Journal of Mathematical Chemistry Vol. 32, No. 4, November 2002 (© 2002)

Computation of the effective volumes of covalently bonded molecules

P.G.L. Leach [1, 2]

GEODYSYC, Department of Mathematics, University of the Aegean, Karlovassi 83 200, Greece

T.A. Ford [2] and P.H. Matabal

School of Pure and Applied Chemistry, University of Natal, Durban 4041, South Africa

David P. White [3]

Beckman Institute for Advanced Science and Technology, University of Illinois, 405 North Mathews Avenue, Urbana, IL 61801, USA

Received 23 July 2002

An analytical algorithm for the calculation of molecular volume based on a model of partially overlapping spheres is presented. The algorithm takes into account the volume of overlapping caps of spheres and includes the contribution to the molecular volume by interstitial voids. The algorithm is constructed in a manner to ensure generality. The working of the algorithm is illustrated by examples of some different types of basic molecular structures. The Mathematica program is available.

1. Introduction

Molecular volumes have been applied to many different disciplines within the chemical sciences [1]. For example, molecular volumes have been used for packing predictions in crystallography [2–5], protein cavity evaluation [6], charge density estimates [7], topological analysis of biologically important molecules [8], materials chemistry [9] and quantification of chirality [10]. More recently molecular volumes have found extensive use in quantitative structure–activity relationships (QSAR). A recent review summarizes the use of molecular volumes in organometallic chemistry [11]. However, most, if not all, of the above studies have used a numerical approach to calculate molecular volumes.

[1] Permanent address: School of Mathematical and Statistical Sciences, University of Natal, Durban 4041, South Africa.

[2] Member of the Centre for Theoretical and Computational Chemistry, University of Natal, South Africa.

[3] Present address: Department of Chemistry, University of North Carolina, Chapel Hill, NC 27515, USA.

0259-9791/02/1100-0361/0 © 2002 Plenum Publishing Corporation

The history of the calculation of molecular volumes is interesting [12]. Two approaches were favoured: calculation of minimum or maximum volume. In the latter interstitial voids are assumed to contribute to the molecular volume whereas in the former they do not. In 1975 an attempt was made to calculate analytically the molecular volume for spherical and pseudospherical molecules [13]. Various terms were added to the equation to take care of the volume of the solvent cage. However, the method remained limited to spherically shaped molecules. An interesting numerical approach was proposed in 1979: overlapping volume analysis (OVA) [14–17]. The method involved circumscribing the molecule by a parallelepiped and then examining all the points within the parallelepiped to determine whether they were within or without the molecule. The numerical algorithm for the calculation of molecular volumes of choice was proposed by Gavezzotti in 1983 [18]. This paper addressed, in a numerical fashion, the shortcomings of the prior analytical method of Bondi [19]. In this paper we present the analytical approach to these shortcomings.

The model we use is that of an assemblage of spheres which may have overlap with neighbouring spheres. We assume that the radius of each sphere is known and that the locations of the centres of all spheres are known with respect to some laboratory frame. There are N spheres labelled $S_i, i = 1, \ldots, N$ with centres $P_i, i = 1, \ldots, N$. Cartesian coordinates in the laboratory frame are denoted by capital letters so that the centre of S_i, P_i, is the point (X_i, Y_i, Z_i). We also use local coordinate systems for which lower case letters are used to denote the Cartesian coordinates of a point as (x, y, z).

The first part of the algorithm separates the constituent atoms into classes, viz. those which have no overlap with any other atom and no voids (conceivably the null set), those which overlap pairwise, those which have triple overlap and so forth, those which are arranged in such a way as to give rise to pseudotetrahedral voids and those arranged in such a way as to give rise to pseudooctahedral voids. Were one dealing with geometrical spheres, one would have to allow for overlap of anything up to N spheres or even inclusion. However, we are using the spheres to aid in the modeling of molecules and so we ignore higher overlap (more than eight overlapping spheres) and inclusion. In the pairwise overlap we can include chains of atoms with nearest neighbour overlap.

We then consider the contributions to the total volume from each of the classes enumerated above and develop formulae to account for the regions of overlap. Equipped with these formulae we present the second part of the algorithm which computes the volume of the molecule.

2. Enumeration of the overlap classes

The spheres are labelled $S_i, i = 1, \ldots, N$. Commencing with S_1 we determine the existence of overlap with any of the other spheres $S_j, j = 2, \ldots, N$. We then pass on to S_2 and so on. The general procedure is as follows. Consider a sphere S_i with centre P_i (X_i, Y_i, Z_i) and radius R_i. Then each sphere $S_j, j = i + 1, \ldots, N$ with cen-

tre $P_j\ (X_j, Y_j, Z_j)$ and radius R_j is tested for overlap. The test consists of the calculation of the separation of P_i and P_j, viz.

$$d_{ij} = \sqrt{\left[(X_i - X_j)^2 + (Y_i - Y_j)^2 + (Z_i - Z_j)^2\right]}, \tag{2.1}$$

and then

$$s_{ij} = d_{ij} - R_i - R_j. \tag{2.2}$$

If $s_{ij} \geqslant 0$, there is no overlap and the next sphere can be considered for comparison.

If $s_{ij} < 0$, there is overlap. All those pairs of atoms which register overlap are recorded. If spheres S_i and S_j show overlap through $s_{ij} < 0$, triple overlap with sphere S_k with attributes $P_k\ (X_k, Y_k, Z_k)$ occurs if $s_{ik} < 0$ and $s_{jk} < 0$. This information is already known and it is merely a matter of identifying those spheres which do have triple overlap. The set of spheres with triple overlap is tested for quadruple overlap. In turn this set is tested for quintuple overlap. The process continues until the null set of n-tuple overlap is reached. The maximum value of n which we consider is eight.

On chemical/physical grounds complete overlap or inclusion cannot occur. This pathological case is identified by

$$s_{ij} < -\max(R_i, R_j). \tag{2.3}$$

Since a result (2.3) would not be expected from physically realistic data, this test would be useful to include if there were any doubts about the validity of the input data.

3. The volume elements

3.1. Single spheres

For the sake of completeness we state the obvious and elementary result that an isolated sphere, S_i at $P_i\ (X_i, Y_i, Z_i)$, i.e. one with no overlap, contributes

$$V_i = \frac{4}{3}\pi R_i^3. \tag{3.1}$$

3.2. Pairwise overlap

Two spheres S_i and S_j with centres at $P_i\ (X_i, Y_i, Z_i)$ and $P_j\ (X_j, Y_j, Z_j)$ and radii R_i and R_j have $s_{ij} < 0$. To make the calculations easier we work in local coordinates. Take P_i as the origin and P_iP_j as the direction of the y axis. The x and z axes are in the plane through P_i normal to P_iP_j (see figure 1). For double overlap no further specification is required. The equation of the surface of the sphere S_i is

$$x^2 + y^2 + z^2 = R_i^2 \tag{3.2}$$

and that of S_j

$$x^2 + (y - d_{ij})^2 + z^2 = R_j^2. \tag{3.3}$$

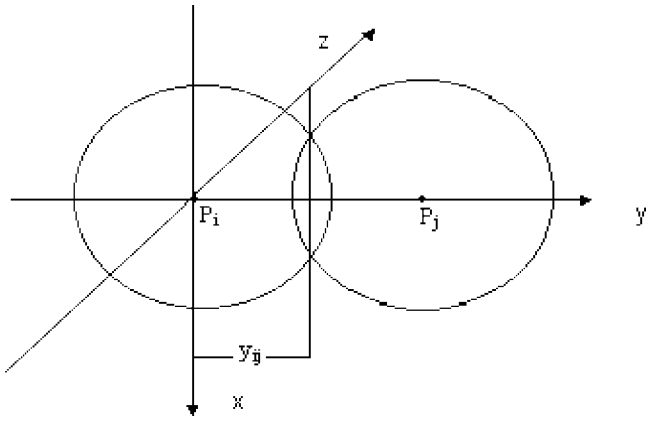

Figure 1. Local coordinate system to calculate the volume of overlap for two intersecting spheres.

Subtraction of (3.3) from (3.2) gives

$$2yd_{ij} - d_{ij}^2 = R_i^2 - R_j^2 \tag{3.4}$$

so that the plane of intersection of the two spheres is given by

$$y = \frac{1}{2}\left\{d_{ij} + \frac{R_i^2 - R_j^2}{d_{ij}}\right\}. \tag{3.5}$$

The volume occupied by the two partially overlapping spheres is expressed as the formula

$$V_{\mathrm{O}ij} = V_i + V_j - V_{\mathrm{C}ij}, \tag{3.6}$$

where, in an obvious notation, $V_{\mathrm{O}ij}$ is the volume occupied by the spheres S_i and S_j, V_i and V_j are the volumes of the spheres S_i and S_j, respectively, and $V_{\mathrm{C}ij}$ is the common volume. V_i and V_j are given by (3.1). For the moment our concern is the calculation of $V_{\mathrm{C}ij}$.

To determine the volume of overlap, $V_{\mathrm{C}ij}$, it is necessary to calculate the volumes of the two caps of spheres into which $V_{\mathrm{C}ij}$ is divided by the plane defined by (3.5).

A thin disc of radius r and thickness $\mathrm{d}z$ (see figure 2) has the volume

$$\mathrm{d}V = \pi r^2\,\mathrm{d}z$$

and the volume of the cap of the sphere is

$$\begin{aligned} V &= \int_b^a \pi r^2\,\mathrm{d}z \\ &= \pi \int_b^a \left(a^2 - z^2\right)\mathrm{d}z \\ &= \frac{1}{3}\pi\,(2a+b)\,(a-b)^2\,. \end{aligned} \tag{3.7}$$

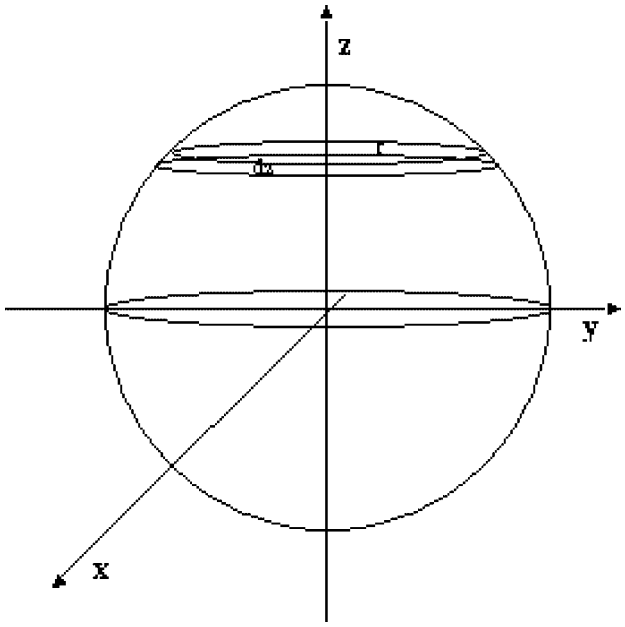

Figure 2. A sphere of radius a and centred at the origin has a cap with base parallel to the xy-plane and height b above it (b can be a negative number).

We use the measurements a and b as they are the ones which occur naturally in the physical calculation, the radius of the sphere and the distance from the centre of the sphere to the plane of intersection of the two spheres.

For the two spheres we have

$$\begin{aligned} a_i &= R_i, \quad b_i = \frac{1}{2}\left\{d_{ij} + \frac{R_i^2 - R_j^2}{d_{ij}}\right\}, \\ a_j &= R_j, \quad b_j = \frac{1}{2}\left\{d_{ij} - \frac{R_i^2 - R_j^2}{d_{ij}}\right\}, \end{aligned} \tag{3.8}$$

and with (3.7) and (3.8) we obtain

$$\begin{aligned} V_{\mathrm{C}ij} = \frac{1}{3}\pi\left\{\left[2R_i + \frac{1}{2}\left(d_{ij} + \frac{R_i^2 - R_j^2}{d_{ij}}\right)\right]\left[R_i - \frac{1}{2}\left(d_{ij} + \frac{R_i^2 - R_j^2}{d_{ij}}\right)\right]^2\right. \\ \left. + \left[2R_j + \frac{1}{2}\left(d_{ij} - \frac{R_i^2 - R_j^2}{d_{ij}}\right)\right]\left[R_j - \frac{1}{2}\left(d_{ij} - \frac{R_i^2 - R_j^2}{d_{ij}}\right)\right]^2\right\}. \end{aligned} \tag{3.9}$$

However, we do not proceed now to calculate $V_{\mathrm{O}ij}$ from (3.6) as it is not necessary to consider the contribution from pairs of overlapping spheres individually.

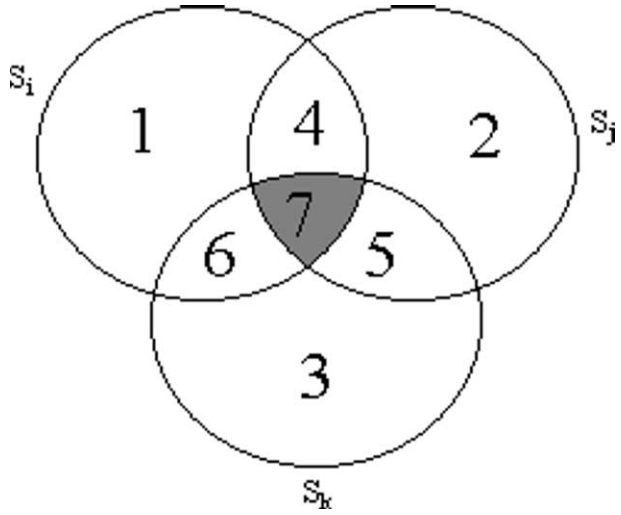

Figure 3. The projection onto a plane of three overlapping spheres in the case that there is a region common to all three spheres. The spheres are S_i, S_j and S_k with attributes labelled accordingly.

3.3. Three overlapping spheres

For those triplets of spheres with a common region of overlap we have to be careful in the construction of the algorithm to avoid the chance of multiple addition or subtraction of the volume of a particular region.

The volume occupied by the three spheres represented in figure 3 is

$$V_{\mathrm{O}ijk} = V_1 + V_2 + V_3 + V_4 + V_5 + V_6 + V_7, \tag{3.10}$$

where V_m refers to the volume of the region marked m ($m = 1, \ldots, 7$) in figure 3. Making a rearrangement of (3.10) we have

$$\begin{aligned} V_{\mathrm{O}ijk} &= (V_1 + V_4 + V_7 + V_6) + (V_2 + V_5 + V_7 + V_4) + (V_3 + V_6 + V_7 + V_5) \\ &\quad - (V_4 + V_7) - (V_5 + V_7) - (V_6 + V_7) + V_7 \\ &= V_i + V_j + V_k - (V_{\mathrm{C}ij} + V_{\mathrm{C}jk} + V_{\mathrm{C}ki}) + V_{\mathrm{C}ijk}. \end{aligned} \tag{3.11}$$

The important term in (3.11) is $V_{\mathrm{C}ijk}$. This is the correction term needed to the idea of subtraction of V_{C} from V when there is a triple overlap.

In the case of the overlap of two spheres only the line P_iP_j was of importance. When there is the overlap of three spheres, the use of the plane $P_iP_jP_k$ simplifies the calculations. We take the origin to be at P_i (the atom we started with), the y-axis to be along P_iP_j (P_j being the next atom treated) and the z-axis to be in the plane of $P_iP_jP_k$ and normal to P_iP_j (see figure 4). The x-axis completes the right orthogonal triad. In this coordinate system the centres of the spheres are located at $(0, 0, 0)$ (P_i), $(0, b_j, 0)$ (P_j) and $(0, b_k, c_k)$ (P_k), where

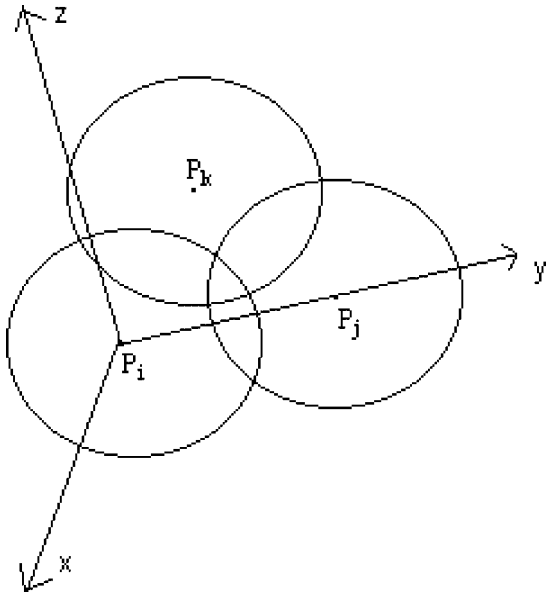

Figure 4. Coordinate axes for three overlapping spheres. The origin is taken at P_i, the y-axis in the direction of P_iP_j and the z-axis normal to P_iP_j in the plane $P_iP_jP_k$. The x-axis completes the right orthogonal triad.

$$\begin{aligned} b_j &= d_{ij}, \\ b_k &= \frac{\mathbf{P}_i\mathbf{P}_k \cdot \mathbf{P}_i\mathbf{P}_j}{d_{ij}}, \\ c_k &= \frac{|\mathbf{P}_i\mathbf{P}_k \times \mathbf{P}_i\mathbf{P}_j|}{d_{ij}}. \end{aligned} \tag{3.12}$$

The vectors $\mathbf{P}_i\mathbf{P}_k$ and $\mathbf{P}_i\mathbf{P}_j$ and the distance d_{ij} are all calculated from global coordinates and we adopt the convention that the centre of S_k is always in the upper half plane in the local coordinate system.

In local coordinates the equations of the surfaces of the spheres are

$$\begin{aligned} S_i&: \quad x^2 + y^2 + z^2 = R_i^2, \\ S_j&: \quad x^2 + (y - b_j)^2 + z^2 = R_j^2, \\ S_k&: \quad x^2 + (y - b_k)^2 + (z - c_k)^2 = R_k^2. \end{aligned} \tag{3.13}$$

Because of the way we have chosen the axes there is symmetry about the yz-plane. The reason for this is that S_i has symmetry about the origin which contains symmetry about the y-axis which is the symmetry which S_j possesses and both S_i and S_j have symmetry about the xy-plane which is what S_k possesses. Hence, common to S_i, S_j and S_k is symmetry about the yz-plane. Although we do not know of an explicit formula to calculate the volume of this region, we can make use of the symmetry about the yz-plane to compute it. The region, the volume of which is to be computed, can be considered as

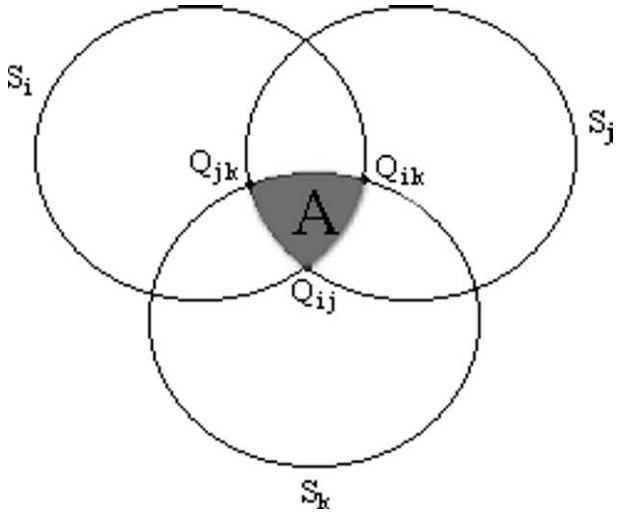

Figure 5. The region for the evaluation of $A_{ijk}(x)$. The points Q_{ij}, Q_{jk} and Q_{ki} are identified as those of the points of intersection of any pair of (small) circles which lies within the third.

a series of thin trianglelike regions. If the area of one of these is given by the function $A_{ijk}(x)$, the common volume is

$$V_{\mathrm{C}ijk} = \int_{-x_0}^{x_0} A_{ijk}(x)\,\mathrm{d}x. \tag{3.14}$$

To find $A_{ijk}(x)$ we fix x which means that we take a plane parallel to the yz-plane. We calculate the points Q_{ij}, Q_{jk} and Q_{ki} which are three of the points of intersection of the three (small when $x \neq 0$) circles (see figure 5). As the circles intersect pairwise in two points, it is necessary to determine which of the two points is to be used. The acceptable point must lie within the third (small) circle. Thus, if $x = x_{\mathrm{c}}$ and $Q_{ij1,2}$ has the coordinates $(x_{\mathrm{c}}, y_{\mathrm{c}ij1,2}, z_{\mathrm{c}ij1,2})$, we select that one of Q_{ij1} and Q_{ij2} which satisfies

$$x_{\mathrm{c}}^2 + (y_{\mathrm{c}ij} - b_k)^2 + (z_{\mathrm{c}ij} - c_k)^2 \leqslant R_k^2. \tag{3.15}$$

Once the three acceptable points are established, the area of the enclosed region is calculated by means of the formula

$$A_{ijk} = \int_{y_{jk}}^{y_{ij}} (z_j - z_k)\,\mathrm{d}y + \int_{y_{ij}}^{y_{ki}} (z_i - z_k)\,\mathrm{d}y, \tag{3.16}$$

where the limits of integration are given by the y-ordinates of Q_{ij}, Q_{jk} and Q_{ki} and

$$\begin{aligned} z_i &= \left[R_i^2 - x_{\mathrm{c}}^2 - y^2\right]^{1/2}, \\ z_j &= \left[R_j^2 - x_{\mathrm{c}}^2 - (y - b_j)^2\right]^{1/2}, \\ z_k &= c_k + \left[R_k^2 - x_{\mathrm{c}}^2 - (y - b_k)^2\right]^{1/2}. \end{aligned} \tag{3.17}$$

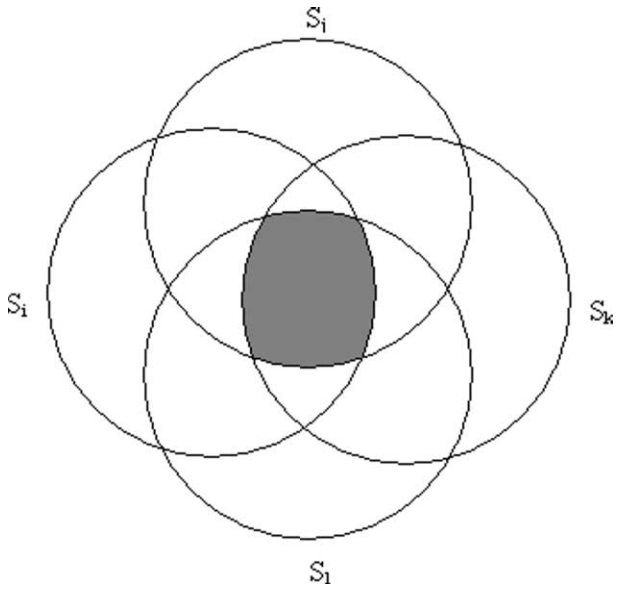

Figure 6. Quadruple overlap. For the purposes of the general algorithm no assumptions are made about any coincidences of the coordinates and radii of the spheres.

Although the integrals in (3.16) and (3.14) can be evaluated analytically by standard methods, the resultant formulae are so complicated as to be meaningless in a general setting due to the necessity to label each representation of coordinates and radii by identifying symbols and we do not write them down.

3.4. Multiple overlap

In the instances of rings such as cyclobutane, cyclopentadiene, benzene and cyclooctadiene the distance between the centroid of the ring and the centre of any of the carbon atoms is less than the van der Waals radius of carbon. Hence, there can be overlap of four, five, six or even eight carbon atoms. Although this presents a complication, the general algorithm is not more complicated than it is in the case of the triple overlap of section 3.3. In figure 6 we depict the situation for the overlap of four spheres. The situation for the other cases mentioned above is similar. A simple counting exercise of the same type as in (3.11) shows that

$$V_{\mathrm{O}ijkl} = V_i + V_j + V_k + V_l - (V_{\mathrm{C}ij} + V_{\mathrm{C}jk} + V_{\mathrm{C}kl} + V_{\mathrm{C}li}) + V_{\mathrm{C}ijkl}. \tag{3.18}$$

This notation is too cumbersome for larger rings. In general, we have

$$V_{\mathrm{O}} = \sum_{\alpha} V_\alpha - \sum_{\alpha\beta} V_{\mathrm{C}\alpha\beta} + V_{\mathrm{C}}, \tag{3.19}$$

where V_{O} is the total space occupied, V_α is the volume of S_α, $V_{\mathrm{C}\alpha\beta}$ is the volume common to S_α and S_β given by (3.9) and V_{C} is the volume common to all spheres.

As before, we use a local coordinate system based on the first three spheres identified as overlapping. We assume that the labelling of the spheres has been done in an anticlockwise fashion. We take a section parallel to and distant x_c from the yz-plane. In the case of identical spheres arranged in a plane the common area will be symmetric with N vertices. However, this is not the case in general and we cannot make use of the partial symmetry that we had in the triple overlap case. If the equations of the (small) circles of the spheres in the plane distant x_c from the yz-plane are

$$(y - b_\alpha)^2 + (z - c_\alpha)^2 = r_\alpha^2, \quad \alpha = 1, \ldots, N, \tag{3.20}$$

where

$$r_\alpha^2 = R_\alpha^2 - (x_c - a_\alpha)^2, \tag{3.21}$$

the coordinates of the vertices of the common region are the real roots of the pairwise solutions of (3.20) (β) and (3.20) ($\beta + 1$) which satisfy

$$(y - b_\alpha)^2 + (z - c_\alpha)^2 \leqslant r_\alpha^2, \quad \alpha \neq \beta, \beta + 1. \tag{3.22}$$

Although in principle it is possible to write an analytic expression for the area of the section, it is not a feasible proposition to do so in a general algorithm. The vertices of the region $Q_{\alpha\beta}$ have been identified. Let them be ordered from the minimum value of y to the maximum in two sequences, one giving the upper vertices of the common section and the other giving the lower vertices. The area of the common section is given by

$$A_C(x_c) = \sum (z^+ - z^-)\delta y, \tag{3.23}$$

where z^+ is on the arc of the small circle which marks the current upper boundary and z^- is on the lower one. The circles contributing the arcs change as y passes through each vertex of the section. Upper and lower arcs change independently. The step length δy can be chosen to give the accuracy desired. After $A_C(x_c)$ is evaluated for a given x_c, the value of x_c is increased and the process repeated.

The volume is given by

$$V_C = \sum A_C(x)\delta x, \tag{3.24}$$

where the summation is best started at $x = 0$ and moved outwards in both positive and negative directions to those values of x for which the testing of the existence of overlap shows that it is no longer multiple or triple.

The procedure described above for the overlap of N dissimilar spheres is heavy on computation because it is necessary to solve the set of equations (3.20) pairwise for each value of x_c and then verify the inequalities (3.22). In the case of like spheres symmetrically arranged we may make use of this symmetry to obtain the volume less expensively. The procedure is common to all cases of multiple overlap which exhibit C_N symmetry, where N is the number of spheres. In figure 7 we illustrate it for cyclobutane.

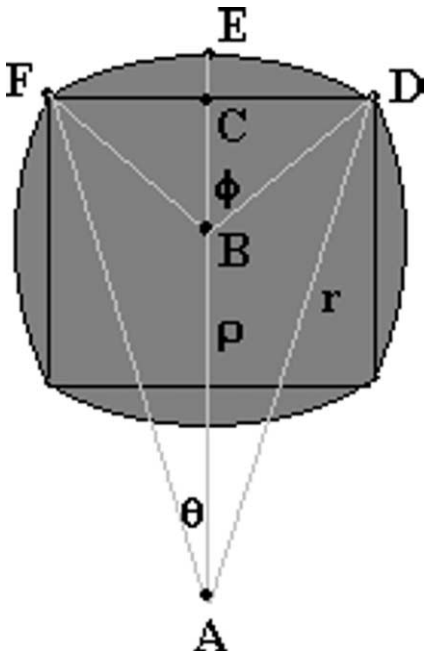

Figure 7. Common section of a four sphere overlap representing the four carbon atoms of cyclobutane. In general the common region is an Ngon plus N equal segments of circles.

Let the radius of each sphere be R and the radius of each small circle distant x_c from the yz-plane be r. In figure 7 $AD = r$ and $AB = \rho$. The area of the segment $CDEF$ is

$$A_S = \frac{1}{2}r^2(2\theta - \sin 2\theta). \tag{3.25}$$

The area of the triangle BCD is

$$A_T = \frac{1}{2}\mathbf{BC}\cdot\mathbf{CD} = \frac{1}{2}(r\cos\theta - \rho)^2\tan\phi. \tag{3.26}$$

From the application of the cosine rule to the triangle ABD

$$BD = \rho\cos\phi + \sqrt{r^2 - \rho^2\sin^2\phi}. \tag{3.27}$$

Application of the sine rule to the triangle ABD gives

$$\sin\theta = \frac{BD}{r}\sin\phi. \tag{3.28}$$

For an Ngon

$$\phi = \frac{\pi}{N}, \tag{3.29}$$

and the total area of the section is

$$A = N(A_S + 2NA_T). \tag{3.30}$$

The common volume is

$$V_C = 2\int_0^{x_1} A\,\mathrm{d}x, \tag{3.31}$$

where

$$x_1 = \sqrt{R^2 - \rho^2}. \tag{3.32}$$

The integral can be evaluated when the substitution $r = \sqrt{R^2 - x^2}$ is made. The distance from ring centroid to sphere centre is ρ.

4. The interstitial void

When groups of atoms clump together, there may be enclosed space amongst them, the interstitial void, which, although it is not material volume, contributes to the total volume occupied by the molecule or portion thereof.

4.1. The tetrahedral interstitial void

The tetrahedral interstitial void occurs when three atoms lie in a plane with another sitting on top and a fifth placed underneath. We assume that the enumeration of overlap classes of section 2 has been performed. We identify the existence of a tetrahedral interstitial void as follows. If S_i has overlap or contact with S_j, S_k, S_l and S_m, likewise for S_j and S_k, and S_l and S_m do not have contact or overlap, there is a tetrahedral interstitial void contained within the five atoms. The coordinate system is as defined by figure 8 with the origin being at P_i, the y-axis along P_iP_j and the z-axis normal to P_iP_j in the plane $P_iP_jP_k$. The x-axis completes the orthogonal triad.

S_i, S_j and S_k are treated for triple overlap (if this is likely to happen) as in section 3.3. This determines whether the integration starts from $x = 0$ if there is no triple overlap or from $x = x_0$ if there is. A plane section of the tetrahedral interstitial void parallel to the yz-plane can have one of the possible shapes depicted in figure 9.

In figure 9(a) the spheres closing off the tetrahedral interstitial void do not intrude into the region and the section is roughly triangular. In figure 9(b) a section of one of them (a small circle) is contained within the triangular region. In figure 9(c) the same section now overlaps one or more of S_i, S_j and S_k. The area of the triangular region $B_{ijk}(x)$ is calculated by a formula very similar to (3.16), viz.

$$B_{ijk}(x_c) = \int_{y_{ki}}^{y_{jk}} (z_k - z_i)\,\mathrm{d}y + \int_{y_{jk}}^{y_{ij}} (z_k - z_j)\,\mathrm{d}y, \tag{4.1}$$

where the limits of integration are given by the y-ordinates of Q_{ij}, Q_{jk} and Q_{ki}, z_i, z_j and z_k are given by (3.17) and x_c is the distance of the plane from the yz-plane. Suppose

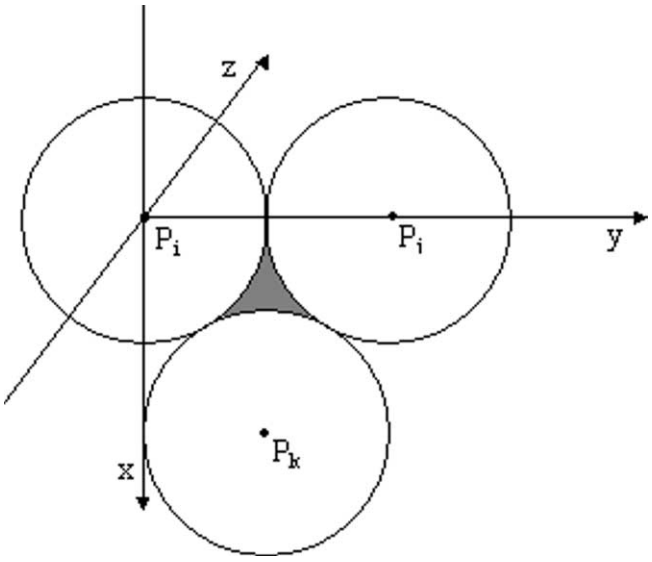

Figure 8. Tetrahedral interstitial void: a cross-section parallel to the yz-plane.

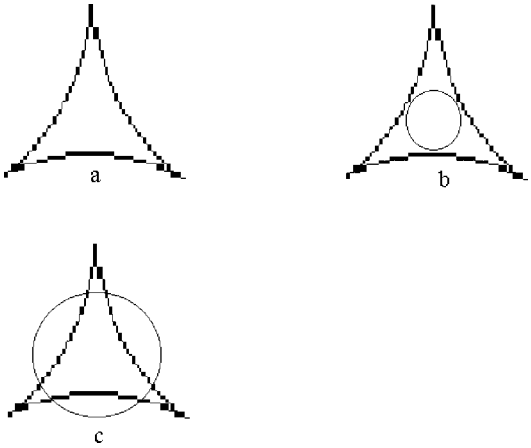

Figure 9. Tetrahedral interstitial void: (a) the closing off sphere does not intrude; (b) it intrudes, but is contained within the region; (c) it overlaps part of the outside region.

that the x-ordinate of P_l is positive and that of P_m is negative. For $x_c > 0$ we test the equation for S_l, viz.

$$(x - a_l)^2 + (y - b_l)^2 + (z - c_l)^2 = R_l^2, \tag{4.2}$$

to determine if S_l intrudes into the triangular region. This occurs when

$$(y - b_l)^2 + (z - c_l)^2 = R_l^2 - (x - a_l)^2 > 0. \tag{4.3}$$

The area of this circle is

$$A_l(x_c) = \pi r_l^2, \tag{4.4}$$

where

$$r_l^2 = R_l^2 - (x - a_l)^2. \tag{4.5}$$

It is necessary to test for possible overlap as depicted in figure 9(c). This is done by finding the points of intersection of

$$(y - b_l)^2 + (z - c_l)^2 = r_l^2 \tag{4.6}$$

with

$$y^2 + z^2 = r_i^2, \tag{4.7}$$
$$(y - b_j)^2 + z^2 = r_j^2, \tag{4.8}$$
$$(y - b_k)^2 + (z - c_k)^2 = r_k^2, \tag{4.9}$$

where r_l, r_j and r_k are defined similarly to r_l in (4.5), in turn. Suppose that these occur at points T_{i1} and T_{i2} on S_i, T_{j1} and T_{j2} on S_j and T_{k1} and T_{k2} on S_k. The portion of the area of the circle over-counted with each sphere is given by

$$C_{l\alpha}(x_c) = \frac{1}{2} r_l^2, \tag{4.10}$$

where

$$\theta_l = \frac{l_{\alpha 12}}{r_l}, \tag{4.11}$$
$$\theta_\alpha = \frac{l_{\alpha 12}}{r_\alpha}, \tag{4.12}$$
$$l_{\alpha 12} = (y_{\alpha 2} - y_{\alpha 1})^2 + (z_{\alpha 2} - z_{\alpha 1})^2, \tag{4.13}$$

and α ranges over the indices i, j and k. The actual area at each value of x_c is given by

$$D_{ijkl}(x_c) = B_{ijk}(x_c) - A_C(x_c) + \sum_{\alpha = i,j,k} C_{l\alpha}(x_c). \tag{4.14}$$

The volume of the tetrahedral interstitial void contributed by the presence of S_l is given by

$$V_{ijkl} = \int_{x_0}^{x_l} D_{ijkl}(x)\,\mathrm{d}x, \tag{4.15}$$

where x_l is the value of x_c for which the points Q_{ij}, Q_{jk} and Q_{ki} are contained within the circle which is tested by the inequality

$$(y_\alpha - b_l)^2 + (z_\alpha - c_l)^2 < r_l^2 \tag{4.16}$$

for $\alpha = i, j, k$ and x_0 is either zero (no triple overlap of S_i, S_j and S_k) or the upper limit of the integral in (3.14). Likewise the volume of the tetrahedral interstitial void contributed by S_m is given by

$$V_{ijkm} = \int_{x_m}^{-x_0} D_{ijkm}(x)\,\mathrm{d}x \tag{4.17}$$

in an obvious notation. Note that there is symmetry about the yz-plane for the x_0 bound of integration, but not the other as S_l and S_m need not have the same radii.

4.2. The octahedral interstitial void

The second type of interstitial void is called the octahedral interstitial void and occurs when four spheres are arranged more or less in a plane and spheres top and bottom close off the region. We must firstly identify when the octahedral interstitial void arises. It occurs when, of a group of six spheres, each one has overlap or contact with four others. The distinction between a tetrahedral interstitial void and an octahedral interstitial void is that the former is a group of five spheres whereas the latter is a group of six spheres.

We establish local coordinates using the centres of S_i, S_j and S_k as before. Unlike the case of a tetrahedral interstitial void there is no real distinction to be made amongst the spheres. Any of the six could be taken for S_i, S_j and S_k and we may as well take S_i, S_j and S_k in the order they are found to be elements of the group. The other spheres are S_l, S_m and S_n. Let $|x_l| < \min|x_m|, |x_n|$ so that S_l is the sphere closest to the plane defined by $P_i P_j P_k$. S_k and S_l are taken in anticyclic order as indicated in figure 10. (This may involve a relabelling of S_k and S_l.) To identify S_m and S_n we take S_m to be the sphere with the x-ordinate of P_m positive and S_n to be the sphere with the x-ordinate of P_n negative, i.e. S_m is on the positive x side of the yz-plane and S_n is on the negative side.

In figure 10 we illustrate a section parallel to the yz-plane and distant x_c from it. The small circles of the spheres intersect at Q_{ij}, Q_{jk}, Q_{kl} and Q_{li}. We can no longer assume that the y-ordinate of the intersection of S_i and S_j provides a suitable division of the region as was the case for the tetrahedral interstitial void in section 4.1. Instead we use the line joining Q_{li} and Q_{jk} with equation

$$\frac{z - z_{jk}}{y - y_{jk}} = \frac{z_{li} - z_{jk}}{y_{li} - y_{jk}}. \tag{4.18}$$

Then the area of the section of the octahedral interstitial void at x_c is given by

$$\begin{aligned} E_{ijkl}(x_\mathrm{c}) = &\int_{y_{li}}^{y_{kl}} (z_l - z_p)\,\mathrm{d}y + \int_{y_{kl}}^{y_{jk}} (z_k - z_p)\,\mathrm{d}y \\ &+ \int_{y_{kl}}^{y_{ij}} (z_p - z_i)\,\mathrm{d}y + \int_{y_{ij}}^{y_{jk}} (z_p - z_j)\,\mathrm{d}y, \end{aligned} \tag{4.19}$$

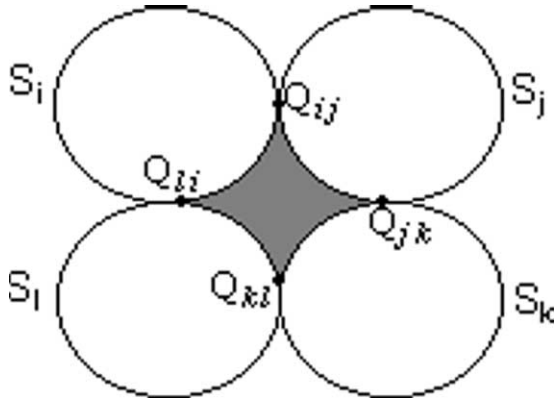

Figure 10. Octahedral interstitial void: a section of the four spheres S_i, S_j, S_k and S_l parallel to the yz-plane.

where z_i, z_j and z_k are defined by (3.17),

$$z_p = z_{jk} + \frac{z_{li} - z_{jk}}{y_{li} - y_{jk}}(y - y_{jk}) \tag{4.20}$$

from (4.18) and

$$z_l = c_l - \left[R_l^2 - (x_c - a_l)^2 - (y - b_l)^2\right]^{1/2}. \tag{4.21}$$

As in the case of the tetrahedral interstitial void, the top and bottom spheres may intrude into this area in a manner similar to that depicted in figure 11(b) and (c). For the top sphere we have the equivalent to (4.4), viz.

$$A_m(x_c) = \pi r_m^2 \tag{4.22}$$

which is added when

$$(y - b_m)^2 + (z - c_m)^2 = R_m^2 - (x_c - a_m)^2 > 0, \tag{4.23}$$

and, as in (4.5), we have

$$r_m^2 = R_m^2 - (x_c - a_m)^2. \tag{4.24}$$

The existence of any overlap of the type illustrated in figure 11(c) is determined by finding whether the small circle

$$(y - b_m)^2 + (z - c_m)^2 = r_m^2 \tag{4.25}$$

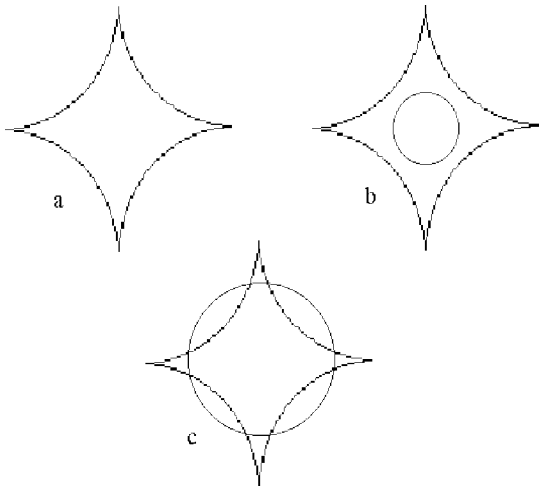

Figure 11. Octahedral interstitial void: the small circles intersect at the points Q_{ij}, Q_{jk}, Q_{kl} and Q_{li} which mark the region belonging to the octahedral interstitial void: (a) the closing off sphere does not intrude into the region; (b) it intrudes into the region, but is contained within the region; (c) it overlaps part of the outside region.

has real points of intersection with each of the small circles

$$\begin{aligned} &y^2 + z^2 = r_i^2, \\ &(y - b_j)^2 + z^2 = r_j^2, \\ &(y - b_k)^2 + (z - c_k)^2 = r_k^2, \\ &(y - b_l)^2 + (z - c_l)^2 = r_l^2, \end{aligned} \tag{4.26}$$

in turn. With the same notation as in section 4.1 the area of overlap is denoted by $C_{m\alpha}(x_c)$, where α now ranges over i, j, k and l.

The actual area of the section of the octahedral interstitial void at each value of x_c is

$$F_{ijklm}(x_c) = E_{ijkl}(x_c) - A_m(x_c) + \sum_{\alpha=i,j,k,l} C_{m\alpha}(x_c). \tag{4.27}$$

In (4.27) $A_m(x_c)$ and $C_{m\alpha}(x_c)$ are taken to be zero when they make no contribution to the area.

Finally the volume contributed by the octahedral interstitial void on the positive x side of the yz-plane is

$$V_{ijklm} = \int_0^{x_m} F_{ijklm}(x)\,\mathrm{d}x. \tag{4.28}$$

In a similar fashion the volume contributed by the octahedral interstitial void on the negative x side of the yz-plane is V_{ijkln}.

Table 1
The contributions to the nett volume of selected molecules.

Molecule	Shape	Volume/10^{-30} m^3				
		Gross volume	Pairwise overlap	Multiple overlap	Interstitial voids	Nett volume
HF	linear	16.48	3.40	0	0	13.08
F_2	linear	22.99	3.54	0	0	19.45
OF_2	angular	34.98	8.14	1.05	0	27.89
BF_3	planar	53.30	12.71	0.79	0	41.38
NF_3	trigonal pyramidal	47.52	10.89	1.97	0	38.60
CF_4	tetrahedral	60.98	12.26	1.05	0	49.77
PF_5	trigonal bipyramidal	84.42	27.11	1.76	0	59.07
SF_6	octahedral	93.39	38.09	7.42	0	62.73

5. The total volume

When all of the volumes of the regions of double, triple and multiple overlaps together with the contributions from any interstitial voids are added, the total volume occupied by the molecule, in an extension of the notation of section 3.4, is given by

$$V = \sum_{\alpha} V_{\alpha} - \sum_{\alpha\beta} V_{C\alpha\beta} + V_{C} + V_{T} + V_{O}, \tag{5.1}$$

where V_T is the total contribution from tetrahedral interstitial voids and V_O is the total contribution from octahedral interstitial voids, i.e. from the sum of the volumes of the spheres is subtracted the total of the pairwise overlaps and added the total of the n-tuple overlaps and the contributions from possible interstitial voids.

6. Conclusion

We illustrate the computations of molecular volumes by means of the algorithm developed here with a series of simple covalently-bonded fluorides. These molecules cover the basic shapes found in simple molecules and radicals. They are, with the geometric shape in brackets after the name, hydrogen fluoride (linear), fluorine (linear), oxygen difluoride (angular), boron trifluoride (planar), nitrogen trifluoride (trigonal pyramidal), carbon tetrafluoride (tetrahedral), phosphorus pentafluoride (trigonal bipyramidal) and sulphur hexafluoride (octahedral). For the purposes of this calculation we used the covalent bond lengths given by Bondi [19].

In the tabulated volumes (see table 1), which are ordered in terms of increasing number of fluorine atoms attached to the "central" atom, one may at first be surprised that the volume of the planar BF_3 is greater than that of the pyramidal NF_3. However, one should bear in mind that the atomic radius of boron is 165 pm while that of nitrogen is 146 pm [19]. Consequently, the volume of BF_3 is almost 50% greater than that of NF_3.

An analytical algorithm for the calculation of molecular volumes has been presented. The implementation in Mathematica may be obtained from one of the authors (PGLL at leachp@nu.ac.za). The algorithm addresses the shortcomings of previous algorithms as it takes into account n-tuple overlap of atoms as well as interstital voids. The final formula, (5.1), is very simple and easy to implement and it allows the practitioner to decide whether or not to exercise the option of inclusion of the calculation of interstitial voids and their contribution to the molecular volume. In the cases of the simple compounds listed above the calculation has shown that interstitial voids do not exist, whence their zero volumes. However, these are particularly simple molecules and one should not be tempted summarily to ignore the probable presence of interstitial voids in more complex molecules.

Acknowledgements

PGLL and TAF thank the National Research Foundation of South Africa and the University of Natal for their continuing support. PHM thanks the Centre for Theoretical and Computational Chemistry and the University of Natal for their support and SASOL for the award of a bursary.

References

[1] D. White and N.J. Coville, Adv. Organomet. Chem. 36 (1994) 95.
[2] G.R. Desiraju and A. Gavezzotti, Acta Crystallogr. B 45 (1989) 473.
[3] A. Gavezzotti, Chem. Phys. Lett. 161 (1989) 67.
[4] A. Gavezzotti, Acta Crystallogr. B 46 (1990) 275.
[5] D. Braga, A. Rodger and B.F.G. Johnson, Inorg. Chim. Acta 174 (1990) 185.
[6] P. Alard and S.J. Wodak, J. Comput. Chem. 12 (1991) 918.
[7] G. Arteca, N.D. Grant and P.G. Mezey, J. Comput. Chem. 12 (1991) 1198.
[8] G. Arteca, A. Hernandez-Laguna, J.J. Randez, Y.G. Smeyers and P.G. Mezey, J. Comput. Chem. 12 (1991) 705.
[9] N.W. Thomas, Acta Crystallogr. B 47 (1991) 588.
[10] A.Y. Meyer and D. Avnir, Struct. Chem. 2 (1991) 475.
[11] D. Braga, Chem. Rev. 92 (1992) 633.
[12] A.Y. Meyer, Chem. Soc. Rev. 15 (1986) 449.
[13] J.T. Edward and P.G. Farrell, Can. J. Chem. 53 (1975) 2965.
[14] I. Motoc, MATCH 4 (1979) 113.
[15] I. Motoc, MATCH 5 (1979) 275.
[16] I. Motoc, R. Vancea and I. Muscutaniu, MATCH 5 (1979) 263.
[17] I. Motoc, O. Dragomir-Fillimonescu and I. Muscutaniu, MATCH 8 (1980) 323.
[18] A. Gavezzotti, J. Am. Chem. Soc. 105 (1983) 5220.
[19] A. Bondi, J. Phys. Chem. 68 (1964) 441.