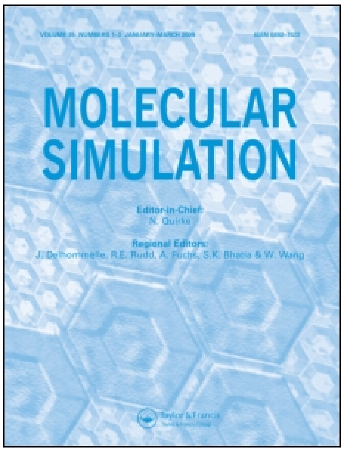

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Two Algorithms for Defining Atoms and Molecules in Condensed Phases

L. Delle Site[a][b]
[a] Atomistic Simulation Group and Irish Centre for Colloid Science, School of Mathematics and Physics, The Queen's University, Belfast, U.K. [b] Theory Group, Max-Planck-Institute for Polymer Research, Mainz, Germany

To cite this Article Site, L. Delle(2001) 'Two Algorithms for Defining Atoms and Molecules in Condensed Phases', Molecular Simulation, 26: 3, 217 — 235

To link to this Article: DOI: 10.1080/08927020108028294

URL: http://dx.doi.org/10.1080/08927020108028294

Molecular Simulation, 2001, Vol. 26, pp. 217-235
Reprints available directly from the publisher
Photocopying permitted by license only

Published by license under
the Gordon and Breach Science
Publishers imprint.
Printed in Malaysia.

TWO ALGORITHMS FOR DEFINING ATOMS AND MOLECULES IN CONDENSED PHASES

L. DELLE SITE*

Atomistic Simulation Group and Irish Centre for Colloid Science, School of Mathematics and Physics, The Queen's University, Belfast BT7 1NN, U.K.

(Received December 1999; accepted December 1999)

Two algorithms for defining single entities in condensed phase based on the calculation of a zero flux surface of the gradient of the electron density are proposed. First the two approaches are applied to a water molecule in a liquid system; the different behaviour of the two algorithms is graphically shown in the planar section of the molecular plane (*i.e.*, HOH plane of the molecule analyzed). Next, using the two algorithms for partitioning electron density, the average molecular dipole moment of the water molecule in liquid phase is calculated averaging over several configurations; analogies and differences with other methods are reported and their physical interpretation is discussed. There exist different approaches for calculating the zero flux surface (or equivalently the volume defined by such a surface); the two described in this work differ from the others because the calculation in both cases is straightforward (*i.e.*, the surface is determined directly from the data corresponding to the electron density) and is characterized by a relatively short computational time.

Keywords: Electron density; Molecular dipole; Modeling

1. INTRODUCTION

The spatial partitioning of the electron density in a condensed system represents a powerful tool for modeling non isolated atoms and molecules. There exist different methods to do so (see for example [1-3]); here we focus our attention on the topological behaviour of the electron density and define the region bounded by a zero flux surface of the gradient of the electron

---

*Present address: Max-Planck-Institute for Polymer Research, Theory Group, P.O. Box 3148, D-55021 Mainz, Germany. e-mail: dellsite@mpip-mainz.mpg.de

density as the spatial definition of a specified atom or molecule in a condensed phase (for more details see [2]). Mathematically the zero flux surface of the gradient of the density is defined as the surface $S(\mathbf{r})$ which satisfies the following condition:

$$\nabla\rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \tag{1.1}$$

for any point $\mathbf{r}$ of the surface S, where $\mathbf{n}(\mathbf{r})$ is the unit vector perpendicular to $S(\mathbf{r})$ at each point $\mathbf{r}$ of the surface S. This means that the lines of gradient (or trajectories of gradient) of the electron density originating from the core of an atom or a molecule will not cross the surface. A trajectory of $\nabla\rho(\mathbf{r})$ is obtained calculating at a given point $\mathbf{r}_0\nabla\rho(\mathbf{r}_0)$ and then moving of an amount $d\mathbf{r}$ from $\mathbf{r}_0$ along the direction indicated by the vector $\nabla\rho(\mathbf{r}_0)$, and repeating the procedure until the trajectory terminates. This criterion gives an exact topological definition of an atom or a molecule from both quantum and classical point of view (see [2, 4]) but the practical calculation of the surface, because of the topological complexity of the electron density and the discrete set of data available in most of the numerical calculations, represents a difficult problem. In this work we propose two different algorithms to calculate the zero flux surface when the electron density has been calculated using a pseudopotential approach (the extension to the case of all-electron density is straightforward) and is distributed on a discrete grid in real space; this is the case of most of the numerical calculations based on a quantum treatment of the electrons. The two methods, which we will refer to as the "downhill" method and the "triangulation" method, are described in the following sections.

2. NUMERICAL CALCULATION OF THE ZERO FLUX SURFACE

2.1. Downhill Method

The discrete nature of ρ makes feasible a convenient implementation of the criterion of zero flux-surface on the grid. Sitting on the points of the grid which correspond to the core of a particular atom or molecule, one natural way to identify, with reasonable approximation, the volume inside the surface of flux zero of $\nabla\rho(\mathbf{r})$ consists in exploring (from each point in the core) each neighbouring point in turn, accepting it if its density is less or equal to that of the first point and if the gradient in the direction is downhill. The gradient is estimated by linear interpolation formula. The procedure is then repeated for each accepted point until a turning point in the function $\rho(\mathbf{r})$ is found. In the region of the core the density is automatically assigned to the atom or to the molecule, so that the criterion in practice, is applied only for points outside the core. There are two reasons for defining a core; in the case of an atom, the pseudopotential used for the calculation generates a minimum for the density at the nuclear site as only the valence electrons are taken into account. This leads to the necessity of defining a region around each nuclear site which is larger than the region of inner electrons where the valence electron is. In the case of a molecule, the region of the core is defined as a volume which contains the core of the atoms forming the molecule and satisfies the condition that the gradient at each point of the surface bounding the core is downhill. This latter requirement avoids the possibility of being trapped in the volume defined by the zero flux-surface of one of the atoms. Points which can be reached from two molecules defines the zero flux-surface which separates the two molecules, those which can be reached from three molecules define lines on this surface, *etc.* The electron density associated with these shared points is counted with the appropriate weight (for example 1/2, 1/3, *etc.*). In practice there are problems in the implementation of this criterion arising from the coarseness of the grid, which means that some other points near the zero flux surface can be reached from different molecules. In fact suppose that the "true" zero flux-surface is located between two points of the grid characterized by the same value of the density. These two points are equally reachable from two different molecules so that instead of a surface, we will have an overlapping region around the true surface. As for the shared points on the surface, these points, in first approximation, can be calculated with the appropriate weight. The second criterion for accepting a neighbouring point (the gradient in the direction is downhill) should solve in part this problem. This method turned out to be reliable and useful for analyzing properties of pure liquid water (see [5]) and solvation (see [6]).

2.2. The Triangulation Method

2.2.1. *Introduction*

The second algorithm we name the "triangulation method" represents an attempt to develop a procedure in which the calculation of the zero flux surface is as independent as possible of the size of the grid spacing. The basic idea is that sitting at the centre of the molecule, we can determine the zero flux surface searching for the turning point of $\rho(\mathbf{r})$ along each angular directions; next we show in details how it is possible to implement this idea in numerical calculations. Consider a unit sphere centred on the centre of a specified molecule. Each point P on the sphere will be identified by the polar and the azimuthal angles θ, ϕ, so that $P = P(\theta, \phi)$. If we are placed at the centre of the sphere, each radial direction is characterized by the value of θ and ϕ. Suppose we move from the centre along a radial direction until we find a turning point for the function $\rho(\mathbf{r})$ at a point $(\mathbf{r}, \theta, \phi)$, if the procedure is repeated for all the radial directions (a continuous set of θ and ϕ) a surface $S = S(\theta, \phi, r(\theta, \phi))$ is obtained which can be identified with the zero flux surface for that specified atom or molecule. Again the problem of the atomic pseudopotential or the molecular core is avoided by searching for the turning point outside the core region as for the case of the "downhill method". In practice the fact that we deal with a density on a grid again requires approximations in the implementation of the procedure. The directions along which the turning points of ρ are determined must be represented by a discrete set of the values θ and ϕ. The range of ϕ is $[0, \pi]$, while for θ is $[0, 2\pi]$. The number of intervals in which θ, ϕ are divided is given by $(2\pi/n)$ and (π/m), where n, m are two integers. These two parameters can be chosen on the base of mathematical or physical considerations; for example one could require the topological regularity of the surface S (mathematical criterion) or could require that a physical quantity (for example the molecular dipole moment) as a function of m and n is characterized by a negligible fluctuation, with respect to the desired accuracy, as m and n are varied of a certain amount (this latter criterion allows one to identify a valid range in which m and n can be chosen).

2.2.2. *Determining the Discrete Zero Flux Surface*

Once the set of discrete directions $(\theta_i, \phi_j, i = 1, m; j = 1, n)$ is chosen, a further step can be done; this section is dedicated to the description of the procedure for determining the points of the zero flux surface along the directions defined by the each pair (θ_i, ϕ_j) while the next two sections are dedicated respectively to the description of the procedure for constructing a continuum surface from the discrete one and to the procedure developed to identify grid points inside this surface. For each pair θ_i, ϕ_j the turning point of $\rho(\mathbf{r})$ along the corresponding direction outside the core region can be determined. As the electron density is calculated on a grid, the interpolation along the specified direction has to be computed. The procedure applied to move along the radial directions consists in choosing a step Δr (usually the choice of Δr equal to the grid spacing is satisfactory) which automatically defines:

$$\begin{aligned} r_{s+1} &= r_s + \Delta r \\ s &= 1,\ s_0 \\ s_0 \cdot \Delta r &= R \end{aligned} \tag{2.1}$$

the parameter R can be fixed choosing a value which cover a region in the space where the molecules is certainly contained, for example an extreme case would be the linear dimension of the simulation box. When the condition:

$$\rho(\theta_0, \phi_0, r_s) \leq \rho(\theta_0, \phi_0, r_s + \Delta r) \tag{2.2}$$

is fulfilled, a point $S(\theta_0, \phi_0, r_s)$ of the zero flux surface is found. The procedure is repeated for all pairs $(\theta_i,\ \phi_j)$ so that a discrete surface is defined. The problem that we still need to solve is how to implement the condition that all the points placed in the volume inside the surface S are assigned to the specified atom or molecule. We can notice that each point of the grid can be identified respect to the centre of the molecule by three quantities: r, θ, ϕ where θ, ϕ have the same meaning as specified before and indicate the direction of the vector obtained joining the centre of the atom or molecule with that point, and r specifies the corresponding distance. If the values of θ and ϕ for a point $B(r,\ \theta_B\phi_B)$ on the grid were the same of a corresponding point on S, $S(r',\ \theta_S\phi_S)$ the condition of assignment of B to the atom (or molecule) would have been reduced to the comparison of r with r'; inside the surface if the former is shorter than or equal to the latter, outside in the other case. In general the situation described above is only a coincidence, so that a general criterion must be specified; to do so a convenient way is to build a continuum surface from the discrete one and as it will be shown, the criterion for accepting or rejecting a grid point will be reduced to a simple geometrical condition.

2.2.3. Building a Continuum Zero Flux Surface

An efficient method to build a continuum geometrical surface from a discrete set of points is represented by the technique of "triangulation of the surface". This method is based on the fact that three points uniquely define a plane and so the discrete surface can be locally approximated by a triangle whose vertices are three points of the discrete surface.

Different kind of triangulation can be implemented choosing different criteria to define the triangles [7–9]. We are going to describe our criterion. Consider the grid point $B(r_B, \theta_B, \phi_B)$, this point can be projected on the unit sphere so that we obtain the point $B^{\mathrm{proj}}(\theta_B, \phi_B)$. Next we identify the value θ_i and ϕ_j (θ_i and ϕ_j belong to the discrete set of angles used to identify the points of the zero flux surface) which are closest to θ_B and ϕ_B. In Figure 1 the points B_a, B_b, B_c, B_d represent the possible locations of B^{proj} once the closest θ_i and ϕ_j are found. Once θ_i and ϕ_j are found we go back to the points of the zero flux surface and pick up the point characterized by the value θ_i and ϕ_j. In this way the first vertex of the triangle P_1 which corresponds to the point $S(r', \theta_i \phi_j)$ of the zero flux-surface, is found (see Fig. 1). Then if ϕ_B is between ϕ_{j-1} and ϕ_j, a second vertex (P_2 in the upper part of Fig. 1) which corresponds to the point $S(r'', \theta_i, \phi_{j-1})$ of the zero flux surface is identified, while in case ϕ_B is between ϕ_j and ϕ_{j+1} this

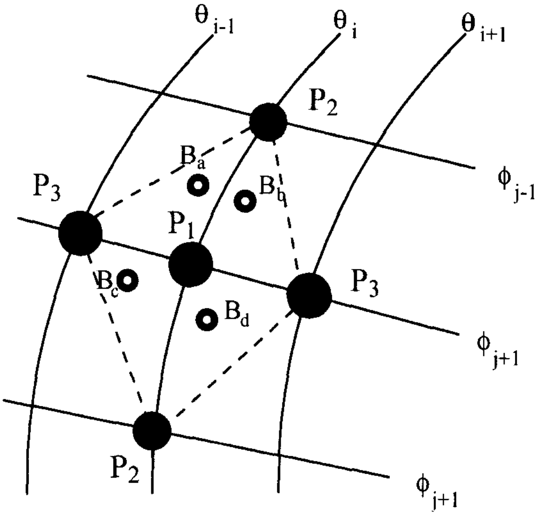

FIGURE 1 The diagram illustrates the possible locations (B_a, B_b, B_c, B_d) of a grid point characterized by angular coordinates θ and ϕ on the unit sphere. Once the angles θ_i and ϕ_j which are closest to θ and ϕ, are determined, the first point of the triangle which will locally approximate the continuum zero flux surface, is obtained (P_1). At this point there are 4 possibilities: (1) ϕ is smaller than ϕ_j and θ is smaller than θ_i. The triangle which approximates the real zero flux surface will be determined by the three vertices P_1, P_2, P_3 which contains the point B_a. (2) ϕ is smaller than ϕ_j and θ is bigger than θ_i. The triangle which approximates the real zero flux surface will be determined by the three vertices P_1, P_2, P_3 which contains the point B_b. (3) ϕ is bigger than ϕ_j and θ is smaller than θ_i. The triangle which approximates the real zero flux surface will be determined by the three vertices P_1, P_2, P_3 which contains the point B_c. (4) ϕ is bigger than ϕ_j and θ is bigger than θ_i. The triangle which approximates the real zero flux surface will be determined by the three vertices P_1, P_2, P_3 which contains the point B_d. The points P_1, P_2, P_3, belongs to the discrete zero flux surface.

point will be $S(r'', \theta_i, \phi_{j+1})$ (P_2 in the lower part of Fig. 1). The same procedure is repeated for θ_B so that a third vertex $S(r''', \theta_{i-1}, \phi_j)$ (P_3 on the left side of Fig. 1) or $S(r'', \theta_{i+1}, \phi_j)$ (P_3 on the right side of Fig. 1) is found. The situation is described by Figure 1. Once the vertices P_1, P_2, P_3 are defined, it is possible to identify the plane on which they lie that locally approximate the zero flux surface.

2.2.4. Condition for Accepting a Grid Point

At this stage we can determine whether or not a grid point is inside the surface by considering two vectors which identify the triangle, the unit vector normal to the plane of the triangle and the vector which specifies the position of the grid point with respect to the centre of the molecule; it follows the description of such a procedure. Consider the vectors OP_1, OP_2, OP_3 (see Fig. 2) where O is the centre of the atom or molecule. The vectors $P_1P_2 = OP_2 - OP_1$ and $P_1P_3 = OP_3 - OP_1$ define a unique plane π. The grid point B is placed along the direction of the vector OM, where M is the intersection of the line along the direction of θ_B, ϕ_B with

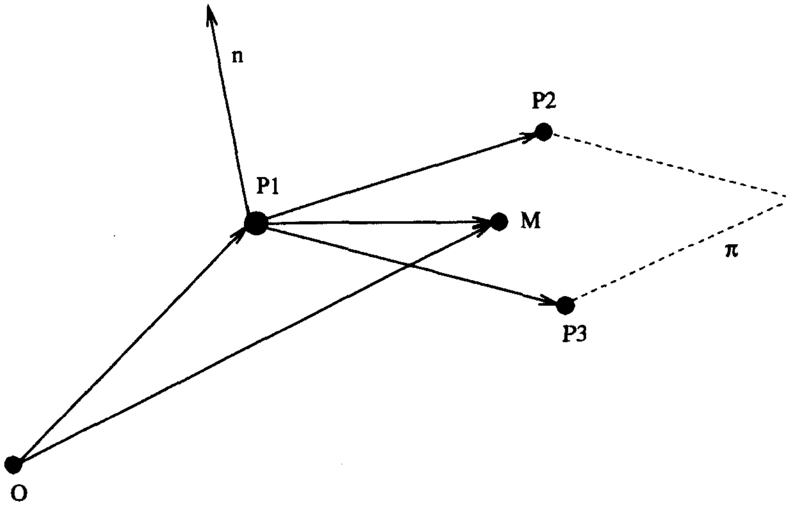

FIGURE 2 The diagram shows the plane π in which the triangle defined by P_1, P_2, P_3 (see the previous figure) lies. The continuum zero flux surface is locally approximated by the planar surface contained in the triangle. n is the unit vector normal to the plane π. O is the centre of the molecule, M is the intersection of the line originating from O and π; it is characterized by the direction given by θ and ϕ. The grid point we are considering is situated along this line. If it is situated in the segment OM then it is contained in the volume defined by the zero flux surface, otherwise is outside.

the plane π (see Fig. 2). In this case three situations are possible:

(a) B coincides with M so it belongs to the zero flux surface and is assigned to the atom or molecule.
(b) B is placed along the line characterized by the vector OM but not in the segment between O and M. This means that the point is located outside the surface of zero flux and so it does not belong to the atom or molecule.
(c) B is located between O and M so it belongs to the atom or molecule.

To understand which one of the three cases we are in, for the particular B considered, we proceed in the following way. We calculate $\mathbf{n}$ the unit vector normal to the plane π (see Fig. 2):

$$\mathbf{n} = \frac{P_1P_2 \times P_1P_3}{|P_1P_2 \times P_1P_3|} \tag{2.3}$$

the vector OM can be written as:

$$OM = \lambda \cdot \mathbf{u}_B \tag{2.4}$$

where $\mathbf{u}_B$ is the unit vector which indicates the direction of the vector OM:

$$\mathbf{u}_B = (\cos\theta_B \sin\phi_B, \sin\theta_B \sin\phi_B, \cos\phi_B). \tag{2.5}$$

As M belongs to π it follows that:

$$\lambda = \frac{\mathbf{n} \cdot OP_1}{\mathbf{n} \cdot \mathbf{u}_B}. \tag{2.6}$$

Once λ is determined we can finally reduce the procedure of assigning a point to the molecule to a simple criterion:

If $r_B \leq \lambda$ the point is assigned to the atom or molecule, otherwise is rejected.

In fact the condition $r_B \leq \lambda$ coincides with the situations described by points (a) and (c), while $r_B > \lambda$ corresponds to the case of point (b).

As for the case of the downhill method, points which are reached from two different atoms or molecule, define the zero flux surface, those from three define lines on this surface and so on. For these points the density is counted with the appropriate weight. As the previous method this one is an approximation for the calculation of the "true" zero flux surface. In this case the "true" lines of gradient are approximated by radial lines of gradient.

3. COMPARISON BETWEEN THE TWO ALGORITHMS

3.1. The Difference

The two approaches developed for defining an atom or a molecule are based both on the same criterion: define the surface of flux zero of the gradient of the density.

The way they differ is due to the fact that we deal with a discrete grid based approximation of the function $\rho(\mathbf{r})$. As stated earlier, the zero flux surface is the surface along which the lines of gradient (or the trajectories of $\nabla\rho$) originated from two different atoms or molecules, are parallel to each other and the cross over is forbidden. What in practice the two methods do, is to follow trajectories of $\nabla\rho$ which are not the "true" lines but represent an approximation. The different way the two methods approximate these lines makes them different. The downhill method based on the principle that neighbouring points are explored and then accepted or rejected, approximates the "true" lines with a sample of stepped paths which terminate at turning points of ρ. The triangulation method determines directly the discrete zero flux surface and then builds in a rather accurate way a continuum surface. In this second case, the lines of forces are approximated by radial lines. It is clear that as the grid spacing becomes smaller the difference between the two methods in determining the volume inside the zero flux surface decreases. This is the crucial question. In the case of plane-waves based *ab initio* calculation for a system of 32 water molecules, in a box of 9.8 Å, an energy cut-off of 90Ry in Fourier space leads to a grid spacing in real space of 0.0817 Å; the cutting process, with both methods, applied to densities generated from calculations performed with different energy cut-offs, from 60Ry up to 90Ry, gives evidence that the methods are not equivalent. This means the grid spacing needs to be much smaller in order to effect the results in a way that the results of the two methods are equivalent. This is a limitation of *ab initio* calculations, because a cut-off of 90Ry is already computationally very expensive. However an encouraging reason to use the two algorithms is the fact that for both, for example in the case of water molecules, the charge associated to each molecule after the cutting process is applied, is very close to the expected one with a fluctuation of about 0.1% for the single molecule, and even less for the whole system. As these methods have no intrinsic requirement for the conservation of the charge, this can be considered as evidence that they capture the essential physics contained in the density.

3.2. Application to a Molecule in a Liquid Water System

In order to understand more about the two methods we analyzed a particular system. Consider a system of 32 classical water molecules (SPC/E model [10]) in a cubic box of dimension 9.845 Å where cubic boundary conditions are applied. The energy cut-off for this calculation is 60Ry which leads to a grid spacing of 0.102 Å. One particular molecule is fixed at the centre of the box, lying with its nuclei on the horizontal plane. The remaining 31 molecules are allowed to move and the system is brought to equilibration. This classical configuration is then used as an input for an *ab initio* wave-function optimization. During the *ab initio* run the molecules do not move but are "dressed" with electron density calculated with the *ab initio* technique. We used the code FEMD, a pseudopotential based *ab initio* approach (for more details see [5, 11, 12]). In the calculation of the water system we kept one molecule fixed in the horizontal plane at the centre of the box and this will allow us to analyze planar sections of the density for that particular molecule so that we can compare the two methods visualizing density contours. Figures 3, 4, 5 show a section of the molecule before cutting 3 and after the two methods are applied 4 (triangulation), 5 (downhill). In particular Figures 4 and 5 show a section of the grid in the region of the molecule we are interested in when the two cutting methods are applied and the plane of section is fulfilled with the points picked up using the two algorithms. The different colours indicate the range of the density, the circles, the squares and the triangles indicate whether the point belongs only to one molecule, is shared by two molecules or by three. The comparison between Figure 3, with Figures 4 and 5 shows that the two methods are rather equivalent in regions where the density is high, but show differences in regions characterized by low density. An encouraging thing to notice is that in both cases they cut the molecule in a reasonable way, in the sense that is what intuitively we would think as a molecule when we analyze the contours

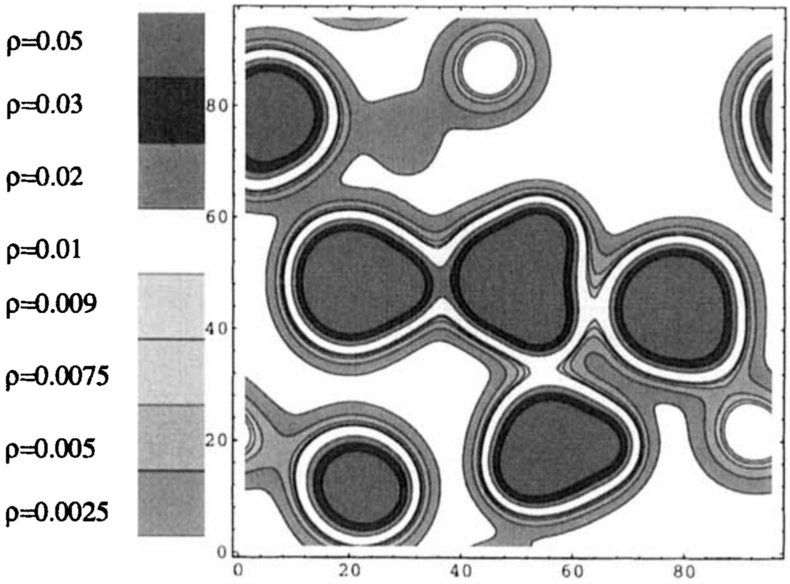

FIGURE 3 Planar section of the electron distribution on the grid, before the cutting process is applied. The plane of section corresponds to the plane in which the specified molecule lies (it contains the hydrogens and oxygen nuclei). The molecule analyzed is placed at the centre of the picture. The red colour indicates the core of the molecule and must be interpreted as the highest value of the electron density in that region. The other colours labeled indicate that the corresponding region is characterized by value of electron density lower than the one indicated by the previous colour and larger than the one indicated by the colour which labels that region. The oxygen nucleus is nearest at grid point (49, 49). (See Color Plate VI).

of the condensed phase. The important thing which emerges from this comparison is that the big differences are in regions of low density located far away from the core. The analysis of the figures also suggests that the triangulation method overestimates the region where the "true" zero flux surface is located particularly at low density. This is due to the fact that the approximation of the true lines of gradient with radial lines can generate a cross over of the zero flux surface. The reasons for this are in the discrete nature of the density and the flatness of the region of low density values. This problem is less important for the case of the downhill method, where cross-over is still possible but the probability that it occurs is limited by the added criterion of the second neighbouring point to estimate the gradient. However in the case of the "downhill" method although a better description of low density regions is furnished, some regions of higher density are not described properly. This is due to the fact that in spite of the added criterion of the second neighbouring point,

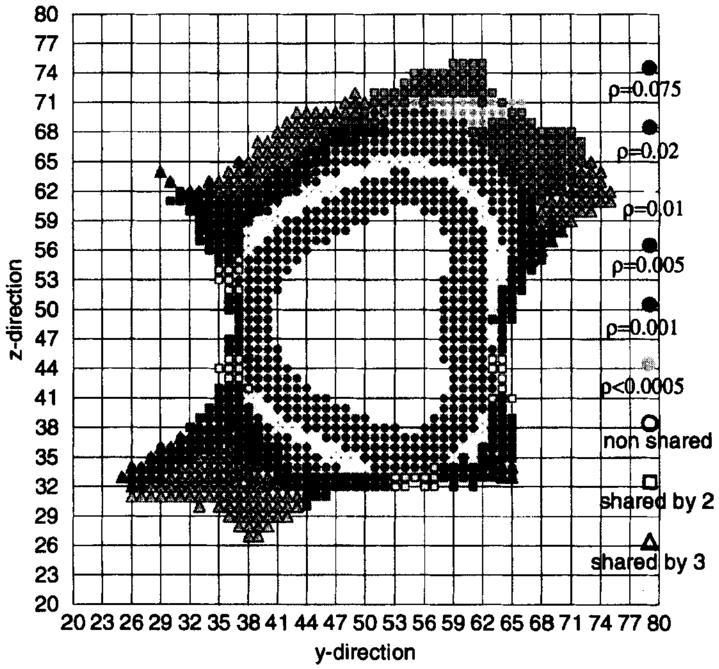

FIGURE 4 Planar section of the molecule on the grid when the triangulation method is applied. Here the grid points are explicitly considered. The colours labeled have the same meaning of the previous figures, the analysis in this case is performed outside the core region of the molecule. The circles indicate that the point is not shared with other molecules, the squares indicate that it is shared by two molecules, and the triangles indicate that it is shared by three molecules. Values of electron density less than 10^{-5}e/(a.u.)^3 are not considered. (See Color Plate VII).

the stepped paths can "walk" around a maximum; however these regions are small compared to the extension of the molecule.

4. COMPARISON WITH OTHER METHODS: THE CASE OF LIQUID WATER

In this section we report and discuss the results obtained for the average molecular dipole moment for liquid water when the two different algorithms are applied. The dipole moment was calculated averaging over three water liquid configurations. We performed *ab initio* molecular dynamics (the nuclei are allowed to move) and applied the two algorithms to the electron density calculated for three different instantaneous configurations. The other details of the run are the same as specified for the water liquid

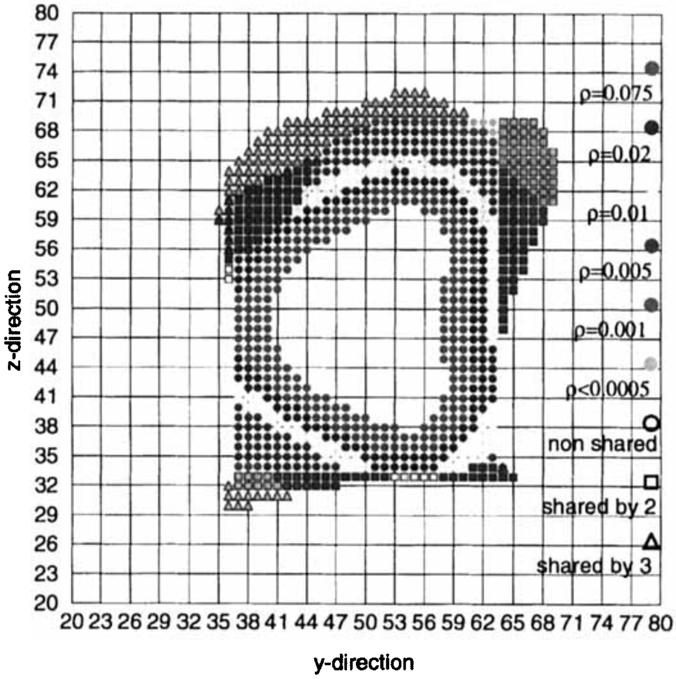

FIGURE 5 Same plot as in the previous figure, but in this case the downhill method is applied. (See Color Plate VIII).

system analyzed in the previous section. The values obtained for the average dipole moment using the two algorithms are respectively 2.55 Debye for the downhill and 2.95 for the triangulation method. This large difference between the two results is the key for understanding the meaning and the limitation of a spatial definition of the water molecule in a condensed phase. In the following part of the paper we focus the attention on this difference. In particular we analyze what this difference implies from the point of view of molecular modeling in the light of results obtained in previous *ab initio* work.

4.1. Anisometric Shape of Water Molecules in Liquid Phase

It is commonly accepted that a water molecule can be modeled as a sphere centred on the oxygen in which the protons and the electron density are contained. This approach is justified by considerations based on the spatial symmetry of the water molecule and implies that the distortion of the electronic charge distribution occurring in a condensed phase is unimportant for calculating most of the physical quantities. Here we show that applying the downhill or the triangulation method the surface which contains the volume of the molecule is not spherical and the external part of the molecule plays an important role which implies that the idea of a spherical shape is questionable. Figure 6 shows the average molecular dipole calculated as a function of a radial cut-off with origin on the oxygen of the molecule (in this way external regions of low density are removed); this means that electron density in spheres with larger radii than the radius of cut-off is not included in the calculation of the dipole of the cut molecule. It is important to notice that when the value of the radius cut-off is reduced the results for the two methods tend to converge. This means that the anisometric shape of the molecule, determined above all by the external part of the electron density plays an important role in determining the value of the dipole moment. There exists a previous *ab initio* work

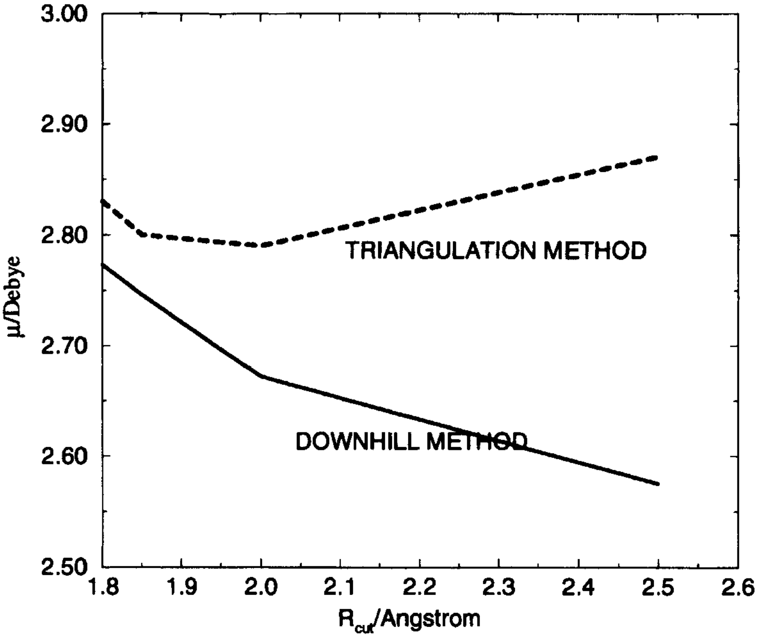

FIGURE 6 The diagram shows the average molecular dipole moment as a function of the cut-off applied to the extension of region in real space where the molecule is defined. When the radius cut-off is chosen, the region outside the sphere with origin the centre of the molecule and radius the radius of cut off, is not included in the calculation. The study is performed when the downhill (solid line) and the triangulation method (dashed line) are applied to define the molecule in condensed phase.

[13] in which the molecular dipole was calculated by including the total electron density in a sphere of a radius equal to half of the oxygen–oxygen distance, centred on the oxygen atom and the result obtained was 2.7 Debye; although in this case the value obtained is numerically reasonable, the possible effects on the dipole moment due to the external part of the electronic charge density are systematically neglected.

4.2. Triangulation Method and Wannier Function Approach

An interesting feature emerges when we analyze the results obtained using the triangulation method. What is interesting is that the triangulation method produces for the average molecular dipole a value of 2.9 Debye. Comparing this result with all the previous results, it seems to be rather high, although this is not true if we compare it with some recent *ab initio* calculations reported in a work of Silvestrelli *et al.* [14]. In that work the value found for the dipole was 2.95 Debye. The approach used is based on the "maximally localized Wannier functions" (see [3]); the wave-functions of the condensed system are projected along the set of Wannier functions, this makes it possible to localize the electron density for each single molecule and the neutrality emerges automatically. There is an interesting analogy with the triangulation method we developed. Figure 7 shows the value of the dipole moment as a function of a cutoff for the electron density, (in the sense that values of density lower that the specified cutoff are not included in the calculation) for the downhill and the triangulation method, the value of the dipole was averaged over three configuration (96 molecules). If we analyze Figure 7 it is possible to notice that while the downhill method does not produce a large variation at low value of the density cutoff, this is not true for the triangulation method. We have shown, consistently with Figure 7, that the triangulation method assigns to the specified molecule regions characterized by low value of electron density and located quite far from the core of the molecule; we also showed that these regions were reachable from more than one molecule. One could expect a similar behaviour when the Wannier functions approach is used. The Wannier functions are extended throughout the box of simulation, decay exponentially (see [3]) and in principle overlap. The large value for the dipole moment obtained in those cases is supported by a recent calculation based on experimental data (see [15]). This rises again the crucial question: how to define the single molecule in a condensed phase in the light of these results.

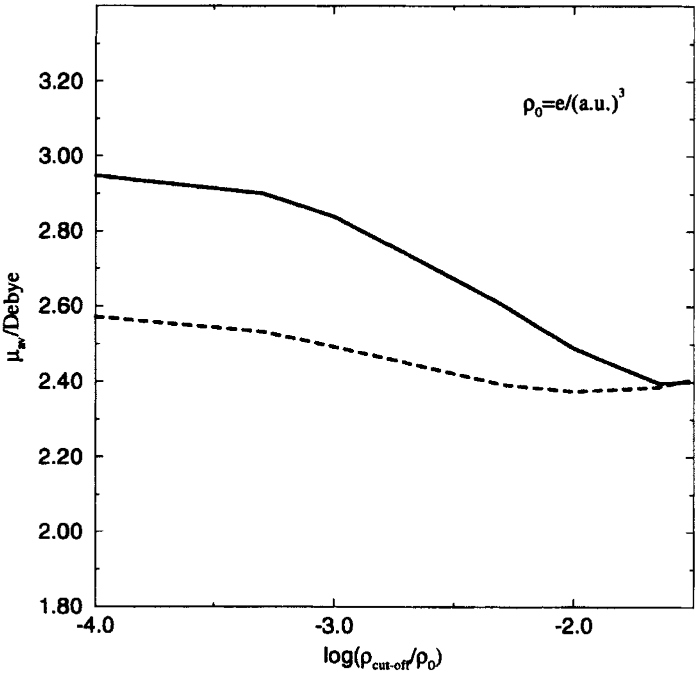

FIGURE 7 The diagram shows the average molecular dipole moment as a function of the electron density cut-off when the downhill (dashed line) and the triangulation (solid line) are used to define a molecule in condensed phase. Once the value of the cut off is fixed, the regions characterized by values of electron density lower than the value of the cut-off are not included in the calculation.

4.3. Implications for Molecular Modeling

As stated earlier a unique definition of molecule in a condensed phase does not exist, but in practical terms it is crucial to give such a definition. This is particularly important for building valid models to use for classical simulation. As a consequence, for example, the question whether the regions of low density located far from the centre of the molecule should be included or not in the calculation for the dipole moment is strictly related to the question of what we mean to represent with the classical model and what we want to investigate using it. Commonly the idea of classical molecule consists in imagining the relevant part of the charge distribution located approximately around the nuclei positions; according to this idea the downhill method is the approach which defines molecules in a more classical sense than the other approaches (Wannier function and triangulation method). On one hand the direct spatial definition of the molecule given by the downhill is, as stated earlier, closer to the classical idea of molecule, on the other hand the description provided by the Wannier function approach furnishes a definition of molecule in Hilbert space which implies that the properties of each single electron are directly taken into account. This means that a molecule can be defined in terms of electron–electron interaction and such a property is not directly captured by the spatial definition of the molecule as provided by the downhill method. The triangulation method if on one hand can be seen as the analogue of the Wannier function approach in real space (in the sense that they show similar behaviour in assigning the electron density to each molecule), on the other hand underlines a general problem of plane waves based *ab initio* calculations: the electron density in real space is not properly described in regions of low density. This is expressed by the fact that fluctuations of $\rho(\mathbf{r})$ occurring in regions of low density are averaged by the linear interpolation; as a consequence longer paths for the gradient trajectories, with respect to the ones determined by the downhill method, are produced; in fact in the case of the downhill method each point is directly tested without performing any manipulation. However both procedures are theoretically well founded and it is difficult to say which one is more realistic. If it is not possible to answer the question of which approach to use for modeling molecules it is at least possible to establish the physical properties which can be captured by using one of the methods. This allows one to choose a method in such a way that the phenomenon analyzed is well described by the physical properties of the resulting model of molecule.

5. CONCLUSION

Two different algorithms for calculating the zero flux surface have been proposed; they can be applied to electron density distributions calculated on a grid. The problem underlined by Gatti *et al.* [16] for density calculated with a pseudopotential based approach has been overcame by defining an atomic or molecular core. As stated earlier evidence for the reliability of the two methods is the conservation of the atomic or molecular charge as both methods have no intrinsic requirement for doing so. The advantage in using the two algorithms is that both approaches directly act on the electron density. This allows one to understand in a simple way how the molecule is numerically defined so that any particular anomaly in the cutting process can be easily understood; in addition the computational time required for the calculation is relatively short. For our standard system of 32 molecules and electron density calculated with an energy cut-off of 90Ry, the computational time required on a common Silicon Graphics is about 30 minutes for the downhill method and about 50 minutes for the triangulation method. Apart from the technical aspects, this work represents also an attempt to throw light on the controversial question of how to define a molecule in a condensed phase and how this question is related to the current research in the field of condensed matter. I did not try to give a definitive answer, but I rather tried to underline the emergent properties, limitations and differences of equally well founded approaches, in particular, in spite of the limitations due to the discrete nature of the electron density distribution, I believe that both the algorithms proposed in this work capture the essential physics of atoms and molecules in a simple and direct way.

Acknowledgements

I would like to thank Professor R. Lynden-Bell for a critical reading of the manuscript, Dr. A. Alavi and Dr. M. Sprik for helpful suggestions and discussions. I also thank EPSRC for financial support (grants GR/K20651 and GR/L08427) and the IFI for financial support for the Irish Centre of Colloids and Biomaterials.

References

[1] Hirshfeld, F. L. (1977). "Spatial partitioning of charge-density", *Israel Journal of Chemistry*, **16**(2–3), 198.

[2] Bader, R. F. W., Atoms in Molecules, Oxford University Press, Oxford, 1994.

[3] Marzari, N. and Vanderbilt, D. (1997). "Maximally localized generalized Wannier functions for composite energy bands", *Phys. Rev. B*, **56**, 12847.

[4] Delle Site, L. (1999). "Partitioning the electronic charge distribution in molecules and in condensed phase: A connection between quantum and classical mechanics", unpublished.

[5] Delle Site, L., Alavi, A. and Lynden-Bell, R. M. (1999). "The electrostatic properties of water molecules in condensed phases: an *ab initio* study", *Mol. Phys.*, **96**, 1683.

[6] Delle Site, L., "*Ab initio* study of Water", *Ph.D. Thesis*, The Queen's University of Belfast, 1999.

[7] Barnette, D. W. (1982). "Generating the triangulations of the projective plane", *J. Comb. Th. B*, **33**, 222.

[8] Barnette, D. W. and Edelson, A. L. (1988). "All orientable 2-manifolds have finitely many minimal triangulations", *Israel Journal of Mathematics*, **62**, 90.

[9] Barnette, D. W. and Edelson, A. L. (1989). "All 2-manifolds have finitely many minimal triangulations", *Israel Journal of Mathematics*, **67**, 123.

[10] Jorgensen, W., Chandrasekhar, J., Madura, J. D., Impey, R. W. and Klein, M. L. (1983). "Comparison of simple potential functions for simulating liquid water", *Chem. Phys.*, **79**, 926.

[11] Alavi, A., Kohanoff, J., Parrinello, M. and Frenkel, D. (1994). "*Ab-initio* molecular dynamics with excited electrons", *Phys. Rev. Lett.*, **73**, 2599.

[12] Alavi, A. (1996). "Path integrals and *ab initio* molecular dynamics", In: *Monte Carlo and Molecular Dynamics of Condensed Matter System*, Binder, K. and Ciccotti, G. Eds., **49**, Bologna, SIF.

[13] Laasonen, K., Sprik, M. and Parrinello, M. (1993). "*Ab-initio* liquid water", *J. Chem. Phys.*, **99**, 9080.

[14] Silvestrelli, P. L. and Parrinello, M. (1999). "Water molecule dipole in the gas and in the liquid phase", *Phys. Rev. Lett.*, **82**, 3308.

[15] Batista, E. R., Xantheas, S. S. and Jonsson, H. (1998). "Molecular multipole moments of water molecules in ice lh", *J. Chem. Phys.*, **109**, 4546.

[16] Gatti, C., Silvi, B. and Colonna, F. (1995). "Dipole moment of the water molecule in the condensed phase-a periodic Hartree-Fock estimate", *Chem. Phys. Lett.*, **247**, 135.