Alpha Shapes — a Survey

Herbert Edelsbrunner[1]

Departments of Computer Science and of Mathematics, Duke University, Durham, North Carolina, USA, Geomagic, Research Triangle Park, North Carolina, USA, and IST Austria (Institute of Science and Technology Austria).
`edels@cs.duke.edu`

Alpha shapes have been conceived in 1981 as an attempt to define the shape of a finite set of point in the plane. Since then, connections to diverse areas in the sciences and engineering have developed, including to pattern recognition, digital shape sampling and processing, and structural molecular biology. This survey begins with a historical account and discusses geometric, algorithmic, topological, and combinatorial aspects of alpha shapes in this sequence.

1 History

The history of alpha shapes started in 1981 during my first trip to the American continent. Studying mathematics and computer science in Austria, I visited David Kirkpatrick at the University of British Columbia in Vancouver and spent a few months of intense research with him and with Raimund Seidel.

Conceptualization

In the late 1970s, Jarvis published a paper in which he generalized his planar convex hull algorithm to a method that generates something like the shape of a finite point set. We recall that this algorithm computes the convex hull one edge at a time, pivoting a line about one endpoint until it hits the other endpoint of the edge [34]. If we replace the line by a line segment of fixed length then we again walk around the point set but venture into sufficiently large cavities [35]. The size requirement on the cavities depends on the chosen length for the pivoting line segment. Sometimes it happens that the line segment gets lost in what we might reasonably call the interior of the shape and does not produce a reasonable description of the boundary.

The idea of alpha shapes was a response to the shortcomings of defining the outline of the shape by a pivoting line segment. Instead we imagined a disk of given radius to define the outline of the shape [24]. We call this the

eraser intuition. To be specific, let S be a finite set of points in the plane and α a positive real number. An open disk is said to be *empty* if it contains no point of S. Then the α-*hull* of S is the set of points that do not lie in any empty open disk of radius α. As illustrated in Figure 1, left, the boundary of the α-hull consists of circular arcs of constant curvature $1/\alpha$. By substituting straight for the circular arcs we get the α-*shape* of S in Figure 1, middle. More precisely, if the bounding circle of an empty open disk passes through two points $p, q \in S$ then we draw the edge connecting p to q and imagine the shape locally on the side of the edge opposite to the center of the circle.
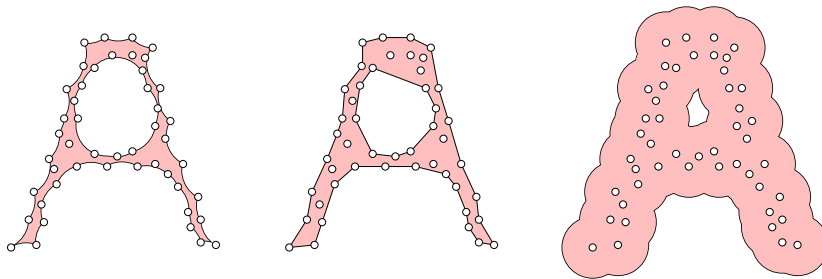

Fig. 1. From left to right: the α-hull of a set of points sampling the shape of the letter 'A' in the plane, the α-shape of the same set, and the union of disks of radius α centered at the points.

The eraser intuition was described in the first paper on α-shapes [24]. There, we also considered empty complements of disks, an extension of the idea that has been largely forgotten. Furthermore, we showed how the α-shape is related to the Delaunay triangulation of S, but more about this topic later.

Software

In the 1980s, alpha shapes were not at the center of anybody's research, also not ours. The interest was rekindled only about a decade later when Ernst Mücke worked on geometric software that culminated in the construction of alpha shapes for finite point sets in $\mathbb{R}^3$ [29]. It is worth revisiting decisive events during his doctoral work.

In the middle of the 1980s, Ernst and I worked on a symbolic perturbation method we referred to as the simulation of simplicity, or SoS for short [28]. This was a response to the general difficulty of turning geometric algorithms into correct software. The source of the difficulty is the mismatch between the high-level logic structure of the algorithm and the numerical evaluation of its geometric primitives. The latter sometimes provides false information,

such as calling a sequence of three almost collinear points a left-turn while they actually turn right. The logic structure can be unforgiving for even slight mistakes because they send it to geometrically impossible configurations for which nobody is prepared. Making the logic structure more tolerant proved to be difficult and incompatible with the desire to retain the efficiency of the algorithm. Alternatively, we can make sure that all primitives give correct information no matter how ambiguous the situation may be. This leads to the use of exact arithmetic and the need to resolve degenerate cases, such as three points collinear or worse. To finesse the latter question, SoS simulates a perturbation and thus reduces every degenerate to a generic case. Importantly, this reduction is done in a consistent manner so that the collection of all decisions made by the algorithm corresponds to a geometrically valid input.

At the time, SoS was a contentious proposal and there was intense competition between proponents of the two schools distinguished from each other by basing all numerical decisions on either inexact or on exact arithmetic. Ernst and I were therefore looking for interesting challenges to demonstrate the utility of SoS in creating geometric software. Three-dimensional Delaunay triangulations and alpha shapes seemed like perfect candidates. More about this later.

Applications

Generalizing the two-dimensional algorithms to three dimensions and turning theoretical algorithms into working code is one thing, then making the software available to the public is another. At the time, there was no world-wide-web but we worked with Ping Fu at the National Center for Supercomputing Applications and Ernst put his software on an ftp server for people to download. The high volume of interest surprised us. Apparently, the alpha shapes software tabbed into a common need. An important factor was the three-dimensionality of the computation, a relevant as well as challenging space to work in. Through these activities we learned something significant, namely the application community's wants. Looking back, I see three categories, each taking alpha shapes in its own interesting direction.

Pattern recognition. How can we characterize point distributions in space, such as galaxies in the universe and the like?
Digital shape sampling and processing. How can we reconstruct a shape from a finite point sample on its surface?
Structural molecular biology. How can we model the structural aspects of proteins and other biomolecules relevant for the functioning of life?

The applications motivated the research that unfolded. While they posed different challenges, they were based on common fundamental mathematical questions. It seemed most productive to address those and build solutions to the applications on answers to the mathematical questions. This is also how we structure this survey.

Geometry. What are the different guises of alpha shapes and how do they relate to each other?

Algorithms. How do we turn the mathematics into working software usable by non-specialists?

Topology. How do we approach the multi-resolution reality of data and how do we define and extract features?

Combinatorics. How do we exploit the rich combinatorial structure of alpha shapes to measure space?

This survey can only scratch the surface of each topic. While there is a lot known that will remain untouched here, we hasten to point out that there are entire fields that are yet unexplored.

2 Geometry

In this section, we describe alternatives to the eraser intuition and generalize alpha shapes beyond two dimensions and to points with weights.

Union of disks

The eraser intuition has a dual view that arises when we take the union of disks centered at the given points. To be specific, let z_1 to z_n be the points in $S \subseteq \mathbb{R}^2$, write $B_i(\alpha)$ for the closed disk with center z_i and radius $\alpha > 0$, and let $U(\alpha) = \bigcup_{i=1}^{n} B_i(\alpha)$ be the union of the disks. In other words, $U(\alpha)$ is the set of points $x \in \mathbb{R}^2$ that are at distance at most α from the set S. This invites an alternative interpretation as a sublevel set of the Euclidean distance function $\varrho_S : \mathbb{R}^2 \to \mathbb{R}$ defined by $\varrho_S(x) = \min_{1 \leq i \leq n} \|x - z_i\|$, namely $U(\alpha) = \varrho_S^{-1}[0, \alpha]$. This interpretation will be useful later. At the time being, we focus on the connection to the eraser intuition. For this, we consider a point $x \in \mathbb{R}^2$ and let $B_x(\alpha)$ be the closed disk with center x and radius α. Clearly, its interior is empty (of points in S) iff x does not belong to the interior of $U(\alpha)$. The boundary of the α-hull consists of points that lie on circles of radius α centered at points x that lie on the boundary of $U(\alpha)$. Generically, there are only two cases:

Type 1: there is a unique closest point z_i at distance α from x;
Type 2: there are two closest points at distance α from x.

Points of Type 1 lie in the interior of the circular arcs that make up the boundary of the union of disks. Points of Type 2 are the endpoints of these arcs. Imagine a point x moving continuously along such an arc, from one endpoint to the other. As x moves, the disk $B_x(\alpha)$ rotates about a point on its boundary, namely the point z_i at the center of the circle that carries the arc. At the endpoints, we get dual arcs in the α-hull, one ending at z_i and the other starting at z_i. The arcs of the α-hull may intersect and in the process

decompose each other into shorter arcs. It follows that a single circle may contribute more than one arc to the boundary of the α-hull or no arc at all. The latter case arises when the pair of points in S define two empty disks of radius α, one centered on each side of the line passing through the two points. In this case, the edge connecting the two points in the α-shape does not bound area on either side.

Voronoi and Delaunay

It is time to introduce the Voronoi diagram and the Delaunay triangulation, named after two Russian mathematicians [51, 12]. The two structures are intimately related to the α-shape. As before, let z_1 to z_n be the points in S and let the *Voronoi cell* of z_i consist of all points for which z_i is the closest,

$$V_i = \{x \in \mathbb{R}^2 \mid \|x - z_i\| \leq \|x - z_j\|, \forall j\}.$$

The points x that satisfy $\|x - z_i\| \leq \|x - z_j\|$ for a fixed $j \neq i$ define a closed half-plane. Hence V_i is the intersection of $n - 1$ half-planes, a closed and possibly unbounded convex polygon. By construction, any two Voronoi cells have disjoint interiors but they may intersect along shared pieces of their boundaries. Generically, there is only one possible case, namely that V_i and V_j share a common side. Indeed, they cannot share more than one side and to share just a corner would require four points of S on a common circle around that corner, a non-generic case. Together, the n Voronoi cells cover the entire plane. The *Voronoi diagram* of S is the set of Voronoi cells, $\text{Vor}\, S = \{V_i \mid 1 \leq i \leq n\}$.

The *Delaunay triangulation*, $\text{Del}\, S$, is dual to the Voronoi diagram. Specifically, whenever two Voronoi cells share a common side then the edge connecting the two corresponding points belongs to the Delaunay triangulation, and whenever three Voronoi cells share a common corner the triangle spanned by the three corresponding points belongs to the Delaunay triangulation. A more formal way of saying the same uses the concept of the *nerve* of the Voronoi diagram, that is, the collection of subsets whose cells have non-empty common intersection,

$$\text{Nrv}\,(\text{Vor}\, S) = \{X \subseteq \text{Vor}\, S \mid \bigcap X \neq \emptyset\}.$$

Generically, the nerve contains singletons, pairs of Voronoi cells with common sides, and triplets of Voronoi cells with common corners. The Delaunay triangulation is the canonical geometric realization of the nerve obtained by mapping each V_i to the generating point, z_i. The edges of $\text{Del}\, S$ are pairwise non-crossing and its triangles decompose the convex hull of the given points. Figure 2 illustrates the definitions by superimposing the Voronoi diagram and the Delaunay triangulation of a set of points in the plane.
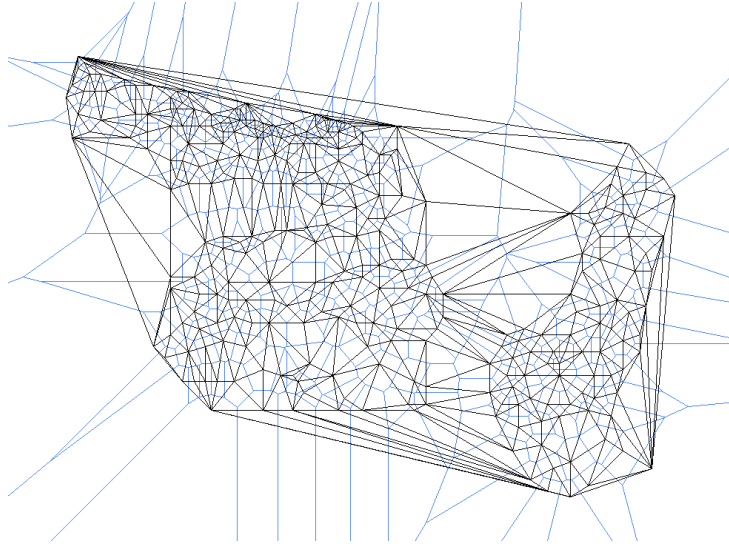
Fig. 2. The points mark trees in the Allerton Park in Monticello, Illinois. Each tree is associated with the region of points for which it is the nearest. Close inspection of the drawing shows that the Voronoi edges are sometimes not exactly halfway between the points. This is because we really see the weighted Voronoi diagram and the weighted Delaunay triangulation in which weights quantify sizes of the trees.

Alpha complex

The Voronoi cells decompose the union of disks into convex sets. Indeed, if $x \in U(\alpha)$ belongs to the Voronoi cell V_i then it also belongs to the disk $B_i(\alpha)$. This implies that $V_i \cap U(\alpha)$ is equal to $R_i(\alpha) = V_i \cap B_i(\alpha)$. The latter is the intersection of half-planes and a disk and therefore convex. For ease of reference, we write $\operatorname{Reg} S(\alpha)$ for the set of regions $R_i(\alpha)$. This definition is illustrated in Figure 3. The decomposition of the union of disks into convex sets sheds light on the relationship between the alpha shape and the Delaunay triangulation. First we observe that $R_i(\alpha) \subseteq V_i$ for each i. If a collection of sets $R_i(\alpha)$ has a non-empty common intersection then so do the corresponding Voronoi cells. In other words, the nerve of $\operatorname{Reg} S(\alpha)$ is isomorphic to a subsystem of the nerve of $\operatorname{Vor} S$. We use the same canonical geometric realization as for Delaunay triangulations and call the result the α-*complex* of S, denoting it by $K(\alpha)$. Specifically, $K(\alpha)$, is the realization of the nerve of $\operatorname{Reg} S(\alpha)$ obtained by mapping $R_i(\alpha)$ to z_i for each i. By construction, the α-complex is a subcomplex of the Delaunay triangulation.

The connection to the α-shape should be clear: it is the union of all simplices in the α-complex. More formally, we introduce the *underlying space*, $|K(\alpha)|$, which is the set of points contained in the simplices of $K(\alpha)$ together
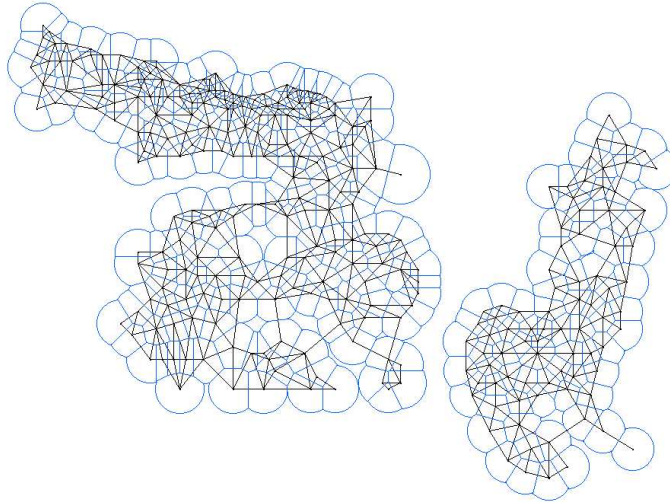
Fig. 3. Every Voronoi cells is restricted to a disk centered at the generating point. The nerve of the resulting convex sets is geometrically realized as the α-complex.

with the subset topology inherited from the Euclidean plane. This is formally what we consider the α-shape of S. There are uncountably many unions of disks, one for each α, but only finitely many alpha complexes. Furthermore, these complexes are totally ordered by inclusion giving rise to what we call a *filtration* of the Delaunay triangulation,

$$\emptyset = K_0 \subset K_1 \subset \ldots \subset K_m = \mathrm{Del}\, S.$$

For example, the first non-empty complex, K_1, is the point set S itself. Then we add edges and triangles until we eventually arrive at the Delaunay triangulation. Importantly, each K_j in the sequence is a complex. This means that whenever we add an edge, the complex also contains the two endpoints as vertices, and whenever we add a triangle, the complex also contains the three sides as edges.

Weights

An important generalization of the concepts uses real weights to control how each point influences its surrounding. Specifically, we let $w_i \in \mathbb{R}$ be the weight of the point z_i and call $\pi_i(x) = \|x - z_i\|^2 - w_i$ the *weighted squared distance* of x from z_i. Assuming $w_i \geq 0$ we may imagine a circle with radius $w_i^{1/2}$ centered at z_i such that $\pi_i(x) < 0$ inside the circle, $\pi_i(x) = 0$ on the circle, and $\pi_i(x) > 0$ outside the circle. The *weighted Voronoi cell* consists of all points for which z_i minimizes the weighted squared distance, $V_i = \{x \in \mathbb{R}^2 \mid \pi_i(x) \leq$

$\pi_j(x), \forall j\}$. The *weighted Voronoi diagram* is the set of weighted Voronoi cells, $\mathrm{Vor}\, S = \{V_i \mid 1 \le i \le n\}$. We use the same notation in the weighted case as in the unweighted case.

The *weighted Delaunay triangulation* should be clear. It is the canonical geometric realization of $\mathrm{Nrv}\,(\mathrm{Vor}\, S)$ obtained by mapping each V_i to the generating point, z_i. Some caution is in order: not every point z_i generates a non-empty weighted Voronoi cell. To describe how a cell can end up being empty, we say a point z with weight w is *orthogonal* to another point y with weight v if $\|y - z\|^2 = w + v$. If $v, w > 0$ this indeed corresponds to two circles that cross at two right angles. Generically, three weighted points have a unique weighted point that is orthogonal to all three. Now consider weighted points z_i, z_j, z_k, z_ℓ such that z_i is in the triangle spanned by the other three and let the point y with weight v be orthogonal to z_j, z_k, z_ℓ. If $\|z_i - y\|^2 > w_i + v$ then V_i is empty and z_i is not a vertex of $\mathrm{Del}\, S$.

Finally, we generalize alpha complexes to the weighted case. Given $\alpha \in \mathbb{R}$, we let $B_i(\alpha)$ be the closed disk with center z_i and radius $(\alpha^2 + w_i)^{1/2}$. If $\alpha^2 + w_i < 0$ then the root is imaginary and $B_i(\alpha)$ is empty, by definition. The union of the disks is $U(\alpha) = \bigcup_{i=1}^n B_i(\alpha)$. Similar to the unweighted case, we have $V_i \cap U(\alpha) = V_i \cap B_i(\alpha)$. Hence, the sets $R_i(\alpha) = V_i \cap B_i(\alpha)$ form again a convex decomposition of the union. The *weighted α-complex*, $K(\alpha)$, is the canonical geometric realization of the nerve of the collection of sets $R_i(\alpha)$.

Weighted Voronoi diagrams are known under various names, including *power diagrams* and *Dirichlet tessellations*; see e.g. [4]. Similarly, weighted Delaunay triangulations are known under different names, including *regular triangulations* and *coherent triangulations*; see e.g. [31]. We prefer to use the adjective 'weighted', which we drop when it is convenient to blur the difference between the weighted and the special, unweighted case.

Convex polyhedra

Voronoi diagrams and Delaunay triangulation can also be viewed as convex polyhedra in $\mathbb{R}^3$. This view is sometimes more natural and high-lights a fundamental symmetry between the two concepts.

Let S be the set of points $z_i \in \mathbb{R}^2$ with weights $w_i \in \mathbb{R}$ for $1 \le i \le n$. Consider the function $\pi : \mathbb{R}^2 \to \mathbb{R}$ defined by $\pi(x) = \min_{1 \le i \le n} \pi_i(x)$. This is a piecewise quadratic function, one piece for each non-empty weighted Voronoi cell. Its restriction to V_i is $\|x - z_i\|^2 - w_i$ and subtracting the squared norm of x turns this into a linear function. Hence, $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x) = \pi(x) - \|x\|^2$ is piecewise linear, again a piece for each non-empty weighted Voronoi cell. Specifically,

$$f(x) = \left(\min_{1 \le i \le n} \pi_i(x) \right) - \|x\|^2 = \min_{1 \le i \le n} \left(\pi_i(x) - \|x\|^2 \right),$$

which shows that f is concave. It is best visualized by its graph. For each i, the graph of $\pi_i(x) - \|x\|^2$ is a plane in $\mathbb{R}^3$, and if $w_i \ge 0$ it intersects the graph

of $-\|x\|^2$ in a curve whose vertical projection to $\mathbb{R}^2$ is the circle with center z_i and radius $w_i^{1/2}$. The graph of f is the lower envelope of these n planes. It is thus the boundary of a convex polyhedron that projects vertically to the weighted Voronoi diagram of S.

There is a similar dual view of the Delaunay triangulation. Specifically, $(z_i, w_i - \|z_i\|^2)$ in $\mathbb{R}^3$ is the *polar point* of the plane that is the graph of $\pi_i(x) - \|x\|^2$. We collect all n polar points, add the point $(0, -\infty)$, and construct the convex hull in $\mathbb{R}^3$. This is again a convex polyhedron. We obtain the weighted Delaunay triangulation by taking its boundary, removing all faces incident to the point $(0, -\infty)$, and projecting the rest vertically to $\mathbb{R}^2$.

In the unweighted case, we get a circumscribed polyhedron for the Voronoi diagram and an inscribe polyhedron for the Delaunay triangulation. Specifically, the face planes of the first polyhedron are all tangent to the graph of $-\|x\|^2$ and the vertices of the second polyhedron all lie on this graph. In the weighted case, we relax the circumscribed/inscribed condition and allow for general convex polytopes. Generically, the polyhedra defined as lower envelopes of planes are simple and the polyhedra defined as convex hulls are simplicial. It is instructive to describe the α-complex in terms of how these two convex polyhedra interact with the graph of $\alpha^2 - \|x\|^2$. We leave this as an exercise to the reader.

Beyond two dimensions

The above definitions are easily extended to weighted points in $\mathbb{R}^d$, for any positive integer d. The Voronoi diagram consists of cells that are convex polyhedra of dimension d. Generically, the common intersection of any $i+1 \leq d+1$ cells is either empty or a $(d-i)$-dimensional convex polyhedron. The nerve thus consists of $(i+1)$-tuplets of Voronoi cells, for $0 \leq i \leq d$. Correspondingly, the Delaunay triangulation consists of simplices of dimension $0 \leq i \leq d$. For a given $\alpha \in \mathbb{R}$, the α-complex is a subcomplex of the Delaunay triangulation. Finally, all diagrams can be viewed as convex polyhedra in $\mathbb{R}^{d+1}$ or as the interaction of such a polyhedron with the graph of the function that maps each point $x \in \mathbb{R}^d$ to $\alpha^2 - \|x\|^2$.

3 Algorithms

There are many different strategies to construct the α-complex of a finite point set and most are based on the Delaunay triangulation. We focus on the most common method which computes the Delaunay triangulation incrementally, adding one point at a time, and then selects the α-complex as a subcomplex. Before we agonize about how fast we can do the computation we ask how big a structure we can expect.

Number of simplices

Assume first that S is a set of n weighted points in $\mathbb{R}^2$. The edges of the De-launay triangulation form a planar graph, that is, they can be drawn without any two crossing each other. Euler's formula for the Euclidean plane states that #vertices $-$ #edges $+$ #faces $=$ #components $+ 1$. For generic input we have #vertices $\leq n$, #faces $=$ #triangles $+ 1$, and #components $= 1$. Hence, #edges $-$ #triangles $\leq n - 1$. But every triangle has three edges and every edge belongs to two faces, either two triangles or one triangle and the outer face. Letting k be the number of edges of the outer face we thus get 2#edges $= 3$#triangles $+ k$. This implies #edges $\leq 3n - 3 - k$ and #triangles $\leq 2n - 2 - k$. In short, the entire Delaunay triangulation has fewer than $6n$ simplices and so has every α-complex. This is true both in the weighted case and in the unweighted case.

The situation is a lot more subtle in $d \geq 3$ dimensions because the number of simplices in the Delaunay triangulation heavily depends on how the points are distributed. The worst case corresponds to so-called cyclic polytopes in $\mathbb{R}^{d+1}$. These are polytopes with n vertices for which every subset of $i + 1$ vertices spans a face of dimension i, for $0 \leq i \leq (d - 1)/2$, see e.g. [33]. For example, there are polytopes in $\mathbb{R}^4$ that have an edge between every pair of vertices and similarly there are Delaunay triangulations in $\mathbb{R}^3$ whose edges form the complete graph on n vertices. According to the Upper Bound Theorem, the cyclic polytopes are indeed the worst case [53] and the weighted Delaunay triangulation for n points in $\mathbb{R}^d$ has at most some constant times $n^{\lceil d/2 \rceil}$ simplices. This is not to say that Delaunay triangulations of this large size are typical. Quite the opposite, large Delaunay triangulations seem very rare but to say anything concrete we would have to decide upon the point distributions we consider. If we pick n points uniformly at random within the unit cube, the expected number of simplices is only some constant times n, where the constant depends exponentially on the dimension [15]. If the points are well distributed on a generic smooth surface in $\mathbb{R}^3$ then the number of simplices is bounded from above by a constant times $n \log_2 n$ [2]. To complicate matters further, it is of course possible that the Delaunay triangulation is large but the α-complex is small. For example, suppose the points are unweighted and no two are closer to each other than some constant times α. Then we can use a packing argument to prove a linear upper bound on the number of simplices in the α-complex while the Delaunay triangulation can still reach worst-case size. This assumption is realistic for the data that describes proteins and other biomolecules in $\mathbb{R}^3$.

Incremental construction

The most popular method for constructing a Delaunay triangulation adds one point at a time. Early versions of this algorithm have been studied in the convex polytope literature [33] and described for Voronoi diagrams and Delaunay

triangulations by Bowyer [5] and by Watson [52]. Bells and whistles needed for an efficient implementation have been added later, as will be discussed shortly. To explain the principle of the algorithm, we return again to the two-dimensional case. Suppose S_i consists of points z_1 to z_i and we construct $\text{Del}\,S_{i+1}$ from $\text{Del}\,S_i$ by adding the point z_{i+1}. Consider first how $\text{Vor}\,S_{i+1}$ is obtained from $\text{Vor}\,S_i$. As illustrated in Figure 4, the point z_{i+1} invades
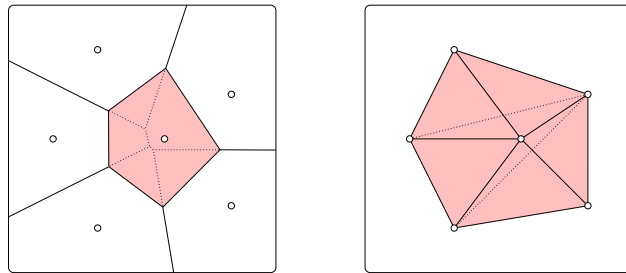

Fig. 4. Left: the addition of the point in the middle creates a new Voronoi cell obtained by stealing pieces from neighboring cells. Right: the Delaunay triangulation evolves by substituting the star of the new point for a collection of simplices that cover the same region in the plane.

some of the existing Voronoi cells to establish its own cell. The generators of the invaded cells are precisely the neighbors of z_{i+1} in the new Delaunay triangulation, $\text{Del}\,S_{i+1}$. Calling the set of simplices incident to z_{i+1} the *star* of the new point, $\text{St}\,z_{i+1}$, and the set of simplices in $\text{Del}\,S_i$ that cover the same region its *prestar*, $\text{Pt}\,z_{i+1}$, we have $\text{Del}\,S_{i+1} = \text{Del}\,S_i - \text{Pt}\,z_{i+1} \cup \text{St}\,z_{i+1}$. Indeed, the triangles in the prestar are exactly the ones whose circumcircles are encroached by z_{i+1} and the open disks they bound are therefore no longer empty. Note that the region covered by the prestar is not necessarily convex but it is always star-shaped since the new triangles in the star are formed by connecting z_{i+1} to the boundary edges of that region.

To turn the incremental algorithm into an efficient implementation we need quick methods to find the neighbors of the new point and to substitute the star for the prestar. Finding neighbors is primarily an algorithmic problem and will not be discussed further.

Randomization

Even if we are able to remove the prestar and add the star in constant time per simplex, the totality of these operations may cost a substantial amount of time. To illustrate this fact, we consider the easy case in which the points are unweighted in $\mathbb{R}^2$ and z_{i+1} lies inside the convex hull of the first i points. The common boundary of the star and the prestar is then a bounded star-shaped polygon. Letting k be its number of sides, we end up removing $k-2$

triangles in the prestar and adding k triangles in the star. In other words, we spend time proportional to $2k - 2$ with a net plus of only two triangles in the triangulation. The number of sides, k, can be as large as i, so the total effort may be as large as $2\sum(i-1)$, which is roughly n^2, while the final triangulation has fewer than $2n$ triangles.

It is indeed possible that the algorithm creates and destroys about n^2 triangles, but this is not very likely. A useful perspective is the randomization over all input sequences. Specifically, we consider all $n!$ permutations of the n points, compute the effort for each, and divide the sum by $n!$ to get the expected effort. This has been advocated as a general strategy in computational geometry by Clarkson and Shor [10]. The analysis has been simplified by Seidel [49], who suggests we look at the entire process going backwards. Indeed, each point is equally likely to be the last one in the input sequence. Since the average number of triangles in the star of a vertex in $\operatorname{Del} S_n = \operatorname{Del} S$ is less than 6, this implies that the combined effort for the respective last points of the $n!$ permutations is less than $6n!$. The same is true for all the other points, so the total effort is less than $6n \cdot n!$. The expected effort for a single input sequence is therefore less than $6n$. In other words, chances are that the algorithm constructs only three times as many triangles as necessary.

As always, the situation is more subtle in three and higher dimensions and we refer to the relevant literature as cited e.g. in [19]. It should be mentioned that randomizing the input sequence has also undesired effects. In particular, the access pattern to the representation of the evolving Delaunay triangulation in memory tends to be non-local, which goes against the grain of current computer hardware which uses fast cache to take advantage of locality in the computation. A compromise between the benefits of random input sequences and locality in computation has been described in [1].

Flipping

Instead of first removing the entire prestar and then adding the entire star, we may prefer to modify $\operatorname{Del} S_i$ one simplex at a time, eventually turning it into $\operatorname{Del} S_{i+1}$ while preserving that we have a triangulation during the entire process. This can be done using *flips*, also known as *bi-stellar operations*. We again simplify the situation by restricting ourselves to the unweighted planar case. An edge flip replaces two triangles sharing an edge by the other two triangles decomposing the same quadrangle. More generally, we define a flip in $\mathbb{R}^2$ in terms of the boundary of a tetrahedron in $\mathbb{R}^3$. Its projection to $\mathbb{R}^2$ is either a quadrangle or a triangle and its upper and lower boundaries provide two triangulations of this projection. Flipping refers to substituting one local triangulation for the other.

The first use of flipping to construct Delaunay triangulations can be found in a paper by Lawson [40]. He starts with an arbitrary triangulation of a set S of unweighted points in $\mathbb{R}^2$ and constructs $\operatorname{Del} S$ in a sequence of edge flips. Each flip is directed in the sense that the two new triangles belong to

the Delaunay triangulation of the four points involved. Although there are configurations that require a quadratic number of such flips, the algorithm is generally quite efficient. There are, however, serious obstacles to generalizing this algorithm to three and higher dimensions. Specifically, there are triangulations of a set of n points in $\mathbb{R}^3$ that are not Delaunay and that do not permit a single flip to bring them closer to the Delaunay triangulation [36]. However, such obstacles do not exist if we add one point at a time, flipping to the Delaunay triangulation before adding the next point. This has first been proved for unweighted points in $\mathbb{R}^3$ [37] and then generalized to weighted points in $\mathbb{R}^d$ [30]. If applied as part of the incremental algorithm, each flip removes one of the d-simplices in the prestar. Hence the number of flips is directly related to the total number of simplices created and destroyed in the process.

Sorting simplices

Given the Delaunay triangulation, we construct the α-complex by selecting the simplices whose Voronoi cells have a non-empty common intersection with the union of balls. To expedite this process, we use the fact that each simplex $\sigma_j \in \operatorname{Del} S$ has a threshold α_j such that $\sigma_j \in K(\alpha)$ iff $\alpha_j \leq \alpha$. This suggests we sort the simplices in a non-decreasing order of the thresholds and retrieve the α-complex as a prefix of this sequence. For some of the simplices, the threshold is the radius of the smallest circumsphere and for others it is the threshold of another simplex in its star. A case analysis complete with analytic formulas for weighted points in $\mathbb{R}^d$ can be found in [16]. Here we illustrate the situation for a set of unweighted points in $\mathbb{R}^2$. The threshold of every vertex is zero and that of every triangle in the Delaunay triangulation is the radius of its circumcircle. For an edge $\sigma_j = \operatorname{conv}\{a, b\}$, there are two cases which can be distinguished by considering the smallest circumcircle, the circle centered at the midpoint and passing through the endpoints of the edge. If the third vertex of each incident triangle lies outside this circle then the threshold is the radius of the circle, $\alpha_j = \|a - b\|/2$. Otherwise the third vertex of one triangle lies inside the circle and the threshold of the edge is the same as that of this triangle.

After computing all thresholds, we sort the simplices into a sequence $\sigma_1, \sigma_2, \ldots, \sigma_m$ such that $\alpha_j \leq \alpha_{j+1}$ for all j. Calling this sequence a *filter*, we get the filtration of the Delaunay triangulation by considering all prefixes of simplices with thresholds smaller than or equal to a real number α. For technical reasons, we require that all prefixes of the filter are complexes, not just those defined by real numbers α. This can be achieved by sorting simplices with equal thresholds by dimension. Guaranteeing this property can be challenging numerically but the advantages are simpler and more efficient algorithms for topological properties of the α-complexes, as we will see later.

Implementation

We aim for an implementation that is robust and reflects the structure of the abstract algorithm without cluttering it with the treatment of special cases. By robust we mean provably correct for all possible inputs. To achieve these two goals, we advocate a disciplined approach in which the numerical aspects are clearly separated from the control flow. We illustrate what we mean by considering the in-circle test for a sequence of four points $a = (a_1, a_2)$, $b = (b_1, b_2)$, $c = (c_1, c_2)$, $x = (x_1, x_2)$ in the plane. The point x lies on the circle defined by a, b, c iff the matrix

$$\Lambda = \begin{bmatrix} 1 & a_1 & a_2 & a_1^2 + a_2^2 \\ 1 & b_1 & b_2 & b_1^2 + b_2^2 \\ 1 & c_1 & c_2 & c_1^2 + c_2^2 \\ 1 & x_1 & x_2 & x_1^2 + x_2^2 \end{bmatrix}$$

has vanishing determinant. If a, b, c form a left-turn then the point x lies inside the circle iff $\det \Lambda < 0$. Finally, a, b, c form a left-turn iff the upper left 3-by-3 submatrix Δ of Λ has positive determinant. We thus get the following boolean function that recognizes when x lies inside the circle defined by a, b, c:

```
boolean IsInCircle (a, b, c, x):
    return det Δ · det Λ < 0.
```

The test is simple but numerically instable if one or both matrices are not clearly full rank. Using exact arithmetic, we can get the correct sign in all cases. It remains to describe how we can think of the non-generic situations defined by $\det \Delta \cdot \det \Lambda = 0$. We propose to disambiguate the many possibilities using a symbolic perturbation that turns every input configuration into a generic configuration nearby. The perturbation is only simulated and does not affect the data. Following [28], we make use of the indices of the input points, writing $a = z_i$ and so forth, and of a sufficiently small but positive real number ε. The perturbation is smaller for larger indices and we define $z_{i,j}(\varepsilon) = z_{i,j} + \varepsilon^{2^{3i-j}}$ for $1 \leq i \leq n$ and $1 \leq j \leq 3$, where $z_{i,3} = z_{i,1}^2 + z_{i,2}^2$. Note that $z_{i,3}(\varepsilon)$ is not equal to $z_{i,1}(\varepsilon)^2 + z_{i,2}(\varepsilon)^2$, which is what Knuth suggested to use [38]. Instead, we perturb an unweighted to a weighted point and this way avoid unnecessary algebraic complications. To see that the method works, we verify that for sufficiently small ε the product of determinants is non-zero no matter how non-generic the input points are chosen. An extreme test case is $a = b = c = x = (0, 0)$. We also note that the perturbation is easy to implement and compute, despite the fact that finding an appropriate ε is prohibitively expensive. Indeed the determinants themselves are polynomials in ε and finding the first non-zero coefficient suffices to decide the sign.

4 Topology

In this section, we focus on how the α-shape is connected and how its connectivity is measured.

Homotopy type

Recall that the α-complex is the nerve of the Voronoi decomposition of the corresponding union of balls. The cells in this decomposition are convex. This situation is covered by the Nerve Theorem which implies that the union of cells and the nerve of the collection of cells have the same homotopy type [42]. In this application, the union of balls has the same homotopy type as the α-complex, $U(\alpha) \simeq K(\alpha)$, which is somewhat weaker than being homeomorphic. Instead of explaining what exactly this means, we discuss a slightly stronger property, namely that $|K(\alpha)|$ is a *deformation retract* of $U(\alpha)$. In other words, there exists a *deformation retraction*, that is, a continuous map $D : U(\alpha) \times [0,1] \rightarrow U(\alpha)$ whose restriction to $U(\alpha) \times \{0\}$ is the identity on $U(\alpha)$, whose restriction to $|K(\alpha)| \times [0,1]$ is the identity on $|K(\alpha)|$, and whose restriction to $U(\alpha) \times \{1\}$ maps the entire $U(\alpha)$ to $|K(\alpha)|$. Following [17], we sketch how this works for the two-dimensional case illustrated in Figure 5.
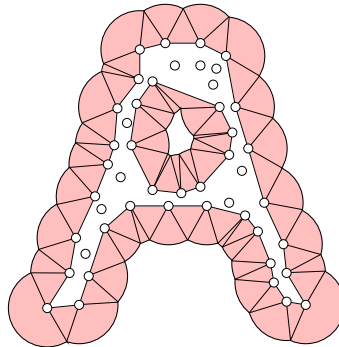

Fig. 5. The decomposition of the union of disks minus the alpha shape into joins. Each join shrinks toward the data points and this way deforms the union into the alpha shape.

First we observe that the α-complex is indeed contained in the union of balls and that the difference can be decomposed into joins of the type $\sigma * s$, where s is a face of the union of balls. In $\mathbb{R}^2$, there are only two types of joins, one defined by a vertex and an arc on a circle and the other by an edge and a point. The first type is a disk sector and the second a triangle. Each join is decomposed into line segments xy with $x \in \sigma$ and $y \in s$. The deformation retraction shrinks the line segments toward x by moving every

point $x_\lambda = (1 - \lambda)x + \lambda y$ to $x_\lambda(t) = (1 - t)x_\lambda + tx$. There are complications because this map is noncontinuous for some points on the boundary of the union but these can be finessed.

Voids and pockets

The deformation retraction does more than deforming the union of balls into the alpha shape, it also deforms the holes of the union into the holes of the shape. For the time being, we only consider two types of holes, *voids* that are the bounded components of the complement, and *pockets* that are not holes at all, at least not to a topologist. Interestingly, this kind of non-hole arises in important everyday phenomena such as ditches in which cars get stuck [7]. As a consequence of $U(\alpha)$ and $K(\alpha)$ having the same homotopy type, they also have the same number of voids. In other words, $\mathbb{R}^d - U(\alpha)$ and $\mathbb{R}^d - |K(\alpha)|$ have the same number of connected components. Moreover, the existence of the deformation retraction implies that each void of $U(\alpha)$ is contained in a unique void of $|K(\alpha)|$ into which is expands during the deformation. We will come back to this topic shortly when we count holes of all dimensions.

As mentioned earlier, pockets are not really holes. They are cavities or depressions near the boundary. The reason we bother in spite of the difficulty to define them topologically is their importance in applications. Indeed, the original motivation for the concept was the interest in proteins and how they interact by partial shape complementarity [21]. In order to produce a well-defined concept, we make use of the filtration instead of considering a single alpha shape. Doing so, we can trace the development of a region in the complement. In a nutshell, a *pocket* is a region in $\mathbb{R}^d - |K(\alpha)|$ that turns into a void before it disappears [22]. To make this more concrete, we consider the vector field, $\mathbf{v} : \mathbb{R}^d \to \mathbb{R}^d$, that assigns to each point x the outward normal of $\partial U(\alpha)$ at x for the value of α at which x does lie on the boundary. This is well-defined in the interior of each Voronoi cell and can be extended to all points by letting $N(x) \subseteq S$ be the collection of closest points and mapping x to $\mathbf{v}(x) = Z(x) - x$, where $Z(x)$ is the center of the smallest enclosing ball of $N(x)$ [43]. We note that in the interior of the Voronoi cells, $2\mathbf{v}(x)$ is the gradient of the squared distance function that maps x to $\varrho_S^2(x) = \min_i \|x - z_i\|^2$. Similarly, in the interior of an intersection of Voronoi cells, $2\mathbf{v}(x)$ is the gradient of $\varrho_S^2(x)$ restricted to that intersection. If $x = Z(x)$ we call x is a *critical point* of ϱ_S^2. We get $d + 1$ types of critical points conveniently indexed from 0 to d, just like in conventional Morse Theory [45]. A point x may flow along this discrete vector field until it ends up at infinity or at a critical point. A *sink*, y, is a critical point of index d. It has a neighborhood in which all points flow toward y. The set of points x whose flow lines ends at y is the *catchment region*, C_y, of the sink. If $C_y - U(\alpha)$ is not yet a void then it has the property that classifies it as a pocket, namely it will become a void before it disappears.

We note that the actual situation is somewhat more complicated but also more interesting than we made it appear. Specifically, the sets $C_y - U(\alpha)$

decompose the complement of the union, except for the outer region whose points flow to infinity. We call a component of this decomposition a pocket. As α increases, the pocket shrinks and breaks up into pieces which are eventually swallowed up by the union. In other words, the pockets are organized hierarchically, which smaller pockets being part of larger pockets. We skip the discussion of how the discrete gradient flow translates into an acyclic relation on the simplices of the Delaunay triangulation and refer instead to [20, 32] where it is used for the purpose of reconstructing the surface from a sampled point set.

Betti numbers

We now expand our interest from voids to holes of all dimensions, basing our approach on *homology groups* [46]. There is one group, H_p, per dimension p. The rank of H_p is called the *p-th Betti number* and can be considered the number of p-dimensional holes of the space. We say "can" because there is an ambiguity what exactly one may consider a p-dimensional hole and turning the situation around we use the homology group to make the concept concrete. We note that even settling on homology groups leaves an ambiguity since they can be defined for different coefficient groups. We find it convenient to use addition modulo 2 since this leads to simple interpretations and fast algorithms.

We now return to the specific case of interest in this paper, namely a finite set of points, S, the Delaunay triangulation, $\mathrm{Del}\,S$, and the filter $\sigma_1, \sigma_2, \ldots, \sigma_m$ of the simplices in $\mathrm{Del}\,S$. As mentioned earlier, we may assume that the faces of a simplex precede the simplex in the filter. If follows that $K_j = \{\sigma_1, \sigma_2, \ldots, \sigma_j\}$ is a complex for every j. Let $\mathsf{H}_p(K_j)$ be the dimension p homology group and $\beta_p(K_j) = \mathrm{rank}\,\mathsf{H}_p(K_j)$ the dimension p Betti number of K_j. We are interested in computing the Betti numbers incrementally. For this, we need to understand how $\beta_p(K_{j+1})$ differs from $\beta_p(K_j)$. There are two cases depending on whether or not σ_{j+1} belongs to a q-dimensional cycle in K_{j+1}, where q is the dimension of σ_{j+1}. For modulo 2 arithmetic, a *p-cycle* is a set of p-simplices whose boundary vanishes, that is, every $(p-1)$-simplex belongs to an even number of p-simplices in the set.

1. If σ_{j+1} belongs to a non-zero q-cycle in K_{j+1} then $\beta_q(K_{j+1}) = \beta_q(K_j) + 1$.
2. Otherwise, $\beta_{q-1}(K_{j+1}) = \beta_{q-1}(K_j) - 1$.

The Betti numbers not affected by the rule remain unchanged. In the first case, we say σ_{j+1} gives *birth* to a q-dimensional homology class. In the second case, we say σ_{j+1} gives *death* to a $(q-1)$-dimensional homology class. Algorithmically, the two cases can be distinguished by reducing the incidence matrices of the complex [46]. For modulo 2 arithmetic, this leads to a cubic time algorithm. For a complex in $\mathbb{R}^3$, there is an alternative method that distinguishes the two cases in almost constant time using the union-find data structure [13]. This leads to an algorithm that computes the Betti numbers

of all m complexes K_j in time proportional to $mA^{-1}(m)$, where A^{-1} is the notoriously slow growing inverse of the Ackermann function [50].

Persistence

In a typical evolution of the Betti numbers from K_0 to K_m, we see cycles come and go in quick sequence. The cycles that stay for a while are the important ones since they capture major shape features of the data set. The author of this paper observed this first for biomolecules, such as the gramicidin protein which is embedded in cell membranes and functions as a channel for ions; see Figure 6. The question is obvious: how can we define the length along
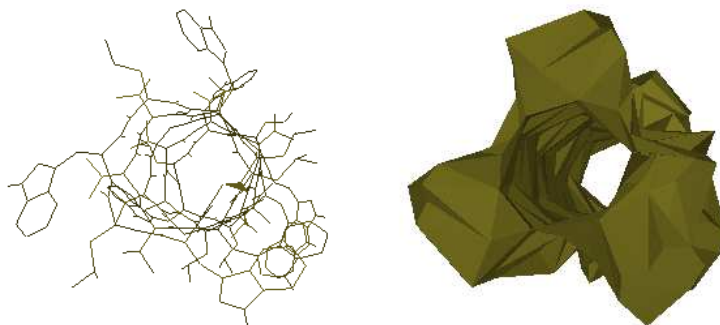

Fig. 6. Two alpha shapes of the gramicidin protein. The tunnel forming the ion channel along the length of the protein is present in both.

which a cycle exists during the filtration. While seemingly obvious, there is a difficulty in the definition because the β_p p-cycles counted by the homology group really generate 2^{β_p} p-cycles and all undergo change as the complex grows. The crucial insight is the pairing of births and deaths that can be done in a canonical fashion [27]. We just need to observe that the inclusion of the complexes $K_i \subseteq K_j$ for $i \leq j$ implies a homomorphic map from $\mathsf{H}_p(K_i)$ to $\mathsf{H}_p(K_j)$ for every p. Indeed, every p-cycle in K_i still exists in K_j although it might now be homologous to others and possibly even trivial. We say a class $c \in \mathsf{H}_p(K_i)$ is *born at* K_i if it is not in the image of the map from $\mathsf{H}_p(K_{i-1})$. Furthermore, this class c *dies entering* K_j if its image in $\mathsf{H}_p(K_j)$ is contained in the image of $\mathsf{H}_p(K_{i-1})$ but its image in $\mathsf{H}_p(K_{j-1})$ is not. The difference between time of birth and time of death is the *persistence* of the class c, $\mathrm{pers}(c) = \alpha_j - \alpha_i$.

While the filtration represents the data at different levels of resolution, the measurement of longevity along the filtration differentiates between more and less important features in the data. Specifically, cycles of very short persistence

may be discarded as noise and suppressed in the analysis of the data. Since our arch-enemy, noise, is ever present, the possibility to filter it out without changing the data is significant. We refer to [23] and [8] for further details.

5 Combinatorics

In this section, we discuss the use of the α-complex to measure the corresponding union of balls. The motivation comes from biochemistry where proteins and other biomolecules are routinely modeled as unions of balls in three dimensions. The method of choice is inclusion-exclusion which we use to compute the volume and surface area as well as their derivatives.

Inclusion-exclusion

Given a collection S of sets in $\mathbb{R}^d$, the d-dimensional volume of the union can be expressed as an alternating sum of volumes of common intersections of subcollections,

$$\mathrm{vol}(\bigcup S) = \sum_{\emptyset \neq \sigma \subseteq S} (-1)^{\dim \sigma} \mathrm{vol}(\bigcap \sigma), \tag{1}$$

where $\dim \sigma = \mathrm{card}\, \sigma - 1$. This notation is convenient as it suggests we think of a collection as an abstract simplex whose dimension is one less than its number of elements. Note that a region contained in $k+1$ of the sets is added $k+1$ times, once for each set, subtracted $\binom{k+1}{2}$ times, once for each pair, added $\binom{k+1}{3}$ times, etc. The equation follows because $\sum_{i=0}^{k}(-1)^i \binom{k+1}{i+1} = 1$. This is known as the *principle of inclusion-exclusion*.

The trouble with the formula is its length which is exponential in the number of sets. For general sets, we cannot do better but in special situations we sometimes have shorter formulas. Call σ *independent* if every face $\tau \subseteq \sigma$ has its own region, that is, $\bigcap \tau - \bigcup(\sigma - \tau) \neq \emptyset$. If all independent subcollections have dimension k or less then there is a formula with integer coefficients whose terms are common intersections of at most $k+1$ sets. To see this, take an abstract simplex of dimension $\dim \sigma \geq k+1$. By assumption, σ is not independent and has therefore faces $\tau \subseteq \sigma$ with $\bigcap \tau \subseteq \bigcup(\sigma - \tau)$. Choosing τ to be maximal with this property implies that each set u in $\sigma - \tau$ has a non-empty intersection with $\bigcap \tau$. If $\tau = \sigma$ then $\bigcap \sigma = \emptyset$ and we drop the corresponding term from (1). Else we apply the principle of inclusion-exclusion to the sets $\bigcap \tau \cap u$, where $u \in \sigma - \tau$. The union of these sets is $\bigcap \tau$, so (1) gives

$$\mathrm{vol}(\bigcap \tau) = \sum_{\emptyset \neq \upsilon \subseteq \sigma - \tau} (-1)^{\dim \upsilon} \mathrm{vol}(\bigcap(\tau \cup \upsilon)).$$

One of the terms on the right hand side is $\mathrm{vol}(\bigcap \sigma)$ and all other terms are volumes of intersections of strictly fewer than $k+1$ sets. Hence, $\mathrm{vol}(\bigcap \sigma)$

can be written as an alternating sum of volumes of common intersections of strictly fewer that $\text{card}\,\sigma$ sets. Starting with the exponential size formula, we can drop terms for non-independent simplices or replace them by integer sums of terms for independent simplices. The claim follows.

The maximum number of independent balls in $\mathbb{R}^d$ is $d+1$ which implies the existence of formulas with terms of at most $d+1$ balls. This has been observed by Kratky in low dimensions [39] and used by Scheraga and collaborators to compute the surface area of proteins [48].

Minimal formulas

Instead of going through a possibly lengthy substitution process, we can use the α-complex to get a short formula directly. To explain how this works, we let S be a finite collection of balls in $\mathbb{R}^d$ and $K = K(0)$ the weighted α-complex for $\alpha = 0$. Recall that the *Euler characteristic* of K is the alternating sum of simplices,

$$\chi(K) = \sum_{\emptyset \neq \sigma \in K} (-1)^{\dim \sigma}.$$

More than in K itself, we are interested in certain subcomplexes of K. Specifically, for each point $x \in \mathbb{R}^d$, we consider the simplices $\sigma \in K$ defined by balls that contain x. The set of such simplices forms a full subcomplex $F(x)$ of K. As it turns out, $F(x)$ is either empty or contractible. In the former case the Euler characteristic vanishes while in the latter case it is one. In other words, $\chi(F(x)) = 0$ if $x \notin \bigcup S$ and $\chi(F(x)) = 1$ if $x \in \bigcup S$. We can therefore integrate and get $\text{vol}(\bigcup S) = \int \chi(F(x))\,\mathrm{d}x$. A simplex $\sigma \in K$ contributes $(-1)^{\dim \sigma}$ for all points $x \in \bigcap \sigma$ and 0 for all other points. Its overall contribution is therefore $(-1)^{\dim \sigma}\text{vol}(\bigcap \sigma)$. Summing over all simplices in the α-complex gives the anticipated result, namely

$$\text{vol}(\bigcup S) = \sum_{\emptyset \neq \sigma \in K} (-1)^{\dim \sigma}\text{vol}(\bigcap \sigma). \tag{2}$$

In words, we get the correct volume if we restrict the inclusion-exclusion formula (1) to subcollections that form simplices in the α-complex, for $\alpha = 0$. A complete proof of (2) and of similar formulas for surface area and other measures of a union of balls can be found in [17]. As proved in [3], (2) is minimal in the sense that no terms can be removed, but there are also other minimal formulas corresponding to other complexes made up of independent simplices. To put the result in perspective, we mention that just truncating (1) to terms of dimension d or less does generally not give the correct volume. Nevertheless, truncating it to the simplices of the Delaunay triangulation of S does give the correct result [47].

Voids

The voids in a protein structure and their sizes are of some interest in bio-
chemistry. We thus modify (2) in such a way that the sum can be decomposed
into geometrically meaningful portions. A key ingredient is the notion of an-
gle inside a d-simplex σ in $\mathbb{R}^d$. Given a sufficiently small ball centered at an
interior point of a face τ of σ, we call the fraction of the ball that lies inside σ
the *angle of σ at τ*, denoted as $\mu_{\sigma,\tau}$. This agrees with the usual notion of angle
except that it is normalized to 1. To make a connection to (2), we consider
a k-simplex τ in the interior of $|K|$. The spheres bounding the $k + 1$ balls
defining τ intersect in a sphere of dimension $d - k - 1$. The d-simplices in the
star of τ define a decomposition of this $(d - k - 1)$-sphere into as many pieces
as there are d-simplices and the $(d - k - 1)$-dimensional volume of the piece
that corresponds to $\sigma \in \mathrm{St}\,\tau$ is $\mu_{\sigma,\tau}$ times the volume of the entire sphere.

We now return to the question of measuring a void in the union of balls,
that is, a bounded component V of $\mathbb{R}^d - \bigcup S$. As suggested by Figure 5,
we compute the volume of the corresponding void in the α-shape, V', and
subtract the volume of the fringe, the part of the union that reaches into it.
However, we decompose the fringe in a way that is different than suggested
in the figure, using angles and inclusion-exclusion. To state the formula we
let $L \subseteq \mathrm{Del}\,S - K$ be the set of simplices that define V'. The d-dimensional
volume of that void is simply the sum of volumes of the d-simplices in L. Some
of the faces τ of a d-simplex $\sigma \in L$ belong to L and others belong to K. We
use the latter to compute the volume of the fringe and get

$$\mathrm{vol}(V) = \mathrm{vol}(V') - \sum_{\sigma} \sum_{\tau} (-1)^{\dim \tau} \mu_{\sigma,\tau} \mathrm{vol}(\bigcap \tau), \qquad (3)$$

where the outer sum is over all d-simplices σ in L and the inner sum is over
all faces τ of σ that belong to K. A formal proof of (3) and of similar formulas
for surface area and other measures of voids can be found in [17].

Derivatives

An important concern in biochemistry is the dynamics of molecules, how they
move and change their shape. A natural approach to this question models the
forces and solves the classical equation of motion for all atoms. Some of the
forces are related to area and volume [26] which motivates our interest in the
derivatives of these measures. The alpha complex is again useful because it
gives ready access to geometric quantities that figure in the expression of the
derivatives. The details for the volume derivative can be found in [25] and
those for the area derivative in [6]. Here we focus on the latter since it is
geometrically more interesting. Letting S be a set of n balls in $\mathbb{R}^3$, we write z_i
for the center and r_i for the radius of the ball B_i, for $1 \leq i \leq n$. In formulating
the derivative, we concatenate the centers to get a vector $\mathbf{z} \in \mathbb{R}^{3n}$ representing
the set S. The surface area of the union is then a map $A : \mathbb{R}^{3n} \to \mathbb{R}$. Its

derivative at the point $\mathbf{z}$ is a linear map $\mathrm{d}A_{\mathbf{z}} : \mathbb{R}^{3n} \to \mathbb{R}$. The *gradient* at a point $\mathbf{z}$ is defined such that $\mathrm{d}A_{\mathbf{z}}(\mathbf{x}) = \langle \mathbf{x}, \nabla A_{\mathbf{z}} \rangle$ for every motion vector $\mathbf{x}$. The derivative, $\mathrm{d}A : \mathbb{R}^{3n} \times \mathbb{R}^{3n} \to \mathbb{R}$, can therefore be represented by $\nabla A : \mathbb{R}^{3n} \to \mathbb{R}^{3n}$ defined by mapping $\mathbf{z}$ to the gradient of A at $\mathbf{z}$. To get a handle on the gradient at $\mathbf{z}$, we decompose the motion vector $\mathbf{x} \in \mathbb{R}^{3n}$ into three-dimensional vectors x_i, where x_i moves the center of the i-th ball to $z_i + x_i$. Since $\mathrm{d}A_{\mathbf{z}}$ is a linear function, we can focus on the infinitesimal change caused by the motion of B_i along x_i and add up the results for different indices i. Similarly, we can focus on the impact of that motion on the interaction of B_i with individual other balls and add up the results to get the infinitesimal change caused by B_i. Here it turns out that we need the interaction of B_i with individual other balls B_j and with pairs of other balls B_j and B_k. The latter is interesting as it induces rotational forces on individual atoms which seem to be consistent with the type of motion typically observed in initial stages of protein folding.

Beginning with the case of two balls, B_i and B_j, we further decompose x_i into a component u_{ij} along the direction defined by the two centers and another component v_{ij} orthogonal to that direction. The magnitude of the contribution of the motion along u_{ij} depends on the circle in which the two spheres bounding B_i and B_j intersect and more specifically the fraction s_{ij} that belongs to the boundary of the union. This fraction can be computed with inclusion-exclusion as explained above. Even though the motion of B_i along v_{ij} keeps the surface area of $B_i \cup B_j$ to first order unchanged, there is a non-zero contribution to the derivative if there in an interaction with a third ball, B_k. To see this, we consider the caps C_{ij} and C_{kj} in which B_i and B_k intersect the sphere bounding B_j. The cap C_{kj} stays of course fixed but C_{ij} moves while its size remains to first order unchanged. The picture is similar to the two-ball case except one dimension lower. Again we decompose v_{ij} into a component u_{ijk} inside the plane spanned by the three centers and another component v_{ijk} orthogonal to that plane. The magnitude of the contribution of the motion along u_{ijk} depends on the line segment connecting the two points at which the circles bounding the two caps intersect and more specifically on the fraction b_{ijk} that forms an edge in the Voronoi decomposition of $\bigcup S$. This fraction can again be computed with inclusion-exclusion. Combining all contributions, we write the i-th coordinate triplet of the gradient as

$$\begin{bmatrix} \mathbf{a}_{3i-2} \\ \mathbf{a}_{3i-1} \\ \mathbf{a}_{3i} \end{bmatrix} = \sum_j (s_{ij} \cdot a_{ij} + \sum_k b_{ijk} \cdot a_{ijk}), \qquad (4)$$

for $1 \leq i \leq n$. Here the sums are over all boundary edges connecting points z_i and z_j in $K = K(0)$ and all triangles connecting this edge to further vertices z_k. The a_{ij} and a_{ijk} are vectors in $\mathbb{R}^3$ that depend on straightforward geometric quantities such as the radii of the balls, the distances between the centers, and the components of the motion vector introduced above. For details we refer to [6].

6 Discussion

In Section 1, we state that the interest in alpha shapes is driven by three application areas: pattern recognition, digital shape sampling and processing, and structural molecular biology. All three require that the theory surveyed in the paper be turned into widely available software. This has indeed been done through a combination of not-for-profit efforts at Universities and commercial software development in industries. There is a place for both since they satisfy complementary needs. As the main contribution to Patton recognition, we see the fast Betti number algorithms and the introduction of persistent homology. The main impact in digital shape sampling and processing has been achieved by the surface reconstruction algorithm based on the idea of discrete flow [20]. The impact in structural molecular biology is based primarily on fast computation of measures and derivatives but we should also mention the construction of surface representations, such as the solvent accessible surface [41], the molecular surface [11], and the molecular skin [18].

The more we know the more we know we don't know. Indeed there are major directions that remain largely unexplored. Work beyond three dimensions has only started and we mention the Čech, Vietoris-Rips, and witness complexes that have been used to study high-dimensional data sets with persistent homology [14]. Probabilistic studies of alpha complexes and persistent homology are also still in their infancy [44]. Finally, there are ideas of creating shape spaces from alpha complexes which have yet unrealized potential [9].

References

1. N. Amenta, S. Choi and G. Rote. Incremental constructions con BRIO. In "Proc. 19th Ann. Sympos. Comput. Geom., 2003", 211–219.
2. D. Attali, J.-D. Boissonnat and A. Lieutier. Complexity of the Delaunay triangulation of points on surfaces: the smooth case. In "Proc. 19th Ann. Sympos. Comput. Geom., 2003", 201–210.
3. D. Attali and H. Edelsbrunner. Inclusion-exclusion formulas from independent complexes. *Discrete Comput. Geom.* **37** (2007), 59–77.
4. F. Aurenhammer. Voronoi diagrams — a study of a fundamental geometric data structure. *ACM Comput. Surveys* **23** (1991), 345–405.
5. A. Bowyer. Computing Dirichlet tessellations. *Computer J.* **24** (1981), 162–166.
6. R. Bryant, H. Edelsbrunner, P. Koehl and M. Levitt. The area derivative of a space-filling diagram. *Discrete Comput. Geom.* **32** (2004), 293–308.
7. R. Casati and A. C. Varzi. *Holes and Other Superficialities.* MIT Press, Cambridge, Massachusetts, 1994.
8. F. Chazal and D. Cohen-Steiner. Geometric inference. This volume, 2008.
9. H.-L. Cheng, H. Edelsbrunner and P. Fu. Shape space from deformation. *Comput. Geom. Theory Appl.* **19** (2001), 191-204.
10. K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry. *Discrete Comput. Geom.* **4** (1989), 387–421.

11. T. H. Connolly. Analytical molecular surface calculation. *J. Appl. Cryst.* **16** (1983), 548–558.
12. B. Delaunay. Sur la sphère vide. *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk* **7** (1934), 793–800.
13. C. J. A. Delfinado and H. Edelsbrunner. An incremental algorithm for Betti numbers of simplicial complexes on the 3-sphere. *Comput. Aided Geom. Design* **12** (1995), 771–784.
14. V. de Silva and G. Carlsson. Topological estimation using witness complexes. *In* "Proc. Sympos. Point-Based Graphics, 2004", 157–166.
15. R. A. Dwyer. Average-case analysis of algorithms for convex hulls and Voronoi diagrams. Ph. D. thesis, Report CMU-CS-88-132, Carnegie-Mellon Univ., Pittsburgh, Pennsylvania, 1988.
16. H. Edelsbrunner. Weighted alpha shapes. Rept. UIUCDCS-R-92-1760, Dept. Comput. Sci., Univ. Illinois at Urbana-Champaign, Illinois, 1992.
17. H. Edelsbrunner. The union of balls and its dual shape. *Discrete Comput. Geom.* **13** (1995), 415–440.
18. H. Edelsbrunner. Deformable smooth surface design. *Discrete Comput. Geom.* **21** (1999), 87–115.
19. H. Edelsbrunner. *Geometry and Topology for Mesh Generation.* Cambridge Univ. Press, England, 2001.
20. H. Edelsbrunner. Surface reconstruction by wrapping finite sets in space. *Discrete and Computational Geometry — The Goodman-Pollack Festschrift*, 379–404, eds. B. Aronov, S. Basu, J. Pach and M. Sharir, Springer-Verlag, Berlin, 2003.
21. H. Edelsbrunner. Biological applications of computational topology. In *Handbook of Discrete and Computational Geometry*, 1395–1412, eds. J. E. Goodman and J. O'Rourke, CRC Press, Boca Raton, Florida, 2004.
22. H. Edelsbrunner, M. Facello and J. Liang. On the definition and the construction of pockets in macromolecules. *Discrete Appl. Math.* **88** (1998), 83–102.
23. H. Edelsbrunner and J. Harer. Persistent homology — a survey. *Surveys on Discrete and Computational Geometry. Twenty Years Later*, eds. J. E. Goodman, J. Pach and R. Pollack, Contemporary Mathematics **453**, 257–282, Amer. Math. Soc., Providence, Rhode Island, 2008.
24. H. Edelsbrunner, D. G. Kirkpatrick and R. Seidel. On the shape of a set of points in the plane. *IEEE Trans. Inform. Theory* **IT-29** (1983), 551–559.
25. H. Edelsbrunner and P. Koehl. The weighted volume derivative of a space-filling diagram. *Proc. Natl. Acad. Sci.* **100** (2003), 2203–2208.
26. H. Edelsbrunner and P. Koehl. The geometry of biomolecular solvation. *Combinatorial and Computational Geometry*, 243–275, eds. J. E. Goodman, J. Pach and E. Welzl, MSRI Publ. **52**, Cambridge Univ. Press, England, 2005.
27. H. Edelsbrunner, D. Letscher and A. Zomorodian. Topological persistence and simplification. *Discrete Comput. Geom.* **28** (2002), 511–533.
28. H. Edelsbrunner and E. P. Mücke. Simulation of simplicity: a technique to cope with degenerate cases in geometric algorithms. *ACM Trans. Graphics* **9** (1990), 66–104.
29. H. Edelsbrunner and E. P. Mücke. Three-dimensional alpha shapes. *ACM Trans. Graphics* **13** (1994), 43–72.
30. H. Edelsbrunner and N. R. Shah. Incremental topological flipping works for regular triangulations. *Algorithmica* **15** (1996), 223–241.

31. I. M. Gelfand, M. M. Kapranov and A. V. Zelevinsky. *Discriminants, Resultants and Multidimensional Determinants.* Birkhäuser, Boston, 1994.
32. J. Giesen and M. John. The flow complex: a data structure for geometric modeling. *In* "Proc. 14th Ann. ACM-SIAM Sympos. Discrete Alg., 2003", 285–294.
33. B. Grünbaum. *Convex Polytopes.* Second edition, Springer-Verlag, New York, 2003.
34. R. A. Jarvis. On the identification of the convex hull of a finite set of points in the plane. *Inform. Process. Lett.* **2** (1973), 18–21.
35. R. A. Jarvis. Computing the shape hull of points in the plane. *In* "Comput. Soc. Conf. Pattern Recognition and Image Processing, 1977", 231–241.
36. B. Joe. Three-dimensional triangulations from local transformations. *SIAM J. Sci. Statist. Comput.* **10** (1989), 718–741.
37. B. Joe. Construction of three-dimensional Delaunay triangulations from local transformations. *Comput. Aided Geom. Design* **8** (1991), 123–142.
38. D. E. Knuth. *Axioms and Hulls.* Springer-Verlag, Heidelberg, Germany, 1992.
39. K. W. Kratky. The area of intersection of n equal circular disks. *J. Phys. A: Math. Gen.* **11** (1978), 1017–1024.
40. C. L. Lawson. Software for C^1 surface interpolation. In *Mathematical Software III*, Academic Press, New York, 161–194.
41. B. Lee and F. M. Richards. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55** (1971), 379-400.
42. J. Leray. Sur la forme des espaces topologiques et sur les points fixes des représentations. *J. Math. Pures Appl.* **24** (1945), 95–167.
43. A. Lieutier. Any open bounded subset of $\mathbb{R}^n$ has the same homotopy type as its medial axis. *In* "Proc. 8th ACM Sympos. Solid Modeling Appl., 2003", 65–75.
44. D. J. Marchette. *Random Graphs for Statistical Pattern Recognition.* John Wiley & Sons, Hoboken, New Jersey, 2004.
45. J. Milnor. *Morse Theory.* Princeton Univ. Press, New Jersey, 1963.
46. J. R. Munkres. *Elements of Algebraic Topology.* Perseus, Cambridge, Massachusetts, 1984.
47. D. Q. Naiman and H. P. Wynn. Inclusion-exclusion-Bonferroni identities and inequalities for discrete tube-like problems via Euler characteristics. *Ann. Statist.* **20** (1992), 43–76.
48. G. Perrot, B. Cheng, K. D. Gibson, J. Vila, A. Palmer, A. Nayeem, B. Maigret and H. A. Scheraga. MSEED: a program for rapid determination of accessible surface areas and their derivatives. *J. Comput. Chem.* **13** (1992), 1–11.
49. R. Seidel. Backwards analysis of randomized geometric algorithms. In *New Trends in Discrete and Computational Geometry*, J. Pach (ed.), Springer-Verlag, Berlin, 1993, 37–67.
50. R. E. Tarjan. *Data Structures and Network Algorithms.* SIAM, Philadelphia, Pennsylvania, 1983.
51. G. Voronoi. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. *J. Reine Angew. Math.* **133** (1907), 97–178 and **134** (1908), 198–287.
52. D. F. Watson. Computing the n-dimensional Delaunay tessellation with application to Voronoi polytopes. *Computer J.* **24** (1981), 167–172.
53. G. M. Ziegler. *Lectures on Polytopes.* Springer-Verlag, New York, 1995.