting selected critical points in the valence shell. Green, purple, yellow and brown spheres represent (3,−3), (3,−1), (3,+1) and (3,+3) critical points respectively. The maxima in $L(\boldsymbol{r})$, or the (3,−3) critical points, are labelled as core(c), bonding(b) or non-bonding(nb). Light grey paths originate at (3,−1) and orange paths terminate at (3,+1) critical points. The nuclear positions are effectively indicated by the three maxima marked as "core" since they practically coincide with them. (b) Representation of the basins in $L(\boldsymbol{r})$: two hydrogen cores, 1(H); one oxygen core, 1(O); two bonding, 2b(OH); and two non-bonding, 2n(O). The basins are bounded by separatrices characterised by scattered dots. (c) Superposition of the graphs in (a) and the basins in (b). The maxima are attractors for the basins they lie within. Note how the orange gradient paths gaze over the separatrices and thereby circumvent the basins.

In Fig. 1b we show the six basins in $L(\boldsymbol{r})$, each of which is dominated by one of the six maxima. We propose a transparent nomenclature to name these basins. For example, 2n(O) denotes a lone pair on oxygen, where the "2" refers to the valence shell of oxygen, which is the $n = 2$ quantum shell (sometime referred to as the "L" shell). The letter "n" designates the non-bonding character of this basin, and the "(O)" obviously refers to the oxygen. The bonding basins are also part of the valence shell, hence the number "2" in 2b(OH), and the "OH" refers to the sharing between oxygen and hydrogen. Basins belonging to the core shell ($n = 1$ quantum shell or "K") are marked as 1(O) for oxygen and 1(H) for hydrogen. Hydrogen is peculiar because in a way it lacks a core, or more strictly, its only electron is involved in its valence shell.

It is by *associating each basin with a domain of the VSEPR model* that we introduce a new view on the physical basis for VSEPR provided by $L(\boldsymbol{r})$'s topology. This association is most natural since it follows from the basic philosophy, applied in QCT, of ascribing any property integrated over the volume of a basin with the attractor maximum that dominates the basin. In the old VSEPR model electron pairs were thought of as points on a sphere centered at the nucleus of the central atom. The points were associated with the maxima in $L(\boldsymbol{r})$ and the sphere with the VSCC shell. The new VSEPR model shifts emphasis from points to the space occupied by a valence shell electron pair, called domain. This shift in emphasis is paralleled by our shift from critical points to basins. The price one pays is abandoning the prime role of the VSCC because a basin consists of regions where $L(\boldsymbol{r}) > 0$ (charge concentration) as well as where $L(\boldsymbol{r}) < 0$ (charge depletion). This means that a basin also contains a zone of charge depletion, where previously only charge concentration was expected in view of the partial pair condensation. The fact that $L(\boldsymbol{r})$ is a function that contains both positive and negative values, each corresponding to a qualitatively different local condition, is often a source of confusion, in contrast to the situation in ρ, which is everywhere positive.

There are two more important matters to note about Fig. 1b. Although we explain them in the context of the water molecule they are not unique to water. The first is an important difference in the topology of ρ and that of $L(\boldsymbol{r})$ for free molecules. In ρ there are three basins, one for each nucleus, which extend to infinity. This is why the basins are often capped by iso-density contours (typically $\rho = 0.001$ au for visual purposes). In $L(\boldsymbol{r})$, however, the six basins that we would logically bracket together to spanning the whole molecule, leave a large electronic population (about 3 e) unassigned. Of course, theoretically we should assign this population to the global attractor at infinity, but it is not clear what the meaning is of this assignment in chemical terms. In summary, we are facing a free water molecule, completely bounded by separatrices, immersed in an infinitely large outer basin that accounts for about a third of the total population. These outer basins are present for all other molecules that we have explored with our new algorithm, whether reported

here or not. Their population fluctuates around 3 *e* for second row hydrides but can amount to 10 *e* for PF_3 for example. The second matter that may appear unusual to readers used to the old way that $L(\boldsymbol{r})$ was linked to the VSEPR model is the 1(H) basin on hydrogen. The current VSEPR model would not account for a domain centered on the hydrogen nucleus in water. However, this basin is a QCT reality we cannot deny. Given its occurrence in many other hydrides this may modify the VSEPR model on long term.

Finally we superimpose the critical points and VSCC and VSCD graphs and the basins in Fig. 1c. This picture clearly shows the relation between the attractor and their domains, but more remarkably the way some gradient paths of the VSCD graph graze over the separatrices and seem to accentuate where they merge. This is an important observation, not made before to the best of our knowledge, since it means that the VSCD graph can give a crude description of the basins and their contacts. The graph is much cheaper to obtain computationally than the actual basin representation.

3. Computational methods

All optimised geometries and corresponding wave functions were obtained by the program GAUSSIAN98[79] at B3LYP/6-311+G(2d,p)//B3LYP/6-311+G(2d,p) level.[80–83] Molecular symmetry was imposed when present. In order for the graphs in $L(\boldsymbol{r})$ to be topologically stable one needs to use at least a triple zeta basis set. This was concluded from a previous study,[78] which isolated so-called atomic *L*-graphs from a set of thirty-one molecules. The critical points in $L(\boldsymbol{r})$ were located with a new algorithm[77] implemented in an extended local version of the program MORPHY98.[84] Figures were generated by the visualization program RASMOL.[85]

Central to the results in this contribution is the "octree algorithm", full details of which can be found elsewhere[15] but the essentials are briefly reviewed here. The driving force for the development of this algorithm stems from the need for general source code able to integrate a property density over the volume of a basin in *any* (other) property density. Such code is vital in the assessment of ever more sophisticated property density functions in support of extracting chemical information from modern wave functions. This task is currently the most feasible and important long-term goals of QCT. Another reason for the development of code along alternative and novel lines is the failure of classic analytical expressions[86] to represent separatrices in $L(\boldsymbol{r})$. A basin integration algorithm was proposed[87] before that operated without analytical described separatrices. However its implementation was specific to the electron density and did not allow for a recursive refinement in the delineation of the separatrices.

The "octree algorithm" is known in computer graphics[88] as the "octal tree search algorithm". The task of this algorithm[89] is to delineate the boundaries of an object in space, which is in our case one or more (contiguous) basins in $L(\boldsymbol{r})$. Firstly one surrounds the topological object with a box called the parent cube. Although it is convenient to refer to this box as a "cube" it is in reality a cuboid because its edges can have different lengths. Then we determine in which basin each of its eight vertices lie. If all vertices lie in the same basin then the algorithm stops. However if at least two vertices lie in a different basin the parent cube is subdivided into eight sub-cubes. Subsequently we determine in which basin the vertices of each sub-cube lie. If any two vertices lie in a different basin the sub-cube is itself subdivided, but if the vertices of the sub-cube lie in identical basins then the algorithm returns to the parent and investigates the next of the eight sub-cubes. This algorithm is obviously recursive and investigates space in trees of successively smaller cubes, hence the name "octal tree", where "octal" refers to the *eight* vertices. The algorithm keeps a record of the number of subdivisions made. It stores this information as the depth of the tree, where the initial parent cube is said to have zero depth. The results reported in this paper have been attained at depth seven. Although some issues have not been discussed here it should be clear that the octree algorithm is simple and robust. The greatest computational effort is assigning in which basin a given vertex lies because the most rigorous way of accomplishing this is by the compute intensive tracing of gradient paths. For this algorithm to become widely used further research is required to speed up this assignment process.

In the current algorithm the integration over a basin's volume is performed *via* a Gauss–Legendre quadrature grid, such as in the MORPHY98 code. Once the boundaries of the basins

have been delineated by the octree algorithm an integration ray sweeps through the full angular space and probes the boundaries (*i.e.* separatrices) of the basin. Once the intersections between an integration ray and the basin boundaries are located, radial integration intervals are determined and radial quadrature can proceed. Our source code independently determined the population of the outer basin in order to check the total population after summing populations of all basins. For practical reasons the outer basin was bounded by an arbitrary iso-density surface (*e.g.* $\rho = 10^{-6}$ au), considered to lie at infinity. The discrepancies between exact and recovered total population are typically of the order of a percent or less. Since the volumes of the outer basins are very large and arbitrary (because of the choice of iso-density surface) we do not report them.

4. Results and discussion

Table 1 shows the volumes and populations of the basins in $L(\boldsymbol{r})$ for the hydrides AH_n, A = C, N and O. As expected, the core basins 1(A) are very small but their populations are much lower than two. The "core" basins of hydrogen are about two orders of magnitude larger than the cores of *A* and have a population of about 0.6 *e*. The bonding basins, 2(AH), substantially decrease in volume as the electronegativity of the central *A* increases, from C to O. This trend is contrary to the (b)-part of the second VSEPR postulate because as the electronegativity of the central atom increases in going from C to O, the bonding domain is supposed to swell.

Previously Bader *et al.* estimated the relative sizes of bonding and non-bonding electron pairs by measuring an area on the VSCC "sphere". This basic approximation involved the average angle subtended at the nucleus by the lines linking the nucleus to the edges of the charge concentration in question. In this approach the boundary of a charge concentration is defined by the position of the neighbouring saddle point in $L(\boldsymbol{r})$. In our approach we can directly and rigorously measure the volume of a basin. It is clear from Table 1 that the volume of the non-bonding basins is considerably larger than that of the bonding basins *in the same molecule*. Hence we conclude that the first VSEPR postulate is corroborated by our computed values.

Table 2 shows the volumes and populations of the basins in $L(\boldsymbol{r})$ for the fluorides AF_3, A = N, P and Cl. This Table has been designed to further investigate the second VSEPR postulate on electronegativity, but first we comment on a few other observations. The core basins are again very small and shrink with increasing atomic number. It is helpful to consult Fig. 2, which shows all basins in (a) NF_3, (b) PF_3 and (c) ClF_3, to understand their relative positions and how they touch (*i.e.* their "synapticity"). The core basin 1(N) in NH_3 and NF_3 have almost the same volume and population. Again we confirm that the core basins of all non-hydrogen elements studied so far (C, N, O, F and Cl) have a population much closer to one than to two (this is also true for S and Si). It is reassuring to see that a non-bonding basin in the central atom is again much larger than its bonding counterparts, in the same molecule. However the populations of the non-bonding basins can vary from around two to more than five. The molecules NF_3 and PF_3 share the feature that the fluorine non-bonding basins at the same side ("*cis*") of the lone pair on N or P are slightly larger and more populated than the basins at the other side ("*trans*").

By comparing NH_3 and NF_3 we can scrutinise the link with the (a) part of the second VSEPR postulate. The volume of the bonding domain increases in going from NH_3 ($v = 2.3$ au) to NF_3 ($v = 2.7$ au). This is contrary to the statement made in (a) because bonding domains are expected to shrink with increasing electronegativity of the ligand. The VSEPR model relies on this postulate

Table 1 Volumes (v) and electronic populations (N) and of $L(\boldsymbol{r})$-basins in the hydrides methane, ammonia, and water

	v	N		v	N		v	N
CH_4	37.9		NH_3	23.0		H_2O	13.2	
1(C)	0.05	0.93	1(N)	0.03	0.96	1(O)	0.01	0.76
1(H)	5.4	0.62	1(H)	4.1	0.58	1(H)	3.2	0.57
2b(CH)	4.1	0.98	2b(NH)	2.3	0.97	2b(OH)	0.8	0.52
			2n(N)	3.8	1.60	2n(O)	2.6	1.83

Table 2 Volumes(v) and electronic populations(N) and of $L(\boldsymbol{r})$-basins in the fluorides NF_3, PF_3 and ClF_3

	v	N		v	N		v	N
NF_3	27.0		PF_3	56.4		ClF_3	42.0	
1(N)	0.03	0.98	1(P)	0.01	0.91	1(Cl)	0.00	0.93
1(F)	0.02	1.02	1(F)	0.02	1.02	$1(F_{ax})^b$	0.2	1.28
2b(NF)	2.7	0.92	3b(PF)	1.7	0.46	$1(F_{eq})^b$	0.02	0.98
2n(N)	5.9	2.29	3n(P)	33.0	5.17	$2(F_{ax})$	2.1	2.85
2n(F,*cis*[a])	2.4	2.68	3n(F,*cis*[a])	2.3	2.61	$2(F_{eq})$	2.1	2.38
2n(F,*trans*[a])	1.9	2.16	3n(F,*trans*[a])	2.0	2.29	$3b(ClF_{eq})$	0.4	1.46
			3n′(P)	4.3	0.69	3n(Cl)	14.2	3.67
			2n(P)	0.1	1.29	2n(Cl)	0.2	3.17
			2n′(P)	0.6	4.76			

[a] With respect to the position of the lone pair. [b] In this T-shaped conformation of ClF_3 there is one equatorial F lying in the plane of the Cl nucleus and the lone pairs (represented as the maxima in the 3n(Cl) basin and its mirror image). There are also two axial Fs lying on the F-Cl-F axis perpendicular to the aforementioned plane.

to predict correctly that the X–A–X bond angle is smaller in NF_3 than in NH_3. So, from this one example we learn that postulate 2(a) is not supported by a QCT analysis of $L(\boldsymbol{r})$.

However, matters are more complicated because, for PF_3 and PH_3, the VSEPR postulate 2(a) *fails* to predict the correct change in X–A–X angle. Gillespie and Robinson note[12] that hydrides of elements of groups 15 (N,P,*etc.*) and 16 (O,S,*etc.*) are exceptions to the rule that the X–A–X bond angle decreases with increasing electronegativity of X. These hydrides do not obey the VSEPR model and therefore the second postulate is not supposed to operate in the first place. Gillespie ascribes this to the unique character of hydrogen in that it has to play the role of both a bonding pair and a nonbonding pair. This is so because, according to Gillespie,[12] the domain of the A–H pair is not confined to the bonding region but extends into the nonbonding region of the hydrogen atom. Clearly this matter needs further investigation by including computations on extra molecules.

By comparing PF_3 and ClF_3 we can test postulate 2(b) again, since the electronegativity of Cl (3.0) far exceeds that of P (2.1). The postulate predicts that the bonding basin should swell in going from P to Cl. Table 2 communicates just the opposite. In summary, postulate 2(b) failed to be confirmed twice and when postulate 2(a) is valid it cannot be confirmed either.

Finally, Table 3 focuses on multiple bonding in an attempt to scrutinise the third postulate. It shows the volumes and populations of the basins in $L(\boldsymbol{r})$ for the single, double and triple CC bond, as it appears in ethane, ethene, and propyne respectively. It is obvious that the domains increase their volume going from a single bond (4.8) over a double bond ($2 \times 6.8 = 13.6$) to a triple bond (20.0). It should be noted that there are two bonding basins involved in the double bond but only

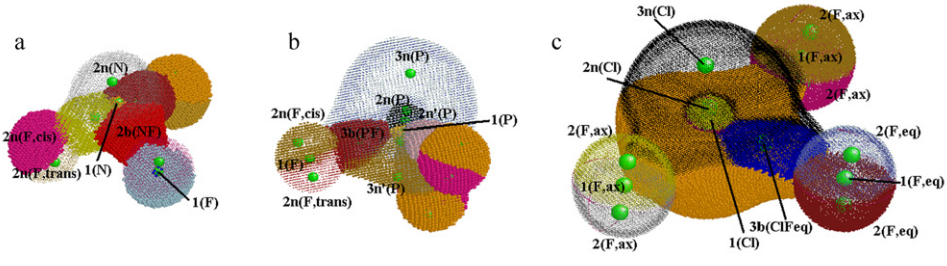

Fig. 2 A representation of $L(\boldsymbol{r})$-basins in (a) NF_3 (b) PF_3 and (c) CF_3. The volumes of populations of these basins are listed in Table 2. Some boundaries have been made semi-transparent in order to show their connectivity (or "synapticity") with other basins, as well as the attractor within, *i.e.* a maximum or (3,−3) critical point. The molecules NF_3 and PF_3 have C_{3v} symmetry while the T-shaped molecule ClF_3 has C_{2v} symmetry.

Table 3 Volumes (v) and electronic populations (N) and of $L(\boldsymbol{r})$-basins in ethane (staggered), ethene and propyne. The listed populations and volumes are averages over symmetry-connected atoms

	v	N		v	N		v	N
C_2H_6	67.1		C_2H_4	53.7		$CH_3C{\equiv}CH$	70.1	
1(C)	0.05	0.90	1(C)	0.06	0.94	1(C)	0.05	0.87
1(H)	5.5	0.63	1(H)	5.2	0.60	1(H)[a]	5.3	0.61
2b(CC)	4.8	1.13	2b(CH)	4.8	1.11	1(H)[b]	4.6	0.54
2b(CH)	4.0	1.00	2b(CC)	6.8	1.60	2b(CH)[c]	4.1	1.02
						2b(CH)[d]	6.0	1.36
						2b(C≡C)	20.0	4.81
						2b(C–C)[e]	6.8	1.46
						2b(C–C)[f]	4.2	1.05

[a] The "core" basin of a hydrogen in the methyl group. [b] The "core" basin of the terminal hydrogen, *i.e.* $CH_3C{\equiv}C\mathbf{H}$. [c] The bonded basin between C and H in the methyl group. [d] The bonded basin between a triply bonded C and the terminal H, *i.e.* CH_3CC–H. [e] There are two bonded basins between the triply bonded C and the methyl C. This is the basin nearest to the triple bond. [f] This is the basin nearest to the methyl group.

one in the triple bond. In summary the third postulate is fully supported. The value quoted in Table 3 for 2b(CC) is referring to a single basin. Doubling this value to 13.6 au would yield the total volume of the double bond region, still well under the 20.0 au of the triple bond. An interesting reflection is the high level of transferability of the volumes and populations of all basins (1(C), 1(H) and 2b(CH)) of the methyl group in methane compared to propyne. Finally we point out that there are two single bonded CC basins between the methyl C and the triply bonded C in propyne.

5. Conclusions

We take advantage of a new algorithm that is able to visualise basins in $L(\boldsymbol{r})$ directly, without using contour plots, and to calculate their volume and populations. For the first time we can address a long overdue question: does the *full* topology of $L(\boldsymbol{r})$, that is, beyond its critical points, support the VSEPR model? Following the QCT principle of assigning properties integrated over basins to their attractors we identify the basins in $L(\boldsymbol{r})$ with the VSEPR domains. Basins extend over regions where $L(\boldsymbol{r})$ is negative and positive, and hence the previous association of the VSCC with electron pairs has been modified. Based on a small set of nine molecules we are able to find support for two of the three subsidiary VSEPR postulates. We confirm that a non-bonding domain is larger than a bonding one, and that a multiple-bond domain is larger than a single-bond one. However we cannot corroborate the effect of the electronegativity of the central atom or ligand on the volume of bonding domains. Our preliminary study yields unexpectedly small populations, nearer to one than to two, for non-hydrogen cores and bonding domains. On the other hand non-bonding domains can have populations much larger than two. Overall we have to face the mixed and perhaps counterintuitive results that the full topology of $L(\boldsymbol{r})$ yields.

Acknowledgements

Gratitude is expressed to EPSRC who sponsors this work *via* grant GR/N04423.

References

1 R. F. W. Bader and T. T. Nguyen-Dang, *Adv. Quantum Chem.*, 1981, **14**, 63.
2 R. F. W. Bader, *Atoms in Molecules. A Quantum Theory*, Oxford University Press, Oxford, GB, 1990.
3 P. L. A. Popelier, *Atoms in Molecules. An Introduction*, Pearson Education, London, GB, 2000.
4 R. J. Gillespie and I. Hargittai, *The VSEPR Model of Molecular Geometry*, Allyn & Bacon and Prentice Hall, 1991.
5 R. J. Gillespie, *Molecular Geometry*, Van Nostrand Reinhold, 1972.

6 R. J. Gillespie and R. S. Nyholm, *Q. Rev. Chem. Soc.*, 1957, 339.
7 N. V. Sidgwick and H. E. Powell, *Proc. R. Soc. London, Ser. A*, 1940, **176**, 153.
8 R. J. Gillespie, *Struct. Chem.*, 1998, **9**, 73.
9 I. Mayer, *Struct. Chem.*, 1997, **8**, 309.
10 J. E. Lennard-Jones, *J. Chem. Phys.*, 1952, **20**, 1024.
11 J. E. Lennard-Jones, *Adv. Sci., London*, 1954, **11**, 136.
12 R. J. Gillespie and E. A. Robinson, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 495.
13 R. F. W. Bader, R. J. Gillespie and P. J. MacDougall, *J. Am. Chem. Soc.*, 1988, **110**, 7329.
14 R. F. W. Bader, P. J. MacDougall and C. D. H. Lau, *J. Am. Chem. Soc.*, 1984, **106**, 1594.
15 N. O. J. Malcolm and P. L. A. Popelier, *J. Comput. Chem.*, 2003, in press.
16 R. J. Gillespie and P. L. A. Popelier, *Chemical Bonding and Molecular Geometry from Lewis to Electron Densities*, Oxford University Press, New York, USA, 2001.
17 R. McWeeny, *Methods of Molecular Quantum Mechanics*, ed. D. P. Craig and R. McWeeny, Academic Press, 2nd edn., 1992.
18 P.-O. Lowdin, *Phys. Rev.*, 1955, **97**, 1474.
19 R. J. Gillespie, D. Bayles, J. Platts, G. L. Heard and R. F. W. Bader, *J. Phys. Chem. A*, 1998, **102**, 3407.
20 R. F. W. Bader and M. E. Stephens, *J. Am. Chem. Soc.*, 1975, **97**, 7391.
21 R. F. W. Bader and G. L. Heard, *J. Chem. Phys.*, 1999, **111**, 8789.
22 M. T. Carroll, C. Chang and R. F. W. Bader, *Mol. Phys.*, 1988, **63**, 387.
23 M. T. Carroll and R. F. W. Bader, *Mol. Phys.*, 1988, **65**, 695.
24 M. T. Carroll, J. R. Cheeseman, R. Osman and H. Weinstein, *J. Phys. Chem.*, 1989, **93**, 5120.
25 T. H. Tang, R. F. W. Bader and P. J. MacDougall, *Inorg. Chem.*, 1985, **24**, 2047.
26 P. J. McDougall, G. J. Schrobilgen and R. F. W. Bader, *Inorg. Chem.*, 1989, **28**, 763.
27 P. L. A. Popelier and P. J. Smith, *Phys. Chem. Chem. Phys.*, 2001, **3**, 4208.
28 R. F. W. Bader and P. J. MacDougall, *J. Am. Chem. Soc.*, 1985, **107**, 6788.
29 T.-H. Tang, W.-J. Hu, D.-Y. Yan and Y. P. Cui, *J. Mol. Struct. (THEOCHEM)*, 1990, **207**, 327.
30 S. T. Howard, J. P. Foreman and P. G. Edwards, *Can. J. Chem.*, 1997, **75**, 60.
31 S. T. Howard and J. A. Platts, *J. Phys. Chem.*, 1995, **99**, 9027.
32 P. J. McDougall and C. E. Henze, *Theor. Chem. Acc.*, 2001, **105**, 345.
33 Y. Aray, J. Rodriguez and R. Lopez-Boada, *J. Phys. Chem. A*, 1997, **101**, 2178.
34 Y. Aray and J. Rodriguez, *Surf. Sci.*, 1998, **405**, 532.
35 K. B. Wiberg, R. F. W. Bader and C. D. H. Lau, *J. Am. Chem. Soc.*, 1987, **109**, 985.
36 P. L. A. Popelier and R. F. W. Bader, *Chem. Phys. Lett.*, 1992, **189**, 542.
37 P. F. Zou and R. F. W. Bader, *Acta Crystallogr., Sect. A*, 1994, **A50**, 714.
38 V. G. Tsirelson, P. F. Zou, T. H. Tang and R. F. W. Bader, *Acta Crystallogr., Sect. A*, 1995, **A51**, 143.
39 R. J. Boyd and S. C. Choi, *Chem. Phys. Lett.*, 1985, **120**, 80.
40 M. E. Eberhart, M. M. Donovan, J. M. McLaren and D. P. Clougherty, *Prog. Surf. Sci*, 1991, **3**, 1.
41 M. E. Eberhart, D. P. Clougherty and J. M. McLaren, *J. Mater. Res.*, 1993, **8**, 438.
42 R. P. Sagar, A. C. T. Ku and V. H. Smith, *J. Chem. Phys.*, 1988, **88**, 4367.
43 Y. Aray and R. F. W. Bader, *Surf. Sci.*, 1996, **351**, 233.
44 Y. Aray, F. Rosillo and J. Murgich, *J. Am. Chem. Soc.*, 1994, **116**, 10 639.
45 Y. Aray, J. Rodriguez, J. Murgich and F. Ruette, *J. Phys. Chem.*, 1993, **97**, 8393.
46 Y. Aray and J. Rodriguez, *Can. J. Chem.*, 1996, **74**, 1014.
47 Y. Aray, J. Rodriguez, J. Rivero and D. Vega, *Surf. Sci.*, 1999, **441**, 344.
48 Y. Aray, J. Rodriguez and D. Vega, *J. Phys. Chem. B*, 2000, **104**, 5225.
49 J. M. Ricart, M. P. Habas, A. Clotet, D. Curulla and F. Illas, *Surf. Sci.*, 2000, **460**, 170.
50 I. Bytheway and M. B. Hall, *Inorg. Chem.*, 1995, **34**, 3741.
51 R. F. W. Bader, R. J. Gillespie and F. Martin, *Chem. Phys. Lett.*, 1998, **290**, 488.
52 R. J. Gillespie, I. Bytheway, T. H. Tang and R. F. W. Bader, *Inorg. Chem.*, 1996, **35**, 3954.
53 R. J. Gillespie, I. Bytheway, R. S. Dewitte and R. F. W. Bader, *Inorg. Chem.*, 1994, **33**, 2115.
54 R. Gillespie and S. A. Johnson, *Inorg. Chem.*, 1997, **36**, 3031.
55 I. Bytheway, R. J. Gillespie, T. H. Tang and R. F. W. Bader, *Inorg. Chem.*, 1995, **34**, 2407.
56 I. Bytheway, P. L. A. Popelier and R. J. Gillespie, *Can. J. Chem.*, 1996, **74**, 1059.
57 J. A. Platts, S. T. Howard and B. R. F. Bracke, *J. Am. Chem. Soc.*, 1996, **118**, 2726.
58 A. Sierraalta, F. Ruette, E. Machado and E. Machado, *Int. J. Quantum Chem.*, 1998, **70**, 113.
59 C. Levit and J. Sarfatti, *Chem. Phys. Lett.*, 1997, **281**, 157.
60 I. P. Hamilton, *Chem. Phys. Lett.*, 1998, **297**, 261.
61 J. M. Maestre, J. P. Sarasa, C. Bo and J. M. Poblet, *Inorg. Chem.*, 1998, **37**, 3071.
62 S. H. Choi and Z. Y. Lin, *J. Organomet. Chem.*, 2000, **608**, 42.
63 M. M. Rohmer, M. Benard and J. M. Poblet, *Chem. Rev.*, 2000, **100**, 495.
64 J. Sarasa, J. M. Poblet and M. Benard, *Organometallics*, 2000, **19**, 2264.
65 M. Blattner, M. Nieger, A. Ruban, W. W. Schoeller and E. Niecke, *Angew. Chem., Int. Ed. Engl.*, 2000, **39**, 2768.
66 D. V. Deubel, *J. Org. Chem.*, 2001, **66**, 2686.
67 R. S. Gopalan, G. U. Kulkarni and C. N. R. Rao, *ChemPhysChem*, 2000, **1**, 127.

68 P. L. A. Popelier, F. M. Aicken and S. E. O'Brien, in *Chemical Modelling: Applications and Theory, RSC Specialist Periodical Report*, ed. A. Hinchliffe, 2000, vol. 1, ch. 3, pp. 143–198.
69 P. L. A. Popelier and P. J. Smith, in *Chemical Modelling: Applications and Theory, RSC Specialist Periodical Report*, ed. A. Hinchliffe, 2002, vol. 2, ch. 8, pp. 391–448.
70 P. L. A. Popelier, *Coord. Chem. Rev.*, 2000, **197**, 169.
71 R. F. W. Bader, T. T. Nguyen-Dang and Y. Tal, *J. Chem. Phys.*, 1979, **70**, 4316.
72 R. F. W. Bader, *Pure Appl. Chem.*, 1988, **60**, 145.
73 R. F. W. Bader, Y. Tal, S. G. Anderson and T. T. Nguyen-Dang, *Isr. J. Chem.*, 1980, **19**, 8.
74 R. F. W. Bader and P. M. Beddall, *J. Chem. Phys.*, 1972, **56**, 3320.
75 R. F. W. Bader and P. M. Beddall, *J. Am. Chem. Soc.*, 1973, **95**, 305.
76 R. F. W. Bader, P. M. Beddall and J. J. Peslak, *J. Chem. Phys.*, 1973, **58**, 557.
77 N. O. J. Malcolm and P. L. A. Popelier, *J. Comput. Chem.*, 2003, **24**, 437.
78 P. L. A. Popelier, J. Burke and N. O. J. Malcolm, *Int. J. Quantum Chem.*, 2003, **92**, 326.
79 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, *GAUSSIAN 98 (Revision A.7)*, Gaussian, Inc., Pittsburgh, PA, 1998.
80 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785.
81 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.
82 R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1980, **72**, 650.
83 J. B. Foresman and A. Frisch, *Exploring Chemistry with Electronic Structure Methods*, Gaussian Inc., Pittsburgh, PA, 1996.
84 P. L. A. Popelier, *MORPHY98*, UMIST, Manchester, GB 1998. http://morphy.ch.umist.ac.uk/.
85 R. Sayle, in *RasMol, a molecular visualisation program*, Glaxo, Stevenage, GB, 1994.
86 P. L. A. Popelier, *Theor. Chim. Acta*, 1994, **87**, 465.
87 P. L. A. Popelier, *Comp. Phys. Commun.*, 1998, **108**, 180.
88 J. D. Foley, A. van Dam, S. K. Feiner, J. F. Hughes and R. L. Philips, *Introduction to Computer Graphics*, Addison-Wesley, USA, 1990.
89 N. O. J. Malcolm and P. L. A. Popelier, *J. Comput. Chem.*, 2003, in press.