Cornell 2014 Honors Thesis: Extremal 1-Cycle Lifespan Behavior under the Vietoris-Rips Complex

Walter Cai

April 30, 2014

Abstract

Given a finite discrete point set $X \subset \mathbb{R}^n$ and distance parameter δ, we may generate the Vietoris-Rips Complex $R_\delta(X)$; the simplicial complex where a simplex $[x_i]_{i \in I} \in R_\delta(X)$ if $|x_i - x_j| \leq \delta$ for all $i, j \in I$. Given arbitrary 1-cycle σ appearing in the Rips Complex, we may define the the birth and death of σ as $\alpha = \min(\delta : \sigma \in C_1(R_\delta(X)))$ and $\gamma = \min(\delta : \sigma \equiv 0 \in H_1(R_\delta(X)))$ respectively.

We seek to maximize γ with respect to α. In order to do so we consider a specific class of finite point sets $X = \{x_i\}_{0 \leq i \leq n}$ where the cycle in question is $\sigma = \sum_{i=0}^{n-1} [x_i, x_{i+1}] + [x_n, x_0]$. We conjecture a possible optimal configuration occurring when the vertices of X are equally spaced on the circle. Our algorithms focus on the manipulation of the original X, incrementally increasing γ while maintaining a constant α. These processes operate by finding subsets of X and through reflection and angular splitting, spread out the vertices in the pursuit of a more circular distribution.

1 Work and Definitions up to the end of DIMACS REU 2012

1.1 Background Material

1.1.1 Simplices

We will begin with an introduction to *Simplices* following the construction by Chazal.

A simplex can be thought of from a geometric point of view as the most basic linear building block in a given dimension. In zero dimensions we consider points; in one dimensions we consider line segments; in two dimensions we consider triangles; in three dimensions we consider tetrahedrons; etc.

An abstract approach to defining a k-simplex begins with taking a set of vertices in n-space.

$$X = \{v_0, v_1, v_2, \dots, v_k\} \subset \mathbb{R}^n$$

This set X is affinely independent if:

$$\left(\sum_{i=0}^{k} \lambda_i v_i = 0, \quad \sum_{i=0}^{k} \lambda_i = 0 \right) \Rightarrow \lambda_i \equiv 0.$$

Definition 1.1. Such a set X defines a k-simplex: $\sigma = [v_0, v_1, \dots, v_k]$, i.e. a simplex of dimension k.

$$\begin{aligned} \sigma &= [v_0, v_1, \dots, v_k] \\ &= \left\{ \sum_{i=0}^{k} t_i v_i \middle| t_i \geq 0, \sum t_i = 1 \right\} \end{aligned}$$

Definition 1.2. We will define a *Face* of the simplex $[v_0, \dots, v_k]$ to be a simplex generated by a subset of $\{v_0, \dots, v_k\}$. Example:

$$\sigma = [\hat{v}_0, v_1, \dots, v_l, \hat{v}_{l+1}, \dots, v_k]$$

Here, the notation $\hat{}$ signifies an omitted term.

1.1.2 Simplicial Complexes

We will continue our introduction with the construction of an abstract *Simplicial Complex*.

Definition 1.3. First, define the *Power Set* to be the set of all subsets of a set X.

$$2^X = \{A \,|A \subset X\}$$

Definition 1.4. We will define a simplicial complex S on a finite set of vertices X to be a set of subsets of X. An equivalent definition would be a subset 2^X. Moreover, for every simplex $\sigma \in S$, every face of σ must also be in S.

Define the dimension of a simplicial complex as the maximal dimension over all simplices in the complex.

Definition 1.5. Define the *Geometric Realization* of a simplicial complex, $\Gamma(S)$ to be the union of simplices defined on elements of S.

$$\Gamma(S) = \bigcup_i [v_{i_0}, \dots, v_{i_l}] \,|\{v_{i_0}, \dots, v_{i_l}\} \in S$$

$$\Gamma(S) \subset \mathbb{R}^n$$

Visual example of the geometric realization of a simplicial complex:

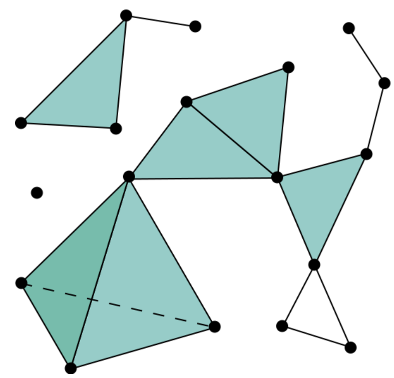

Figure 1: Visual illustration of a Simplicial Comples

1.1.3 Vietoris-Rips Complex

Definition 1.6. Let $X = \{v_0, v_1, \ldots, v_k\} \subset \mathbb{R}^n$ be a set of vertices. We can define the *Vietoris-Rips Complex* at value $\epsilon \geq 0$; $V_\epsilon(X)$, to be the simplicial complex defined on the set X given the following condition:

$$\forall v_a, v_b \in \{v_{i_0}, v_{i_1}, \ldots, v_{i_l}\} \subset X, \quad d(v_a, v_b) \leq \epsilon$$
$$\Rightarrow [v_{i_0}, v_{i_1}, \ldots, v_{i_l}] \in V_\epsilon(X).$$

1.1.4 Space of k-Chains

We will construct k-chains following Chazal.

Definition 1.7. Consider a simplicial complex, S, of dimension d. For $k \in \{0, 1, 2, \ldots, d\}$ we can define the *k-Skeleton* of the simplicial complex, S^k, to be the set of j-dimensional simplices in the complex for $j \leq k$. Thus, the k-skeleton is a simplicial complex on its own.

Definition 1.8. For each k, we can define the set of k-dimensional faces of S to be the finite set $F_k(S)$, where each $\sigma_i \in F_k(S)$ is a distinct k-simplex appearing in the simplicial complex.

Definition 1.9. Define a the *Space of k-Chains* as follows:

$$C_k(S) = \left\{ \sum_{i=0}^{m} \gamma_i \sigma_i \quad | \quad \sigma_i \in F_k(S), \gamma_i \in \mathbb{Z}/2 \right\}.$$

Define summing and scalar multiplication of k-chains as follows:

$$c = \sum_{i=0}^{m} \gamma_i \sigma_i, \quad c' = \sum_{i=0}^{m} \gamma_i' \sigma_i, \quad \lambda \in \mathbb{Z}/2$$

$$c + c' = \sum_{i=0}^{m} (\gamma_i + \gamma_i') \sigma_i \quad \lambda c = \sum_{i=0}^{m} \lambda \gamma_i \sigma_i.$$

Therefore, the space of k-chains, $C_k(S)$, is a $\mathbb{Z}_2$ vector space.

1.1.5 Boundary Function

We will define a function mapping the space of k chains to the space of $k-1$ chains following Chazal.

Definition 1.10. Define the *Boundary Function*; ∂, of a k-simplex to be the sum of all $k-1$ dimensional faces:

$$\partial : C_k(S) \to C_{k-1}(S)$$

$$\partial[v_0, v_1, v_2, \ldots, v_k] = \sum_{i=0}^{k} [v_0, \ldots, \hat{v}_i, \ldots, v_k]$$

Definition 1.11. If we consider a simplicial complex of dimension k; S, we can describe the behavior of ∂ using a *Chain Complex*:

$$0 \xrightarrow{\partial_{k+1}} C_k(S) \xrightarrow{\partial_k} C_{k-1}(S) \xrightarrow{\partial_{k-1}} \cdots \xrightarrow{\partial_2} C_1(S) \xrightarrow{\partial_1} C_0(S) \xrightarrow{\partial_0} 0.$$

A useful property of the boundary function is that when twice applied the result is always 0.

$$\begin{aligned}
\partial \circ \partial(\sigma) &= \partial \circ \partial[v_0, v_1, v_2, \ldots, v_k] \\
&= \partial \left(\sum_{i=0}^{k} [v_0, \ldots, \hat{v}_i, \ldots, v_k] \right) \\
&= \sum_{i=0}^{k} \left(\partial[v_0, \ldots, \hat{v}_i, \ldots, v_k] \right) \\
&= \sum_{i=0}^{k} \left(\sum_{j=1}^{i-1} [v_0, \ldots, \hat{v}_j, \ldots, \hat{v}_i, \ldots, v_k] + \sum_{j=i+1}^{k} \partial[v_0, \ldots, \hat{v}_i, \ldots, \hat{v}_j, \ldots, v_k] \right) \\
& \text{Thus, the term that omits } v_a, v_b \text{ will appear twice (when } (i,j) = (a,b),(b,a)). \\
& \text{By the binary property of the } k\text{-chains:} \\
&= 0
\end{aligned}$$

Definition 1.12. Define the image and kernel of boundary function as follows:

$$\begin{aligned}
B_k &= \partial(C_{k+1}(S)) = \operatorname{im}(\partial_{k+1}) \\
Z_k &= \ker(\partial_k).
\end{aligned}$$

Because the composition of the boundary function always maps to 0, We know that the image of the boundary function on the $(k+1)$-chain space; B_k, is always in the nullspace of the boundary function on the k-chain space; Z_k: $B_k \subset Z_k$.

1.1.6 Betti Numbers and Cycles

We can now define a method of determining topological properties of a given simplicial complex by computing the *Homology* of a complex, S following Chazal.

Definition 1.13. Define a *Cycle*, c as a k-chain in Z_k. This indicates that each boundary element is included an even number of times by elements of c suggesting some manner of k dimensional loop. For example, a 1-cycle could be a completed loop of 1-simplices inside S.

Definition 1.14. We say that two cycles, $c, c' \in Z_k$ are *Homologous* if $c + c' \in B_k$. This implies that c and c' are in the same *Homology Class*; a set of homologous cycles. Intuitively, this suggests that the difference between two homologous cycles is a section of the complex that is filled in with higher dimensional simplices.

Definition 1.15. Define the k^{th} *Homology Group* of a complex to be the quotient space of the image and kernel:

$$H_k = Z_k / B_k$$

We can interpret elements of H_k as distinct homology classes of k-cycles, which in turn can be thought of as the space of distinct $k+1$ dimensional holes in the complex.

Definition 1.16. Define the *Betti Numbers* of a complex, S, as the dimension of the homology groups:

$$\beta_k(S) = \dim(H_k) = \dim Z_k - \dim B_k$$

This produces a sequence β_i that describes the number of non-homologous cycles. β_0 describes the number of disconnected components in S. β_1 describes the number of 2-dimensional holes in S. β_2 describes the number of 3-dimensional holes, or completely walled off hollow area, they are often referred to as *Voids*.

Example: Consider a non-hollow donut: it is all connected so $\beta_0 = 1$. There exists a single hole, so $\beta_1 = 1$. Lastly, there are no hollow sections of the torus, so $\beta_2 = 0$.

In contrast, if we consider a hollow 3-dimensional ball simplex, a complex homeomorphic to the 2-sphere. We would have $(\beta_0, \beta_1, \beta_2) = (1, 0, 1)$.

1.2 Hypothesis

1.2.1 Conditions

Consider a finite set of vertices $X = \{v_0, v_1, v_2, \ldots, v_k\} \subset \mathbb{R}^n$ displaying the following characteristics:

- The Vietoris-Rips Complex begins displaying a 1-cycle at $\epsilon = \alpha$ and not before:
$$\beta_1(V_\epsilon(X)) = 0 \text{ for } \epsilon < \alpha, \quad \beta_1(V_\alpha(X)) = 1.$$
- For every $v_i \in X$, there exist precisely two distinct elements $v_j, v_k \in X, v_j, v_k \neq v_i$ such that $d(v_i, v_j), d(v_i, v_k) \leq \alpha$.
- At $\epsilon = \alpha$ there exists only one component:
$$\beta_0(V_\alpha(X)) = 1.$$
Note: these three conditions necessitate that the $V_\alpha(X)$ is simply a loop of 1-simplices that is homeomorphic to S^1.

Definition 1.17. Define α as the *Birth*-value of the 1-cycle.

Definition 1.18. Define the *Death*-value of the 1-cycle as:

$$\gamma = \min(\epsilon > \alpha : \beta_1(V_\epsilon(X)) = 0).$$

The death-value should be interpreted as the smallest value of ϵ when the 1-cycle born at α is filled in with 2-simplices.

Definition 1.19. Define the *Lifespan* of the 1-cycle to be $l = \gamma - \alpha$.

Definition 1.20. Define S_k^1 to be set of vertices resulting from the discretization of the 1 dimensional circle S^1 into k points where each point is precisely distance α away from its two neighbors. Also, let the death-value of S_k^1 be the value γ_k.

1.2.2 Claim:

Given a set of vertices X that satisfies the above properties, we will have the following result:

$$\gamma \leq \gamma_{k+1}.$$

1.2.3 Lifespan of the Discretized Circle

The objective of this section is to find a formula for γ_k in terms of α and k.

Recall the definition of $S^1_k = \{v_0, \ldots, v_{k-1}\}$ is by taking a subset of $k \geq 3$ vertices equally spaced apart on a circle. Enumerate the indices of the vertices to reflect order along the circle. The euclidean distance between any point, and its neighbors $d(v_i, v_{i\pm 1}) = \alpha, \quad i \in \mathbb{Z}/k$. Thus, we know that the Vietoris-Rips complex for S^1_k will display a 1-cycle beginning at $\epsilon = \alpha$ birth-value.

Since any 2-simplex with vertices in S^1_k must lie in the plane spanned by S^1_k we know that for the 1-cycle to disappear, 2-simplices must eventually contain the center of the discretized circle, y.

Definition 1.21. Consider the 2-simplex, Δ, with vertices $v_a, v_b, v_c \in S^1_k \quad (a < b < c)$ and boundary 1-simplices $[v_a, v_b], [v_b, v_c], [v_c, v_a]$. Define the gap function as follows:

$$\begin{aligned} g(\Delta) \quad &: \quad \{[v_a, v_b, v_c] \mid v_i \in S^1_k\} \to \mathbb{Z}/k \oplus \mathbb{Z}/k \oplus \mathbb{Z}/k \\ g(\Delta) \quad &= \quad (g_{a,b}(\Delta), g_{b,c}(\Delta), g_{c,a}(\Delta)) \\ g([v_a, v_b, v_c]) \quad &= \quad (b-a, c-b, k-(c-a)). \end{aligned}$$

The gap function returns the number of gaps between the two vertices associated to a specific $[v_i, v_j] \subset \Delta$. Note also for a given Δ:

$$\sum g_{i,j}(\Delta) = k.$$

Definition 1.22. Define the function $l([v_i, v_j])$ to be the euclidean length of the 1-simplex, $[v_i, v_j]$.

We can claim that y will be in Δ if and only if $\nexists [v_i, v_j] \subset \Delta$ such that $g_{i,j}(\Delta) > k/2$. This condition is both necessary and sufficient since it guarantees that y will not appear on the exterior of Δ relative to any of the boundary 1-simplices. Define the set of such Δ to be Y. Thus, since $g_{i,j}(\Delta) \leq k/2 \quad \forall a, b, c$ (not necessarily ordered), we have: $g_{a,b}(\Delta) \leq g_{b,c}(\Delta) \Leftrightarrow l([v_a, v_b]) \leq l([v_b, v_c])$.

A 2-simplex $\Delta = [v_a, v_b, v_c] \in Y$ will appear in the Vietoris-Rips complex when $\epsilon \geq \max l([v_i, v_j])$. Thus, the earliest such possible 2-simplex will appear at:

$$\epsilon = \min_{\Delta \in Y} \left(\max l([v_i, v_j])\right).$$

Due to the equivalence of the g and l functions described above, this will occur when the following is achieved:

$$\min_{\Delta \in Y} \left(\max g_{i,j}(\Delta)\right)$$

We know for all Δ that $\sum g_{i,j}(\Delta) = k$. Therefore, $\max g_{i,j}(\Delta) \geq k/3$. Therefore;

$$\min_{\Delta \in Y} \left(\max g_{i,j}(\Delta)\right) \geq \frac{\sum g_{i,j}(\Delta)}{3} = \frac{k}{3}.$$

However, $g_{i,j}(\Delta)$ lives in $\mathbb{Z}/k$. Therefore;

$$\min_{\Delta \in Y} \left(\max g_{i,j}(\Delta)\right) \geq \lceil k/3 \rceil.$$

This possibility can always be achieved using: $a = 0, b = \lceil k/3 \rceil, c = \lceil 2k/3 \rceil$. For these specific a, b, c we can construct a set of $k-3$ many 2-simplices;

$$B = \{\hat{\sigma}_0, \ldots, \hat{\sigma}_{\lceil k/3 \rceil}, \ldots, \hat{\sigma}_{\lceil 2k/3 \rceil}, \ldots, \sigma_{k-1}\}.$$

This set will satisfy:

$$\partial \left(\left[v_0, v_{\lceil k/3 \rceil}, v_{\lceil 2k/3 \rceil}\right] + \sum_{\sigma_i \in B} \sigma_i \right) = \sum_{i=0}^{k-1} [v_i, v_{i+1}].$$

We will construct σ_i by describing its three boundary 1-simplices. When constructing σ_i, for $i \in \{1, \ldots, \lceil k/3 \rceil - 1\}$ take $l = 0$, for $i \in \{\lceil k/3 \rceil + 1, \lceil 2k/3 \rceil - 1\}$ take $l = \lceil k/3 \rceil$, for $i \in \{\lceil 2k/3 \rceil + 1, k-1\}$ take $l = \lceil 2k/3 \rceil$. Define the three boundary of σ_i as follows:

$$\partial \sigma_i = [v_i, v_{i+1}] + [v_l, v_i] + [v_l, v_{i+1}]$$

Because of the geometry of the S^1_k, we know that these 2-simplices must appear in the vietoris-rips complex at a smaller value of ϵ then that value necessary for the appearance of the $\left[v_0, v_{\lceil k/3 \rceil}, v_{\lceil 2k/3 \rceil}\right]$ simplex. Thus, through this construction, in the equation $\partial \left(\left[v_0, v_{\lceil k/3 \rceil}, v_{\lceil 2k/3 \rceil}\right] + \sum \sigma_i \right)$, every interior 1-simplex will appear twice while the 1-simplices between neighboring vertices will only appear once: the 1-cycle will have been filled in.

Therefore, the first possible Δ that contains y will have a longest edge $[v_i, v_j]$, where $g_{i,j}(\Delta) = \lceil k/3 \rceil$. Basic geometry provides:

$$l([v_i, v_j]) = \frac{\sin\left(\frac{\lceil k/3 \rceil}{k} \cdot \pi\right) \cdot \alpha}{\sin\left(\frac{\pi}{k}\right)}.$$

1.3 Partial Results

1.3.1 General 1-Cycles to Convex 1-Cycles

The purpose of this section is to describe and verify a method to show that maximal lifespan of a 1-cycle in $\mathbb{R}^2$ can be done by considering only convex cycles.

Let $X = \{v_0, \ldots, v_k\} \subset \mathbb{R}^2$, and let X satisfy the conditions outlined in section 2.1. It is our goal to manipulate the vertices through repetition of the same process that will eventually result in a planar set of vertices, Y, for which $\Gamma(V_\alpha(Y))$ displays a convex polygon with the same birth value, α, and a death value γ that has only increased.

Definition 1.23. Note: we will abuse terminology by saying a set of vertices X displaying the conditions in section 2.1. is *Convex* if the geometric realization of the rips complex at value α: $\Gamma(V_\alpha(X))$ is a convex polygon.

The process begins by assuming that $\Gamma(V_\alpha(X))$ displays a non-convex polygon. It is an equivalent statement then that we are able to choose two non-neighboring vertices $v_a, v_b \in X; a \neq b \pm 1$, such that the line intersecting both vertices, L, is not contained within $\Gamma(V_\alpha(X))$ and all other vertices v_i are on "one side" of the L. Note: it is possible there exists $v_i \in L : l \neq a, b$, however this would be a trivial example that would non be affected by later steps of the process.

We also know that since v_i, v_j are non-neighboring, they split $\Gamma(V_\alpha(X))$ into two pieces, both of which containing their own set of vertices $v_i \in X_1$, and $v_j \in X_2$. Note that we will set $v_a, v_b \notin X_1, X_2$. We will then reflect all points from one side of X, WLOG choose X_1, across L. Call the new set of vertices $X' = \{v'_0, \ldots, v'_k\}$.

After reflecting, pairwise distance between any two neighbors has remained unchanged:

$$d(v_i, v_{i+1}) = d(v'_i, v'_{i+1}).$$

Moreover, we also know that pairwise distance between any two points has only increased. Let $v_i, v_j \in X_1$, we will have $d(v_i, v_j) = d(v'_i, v'_j)$. Let $v_i \in X_1$. Let $v_i, v_j \in X_2$, we will have $d(v_i, v_j) = d(v'_i, v'_j)$. Let $v_i \in X_1$; $d(v_i, v_a) = d(v'_i, v'_a)$,$d(v_i, v_b) = d(v'_i, v'_b)$. Let $v_i \in X_1, v_j \in X_2$. Consider the line segment between between v'_i and v_j, let it be H. let $h = H \cap L$ we have $d(h, v_i) = d(h, v'_i)$. Therefore, by the triangle inequality:

$$\begin{aligned} d(v'_i, v'_j) &= d(v'_i, v_j) \\ &= d(v'_i, h) + d(h, v_j) \\ &= d(v_i, h) + d(h, v_j) \\ &\geq d(v_i, v_j). \end{aligned}$$

Thus, pairwise distance has only ever increased between vertices after the reflection process. Thus, we know that $V_\alpha(X')$ displays a 1-cycle with birth value α.

Define 2-simplices as follows: $A = [v_i, v_j, v_k]$, and let $A' = [v'_i, v'_j, v'_k]$. Define $\delta = \min(d(v_i, v_j), d(v_j, v_l), d(v_l, v_i))$, and $\delta' = \min(d(v'_i, v'_j), d(v'_j, v'_l), d(v'_l, v'_i))$. Because pairwise distance has only increased we know $\delta' \leq \delta$. Therefore, any 2-simplex will take longer to appear in the Vietoris-Rips Complex of X' than for X.

Because every vertex in X' corresponds to a vertex in X and vice versa, we know that any 2-chain in $V_\epsilon(X)$ corresponds to a 2-chain in $V_\epsilon(X)$. In order for a 1-cycle in X, X' to be filled in, there must exist a 2-chain, c, c', respectively, s.t.:

$$\partial c = \sum v_i \quad \partial c' = \sum v'_i.$$

Because of the correspondence between vertices in X and X', we know that if $\partial c = \sum v_i$, then $\partial c' = \sum v'_i$, and vice-versa. However, we know that the necessary ϵ for c to appear $V_\epsilon(X)$ is less than or equal to the ϵ' necessary for c' to appear in $V'_\epsilon(X')$. Therefore, we know that the death value of $V'_\epsilon(X')$; $\gamma' \geq \gamma$.

The strategy is that through multiple iterations of this process we will achieve a set of vertices Y where $\Gamma(V_\alpha(Y))$ is a convex k-polygon. By the Jordan Curve Theorem, we will be able to define a specific interior and exterior section of the plane. Thus, through every iteration the interior area of the curve will only have increased. Since the arc length of $\Gamma(V_\alpha(X'))$ remains constant, by the isoperimetric inequality, we know that the area inside is bounded. Therefore, the area of the subsequent iterations of the reflecting process forms a monotonic bounded sequence, which must have a limit. This limit corresponds to the area of the curve on which we can no longer run the reflecting process and therefore, must be convex. Thus, maximal lifespan can be achieved by considering only sets X where $\Gamma(V_\alpha(X))$ is a convex k-polygon.

1.3.2 Convex 1-Cycles to Uniform Convex 1-Cycles

We have shown that maximal lifespan can be achieved by considering only convex 1-cycles; The purpose of this section is to show that we can further specify by considering only regular convex 1-cycles.

Consider X satisfying the conditions in 2.1., we will say that $V_\alpha(X)$ is regular if for any $v_i \in X$, $d(v_i, v_{i+1}) = \alpha$.

Let X be an irregular convex. This implies that in the k-polygon $\Gamma(V_\alpha(X))$ there exists v_i, v_{i+1} where $d(v_i, v_{i+1}) < \alpha$.

Pick another vertex away from this 1-simplex: choose v_j where $j = i + \lceil k/2 \rceil$. We know that the outer angle at v_j, θ_j must be nonnegative due to the convexity of $\Gamma(V_\alpha(X))$. We can bend the complex at that vertex until either $d(v_i, v_{i+1}) = \alpha$ or $\theta_j = 0$. If $\theta_j = 0$ happens first move on to the vertices $v_{j\pm1} \neq v_i, v_{i+1}$, and continue to bend the complex until $d(v_i, v_{i+1}) = \alpha$ or $\theta_{j\pm1} = 0$. If $\theta_{j\pm1} = 0$ first then move on to $v_{j\pm2}$ and repeat.

note through every iteration of this process, i.e. every time a vertex is bent out so the outer angle is 0, that pairwise distance between any two points can only increase. This can be seen through simple triangle inequalities between the two points and the most recent bent out vertices. This is due to the fact that an outer-angle at a vertex is never bent to the point where it is negative, just to 0 and then left alone, thus always maximizing distance. Therefore, by the same logic put forth in section 3.1. the death value of the 1-cycle can only increase.

It is also impossible for the complex to run out of vertices at which to bend since it would imply that $d(v_i, v_{i+1}) > \alpha$. Therefore, this process can be run for every 1-simplex in $V_\alpha(X)$ producing a regular convex 1-cycle with a larger death value and birth value α. Therefore, maximal lifespan on a planar set can be achieved by considering only uniform convex sets.

2 Results Following the Conclusion of DIMACS REU 2012

2.1 Strategy to Find Non-neighboring Pair of Vertices in n-Dimensional Space

Consider a finite set of vertices $X \subset \mathbb{R}^n$, $n \geq 3$ where $\Gamma(V_\alpha(X))$ is homeomorphic to $\mathbb{S}^1$.

We wish to choose a pair of non-neighboring vertices $v_i, v_j \in X$ s.t. we can draw an $n-1$ dimensional hyperplane H, where $v_i, v_j \in H$ and that $\forall v_l \in X$, v_l is on one side of H. We can define sides of H since the n-dimensional space is divided into two sections by H.

Definition 2.1. A hyperplane in $\mathbb{R}^n$ is a solution set $\{x \in \mathbb{R}^n \mid a \cdot x + c = 0\}$ where $a \in \mathbb{R}^n \setminus \{0\}$, $c \in \mathbb{R}$, and $\cdot$ is the dot product, or inner product on $\mathbb{R}^n$.

Claim 2.1. *Let X be a finite subset of $\mathbb{R}^n$ so that $n \geq 3$, and $\Gamma(V_\alpha(X)) \cong \mathbb{S}^1$. Suppose further that X does not lie in any hyperplane. Then there exists a hyperplane H given by $a \in \mathbb{R}^n \setminus \{0\}$, and $c \in \mathbb{R}$ so that:*

1. *if $x \in X$, then $x \cdot a + c \geq 0$;*
2. *there exist vertices $v_i, v_j \in X$ which are not neighbors, and $v_i, v_j \in H$.*

Let Y be the the convex hull of X, i.e. the smallest convex subset of $\mathbb{R}^n$ s.t. $X \subset Y$. Because X is a finite point set, we know Y will be a convex polytope living in $\mathbb{R}^n$;

$$S = \left\{ x = \sum_{k=1}^{n} \alpha_k x_k \in \mathbb{R}^m \,\middle|\, \sum_{k=1}^{n} \alpha_k = 1, \alpha_k \geq 0 \right\}. \qquad \text{(Convex Hull)}$$

Claim 2.2. *Let $v_l \in X \cap \partial Y$ where ∂Y denotes the boundary of Y. Consider the collection $E_l = \{[v_l, v_j] \mid v_j \in \partial Y, [v_l, v_j] \subseteq \partial Y\}$. Then E_l contains more than two elements.*

Proof. Consider any $v_l \in X \cap \partial Y$. Then the set $\{v_j - v_l \mid [v_l, v_j] \in E_l\}$ spans $\mathbb{R}^n$. This follows because $X \subseteq \text{Span}\{v_j - v_l \mid [v_l, v_j] \in E_l\}$, and we assume that X does not lie in any hyperplane. □

We know that in the original complex $V_\alpha(X)$, the vertex v_l, has precisely two neighbors, therefore we have connected v_l with $n > 2$ different vertices with outer edges in ∂Y. Hence, it must be possible to choose another vertex, $v_m \in X \cap \partial Y$ which does not neighbor v_l. Then $[v_l, v_m] \subseteq \partial Y$.

Therefore, we may always choose two non-neigboring vertices in X and an $n-1$-dimensional hyperplane H where all of X is on one side of H. Our object is to run an analogous reflecting algorithm as used in the 2-dimensional case and again applying the triangle inequality realize the lifespan can only increase in the newer version of the version of X.

2.2 Strategy to Eliminate Dimensions

Consider a subset $\{x_1, \dots, x_n\} = X \subset \mathbb{R}^m, m \geq 3$ s.t. $\text{Span}_{\mathbb{R}}(X) = \mathbb{R}^m$ that satisfy the necessary conditions to form a 1-cycle homeomorphic to S^1 under the Vietoris-Rips complex at radius α. Moreover, the indices of the vertcies reflect the order of the vertices in the 1-cycle.

Definition 2.2. Let ρ_H be the natural surjection $\rho_H : \mathbb{R}^m \to \mathbb{R}^m/H$ for some subspace $H \subset \mathbb{R}^m$.

Definition 2.3. Let H be an $m-1$ dimensional subspace, $\mathbb{R}^m/H$ is isomorphic to $\mathbb{R}$.
Define the vector space isomorphism $f_H : \mathbb{R}^m/H \to \mathbb{R}$.

Definition 2.4. Define $\phi_H = f_H \circ \rho_H : \mathbb{R}^m \to \mathbb{R}$.

Lemma 2.1. *There exist x_i, x_j, $j \neq i \pm 1 \mod n$ and $m-1$ dimensional subspace $H \subset \mathbb{R}^m$ where $\rho_H(x_i) = \rho_H(x_j)$ and for all $x \in X$, $\phi_H(x) \leq \phi_H(x_i)$ or $\phi_H(x) \geq \phi_H(x_i)$. Furthermore, there exists at least one $y \in X$ s.t. $\phi_H(y) \neq \phi_H(x_i)$.*

Visually, this posits that if the set of points from which the 1-cycle is constructed spans all of $\mathbb{R}^m$ then there exist non-neighboring vertices s.t. there exists an $m-1$ dimensional hyperplane containing the vertices where all other elements of X do not appear on one side of the hyperplane.

Lemma 1.1. Again, consider the convex hull of X:

$$S = \left\{ x = \sum_{k=1}^{n} \alpha_k x_k \in \mathbb{R}^m \;\middle|\; \sum_{k=1}^{n} \alpha_k = 1, \alpha_k \geq 0 \right\}. \quad \text{(Convex Hull)}$$

This set must be a polytope in $\mathbb{R}^m$. Since $X \subset S$, $\text{Span}_{\mathbb{R}}(X) = \mathbb{R}^m \subset \text{Span}_{\mathbb{R}}(S) \subset \mathbb{R}^m \implies \text{Span}_{\mathbb{R}}(S) = \mathbb{R}^m$. Hence, ∂S must be tiled with $m-1$ dimensional hyperplanes. Since at least three elements of X are required to define an $m-1$ hyperplane, call them $\{y_1, y_2, y_3, \dots\} \subset X$ there must exist a pair that are non-neighbors.

This is easily justified: Consider first y_1, y_2, if they are non-neighbors, then we are done. If they are neighbors, consider instead y_1, y_3. Again, if they are non-neighbors we are done, if not then consider y_2, y_3. Since any element of X has at most two neighbors we know y_2 has neighbors y_1, y_3. Hence if y_1, y_3 are neighbors then X has only three elements: $X = \{y_1, y_2, y_3\}$. This case is trivial since $\text{Span}_{\mathbb{R}}(X) \subset \mathbb{R}^2$. Hence there must exist a non-neighboring pair for any of the $m-1$ dimensional faces in ∂S. □

Definition 2.5. For $x_{k_1}, x_{k_2} \in X$, assume $k_1 < k_2$, define the two **arms** of X relative to x_{k_1}, x_{k_2} by sets:

$$\{x_k \in X : k_1 < k < k_2\}, \quad \{x_k \in X : k < k_1, k > k_2\}$$

Every hyperplane, P, in ∂S must contain a pair of non-neighbors from X. These two points define two arms, and since X cannot be a subset of P , there must exist an arm containing at least one point outside of P.

Our strategy is to reflect this arm across P, then repeat the above process with the newer version of X.

- Pairwise distance between neighbors remains constant, hence α never changes.
- Pairwise distance between elements of the same arm never change.
- Pairwise distance between elements of different arms can only increase.

Proof. Take $x, y \in X$ in different arms, where x is reflected to x' across a hyperplane P. Consider the line segment, l, connecting x', y. Define z as the intersection of l and P:

$$\begin{aligned} |x' - y| &= |x' - z| + |z - y| \\ &= |x - z| + |z - y| \\ &\geq |x - z| \end{aligned}$$

□

Hence, pairwise distance between between any two points can only increase and therefore, by construction, the lifespan of the newer versions of X can only increase while the α value remains constant. We will refer to this as the **Reflecting Process** with respect to P, x_i, x_j. We will denote the reflecting process with "RP" for brevity. Note that the choice of arm relative to P, x_i, x_j is insignificant since either choice of arm will result in the same loop.

Definition 2.6. For any 1-cycle derived from finite subset $X \subset \mathbb{R}^m$, define the total pairwise distance function;

$$D(X) = \sum_{x,y \in X} |x - y|$$

Suppose the RP with respect to hyperplane P, and points $x_i, x_j \in X$ yields the new collection of points $X' \subset \mathbb{R}^m$.

For any P, x_i, x_j, since pairwise distance can only increase, $D(X') \geq D(X)$.

Hence, for any satisfactory $X \subset \mathbb{R}^m$, that spans $\mathbb{R}^m$, we have a finite choice of hyperplanes P lying in ∂S and associated non-neighboringx_i, x_j. We may construct a rooted tree $\mathcal{T}$ representing every possible outcome of the RP, and where each node $T \in \mathcal{T}$ represents a new version of X after the RP relative to some P, x_i, x_j. The root represents X. Hence any single node will directly spawn at most finite number of child nodes.

Let Y represent the version of X after some arbitrary number of applications of the RP, and $T_Y \in \mathcal{T}$ the node representing Y. Since α remains constant under the RF, we know D is bounded over all of $\mathcal{T}$:

$$D(Y) < n^3 \alpha$$

We may define a finite supremum.

$$\sup_{T_Y \in \mathcal{T}} D(Y) < \infty$$

When traveling along a path from the root, the value D only increases, hence, there must exist some path in the tree $\{T_X, T_{X'}, T_{X''}, \ldots\} = \{T_{X^{(k)}}\} \subset \mathcal{T}$ where

$$\lim_k D(X^{(k)}) = \sup_{T_Y \in \mathcal{T}} D(Y)$$

3 Conclusion

Although these results represent some strong new knowledge on how to treat the problem of maximal lifespan, the answer is currently very incomplete due to a lack of a generalization from 1-cycles in $\mathbb{R}^n$ into planar 1-cycles and more glaringly the lack of a final statement regarding the identity of the planar 1-cycle shape that achieves maximal lifespan for a given $k + 1$ and α.

My mentor, Glen Wilson, and I will continue to pursue strategies to try and achieve a concrete answer in both. Naturally, this attempt will require more work over the coming weeks and months. Currently, we plan to stay in correspondence with email and TeX documents.

Another direction this project could follow is the case of general l-cycles and the construction of the most efficient l-cycle, given $l, \alpha, k + 1$.

This question arose as a result of the initial question prompted to my by Glen: is there a way to discover links in 1-cycles under the vietoris-rips complex using the barcode; a compilation of persistence intervals for cycles under the rips complex. My plan of attack would be to try and isolate 1-cycles by only considering simplicial complexes S on sets X that displayed 2 nonintersecting components each of which displayed a single 1-cycle. Next, one could separate the set of vertices into two mutually exclusive subsets of vertices X_1, X_2 both of which displaying their own 1-cycle. One could then compute and analyze three different barcodes, one for each of the 1-cycles by themselves and one together in order to track how their interaction alters the lifespan of any single 1-cycle.

My thought was to produce sets of vertices that would force a link to occur. This was most easily achieved if we could say concretely that a discretized circle maximized lifespan for a given α, k. In this way one could force both 1-cycles to be discretized circles and then more easily force them to be linked or unlinked given the rigidity of their construction. One could track this interaction with the barcodes for X_1, X_2 and how they were altered with the added presence of the other 1-cycle observed in the barcode for the original X.

The highlights of this research appeared under the title *On the optimization of 1-cycle persistence under the Vietoris-Rips complex.* at the 2014 Joint Math Meeting in the AMS Session on Undergraduate Research in Analysis and Topology.

4 Acknowledgments

I would like to thank my research advisor Glen Wilson for his assistance throughout the research process, both during the program and following its end. His continued guidance and patience is greatly appreciated. I would also like to thank Rutgers University, DIMACS, and specifically Associate Director Gene Fiorini for providing me with the opportunity and environment to pursue this research. Finally, I would also like to thank the NSF for their generous funding of my research.