Proceedings of ALGORITMY 2009
pp. 392–401

DETECTING CAVITIES IN A SYSTEM OF OVERLAPPING SPHERES USING ENVELOPING TRIANGULATION*

JÁN BUŠA†, EDIK HAYRYAN‡, SHURA HAYRYAN§, CHIN-KUN HU§, JAROSLAV SKŘIVÁNEK†, AND MING-CHYA WU§

Abstract. The existence of cavities should be taken into account when solving different problems connected to the molecular properties. In the presented paper some special triangulation enclosing possible cavities points of a system of overlapping spheres is considered. An algorithm of the construction of such triangulation is discussed. It can be used for detecting the internal cavities and their positions.

Key words. cavities, overlapping spheres union, macromolecules, solvent-excluded volume, solvent accessible surface

AMS subject classifications. 51M04, 5104, 51M25

1. Introduction. Systems of overlapping spheres are widely used in macromolecular modeling, where atoms are represented by spheres. Study of geometric properties of such systems, like the surface area, the volume or the existence of internal cavities is important because of their physical applications.

The problem of computation of the volume and the surface area of the union of overlapping spheres has been the subject of methods both numerical (Pavani and Ranghino [1]; Gavezzotti [2]) and analytical (Richmond [3]; Kang, Nemethy, and Scheraga [4]; Gibson and Scheraga [5, 6, 7]; Guerrero-Ruiz, Ocadiz-Ramirez, and Garduno-Juarez [8]; Petitjean [9], Hayryan et al [10, 11]). More information can be found at `www.netsci.org/Science/Compchem/feature14.html`, where an excellent overview, written by Connolly, author of [12, 13, 14, 15] is presented. The GEPOL package created by Silla et al. [16, 17] for computing the molecular area and volume is referred there.

Lee and Richards [18] used a modified version of their area and volume computation method to calculate cavities. Rashin et al. [19] calculated cavities using modified Shrake and Rupley algorithm. Programs for identification and location protein cavities were presented in [20, 21]. In [22, 23] an algorithm for accurate computation of the location, shape, and size of internal cavities in proteins was presented. Though the triangulation defined in this paper looks like a part of Delaunay complex used in [23], its construction and usage described below are different.

2. Cavities enveloping triangulation. The decision whether some point lies inside or outside of the molecule envelope is not easy. Below we will describe the

*This work was supported by the Slovak Grant Agency under Grant No. 1/0819/08-28, by the National Science Council of the Republic of China (Taiwan) under Grant No. NSC 95-2111-M 001-003-MY3, National Center for Theoretical Sciences in Taiwan, and Academia Sinica (Taiwan) under Grant Nos. AS-92-TP-A09 & AS-95-TP-A07, and by Russian RFBR Grants 08-01-00800-a and 07-01-00738-a.

†Department of Mathematics, FEI, Technical University, 040 01 Košice, Slovak Republic (`Jan.Busa@tuke.sk, Jaroslav.Skrivanek@tuke.sk`)

‡Joint Institute for Nuclear Research, 141 980 Dubna, Moscow Region, Russia (`ayrjan@jinr.ru`)

§Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan (`shura@phys.sinica.edu.tw, huck@phys.sinica.edu.tw, mcwu@phys.sinica.edu.tw`)

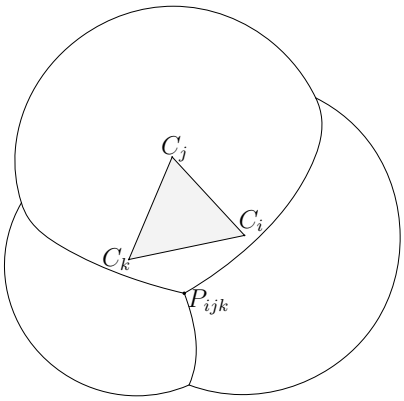

FIG. 2.1. *Wall triangle*

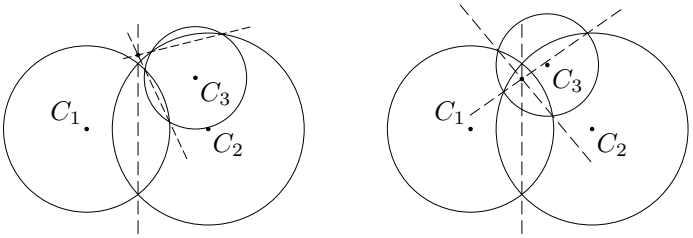

FIG. 2.2. *Nonwall and wall triangles*

algorithm of solving this problem, using the sum of oriented spherical angles for a point with respect to the oriented three-dimensional space triangles assembling some special triangulation. Similar triangulation has been used by Sanner et al. [24].

2.1. Introductory notions. General problem we confront is to decide for any point of the three-dimensional Euclidean space E_3, if it is or it is not inside of the possible cavity of the system composed of a finite number of overlapping spheres.

Let $\mathcal{M} = \bigcup_i \mathcal{S}_i$ be a finite union of spheres S_i in E_3. The cavity $\mathcal{C}$ is a nonempty, bounded, and closed domain in E_3, which whole boundary consists of the part of molecule's boundary, and which interior $\operatorname{Int}\mathcal{C}$ has no common points with the molecule $\mathcal{M}$.

The basic idea of proposed method consists in construction of a special *triangulation* $\Delta\mathcal{M}$ in $\mathcal{M}$ such that the conclusion whether a point is or is not in a cavity of $\mathcal{M}$, depends only on the relation between the point and the triangulation $\Delta\mathcal{M}$. Triangulation $\Delta\mathcal{M}$ as a set of triangles may be *nonempty* even if there exists no cavity. However, we can construct a minimal volume triangulation for each cavity. The final triangulation will then be empty, if there exists no cavity.

If we imagine a closed surface without the holes, composed only of spheres, the crucial role will play the spheres triplets, which we introduce below.

DEFINITION 2.1. *The triangle $\Delta C_1C_2C_3$ of centers of spheres of $\mathcal{M}$ is a* wall triangle *in $\mathcal{M}$ if and only if the intersection $\delta\mathcal{S}_1 \cap \delta\mathcal{S}_2 \cap \delta\mathcal{S}_3$ of the surfaces of the spheres consists just of one or two points, and at least one of them is external to $\mathcal{M}$ (does not lie in interior of $\mathcal{M}$, $\operatorname{Int}\mathcal{M}$).*

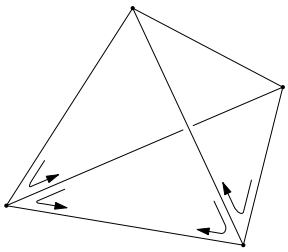

FIG. 2.3. *Coincident inward orientation of four triangles*

REMARK 1. *Fig. 2.1 shows the wall triangle in 3D space. Fig. 2.2 shows two different cases when the intersection of three spheres is not empty. Only spheres in the picture on the right can generate a* wall triangle*, the spheroids intersection* $\delta\mathcal{S}_1 \cap \delta\mathcal{S}_2 \cap \delta\mathcal{S}_3$ *is for the left case empty.*

PROPOSITION 2.2. *Let* $\Delta C_1C_2C_3$ *be a wall triangle in* $\mathcal{M}$ *with corresponding spheres* $\mathcal{S}_1$, $\mathcal{S}_2$ *and* $\mathcal{S}_3$*. Then the planes of intersections of all pairs of spheroids* $\delta\mathcal{S}_1$, $\delta\mathcal{S}_2$*, and* $\delta\mathcal{S}_3$ *intersect with the plane of centers* C_1*,* C_2*, and* C_3 *in the only point* $\boldsymbol{x}$ *which is an element of each of these spheres. So,* $\boldsymbol{x} = (x, y, z)$ *is the unique solution to the linear system*

$$
\begin{aligned}
&(x_2 - x_1)x + (y_2 - y_1)y + (z_2 - z_1)z = \\
&\qquad \frac{1}{2}(r_1^2 - r_2^2 - x_1^2 + x_2^2 - y_1^2 + y_2^2 - z_1^2 + z_2^2) \\
&(x_3 - x_1)x + (y_3 - y_1)y + (z_3 - z_1)z = \\
&\qquad \frac{1}{2}(r_1^2 - r_3^2 - x_1^2 + x_3^2 - y_1^2 + y_3^2 - z_1^2 + z_3^2)
\end{aligned}
$$

$$
\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0, \tag{2.1}
$$

where $C_i = (x_i, y_i, z_i)$*,* $i = 1$*, 2, and 3 are centers and* r_i *radii of the spheres.*

DEFINITION 2.3. *A* cavities enveloping triangulation $\Delta\mathcal{M}$ *(briefly* envelope triangulation*) of a finite system of spheres* $\mathcal{M}$ *is either an empty set or it is such a set of* wall triangles *in* $\mathcal{M}$ *with coincident orientation that*

(i) $\Delta\mathcal{M}$ *forms a family of polyhedral* $\mathcal{P}_i$ *with wall triangles at post of faces oriented inward such that* $Int\mathcal{P}_i \cap Int\mathcal{P}_j = \emptyset$*, for all* $i \neq j$*,*

(ii) any cavity point lies inside of some of these polyhedrons,

(iii) if a point from $E_3 - \mathcal{M}$ *is not included in any cavity of* $\mathcal{M}$*, it lies outside of the triangulation* $\Delta\mathcal{M}$*.*

Fig. 2.3 shows the coincident orientation of wall triangles for an *envelope triangulation.* If we are looking at any triangle from inside, the counter-clockwise orientation defines the positive (inward) triangle orientation.

REMARK 2. *For the 2D example of points included or not included in a cavity see Fig. 2.4 — points* P_1 *(with the sum of oriented angles along the triangulation equal to* 2π*) or* P_2 *(with the sum of angles equal to* 0*), respectively.*

The crucial question connected to the definition of the *enveloping triangulation* is the question about the existence of such triangulation. It seems to be evident, that if cavity exists, than the *cavities enveloping triangulation* exist.

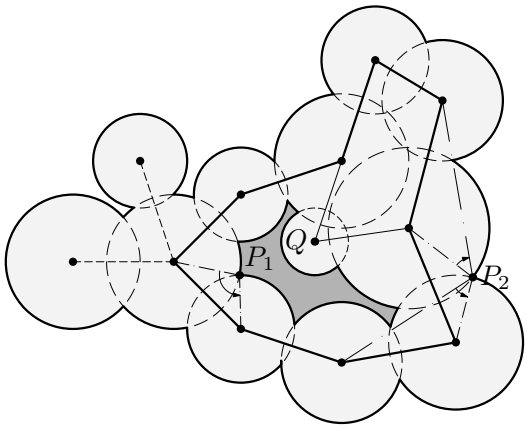

FIG. 2.4. *2D analogy for triangulation*

2.2. 2D inspiration of triangulation. Let us consider a 2D problem of computing the accessible "area" of a molecule consisting of atoms, represented by full circles (see Fig. 2.4). The boundary of the molecule consists of two parts. The outer part is the envelope of the molecule, the inner part is the envelope of a cavity, which is inaccessible for a solvent from outside. So in this case, the accessible "area" consists only of circular arcs of the outer part of the boundary, which is shown as a solid line on Fig. 2.4. If we connect the centers of the circles of the intersecting "spheres" (with some intersection point lying on the surface) by straight lines we get a special two-dimensional net. (In three-dimensional case we consider the triangles connecting three intersecting spheres, so each connection here represents a "triangle"). The connection, which has neighbours only at one side, cannot affect the closure of the boundary, and can be excluded from the mentioned net. We will denote such connections as "lugs". If we exclude some "lug" from the net, new lug can appear.

Our goal is to construct a minimal envelope of connections, containing all arcs, assembling all cavities boundaries. Such envelope is shown by the bold solid straight lines cycle on Fig. 2.4. There exists one more – *minimal* cycle in Fig. 2.4.

Some connections, which are not "lugs", may lie inside this envelope, and will be excluded from the net. At the end we obtain one or more straight line cycles, which together compose the "cavities enveloping triangulation" of a molecule.

3. Algorithm of the construction of the enveloping triangulation for $\mathcal{M}$. Below we briefly describe the algorithm of the construction of the *outer cavities enveloping triangulation.*

3.1. Collecting the preliminary wall triangles list. According to the definition of *wall* triangles we first assemble a list of all possible wall triangles included in molecule $\mathcal{M}$. For this purpose we start with the study of neighbourhood relations of sphere pairs. At the moment we check all pairs, in the future we would like to adopt the BSD Tree technique discussed in [24].

Using neighbourhood's relations we check all possible sphere triplets, if they have a unique solution to the system (2.1), and if at least one of three spheroids' intersection points lies not inside of any sphere of $\mathcal{M}$. For a better imagination see Fig. 3.1. Only 2D "wall triangles" are depicted by the centers connections.

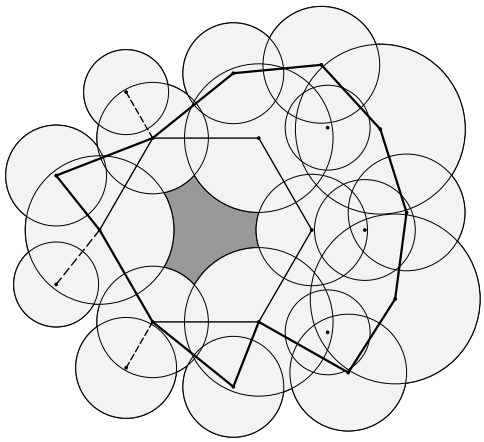

FIG. 3.1. *2D preliminary "wall triangles"*

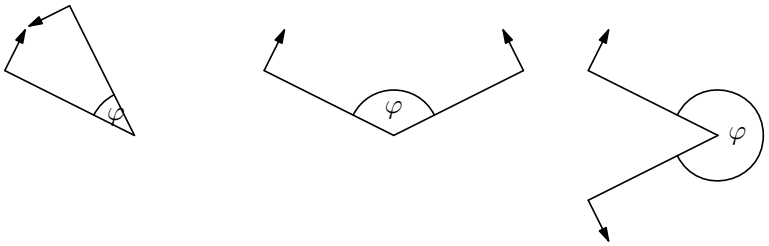

FIG. 3.2. *The angle of two adjacent triangles*

3.2. Creating new envelope triangulation component. Below will be described our algorithm of constructing the envelope triangulation. As we have mentioned above, the triangulation may be not unique, and it is possible to use different algorithms for its construction.

3.2.1. Lugs cutting. Any wall triangle in the list of wall triangles, which has a side not occupied by a neighbour (free side) must be "deleted" from the list of wall triangles. After the deletion of some lug-triangle, new one with some free side may appear. This lugs cutting should be repeated as long as any lug-triangle exists. After the lugs cutting, each triangle from the wall triangles list has neighbour at all sides.

3.2.2. Determination of a unique outer neighbour for each side. Let us consider an oriented triangle (see Fig. 2.3). Except some special sphere configurations, each triangle, which forms the part of the outer triangulation (its face is exposed to the outer space with respect to the set $\mathcal{M}$) has at each side only one neighbouring triangle, which includes the maximal angle with the actual triangle from the point of view of the inward orientation (see Fig. 3.2), and which constitutes a triangulation's part, too.

In the beginning, we do not know which triangle orientation is the inward orientation. Therefore, we define the neighbouring triangles for each side for both orientations. This information is stored in the matrix whose rows represent triangles.

At this stage we delete multiple triangles, which include the maximal angle with

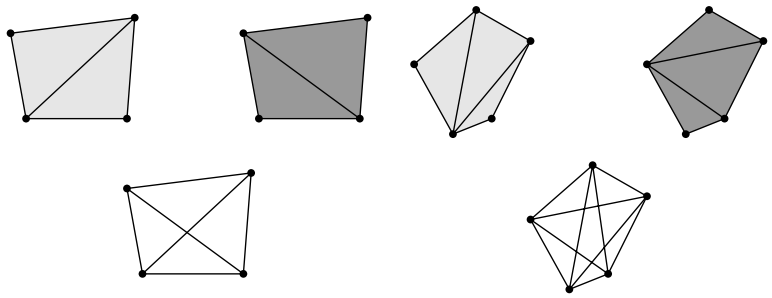

FIG. 3.3. *Special spheres configurations*

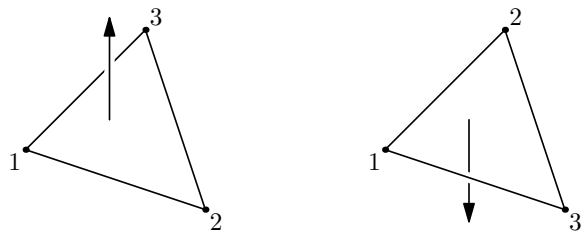

FIG. 3.4. *Positive (left) and negative (right) inward orientation of the southernmost triangle*

the actual triangle, we keep only one such triangle for each side. Such special case of multiple neighbours may appear, if more than three spheres' centers produce the triangles set, for which sides compose a complete graph (see Fig. 3.3 bottom). We may choose which triangle will be left, and which will be deleted. The deletion of some triangle activates the deletion of some of its neighbours (see the top pictures in Fig. 3.3).

3.2.3. Start of the new envelope component assembling. We need to start with the *wall* triangle from the preliminary list of wall triangles. The appropriate choice is e.g. the southernmost triangle (for which the South Pole of one vertex spheres is the southernmost).

When the choice of the first southernmost triangle was done, the (positive) inward orientation of this triangle is defined (see Fig. 3.4 left).

3.2.4. Looking for a new neighbour. The triangles list of the next envelope triangulation component is built in successive steps. The first triangle of this list is the southernmost triangle, described above. It will be the first actual triangle. Next one new neighbour triangle from the preliminary wall triangles list, which is the neighbour to the actual triangle adjoining his vacant side with corresponding orientation is added to the list. The indices of vertices of the new triangle are first written in increasing order. If the orientation should be changed (to become inward), we replace the indices of the second and the third vertex.

When all sides of the actual triangle are occupied by triangles of its neighbours, new actual triangle from component's list, with at least one free side is chosen.

3.2.5. Completion of the enveloping triangulation component. At some moment all triangles in the list of actual component will have all sides occupied. At this moment new envelope triangulation component is complete. All triangles creating

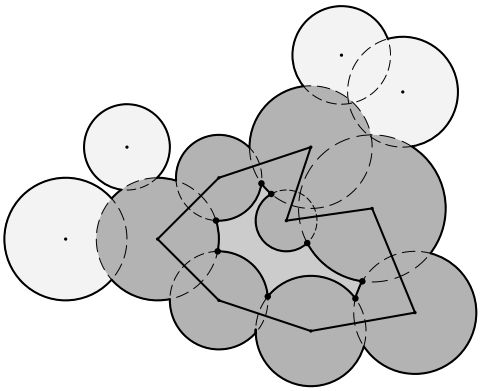

FIG. 3.5. *Minimal cavity triangulation*

the component are written to the output list. All other triangles from the preliminary wall triangles' list are checked, whether some vertex lies inside the actual envelope triangulation component. If yes, such triangle is removed from the preliminary wall triangles' list (its status is set to 0).

After the component is constructed, we can check if it contains a cavity or not. According to the Definition 2.1, at least one of the intersection points of three sphere surfaces of each wall triangle is external for Int $\mathcal{M}$. During the construction of the preliminary wall triangles list we save all these intersection points. If any of these points lies inside the triangulation polyhedral segment, then a cavity belonging to this triangulation component exists, otherwise, there is no cavity inside. The corresponding test is explained below in the Section 4.

After each cycle of assembling the envelope triangulation component, the number of triangles from the preliminary list becomes smaller and smaller. At some moment, it will contain less then 4 triangles and the *outer envelope triangulation* will be ready in hand.

3.3. The determination of a minimal volume triangulation and the cavities' boundaries. As we have mentioned above, at the end of the segment construction we are looking for the intersection points of three spheroids determining the wall triangles which belongs to the interior of the actual triangulation segment. If we find at least one such point, we know, that there exists at least one cavity. If we will find not only the first such point, but all points internal to the segment's *outer triangulation*, we will get the list of all wall triangles and corresponding spheres, which are partially exposed to some cavity. From such spheres a set of closed domains can be created, each of them containing inside exactly one cavity.

Corresponding wall triangles create together a set of *minimal volume boundary triangulations* for all cavities inside the actual *outer triangulation* segment. An example of 2D minimal volume cavity's triangulation with denoted internal vertices is shown in Fig. 3.5. It can be compared with the outer triangulation shown in Fig. 2.4. The minimal volume cavity's triangulation is also shown in Fig. 3.1 (see the internal cycle).

4. Calculating the spherical angles and the cavity index. As we have mentioned above, in 2D case the sum of oriented angles can be used to make decisions, if a point belongs to the interior or to the exterior of some closed curve. Similarly,

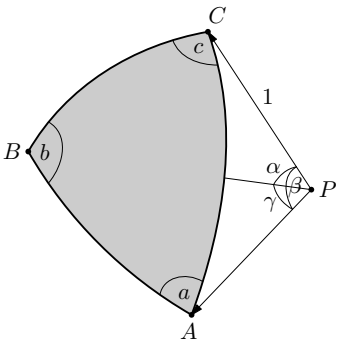

FIG. 4.1. *The surface area of the spherical triangle*

in 3D case, the sum of spherical angles can be used to make decisions, if a point lies inside the closed polyhedral, or outside of it.

4.1. Spherical triangle area. A spherical triangle is a figure formed on the surface of a sphere by three great circular arcs intersecting pairwise in three vertices.

Let ΔABC be a spherical triangle on the *unit sphere* with angles a, b, and c (measured in radians at the vertices along the surface of the sphere, see Fig. 4.1). Then the surface area σ of the spherical triangle is

$$\sigma(\Delta ABC) = a + b + c - \pi. \tag{4.1}$$

The arc (angular) lengths $\alpha = \angle BPC$, β, γ and vertex angles a, b, and c of the unit sphere triangle in Fig. 4.1 are related by the following cosine formula

$$\cos\alpha = \cos\beta\,\cos\gamma + \sin\beta\,\sin\gamma\,\cos a \tag{4.2}$$

and its next two permutations by variables (see [25, 26, 27]).

4.2. Spherical angle calculation. An oriented triangle $\Delta V_1V_2V_3$ in E_3 is seen from a point P (different from vertices) under spherical angle that is equal to the surface area of the projection of the triangle onto unit sphere with the center P (see Fig. 4.2). We denote this *oriented spherical angle* as $\sigma_P(\Delta V_1V_2V_3)$. The absolute value of $\sigma_P(\Delta V_1V_2V_3)$ one gets using Eq. (4.2) and Eq. (4.1). The sign of $\sigma_P(\Delta V_1V_2V_3)$ of this spherical angle is positive if the point P is on the inward side of the triangle, otherwise it is negative.

4.3. Cavity index. We now consider some point P, which does not lie on the triangulation surface. All vertices of cavity triangulation are projected on the unit sphere with the center in this point P (see Fig. 4.2).

DEFINITION 4.1. *Let P be a point of E_3 not belonging to the cavity triangulation $\Delta\mathcal{M}$. The* index *of P with respect to the cavity triangulation $\Delta\mathcal{M}$ is the number*

$$\chi_P(\Delta\mathcal{M}) = \frac{1}{4\pi}\sum_{\delta\in\Delta\mathcal{M}}\sigma_P(\delta).$$

PROPOSITION 4.2. *An envelope triangulation $\Delta\mathcal{M}$ encloses the point P (different from vertices of triangulation) if and only if $\chi_P(\Delta\mathcal{M}) = 1$.*

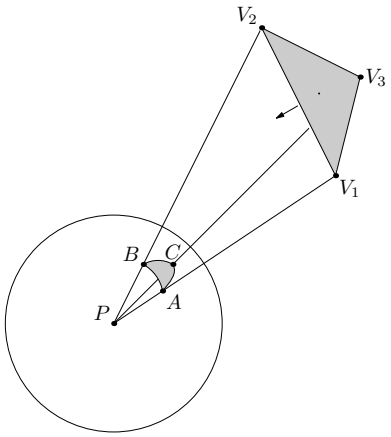

FIG. 4.2. *Cavity indexing*

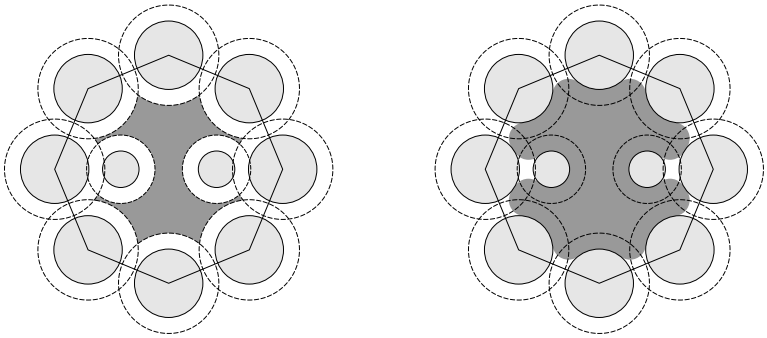

FIG. 5.1. *The central and the water cavity volumes*

5. Cavity volume definitions. There are several possibilities to define the volume of a cavity for a given radius of a probe sphere, which for the $r_p = 0$ will give the true value of the cavities volume.

Richmond [3] has defined the solvent-excluded volume to mean the volume contained within the solvent accessible surface, i.e. the volume which is inaccessible to the centers of solvent particles. The *central cavity volume* in Fig. 5.1 left should be added to the molecule (with enlarged radii) volume to get the solvent-excluded volume.

The domain accessible to probe sphere put into the cavity, is shown in the Fig. 5.1 right. The volume of this domain we may call the *water cavity volume*.

Conclusions. We have presented algorithm for detection of the existence of the cavities and for the study of their properties such as localization, boundary "atoms", central volume and surface area. It allows us also to compute the solvent-excluded volume and the solvent accessible surface area of a molecule.

REFERENCES

[1] R. PAVANI AND G. RANGHINO, *A method to compute the volume of a molecule,* Computers and Chemistry 6, 133–135 (1982).

[2] A. Gavezzotti, *The calculation of molecular volumes and the use of volume analysis in the investigation of structured media and of solid-state organic reactivity,* J. Am. Chem. Soc. 105, 5220–5225 (1983).
[3] T. J. Richmond, *Solvent accessible surface area and excluded volume in proteins: Analytical equations for overlapping spheres and implications for the hydrophobic effect,* Journal of Molecular Biology, No. 178, pp. 63–89 (1984).
[4] Y. K. Kang, G. Nemethy, and H. A. Scheraga, *Free energy of hydration of solute molecules. 1. Improvement of the hydration shell model by exact computations of overlapping volumes,* J. Phys. Chem. 91, 4105–4109 (1987).
[5] K. D. Gibson and H. A. Scheraga, *Exact calculation of the volume and surface area of fused hard-sphere molecules with unequal atomic radii,* Molec. Phys. 62, 1247–1265 (1987).
[6] K. D. Gibson and H. A. Scheraga, *Volume of the intersection of three spheres of unequal size. A simplified formula,* J. Phys. Chem. 91, 4121–4122 (1987).
[7] K. D. Gibson and H. A. Scheraga, *Surface area of the intersection of three spheres with unequal radii, a simplified analytical formula,* Molec. Phys. 64, 641–644 (1988).
[8] G. Guerrero-Ruiz, A. Ocadiz-Ramirez, and R. Garduno-Juarez, *ES FERA: A program for exact calculations of the volume and surface area of fused hard-sphere molecules with unequal atomic radii,* Computers Chem. 15, 351–352 (1991).
[9] M. Petitjean, *On the analytical calculation of van der Waals surfaces and volumes: Some numerical aspects,* J. Comp. Chem. 15, 507–523 (1994).
[10] S. Hayryan, C.-K. Hu, J. Skřivánek, E. Hayryan, and I. Pokorný, *A new analytical method for calculating the solvent accessible surface area of macromolecules and its gradients*, J. Comp. Chem. 26, 334 (2005).
[11] J. Buša, J. Džurina, E. Hayryan, Sh. Hayryan, Ch.-K. Hu, J. Plavka, I. Pokorný, J. Skřivánek, and M.-Ch. Wu, *ARVO: A Fortran package for computing solvent accessible surface area and volume of overlapping spheres via analytic equations*, Comput. Phys. Commun., 165, 59–96 (2005).
[12] M. L. Connolly, *Analytical molecular surface calculation,* J. Appl. Cryst. **16**, pp. 548–558 (1976).
[13] M. L. Connolly, *Computation of molecular volume,* J. Am. Chem. Soc. 107, 1118–1124 (1985b).
[14] M. L. Connolly, *Analytical molecular surface calculation,* J. Appl. Cryst. 16, 548–558 (1983).
[15] M. L. Connolly, *Molecular surface triangulation,* J. Appl. Cryst. 18, 499–505 (1985).
[16] E. Silla, F. Villar, O. Nilsson, J. L. Pascual-Ahuir, and O. Tapia, *Molecular volumes and surfaces of biomacromolecules via GEPOL: A fast and efficient algorithm,* J. Mol. Graphics 8, 168–172 (1990).
[17] E. Silla, I. Tun, and J. L. Pascual-Ahuir, *GEPOL: An improved description of molecular surfaces. II. Computing the molecular area and volume,* J. Comp. Chem. 12, 1077–1088 (1991).
[18] B. Lee and F. M. Richards, *The interpretation of protein structures: Estimation of static accessibility,* J. Mol. Biol. 55, 379–400 (1971).
[19] A. Rashin, M. Iofin, and B. Honig, *Internal cavities and burried waters in globular proteins,* Biochemistry 25, 3619–3625 (1986).
[20] P. Alard and S. J. Wodak, *Detection of cavities in a set of interpenetrating spheres,* J. Comp. Chem. 12, 918–922 (1991).
[21] D. G. Levitt and L. J. Banaszak, *POCKET: A computer graphics method for identifying and displaying protein cavities and their surrounding amino acids,* J. Mol. Graphics 10, 229–234 (1992).
[22] J. Liang, H. Edelsbrunner, P. Fu, P. V. Sudhakar, and S. Subramaniam, *Analytical shape computation of macromolecules: I. Molecular area and volume through alpha shape*, Proteins: Structure, Function, and Genetics 33, 1–17 (1998).
[23] J. Liang, H. Edelsbrunner, P. Fu, P. V. Sudhakar, and S. Subramaniam, *Analytical shape computation of macromolecules: II. Inaccessible cavities in proteins*, Proteins: Structure, Function, and Genetics 33, 18–29 (1998).
[24] M. F. Sanner, A. J. Olson, and J.-C. Spehner, *REDUCED SURFACE: an efficient way to compute molecular surfaces*, Biopolymers 38, (3), 305–320 (1996).

WWW pages:
[25] `http://mathworld.wolfram.com/SphericalTriangle.html` (October 31, 2008)
[26] `http://mathworld.wolfram.com/SphericalTrigonometry.html` (October 31, 2008)
[27] `http://www.shef.ac.uk/~phys/people/vdhillon/Teaching/` The Course of Spherical Geometry (October 31, 2008)