For geometric and topological analysis and for understanding the nonbond interactions between atoms a protein is modeled as a set of spheres in $\mathbb{R}^3$, each sphere corresponding to one atom of the protein. This model, known as the space-filling diagram, assumes that the exact position of each atom is known. We thus represent a protein $P$ as a set $\{B_1, \ldots, B_n\}$ of balls in $\mathbb{R}^3$. Let $z_i$ (resp. $r_i$) denote the center (resp. radius) of $B_i$. The union $\bigcup_i B_i$ is the space filled by $P$. The coordinates of the centers are typically determined using X-ray crystallography and NMR methods. The radius $r_i$ is typically the Van der Waals radius, defining the impenetrable volume of an atom. The smallest distance between neighboring atoms (in the crystalline state) that are not covalently bonded is the sum of their van der Waals radii. A range of values is assigned to van der Waals radii because it depends on how the atom is covalently bonded. Typical values are given in Table 22.1, which are originally taken from [1].

<table>
<thead>
<tr>
<th>Atom</th>
<th>Observed range (Å)</th>
<th>Singly bonded radius (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>1.0–1.54</td>
<td>1.17</td>
</tr>
<tr>
<td>O</td>
<td>1.4–1.7</td>
<td>1.40</td>
</tr>
<tr>
<td>N</td>
<td>1.55–1.6</td>
<td>1.55</td>
</tr>
<tr>
<td>C</td>
<td>1.70–1.78</td>
<td>1.75</td>
</tr>
<tr>
<td>S</td>
<td>1.75–1.80</td>
<td>1.80</td>
</tr>
</tbody>
</table>

Table 22.1: Van der Waals radii of atoms found in proteins.
If the radii of spheres is chosen to be van der Waals radii, the boundary of $\bigcup_i B_i$ is called the *van der Waals* surface (see Figure 22.1).

The biological activity of proteins are affected by their interactions with the environment: with water, salts, membranes, etc. The folded conformations of proteins usually occur only in a liquid-water environment or in membranes. In such cases the van der Waals surface is not particularly useful because part of this surface may not be accessible to water molecules. A more useful concept is the so called *solvent accessible surface*. It is typically defined by the center of a solvent molecule of radius 1.4Å, representative of a water molecule, in contact with the van der Waals surface. More precisely, we roll a sphere $D$ of radius $r$ along $V$ so that $D$ does not intersect the interior of any atom in $P$ but it remains in contact with $V$ all the time. The locus of the center of $D$ specifies the solvent accessible surface, which we denote by $SA$. If we grow the radius of each atom in $P$ to $r_i + r$, then the boundary of the union $\Sigma$ of resulting spheres is $SA$. Let $M$ denote the boundary of the region $\bigcup x \not\in \Sigma D + x$, i.e., $M$ is the envelope of the region that $D$ can access as we move it around without intersecting $P$. $M$ is called the *molecular, Connolly, or reentrant* surface (see Figure 22.1).

To understand the combinatorial and geometric structure of $\bigcup B_i$, we first study a simplified version in $\mathbb{R}^2$ and then extend the ideas to $\mathbb{R}^3$.

---

**Space-filling diagrams in 2D.** Let $\{B_1, \ldots, B_n\}$ be a set of disks in $\mathbb{R}^2$. The boundary curve $V = \partial(\bigcup B_i)$ consists of a set of circular arcs whose endpoints are intersection points between the boundaries of two disks. $V$ may consist of many connected components. If a connected component consists of a single circle, it does not have a vertex. It can be shown that the total number of arcs in $V$ is at most $6n - 12$ [5]. If $\bigcup B_i$ is simply connected, i.e., $V$ consists of a single cycle, as the case in Figure 22.2, then the number of arcs in $V$ can not exceed $2n - 2$. Both of these bounds are tight in the worst case.

We can obtain the curve $SA$ by growing the radius of each disk by $r$ (the outer boundary in Figure 22.2 (ii)). Like $V$, each connected component of $SA$ also consists of a sequence of circular arcs, each being a portion of one of the expanded disks. The structure of $M$ is however different (inner boundary in Figure 22.2 (ii)). $M$ consists of portions of $V$ and a set of circular arcs that are portions of the disk $D$. The latter arcs correspond to the (reflex) vertices of $SA$—let $v$ be a reflex vertex of $SA$, and let $p_1, p_2$ be the two points at which $D + v$ touches $V$. Then the portion of $V$ between $p_1$ and $p_2$ is replaced with the portion of $\partial D + v$ between those points. This is illustrated in Figure 22.3.
Space-fillig diagrams in 3D. Let \( \{B_1, \ldots, B_n\} \) be a set of balls in \( \mathbb{R}^3 \). We now analyze the combinatorial structure of the space-filling diagram model of a protein molecule in \( \mathbb{R}^3 \). Define \( \bigcup_i B_i \) as a finite set of balls as before. Figure (i) illustrates such a set of balls that is a model representation of gramicidin, a small protein with approximately 300 atoms. The combinatorial structure of \( \bigcup_i B_i \) is more complex than the union of disks in \( \mathbb{R}^2 \). To analyze the surface bounding \( \bigcup_i B_i \), we fix one ball \( B_i \), and examine the intersection between the other balls and \( \partial B_i \). The intersection \( D_{ij} = \partial B_i \cap B_j \) is a (possibly empty) disk on \( \partial B_i \). There are at most \( n - 1 \) such disks on \( \partial B_i \). \( \partial B_i \setminus \bigcup_{j \neq i} D_j \) is the portion of \( \partial B_i \) that appears on the Van der Waals surface \( V \) of \( \mathbb{P} \). The disks on \( B_i \) form essentially the same combinatorial structures as \( O(n) \) disks on a plane. Based on prior discussion in \( \mathbb{R}^2 \), we know that there are linear number of vertices on \( \bigcup_{j \neq i} D_j \). It follows that there are \( O(n^2) \) vertices on \( V \). By Euler’s formula, the number of circular arcs and the number of spherical patches in \( V \) is also \( O(n^2) \). This bound is known to be tight in the worst case. However, the packing density of atoms in a protein implies that a sphere can intersect only \( O(1) \) other spheres, so the complexity of \( V \) is only \( O(n) \).

The surface of the union of balls \( \bigcup_i B_i \) can again be made smooth by rolling a ball \( D(r) \) of radius \( r \) outside...
Union $B_i$ around its boundary. During the motion, ball $D(r)$ keeps contacting $\bigcup_i B_i$, but never intersects the interior of $\bigcup_i B_i$. As in the two-dimensional case, the center of the rolling ball $D$ describes the solvent accessible surface of $\bigcup_i B_i$. The front of $D$ describes its molecular surface. In the molecular surface, a reflex sphere patch replaces a vertex in the surface bounding $\bigcup_i B_i$; a reflex torus patch replaces a circular arc; a convex sphere patch replaces a convex sphere face. Figure (ii) illustrates the molecular surface of the gramacidin model.

### 22.1 Voronoi and Power Diagrams

![Figure 22.5: (i) Voronoi diagram of a set of disks. (ii) Power diagram of a set of disks. The figures are taken from [3].](image)

Observe that the combinatorial structure of the van der Waals surface of $\bigcup_i B_i$ and that of the solvent accessible surface are not necessarily the same. A natural question to ask is how to capture the structural change as we grow the radii of the balls comprising $\bigcup_i B_i$. To this end, we introduce the concept of \textit{weighted Voronoi diagram}. Each ball $B_i$ can be viewed as a point $z_i$ in $\mathbb{R}^3$ with weight $r_i$. The \textit{weighted distance} from a point $x \in \mathbb{R}^3$ to $z_i$ is defined as

$$d(x, B_i) = \|x - z_i\| - r_i.$$ 

If $x$ lies outside $B_i$, then $d(x, B_i)$ is the minimum distance between $x$ and a point in $B_i$; see Figure 22.1. The \textit{weighted Voronoi cell} $C_i$ of $z_i$ is:

$$C_i = \{x \mid d(x, B_i) \leq d(x, B_j), \forall j\}$$

The weighted Voronoi diagram of $\mathbb{P}$ represents the decomposition of space into weighted Voronoi cells so that the same ball is closest to all the points within each cell. Figure 22.1 (i) illustrates a weighted Voronoi diagram in $\mathbb{R}^2$.

In $\mathbb{R}^3$ the intersection of two Voronoi cells defines a two-dimensional patch. The intersection of three Voronoi cells defines a one-dimensional curve. The intersection of four Voronoi cells defines a zero-dimensional point. As we grow the balls in $\bigcup B_i$, the boundary of $\bigcup_i B_i$ sweeps through the Voronoi cells. Each time a new element in the Voronoi diagram is generated, there is a change in the combinatorial structure of $\bigcup_i B_i$. 

Next we define the power distance between a point \( x \in \mathbb{R}^3 \) and a ball \( B_i \) as

\[
\pi(x, B_i) = \|x - z_i\|^2 - r_i^2.
\]

If \( x \not\in B_i \), then \( \pi(x, B_i) \) is the square of the length of the tangent segment from \( x \) to \( B_i \); see Figure 22.1. The function \( \pi(x, B_i) = \pi(x, B_j) \) defines a plane in \( \mathbb{R}^3 \). The points defined by \( \pi(x, B_i) = \pi(x, B_j) \) satisfy

\[
0 = \|x - z_i\|^2 - r_i^2 - \|x - z_j\|^2 - r_j^2 + 2\langle x, z_j - z_i \rangle + \|z_i\|^2 - \|z_j\|^2 - r_i^2 + r_j^2.
\]

The above equation describes a plane in \( \mathbb{R}^3 \). Figure 22.1 illustrates the three possible cases in \( \mathbb{R}^2 \).

The power cell of a ball \( B_i \) is

\[
\Pi(B_i) = \{ x \mid \pi(x, B_i) \leq \pi(x, B_j), \forall j \}.
\]

The power diagram \( \Pi(P) \) is the decomposition of \( \mathbb{R}^3 \) induced by the power cells of balls in \( P \). It is well known that the number of vertices, edges, faces, and cells in \( \Pi(P) \) is \( \Theta(n^2) \) in the worst case. A two-dimensional power digram is shown in Figure 22.1 (ii). Note the difference between the power diagram and the weighted Voronoi diagram in Figure 22.1.

### 22.2 Delaunay Triangulation

Let \( P \) be a set of \( n \) spheres in \( \mathbb{R}^3 \). Delaunay triangulation \( DT(P) \) is the dual structure of the power diagram \( \Pi(P) \) of \( P \). Formally speaking,

\[
DT(P) = \{ R \subseteq P \mid \bigcap_{B \in R} \Pi(B) \neq \emptyset \}.
\]
See Figure 22.2. Note that two centers $z_i$ and $z_j$ are connected by an edge in $DT(\mathbb{P})$ if the power cells $B_i$ and $B_j$ share a face; three centers are connected by a triangle if the power cells of the three corresponding balls share an edge, and four centers are connected by a tetrahedron if the corresponding power cells share a vertex. $DT(\mathbb{P})$ decomposes the convex hull of the centers in $\mathbb{P}$ into simplices. The complexity of $DT(\mathbb{P})$ is the same as that of $\Pi(\mathbb{P})$.

References