SUPPORT OF COHOMOLOGY OF DISK CONFIGURATION SPACES

HANNAH ALPERT

ABSTRACT. The configuration space of \( n \) ordered disks of radius \( r \) inside the unit disk is denoted \( \text{Conf}_{n,r}(D^2) \). We study how the cohomology of this space depends on \( r \). In particular, given a cohomology class of \( \text{Conf}_{n,0}(D^2) \), for which \( r \) does the corresponding pullback cohomology class of \( \text{Conf}_{n,r}(D^2) \) vanish? The paper consists of a collection of partial results, and it contains many questions and conjectures.

[Note: This manuscript describes a project that I have not finished yet. Please check for updates.]

1. Introduction

One well-studied topological space is the space of configurations of \( n \) distinct points in the plane. We study a variant of this space. For each radius \( r < 1 \), we study the space of configurations of \( n \) disjoint disks of radius \( r \) in the unit disk \( D^2 \). (If \( r \) is too large, then this space of configurations is empty.)

Definition. \( \text{Conf}_{n,r}(D^2) \) is defined as the subspace of \( (D^2)^n \) consisting of all \( n \)-tuples \( (x_1, \ldots, x_n) \) of points in the unit disk \( D^2 \) such that the open disks of radius \( r \) with centers \( x_1, \ldots, x_n \) are disjoint and contained in \( D^2 \).

According to this definition, as \( n \) remains fixed and \( r \) varies, the spaces \( \text{Conf}_{n,r}(D^2) \) are totally ordered by inclusion; as \( r \) gets larger, the space \( \text{Conf}_{n,r}(D^2) \) gets smaller. The union of these spaces over all \( r \) is denoted by \( \text{Conf}_n(D^2) \) and is the space of all \( n \)-tuples of distinct points in the unit disk \( D^2 \).

The goal is to study how the topology of \( \text{Conf}_{n,r}(D^2) \) changes as the radius \( r \) varies. This general question was introduced in the paper [2] by Baryshnikov, Bubenik, and Kahle. They defined radius \( r \) configuration spaces for arbitrary \( d \)-dimensional regions, and studied the case where the region is a rectangular box. The papers [3] and [6] are also based on this framework.

In this paper we start with the general question of how the topology of \( \text{Conf}_{n,r}(D^2) \) depends on the radius \( r \), and then develop the question in several different directions. Rather than giving one main theorem we uncover a large number of open questions and explore possible methods for approaching them. The remainder of the introduction contains some intuition for the original problem and an overview of the ideas in the sections that follow.

1.1. Example: three disks, connectivity, and critical values. As a first example we consider the case of three disks. How does the topology of \( \text{Conf}_{3,r}(D^2) \) depend on the radius \( r \)? The purpose of this example is to give some intuition for how the topology changes only at special values of \( r \), rather than changing continuously.
Figure 1. As $r$ increases past $\frac{1}{3}$, the space $\text{Conf}_{3,r}(D^2)$ splits into two connected components. As $r$ increases past $\left(1 + \frac{2}{\sqrt{3}}\right)^{-1}$, the space $\text{Conf}_{3,r}(D^2)$ becomes empty. Away from these two critical values, the topology of $\text{Conf}_{3,r}(D^2)$ is locally constant in $r$.

One change in topology occurs when $\text{Conf}_{3,r}(D^2)$ changes between non-empty and empty. This change occurs at $r = \left(1 + \frac{2}{\sqrt{3}}\right)^{-1}$, for which the three small disks form an equilateral triangle, shown in Figure 1.

The space $\text{Conf}_{3,r}(D^2)$ also changes between connected and disconnected. For radii near 0, the configuration space is connected, but for radii near the maximum $\left(1 + \frac{2}{\sqrt{3}}\right)^{-1}$, there are two connected components, which correspond to the clockwise and counter-clockwise orderings of the three disks. The change occurs at $r = \frac{1}{3}$, for which the three disks may be tangent along a diameter.

Another way to look at the topological change at $r = \frac{1}{3}$ is through the fundamental group. The fundamental group $\pi_1(\text{Conf}_3(D^2))$ is called the pure braid group on three strands. It consists of the braids for which the three strands are ordered in the same way at both ends of the braid. As $r$ increases starting at 0, the basic crossings that generate these braids remain possible as long as $r < \frac{1}{3}$, so for all such $r$ the fundamental group of $\text{Conf}_{3,r}(D^2)$ is equal to $\pi_1(\text{Conf}_3(D^2))$. However, for $r > \frac{1}{3}$ the only achievable braids consist of twisting all three strands together around to their starting positions, so for such $r$ the fundamental group of each connected component of $\text{Conf}_{3,r}(D^2)$ is $\mathbb{Z}$.

The tools introduced by Baryshnikov, Bubenik, and Kahle in [2] prove that for $n = 3$ in fact there are no other changes in topology besides at $r = \frac{1}{3}$ and $r = \left(1 + \frac{2}{\sqrt{3}}\right)^{-1}$. That is, if two radii $r_1 < r_2$ are such that the closed interval $[r_1, r_2]$ does not contain either of these two special values of $r$, then the inclusion of $\text{Conf}_{3,r_2}(D^2)$ into $\text{Conf}_{3,r_1}(D^2)$ is a homotopy equivalence. The method is to look at the \textbf{tautological function}

$$\tau : \text{Conf}_n(D^2) \to \mathbb{R},$$
1.2. Example: four disks, the clock map, and cohomology. In the next example, we show that the 4-disk configuration space $\text{Conf}_{4,r}(D^2)$ changes topology as the radius $r$ crosses $\frac{1}{4}$. The purpose of this example is to motivate the use of cohomology, which is the topological invariant we use in the rest of the paper.

As the radius $r$ crosses $\frac{1}{4}$, the number of connected components of $\text{Conf}_{4,r}(D^2)$ does not change, and neither does the fundamental group. However, the higher-dimensional topology does change, as illustrated by the following 3-dimensional construction, depicted in Figure 2. In the unit disk we first draw two disjoint medium-sized disks of radius $\frac{1}{2}$. To fit, they must be tangent to each other and to the boundary circle. Then inside them we draw four disjoint small disks of radius $\frac{1}{4}$: disks 1 and 2 in the first medium-sized disk, and disks 3 and 4 in the second medium-sized disk. The result is a configuration in $\text{Conf}_{4,\frac{1}{4}}(D^2)$. The set of all configurations that can be obtained in this way forms a 3-parameter family, parametrized by a map

$$c : (S^1)^3 \to \text{Conf}_{4,\frac{1}{4}}(D^2),$$

which we call the “clock map”. One of the parameters determines the angle of the vector from the first medium-sized disk to the second, and the others determine the angles of the vectors between the pairs of small disks—from disk 1 to disk 2 and from disk 3 to disk 4.
In this construction, the four small disks cannot be made larger without overlapping. We may ask, is this clock map homotopic to another map for which the small disks can be made larger? In particular, the image of \( c \) contains configurations with all four disks lined up along a diameter; can a homotopic map avoid all such configurations? An answer of “no” would prove that the topology of \( \text{Conf}_{4,r}(D^2) \) changes as \( r \) crosses \( \frac{1}{4} \).

Indeed, we show next that any map \( c' \) with image contained in \( \text{Conf}_{4,r}(D^2) \) for some \( r > \frac{1}{4} \) cannot be homotopic to the original clock map \( c \). We use another map

\[
t : \text{Conf}_{4,\frac{1}{4}}(D^2) \to (S^1)^3,
\]

which we call the “torus map”. It records the angles of three vectors: from disk 1 to disk 2, from disk 2 to disk 3, and from disk 3 to disk 4. Briefly, the argument is that the composition \( t \circ c : (S^1)^3 \to (S^1)^3 \) has degree 1, whereas \( t \circ c' \) has degree 0, and so \( c \) and \( c' \) cannot be homotopic.

First we show that \( t \circ c \) has degree 1, by showing that it is homotopic to the identity map on \((S^1)^3\). In order for this to work, in the definition of \( c \) we order the parameters so that the angle between the medium-sized disks is the second parameter, and the first and third parameters are the angles from disk 1 to disk 2 and from disk 3 to disk 4. Then the first and third coordinates of \( t \circ c \) are the same as in the identity map. Only the second coordinate needs to be modified. To do this we observe that the vector between the medium-sized disks always forms an acute angle with the vector from disk 2 to disk 3, so we can interpolate linearly between those two directions to create the desired homotopy.

Next we show that \( t \circ c' \) has degree 0, by showing that it is not surjective. If \( r > \frac{1}{4} \) then \( \text{Conf}_{4,r}(D^2) \) does not contain any configurations in which the four disks are all collinear, so if the image of \( c' \) is in \( \text{Conf}_{4,r}(D^2) \) then the image of \( t \circ c' \) excludes every point of \((S^1)^3\) that has all three coordinates equal. Thus \( t \circ c' \) must have degree 0, implying that the homotopy class of \( c \) in \( \text{Conf}_{4,r}(D^2) \) disappears as \( r \) increases past \( \frac{1}{4} \).

This argument in terms of degrees of maps can be rephrased in the language of cohomology, as follows. We consider the cohomology class on \( \text{Conf}_{4,\frac{1}{4}}(D^2) \) obtained by using \( t \) to pull back the fundamental cohomology class of \((S^1)^3\). We denote this fundamental class by \([ (S^1)^3 ] \), so that the pullback is \( t^*[(S^1)^3] \). For \( r > \frac{1}{4} \), the restriction of \( t^*[(S^1)^3] \) to \( \text{Conf}_{4,r}(D^2) \) is zero, because the restriction \( t|_{\text{Conf}_{4,r}(D^2)} \) is not surjective so it factors through a subspace of \((S^1)^3\) that has no cohomology in degree 3. However, the class \( t^*[(S^1)^3] \) on \( \text{Conf}_{4,\frac{1}{4}}(D^2) \) must be nonzero, because the fact that \( t \circ c \) has degree 1 implies that the pullback class \( c^*t^*[(S^1)^3] \) must be equal to \([ (S^1)^3 ] \), which is nonzero.

1.3. Main questions and paper outline. In this paper our preferred invariant for detecting topological changes in \( \text{Conf}_{n,r}(D^2) \) is cohomology. Specifically, if

\[
i^n_{n,r} : \text{Conf}_{n,r}(D^2) \to \text{Conf}_n(D^2)
\]

denotes the inclusion map, then \( i^n_{n,r} \) induces a map on cohomology

\[
i^n_{n,r} : H^*(\text{Conf}_n(D^2)) \to H^*(\text{Conf}_{n,r}(D^2)).
\]

In the preceding example, we identified a class in \( H^3(\text{Conf}_4(D^2)) \) that is in \( \ker i^n_{n,r} \) for \( r > \frac{1}{4} \) but not for \( r < \frac{1}{4} \). In general, we study how the subspace \( \ker i^n_{n,r} \) of \( H^*(\text{Conf}_n(D^2)) \) grows with \( r \).
In the first few sections we study the changes in \( \ker i_{n,r}^* \). The resulting questions suggest the following simpler construction. Let \( \text{Seg}_{n,r}(D^2) \) denote the space of configurations of \( n \) labeled, oriented segments of length \( r \), arranged so that they are disjoint inside the unit disk \( D^2 \). Informally, we ask how large the length \( r \) can be while allowing the segments to rotate independently. For instance, if \( r \) is a small constant multiple of \( \frac{1}{\sqrt{n}} \) it is possible to fit \( n \) disjoint medium-sized disks of diameter \( r \) inside the unit disk \( D^2 \), so that each segment can rotate independently inside its own medium-sized disk. How much can we increase \( r \) and still be able to make a construction like this? This question can be viewed as a multi-needle variant of the classical Kakeya needle problem, which asks for the infimal area of a set in the plane in which one segment can rotate through an entire circle of angles.

More formally, there is a “torus map”

\[
t : \text{Seg}_{n,r}(D^2) \to (S^1)^n
\]

recording the direction of each segment. We may ask whether the pullback by \( t \) of the fundamental cohomology class of \( (S^1)^n \) is zero on \( \text{Seg}_{n,r}(D^2) \). For large \( r \) it is zero, and for small \( r \) (proportional to \( \frac{1}{\sqrt{n}} \), for instance) it is nonzero; for each \( n \) we denote by \( r_{\text{crit}}(n) \) the threshold value of \( r \) above which the pullback class is zero and below which it is nonzero. We hope to describe the asymptotic behavior of \( r_{\text{crit}}(n) \) as \( n \) grows large.

Sections 2 through 5 address the original question about disks. In Section 2 we give precise definitions and find the least radius \( r \) for which the cohomology changes. In Section 3 we study in detail the case of four disks of radius \( r \) inside the unit disk. In Section 4 we see how these methods can generalize to the case of \( n \) disks when \( r \) is not too big relative to \( n \). In Section 5 we introduce a conjecture about the asymptotic behavior of cohomology as \( n \) grows large. Sections 6 and 7 address the new question about segments. In Section 6 we give a conjecture for segments analogous to the conjecture about disks from Section 5, and in Section 7 we try out some approaches to this conjecture.

The paper includes a large number of questions and conjectures, which appear at the ends of the various sections.

Acknowledgments. I would like to thank my advisor Larry Guth, who introduced this problem to me by describing the clock map that appears as the second example in the introduction.

2. Background and first critical value

We begin this section by reviewing the setup introduced by Baryshnikov, Bubenik, and Kahle in \([2]\). Then we show that the homotopy type of the space of \( n \) balls of radius \( r \) in the unit ball changes as \( r \) passes \( \frac{1}{n} \), but is the same for all smaller \( r \).

For any bounded region \( U \subseteq \mathbb{R}^d \) we let \( \text{Conf}_{n}(U) \) denote the set of ordered \( n \)-tuples of distinct points in \( U \). For each such \( n \)-tuple \( \vec{x} = (x_1, \ldots, x_n) \), there is a supremal radius \( r \) such that the balls of radius \( r \) centered at \( x_1, \ldots, x_n \) are disjoint and contained in \( U \); we denote this \( r \) by the function \( \tau(\vec{x}) \), called the tautological function. The space \( \text{Conf}_{n,r}(U) \) is defined to be \( \tau^{-1}[r, \infty) \), which is the subspace of \( \text{Conf}_n \) containing those configurations that can be the centers of disjoint balls of radius \( r \) in \( U \). The goal is to study how the topology of \( \text{Conf}_{n,r}(U) \) changes with \( r \).
The paper [2] shows that the tautological function $\tau$ is like a Morse function: between critical values (under some suitable definition), the topology of the superlevel sets $\tau^{-1}[r, \infty)$ does not change. The critical values are defined according to the following description of the critical points. For $\vec{x} \in \text{Conf}_{n,r}(U)$, the stress graph has the following vertex set: every center $x_1, \ldots, x_n$, plus every point $y$ of the boundary $\partial U$ at distance exactly $r$ from one of the centers. The points $x_1, \ldots, x_n$ are called internal points, and the vertices $y$ are called boundary points. The edges of the stress graph are some nonempty collection of the pairs $\{x_i, x_j\}$ at distance exactly $2r$ and the pairs $\{x_i, y\}$ at distance exactly $r$, drawn as segments in $\mathbb{R}^d$. Every edge is assigned a positive weight, and we interpret the weighted graph as a system of mechanical stresses, in which the weight of an edge is the amount of force pushing both endpoints outward. A stress graph is balanced if the weights satisfy the following conditions:

- The mechanical stresses at each internal point sum to zero; and
- On each connected component, the mechanical stresses on the boundary points sum to zero.

A balanced configuration is any configuration admitting a balanced stress graph.

**Theorem 1** ([2]). Suppose that $r_1 < r_2$ are numbers such that if $r \in [r_1, r_2]$ then there are no balanced configurations in $\text{Conf}_{n,r}(U)$. Then $\text{Conf}_{n, r_2}(U)$ is a deformation retract of $\text{Conf}_{n, r_1}(U)$.

This theorem says that in order to study the topology of $\text{Conf}_{n, r}(U)$, it suffices to study what happens as $r$ passes each radius with a balanced configuration. The paper [2] identifies the smallest such radius when $U$ is a rectangular box in $\mathbb{R}^d$ and shows that the homotopy type of $\text{Conf}_{n, r}(U)$ does indeed change as $r$ crosses that critical radius. In the remainder of this section we prove an analogous result for the case where $U$ is the unit ball $B^d$ in $\mathbb{R}^d$. Theorem 2 shows that the smallest critical radius is $\frac{1}{n}$, and Theorem 6 shows that some cohomology is lost as $r$ increases past $\frac{1}{n}$.

**Theorem 2.** The least $r$ for which $\text{Conf}_{n,r}(B^d)$ has a balanced configuration is $r = \frac{1}{n}$.

The balanced configuration in $\text{Conf}_{n, \frac{1}{n}}(B^d)$ has all $n$ balls lined up along a diameter. To show that no smaller radius has a balanced configuration, we use the following two easy lemmas.

**Lemma 3.** Let $T$ be a tree embedded in $\mathbb{R}^d$ with straight edges. If the total length of $T$ is $L$, then there is a closed ball containing $T$ with radius at most $\frac{L}{2}$.

**Proof.** We use induction on the number of edges in $T$. The base case is when $T$ has one vertex and no edges. Otherwise, we let $v$ be a vertex of $T$ incident to only one edge $e$, of length $|e|$. By the inductive hypothesis there is a ball of radius at most $\frac{L - |e|}{2}$ containing all of $T$ except $e$. There is also a ball of radius $\frac{|e|}{2}$ containing $e$, and the two balls intersect at the other vertex of $e$, so there is a ball of radius at most $\frac{L - |e|}{2} + \frac{|e|}{2} = \frac{L}{2}$ containing both balls and therefore containing $T$. \qed

**Lemma 4.** Let $A$ be any collection of points on a sphere $S$ in $\mathbb{R}^d$ of radius $\rho$, not contained in any open hemisphere. Then every ball containing $A$ has radius at least $\rho$. 

Proof. Let $B$ be any ball of radius less than $\rho$. Then $B \cap S$ either is empty or is contained in an open hemisphere of $S$, so it cannot contain $A$. 

Proof of Theorem 2. Suppose for contradiction that $r < \frac{1}{n}$ has a balanced configuration, and assume without loss of generality that the stress graph is connected. The stress graph has two parts: inside the ball of radius $1 - r$, it is a graph on the interior vertices with all edges of length $2r$; outside the ball of radius $1 - r$, from every boundary vertex there is one edge of length $r$ extending radially to the corresponding interior vertex.

Let $B_1$ denote the ball of radius $1 - r$. We apply Lemma 3 to a spanning tree of the portion of the stress graph inside $B_1$. There are at most $n$ vertices and every edge has length $2r$, so the total length is at most $2r(n - 1)$, and so the lemma implies that there is a ball $B_2$ of radius at most $r(n - 1)$ containing the spanning tree.

On the other hand, we know that the boundary vertices do not all lie in an open hemisphere of the unit sphere, because if they did, their mechanical stresses could not add up to zero. Therefore their neighboring interior vertices do not all lie in an open hemisphere of the sphere $\partial B_1$ of radius $1 - r$. We apply Lemma 4 where $S$ is $\partial B_1$ and $A$ is the set of interior vertices on $\partial B_1$. The lemma implies that every ball containing $A$ has radius at least $1 - r$, but from above we know that $B_2$ contains $A$ and has radius at most $r(n - 1)$. Because $r < \frac{1}{n}$, we have $r(n - 1) < 1 - r$, giving a contradiction.

In order to show that the homotopy type of $Conf_{n,r}(B^d)$ changes as $r$ crosses $\frac{1}{n}$, we identify a nonzero element of $H^*(Conf_n(B^d))$ that vanishes when restricted to $Conf_{n,r}(B^d)$ if $r > \frac{1}{n}$. The cohomology $H^*(Conf_n(B^d))$ has a nice description, the proof of which is given by Arnol’d in [1]. For each directed graph $G$ on the vertex set $\{1, 2, \ldots, n\}$ there is a corresponding map

$$t_G : Conf_n(B^d) \to (S^{d-1})^{\left| E(G) \right|}$$

that, for each $\vec{x} \in Conf_n(B^d)$ and each edge $i \to j$ of $G$, records the unit vector $\frac{x_j - x_i}{|x_j - x_i|}$ in $S^{d-1}$. We refer to $t_G$ as a torus map because the target space is a torus when $d = 2$. Every torus map determines a cohomology class in $H^*(Conf_n(B^d))$ obtained by pulling back the fundamental class of $(S^{d-1})^{\left| E(G) \right|}$ by $t_G$. It turns out that $H^*(Conf_n(B^d))$ is generated by these pullback classes and is in fact freely generated by a particular family of them. We say that $G$ is an ordered forest if every edge $i \to j$ has $i < j$ and every vertex has in-degree at most 1.

Theorem 5 ([1]). As $G$ ranges over all ordered forests on $n$ vertices, the pullback classes $t_G^*[(S^{d-1})^{\left| E(G) \right|}] \in H^*(Conf_n(B^d))$ form a free basis for $H^*(Conf_n(B^d))$.

With this notation for pullback cohomology classes, we can state the theorem saying that the cohomology changes as $r$ crosses $\frac{1}{n}$. Let $P_n$ denote the ordered forest that is the path $1 \to 2 \to \cdots \to n$, and let $i_{n,r}$ denote the inclusion of $Conf_{n,r}(B^d)$ into $Conf_n(B^d)$.

Theorem 6. For any $r > \frac{1}{n}$, the inclusion

$$i_{n,r} : Conf_{n,r}(B^d) \to Conf_n(B^d)$$

is not a homotopy equivalence. In particular, $\ker i_{n,r}^* \subseteq H^*(Conf_n(B^d))$ contains the nontrivial element $t_{P_n}^*[(S^{d-1})^{n-1}]$. 


Proof. The nontriviality of $t_{P_n}^*([S^{d-1}]_E)$ is part of Theorem 5 above. It remains to show that $i_{n,r}^* t_{P_n}^*([S^{d-1}]_{n-1}) = 0$. Indeed, the composition

$$t_{P_n} \circ i_{n,r} : \text{Conf}_{n,r}(B^d) \to (S^{d-1})^{n-1}$$

is not surjective for $r > \frac{1}{n}$ because its image does not contain any point for which all $n - 1$ components are equal in $S^{d-1}$. The pullback of the fundamental class $[S^{d-1}]_{n-1}$ along a non-surjective map $t_{P_n} \circ i_{n,r}$ must be zero. □

3. Computation with four disks

For the rest of the paper we focus on the case of dimension $d = 2$ and consider configurations in the unit disk $D^2$. As above we let

$$i_{n,r} : \text{Conf}_{n,r}(D^2) \hookrightarrow \text{Conf}_n(D^2)$$

denote the inclusion. We use the description of $H^*(\text{Conf}_n(D^2))$ in terms of ordered forests from Theorem 5, and ask how the kernel $i_{n,r}^* \subseteq H^*(\text{Conf}_n(D^2))$ changes with $r$. In this section we consider the case $n = 4$ and compute $\ker i_{4,r}^*$ for all $r$. (The simpler cases of $n = 2$ and $n = 3$ are easily computed by the same methods.) The results of the computation are presented as a series of lemmas.

Lemma 7. For all $r \leq \frac{1}{4}$, we have $\ker i_{4,r}^* = 0$.

Proof. This is an immediate consequence of Theorem 2. □

For $r > \frac{1}{4}$, we use the following sublemma to show that $\ker i_{4,r}^*$ contains all of $H^3(\text{Conf}_n(D^2))$ and some of $H^2(\text{Conf}_n(D^2))$.

Sublemma 8. Let $(S^1)^3 \setminus \{\theta_1 = \theta_2 = \theta_3\}$ denote the subspace of $(S^1)^3$ where the three coordinates $\theta_1, \theta_2, \theta_3$ are not equal, and let $i$ denote the inclusion of this subspace into $(S^1)^3$. Then $\ker i^* \subseteq H^*(S^1)^3$ is generated by $d\theta_1 \wedge d\theta_2 \wedge d\theta_3$ (in degree 3) and $d\theta_2 \wedge d\theta_3 - d\theta_1 \wedge d\theta_3 + d\theta_1 \wedge d\theta_2$ (in degree 2).

Proof. Thinking of $S^1$ as $\mathbb{R}/\mathbb{Z}$, we make the following change of coordinates:

$$\varphi_1 = \theta_1,$$

$$\varphi_2 = \theta_2 - \theta_1 + \frac{1}{2},$$

$$\varphi_3 = \theta_3 - \theta_2 + \frac{1}{2}.$$  

We use the cell decomposition of $(S^1)^3$ into eight cells, each obtained by setting some subset of $\varphi_1, \varphi_2,$ and $\varphi_3$ to zero. The set $\{\theta_1 = \theta_2 = \theta_3\}$ is the same as the set $\{\varphi_2 = \varphi_3 = \frac{1}{2}\}$, and the subspace $(S^1)^3 \setminus \{\varphi_2 = \varphi_3 = \frac{1}{2}\}$ deformation retracts onto the space $\{\varphi_2 = 0\} \cup \{\varphi_3 = 0\}$; this space is obtained by deleting the 3-dimensional cell spanned by $\varphi_1, \varphi_2,$ and $\varphi_3$ and the 2-dimensional cell spanned by $\varphi_2$ and $\varphi_3$. Thus ker $i^*$ is generated by $d\varphi_1 \wedge d\varphi_2 \wedge d\varphi_3 = d\theta_1 \wedge d\theta_2 \wedge d\theta_3$ and $d\varphi_2 \wedge d\varphi_3 = d\theta_2 \wedge d\theta_3 - d\theta_1 \wedge d\theta_3 + d\theta_1 \wedge d\theta_2$. □

In what follows, we use "ordered forest" to refer both to the combinatorial objects themselves and to the resulting generators of $H^*(\text{Conf}_n(D^2))$. These generators are a priori only determined up to a sign, because the torus maps $t_{\varepsilon}$ are only determined up to reordering the coordinates of the target. However, in any ordered forest we can order the edges according to their terminal endpoints (which are all different).
Under this ordering every ordered forest corresponds uniquely to a generator of $H^*(\text{Conf}_n(D^2))$.

**Lemma 9.** For all $r > \frac{1}{4}$, the kernel $\text{ker} \ i_*^{4,r}$ contains all six 3-edge ordered forests, and, for each 3-edge ordered forest, the signed sum of its three 2-edge ordered subforests, as shown in Figure 3. These twelve elements of $\text{ker} \ i_*^{4,r}$ are linearly independent.

**Proof.** Let $G$ be any 3-edge ordered forest on 4 vertices. Then for $r > \frac{1}{4}$, the composition
$$t_G \circ i_{4,r} : \text{Conf}_{4,r}(D^2) \hookrightarrow \text{Conf}_4(D^2) \to (S^1)^3$$
has image in $(S^1)^3 \setminus \{\theta_1 = \theta_2 = \theta_3\}$, so it factors through the inclusion
$$i : (S^1)^3 \setminus \{\theta_1 = \theta_2 = \theta_3\} \hookrightarrow (S^1)^3$$
from Sublemma 8. Thus, when we pull back the elements $d\theta_1 \wedge d\theta_2 \wedge d\theta_3$ and $d\theta_2 \wedge d\theta_1 \wedge d\theta_3 - d\theta_1 \wedge d\theta_2 + d\theta_1 \wedge d\theta_2$ to $H^*(\text{Conf}_{4,r}(D^2))$, the result must be zero.

Thus, the desired twelve elements are indeed in $\text{ker} \ i_*^{4,r}$. Their linear independence is a consequence of the linear independence of ordered forests (Theorem 5) and the fact that when they are written out as in Figure 3, the first term of each element is not a term of any other element. \[\square\]

**Lemma 10.** For all $r \leq \frac{1}{3}$, the kernel $\text{ker} \ i_*^{4,r}$ does not contain anything outside the span of the twelve elements from Lemma 9.

**Proof.** In degree 3 there is nothing to show, because the six 3-edge forests from Lemma 9 span $H^3(\text{Conf}_4(D^2))$. In degree 0 there is also nothing to show. To address degrees 1 and 2 we observe that it is possible to turn two disks around each other while leaving the other two fixed, as shown in Figure 4. That is, for every pair of indices $a$ and $b$, we have a map
$$h_{a \to b} : S^1 \to \text{Conf}_{4,\frac{1}{3}}(D^2),$$
defined so that the composition with the torus map $t_{a \to b}$ is the identity on $S^1$. The corresponding homology classes $h_{a \to b,*}[S^1]$ in $H_1(\text{Conf}_4(D^2))$ form a dual basis to the collection of 1-edge forests in $H^1(\text{Conf}_4(D^2))$. Thus for $r < \frac{1}{3}$ we see that $i_*^{4,r}$ must be injective on $H^1(\text{Conf}_n(D^2))$. 

Figure 3. For all $r > \frac{1}{4}$, the kernel $\text{ker} \ i_*^{4,r}$ contains the span of these twelve elements, which are linearly independent because the first term of each element is not a term of any other element.
To see what happens in degree 2, we use the maps \( \hat{h}_{a \to b} \) to construct the maps

\[
\hat{h}_{a \to b} : (S^1)^2 \to \text{Conf}_{4, \frac{1}{r}}(D^2)
\]
as follows. We define \( \hat{h}_{a \to b}(\theta_1, \theta_2) \) to be the rotation of the configuration \( h_{a \to b}(\theta_2 - \theta_1) \) by the angle \( \theta_1 \), so that the vector from disk \( a \) to disk \( b \) points in the direction \( \theta_2 \). To pair a homology class \( \hat{h}_{a \to b,*}[(S^1)^2] \in H_2(\text{Conf}_{4}(D^2)) \) with a 2-edge ordered forest \( G \), we find the degree of the map \( t_G \circ \hat{h}_{a \to b} : (S^1)^2 \to (S^1)^2 \). If \( a \to b \) is not an edge of \( G \), then the degree is 0; if \( a \to b \) is an edge of \( G \), then the degree is 1 or \(-1 \) according to whether \( a \to b \) is the second edge of \( G \) or the first edge.

To show that \( \ker i^*_r \) in degree 2 contains only the span of the six degree-2 elements from Lemma 9, we show that \( i^*_r \) is injective on a complementary subspace to that span. The complementary subspace we choose is the span of the five 2-edge forests that have \( 1 \to 2 \) as an edge. A dual basis in homology to these five generators is given by the homology classes corresponding to \( \hat{h}_{a \to b} \) where \( a \to b \) ranges over the five pairs other than \( 1 \to 2 \). Thus \( i^*_r \) must be injective on the span of these five ordered forests.

\[ \square \]

**Lemma 11.** For all \( r > \frac{1}{3} \), the kernel \( \ker i^*_r \) contains all of \( H^2(\text{Conf}_{4}(D^2)) \) and the subspace of \( H^1(\text{Conf}_{4}(D^2)) \) for which the coefficients of the generators sum to zero.

**Proof.** First we make an observation similar to Sublemma 8. If \( G \) is a 2-edge ordered forest for which the two edges share a vertex, then for \( r > \frac{1}{3} \) the image of the torus map \( t_G \) lies in \( (S^1)^2 \setminus \{ \theta_1 = \theta_2 \} \). If the inclusion of this subspace into \( (S^1)^2 \) is denoted by \( i \), then \( i_* \) is generated by \( d\theta_1 \wedge d\theta_2 \) and \( d\theta_1 - d\theta_2 \), so the pullbacks of those classes by \( t_G \) are in \( \ker i^*_r \).

Thus, in degree 2 the kernel \( \ker i^*_r \) contains every 2-edge ordered forest for which the two edges share a vertex. The span of these and the six degree-2 generators from Lemma 9 contains the remaining three 2-edge forests, so it contains all of \( H^2(\text{Conf}_{4}(D^2)) \). In degree 1, the span of the classes \( t_G^*(d\theta_1 - d\theta_2) \), where \( G \) ranges over the 2-edges forests for which the two edges share a vertex, is indeed the subspace for which the sum of coefficients is zero.

\[ \square \]

**Lemma 12.** For all \( r \leq \frac{1}{1+\sqrt{2}} \), the kernel \( \ker i^*_r \) does not contain any class in \( H^1(\text{Conf}_{4}(D^2)) \) for which the sum of coefficients of the generators is nonzero. If \( r > \frac{1}{1+\sqrt{2}} \), then \( \text{Conf}_{4,r}(D^2) \) is empty, so \( \ker i^*_r \) contains all of \( H^*(\text{Conf}_{4}(D^2)) \).
Proof. The radius \( r = \frac{1}{1+\sqrt{2}} \) corresponds to the balanced configuration for which the four disks are arranged in a square and are tangent to each other and to the boundary circle. There is a map

\[ h : S^1 \to \text{Conf}_4 \left( \frac{1}{1+\sqrt{2}} \right) (D^2) \]

given by spinning the balanced configuration, and the pairing of the corresponding homology class \( h_*[S^1] \) with any element of \( H^1(\text{Conf}_4(D^2)) \) is the sum of coefficients of the generators. Thus if \( r < \frac{1}{1+\sqrt{2}} \) any degree-1 class with a nonzero sum of coefficients cannot be in \( \ker i^*_r \).

We check that \( \text{Conf}_{4,r}(D^2) \) is empty when \( r > \frac{1}{1+\sqrt{2}} \). Suppose for contradiction that \( \vec{x} \in \text{Conf}_{4,r}(D^2) \). Then the centers \( x_1, x_2, x_3, \) and \( x_4 \) are in the disk of radius \( 1-r \). For each pair of indices \( i \) and \( j \), the angle formed by the segments from the origin to \( x_i \) and \( x_j \) must be obtuse, otherwise the distance from \( x_i \) to \( x_j \) is less than \( 2r \). But it is impossible for every consecutive pair of centers to form an obtuse angle.

These Lemmas 7 through 12 compute \( \ker i^*_{n,r} \) for all \( r \). For \( r \in [0, \frac{1}{3}] \) the kernel is 0. For \( r \in (\frac{1}{3}, \frac{1}{1+\sqrt{2}}] \) the kernel is all 6 dimensions of \( H^3(\text{Conf}_4(D^2)) \) and 6 of the 11 dimensions of \( H^2(\text{Conf}_4(D^2)) \). For \( r \in (\frac{1}{3}, \frac{1}{1+\sqrt{2}}] \) the kernel is all 6 dimensions of \( H^3(\text{Conf}_4(D^2)) \), all 11 dimensions of \( H^2(\text{Conf}_4(D^2)) \), and 5 of the 6 dimensions of \( H^1(\text{Conf}_4(D^2)) \). For \( r \in (\frac{1}{1+\sqrt{2}}, \infty) \) the kernel is all 6 dimensions of \( H^3(\text{Conf}_4(D^2)) \), all 11 dimensions of \( H^2(\text{Conf}_4(D^2)) \), all 6 dimensions of \( H^1(\text{Conf}_4(D^2)) \), and all 1 dimension of \( H^0(\text{Conf}_4(D^2)) \).

**Question 1.** Is there a systematic way to compute \( \ker i^*_{n,r} \) for all \( r \) for the next few values of \( n \), like 5, 6, and 7?

### 4. Small Critical Values

For large \( n \) we do not expect to be able to compute \( \ker i^*_{n,r} \) precisely for all \( r \). In particular, when \( r \) is large compared to \( n \), describing \( \text{Conf}_{n,r}(D^2) \) becomes a circle-packing problem. In this section we describe how when \( r \) is fairly small compared to \( n \), the techniques from the \( n = 4 \) computation can carry over. We show that the first several balanced configurations consist of disks lined up along a diameter. Then we construct homology classes that can be used to prove upper bounds on the size of \( \ker i^*_{n,r} \).

**Theorem 13.** Suppose that \( n \) disks of radius \( r \) form a balanced configuration in the unit disk \( D^2 \), such that the stress graph is not just a diameter. Then the radius \( r \) satisfies the inequality

\[ r > \frac{3}{2n+3}. \]

In other words, if there is a balanced configuration with radius \( r \leq \frac{3}{2n+3} \), then \( r = \frac{1}{k} \) for some integer \( k \leq n \) and the configuration has \( k \) disks lined up along a diameter.

**Proof.** Let \( v_1, \ldots, v_m \) denote the boundary points of the stress graph. The edges from \( v_1, \ldots, v_m \) extend along the radii containing them; let \( w_1, \ldots, w_m \) denote the vertices of the stress graph of degree greater than 2 that are closest to \( v_1, \ldots, v_m \) along their radii.
Figure 5. We consider the outer boundary of the stress graph, minus all but three of the radial portions. This subgraph, called $H$, is shown on the right. Its length is at least 3 and is less than $r(2n + 3)$.

For each consecutive pair of boundary points $v_i$ and $v_{i+1}$, the outer boundary of the stress graph traces a concave piecewise-linear curve that begins with the segment from $v_i$ to $w_i$, ends with the segment from $w_{i+1}$ to $v_{i+1}$, and stays within the sector determined by $v_i$ and $v_{i+1}$. Figure 5 depicts an example balanced configuration with its stress graph.

If $w_i$ happens to be at the origin, then the concavity implies that $w_{i+1}$ must also be at the origin; by the same reasoning all $w_j$ must be at the origin. In this case the stress graph consists of at least three radii. We consider the subgraph $H$ containing exactly three of the radii, and suppose that its number of interior vertices is $k$. The total length of $H$ is 3 but is also $r(2k + 1)$, so we have

$$r = \frac{3}{2k + 1} \geq \frac{3}{2n + 1} > \frac{3}{2n + 3}.$$ 

Thus the lemma is proven in the case where some $w_i$ is at the origin.

If all $w_i$ are not at the origin, then the concave curves from each $w_i$ to the next $w_{i+1}$ form a cycle. We consider the subgraph $H$ containing this cycle plus three radial segments, as follows. The boundary points $v_1, \ldots, v_m$ are not contained in an open hemisphere (that is, semicircle), so we can choose three of them, denoted $v_{i_1}, v_{i_2},$ and $v_{i_3}$, such that the origin is in the triangle they determine, possibly on the boundary. Then $H$ contains the cycle from above as well as the segments from $v_{i_1}$ to $w_{i_1}$, from $v_{i_2}$ to $w_{i_2}$, and from $v_{i_3}$ to $w_{i_3}$. Figure 5 includes a picture of $H$.

If $k$ is the number of interior vertices of $H$, then the total length of $H$ is $r(2k + 3)$. To finish the proof of the lemma, we show that the total length of $H$ is also greater than 3, giving the desired inequality

$$r > \frac{3}{2k + 3} \geq \frac{3}{2n + 3}.$$ 

First we replace $H$ by a graph $H_1$ of shorter total length, and then we replace $H_1$ by the graph $H_2$ that contains exactly the three radii at $v_{i_1}, v_{i_2},$ and $v_{i_3}$. We show that $H_2$ has shorter total length than $H_1$, so $H$ has total length greater than 3.

To construct $H_1$, we replace the concave cycle between the points $w_i$ by the triangle with vertices $w_{i_1}, w_{i_2},$ and $w_{i_3}$. In other words, we view this concave cycle
as the disjoint union of curves from \( w_{i_1} \) to \( w_{i_2} \), from \( w_{i_2} \) to \( w_{i_3} \), and from \( w_{i_3} \) to \( w_{i_4} \). Straightening these curves to produce \( H_1 \) shortens the total length.

Then to construct \( H_2 \), we replace the triangle by the segments from the origin to \( w_{i_1}, w_{i_2}, \) and \( w_{i_3} \), so that \( H_2 \) consists of the three radii. If we view \( w_{i_1}, w_{i_2}, \) and \( w_{i_3} \) as points in the normed vector space \( \mathbb{R}^2 \), then the statement that \( H_2 \) is shorter than \( H_1 \) may be rephrased as the inequality

\[
|w_{i_1} - w_{i_2}| + |w_{i_2} - w_{i_3}| + |w_{i_3} - w_{i_1}| > |w_{i_1}| + |w_{i_2}| + |w_{i_3}|
\]

This inequality is the content of Lemma 14 which appears next, concluding the proof of Theorem 13. □

**Lemma 14.** Let \( w_1, w_2 \) and \( w_3 \) be noncollinear points of \( \mathbb{R}^2 \), and let \( p \) be a point in the triangle with vertices \( a, b, \) and \( c \), possibly on the boundary. Then the distances between the points satisfy the inequality

\[
|w_1 - w_2| + |w_2 - w_3| + |w_3 - w_1| > |p - w_1| + |p - w_2| + |p - w_3|
\]

**Proof.** Consider the function \( p \mapsto |p - w_1| \), for \( p \in \mathbb{R}^2 \). This function satisfies the following convexity property: for any two points \( p, p' \in \mathbb{R}^2 \) and any \( t \in [0, 1] \), the convex combination \( tp + (1 - t)p' \) satisfies the inequality

\[
|tp + (1 - t)p' - w_1| \leq t|p - w_1| + (1 - t)|p' - w_1|
\]

The same convexity property applies to \( |p - w_2| \) and \( |p - w_3| \), so the sum \( |p - w_1| + |p - w_2| + |p - w_3| \), viewed as a function of \( p \), has no local maxima when restricted to any line in \( \mathbb{R}^2 \). Thus, the maximum value of \( |p - w_1| + |p - w_2| + |p - w_3| \) as \( p \) ranges over the triangle must occur when \( p \) is equal to one of the vertices, so the maximum is the sum of the greatest two of the three distances \( |w_1 - w_2|, |w_2 - w_3| \), and \( |w_3 - w_1| \). □

Theorem 13 suggests that it may be possible to compute \( \ker i_{n,r}^* \) whenever \( r \leq \frac{3}{2n+1} \). The rest of this section concerns a map \( \varphi_n \) into \( \text{Conf}_{n,\frac{1}{n}}(D^2) \) that we construct next. We show that \( \varphi_n \) gives rise to homology classes that span the dual space to \( H^{n-1}(\text{Conf}_n(D^2)) \). The map \( \varphi_n \) is a useful tool for trying to compute \( \ker i_{n,r}^* \) because modifications of \( \varphi_n \) can produce many different homology classes on \( \text{Conf}_{n,r}(D^2) \); these classes are in the dual space to the image of \( i_{n,r}^* \).

We construct the map

\[
\varphi_n : (S^1)^{n-1} \to \text{Conf}_{n,\frac{1}{n}}(D^2)
\]

recursively in \( n \). For \( n = 1 \) there is nothing to say, because the source and target spaces are each one point. For \( n > 1 \), we would like to construct the configuration \( \varphi_n(\theta_1, \ldots, \theta_{n-1}) \). Inside the unit disk, we draw a medium-sized disk of radius \( \frac{n-1}{n} \) and the \( n \)th disk of radius \( \frac{1}{n} \) such that they are disjoint and tangent, and such that the vector from the center of the medium-sized disk to the center of the \( n \)th disk has angle \( \theta_{n-1} \). Figure 6 depicts this process. Then, inside the medium-sized disk, we draw an \( \frac{n-1}{n} \)-scaled copy of \( \varphi_{n-1}(\theta_1, \ldots, \theta_{n-2}) \), which has \( n - 1 \) disks of radius \( \frac{1}{n-1} \). The resulting picture is \( \varphi_n(\theta_1, \ldots, \theta_{n-1}) \).

The map \( \varphi_n \) corresponds to a homology class \( (i_{n,\frac{1}{n}})_* (\varphi_n)_* [(S^1)^{n-1}] \) on \( \text{Conf}_n(D^2) \) of dimension \( n - 1 \). To construct more homology classes, we modify \( \varphi_n \) by permuting the disks. For each permutation \( \sigma \in S_n \), there is a map

\[
\sigma \circ \varphi_n : (S^1)^{n-1} \to \text{Conf}_{n,\frac{1}{n}}(D^2)
\]
and a corresponding homology class \((i_{n,1})_* \sigma_* (\varphi_n)_*[(S^1)^{n-1}]\). We denote this homology class by \(\sigma \circ \varphi_n\) for short.

**Theorem 15.** A free basis of the dual space to \(H_{n-1}(\text{Conf}_n(D^2))\) is given by the homology classes \(\sigma \circ \varphi_n\), where \(\sigma\) ranges over permutations in \(S_n\) with \(\sigma(1) = 1\).

To prove the theorem we need to compute the pairing of each \(\sigma \circ \varphi_n\) with each \((n-1)\)-edge ordered forest \(G\). The pairing is denoted by \(\langle G, \sigma \circ \varphi_n \rangle\) and is given by the degree of the composition

\[
t_G \circ i_{n,1} \circ \sigma \circ \varphi_n : (S^1)^{n-1} \to \text{Conf}_{n,\frac{1}{n}} \to \text{Conf}_{n,\frac{1}{n}} \to \text{Conf}_n \to (S^1)^{n-1}.
\]

**Lemma 16.** Let \(\sigma \in S_n\) be a permutation such that \(\sigma(1) = 1\), and let \(G\) be an \((n-1)\)-edge ordered forest. The pairing \(\langle G, \sigma \circ \varphi_n \rangle\) has value 0 if \(\sigma\) reverses the order of any pair \(i,j\) such that \(i \to j\) is an edge of \(G\), and has value \(\text{sign}(\sigma)\) if \(\sigma\) preserves the order of every such pair.

**Proof.** It is equivalent and more intuitive to think of \(\sigma\) acting on \(G\) instead of \(\varphi_n\). Let \(\sigma(G)\) be the graph that has edge \(i \to j\) whenever \(G\) has an edge \(\sigma(i) \to \sigma(j)\), so that we have

\[
\langle G, \sigma \circ \varphi_n \rangle = \text{sign}(\sigma) \langle \sigma(G), \varphi_n \rangle.
\]

The sign comes from reordering the edges, which changes the orientation of the torus map.

First we show that the pairing has value 0 if \(\sigma\) reverses the order of some pair of vertices that are adjacent in \(G\). In this case \(\sigma(G)\) has an edge \(j \to i\) with \(i < j\). Using the fact that \(\sigma(1) = 1\), we consider the (directed) path in \(\sigma(G)\) from 1 to \(i\), and locate the greatest vertex \(k\) on this path. Then the two neighbors of \(k\) on this path are both less than \(k\); we denote them by \(a\) and \(b\) so that \(\sigma(G)\) has edges \(a \to k\) and \(k \to b\). Under \(\varphi_n\) the angle between the segments from disk \(a\) to disk \(k\) and from disk \(k\) to disk \(b\) never equals \(\pi\). (Disks \(a\) and \(b\) are contained in a medium-sized disk that the \(k\)th disk is outside.) Thus, the composition \(t_{\sigma(G)} \circ i_{n,\frac{1}{n}} \circ \varphi_n\) is not surjective, so it must have degree 0.
Then we show that the pairing $\langle \sigma(G), \varphi_n \rangle$ has value 1 if $\sigma$ preserves the order of every pair of edges that are adjacent in $G$—in other words, if $\sigma(G)$ is an ordered forest. In fact the corresponding map $(S^1)^{n-1} \to (S^1)^{n-1}$ is homotopic to the identity. The $j$th coordinate of the composition is the following angle: we find the unique $i$ such that $i \to j + 1$ is an edge of $\sigma(G)$, and take the unit vector pointing from the center of disk $i$ toward the center of disk $j$. This $j$th coordinate always forms an acute angle with the $j$th coordinate $\theta_j$ of the input, so we can homotope each coordinate to the identity map. Thus in this case $\langle \sigma(G), \varphi_n \rangle = 1$ and $\langle G, \sigma \circ \varphi_n \rangle = \text{sign}(\sigma)$. □

Proof of Theorem 15. There are $(n-1)!$ ordered forests with $n-1$ edges, because the unique parent of every vertex is an arbitrary lesser vertex. And, there are $(n-1)!$ permutations in $S_n$ fixing the element 1. So, it suffices to show that every element of the dual basis to the set of $(n-1)$-edge ordered forests is a $\mathbb{Z}$-linear combination of elements $\sigma \circ \varphi_n$.

The proof is by induction on $n$. The base case is $n = 2$, for which there is only one ordered forest and one permutation. Suppose $n > 2$. We introduce notation for the ways to add an $n$th vertex to an ordered forest or an $n$th element to a permutation. If $G$ is an ordered forest on $n-1$ vertices and $k$ is one of those vertices, then we let $G^{(k)}$ denote the ordered forest on $n$ vertices obtained by adding the edge $k \to n$ to $G$. Similarly, if $\sigma \in S_{n-1}$ is a permutation and $l$ is between 2 and $n$, we let $\sigma^{(l)}$ denote the permutation on $n$ elements such that

$$
\sigma^{(l)}(i) = \begin{cases} 
\sigma(i), & \text{if } i < n \text{ and } \sigma(i) < l; \\
\sigma(i) + 1, & \text{if } i < n \text{ and } \sigma(i) \geq l; \\
l, & \text{if } i = n;
\end{cases}
$$

Then we have

$$
\langle \sigma^{(l)}(G^{(k)}), \varphi_n \rangle = \begin{cases} 
\langle \sigma(G^{(k)}), \varphi_{n-1} \rangle, & \text{if } \sigma(k) < l; \\
0, & \text{if } \sigma(k) \geq l,
\end{cases}
$$
or equivalently,

$$
\langle G^{(k)}, \text{sign}(\sigma^{(l)}) \cdot \sigma^{(l)} \circ \varphi_n \rangle = \begin{cases} 
\langle G, \text{sign}(\sigma) \cdot \sigma \circ \varphi_{n-1} \rangle, & \text{if } \sigma(k) < l; \\
0, & \text{if } \sigma(k) \geq l.
\end{cases}
$$

For any ordered tree $G^{(k)}$ on $n$ vertices, we apply the inductive hypothesis to write the dual basis element $G^*$ as the linear combination

$$
G^* = \sum_{\sigma \in S_{n-1}, \sigma(1) = 1} a_\sigma \cdot (\text{sign}(\sigma) \cdot \sigma \circ \varphi_{n-1}),
$$

with coefficients $a_\sigma \in \mathbb{Z}$. Then the dual basis element $(G^{(k)})^*$ is the linear combination

$$
(G^{(k)})^* = \begin{cases} 
\sum_{\sigma \in S_{n-1}, \sigma(1) = 1} a_\sigma \cdot (\text{sign}(\sigma^{(k)}(k+1)) \cdot \sigma^{(k)}(k+1) \circ \varphi_{n-1} - \text{sign}(\sigma^{(k)}(k)) \cdot \sigma^{(k)}(k) \circ \varphi_n), & \text{if } k > 1; \\
\sum_{\sigma \in S_{n-1}, \sigma(1) = 1} a_\sigma \cdot (\text{sign}(\sigma^{(k)}(k+1)) \cdot \sigma^{(k)}(k+1) \circ \varphi_{n}), & \text{if } k = 1.
\end{cases}
$$

In this way, we see that the various homology classes $\sigma^{(l)} \circ \varphi_n$ do span the dual space to $H^{n-1}(\text{Conf}_n(D^2))$. □
By itself, Theorem 15 is not very powerful. It merely implies that \( i^*_{n,\frac{1}{n}} \) is injective, which we already know from Theorem 2. However, the same strategy can prove statements about \( \ker i^*_n \) for larger \( r \). We can construct other maps similar to \( \varphi_n \) that map a product of circles into \( \text{Conf}_{n,r}(D^2) \), and compute the pairings of the corresponding homology classes with the ordered forests. For instance, instead of having a medium-sized disk with \( n-1 \) small disks inside and one small disk outside, we could have a smaller medium-sized disk with \( n-2 \) small disks inside and two small disks fixed outside; this new map would produce an \((n-2)\)-dimensional homology class instead of an \((n-1)\)-dimensional class. Or, we could have multiple medium-sized disks of different sizes, each with some number of small disks moving inside.

**Question 2.** Is there a combinatorial description of these recursively constructed maps \((S^1)^j \to \text{Conf}_n(D^2)\) that makes it easy to pair their corresponding homology classes with the ordered forests?

For \( r \leq \frac{3}{2n+3} \), it seems that the only obstruction to the injectivity of \( i^*_n \) is the fact that no \( k \) disks can be collinear if \( r > \frac{1}{k} \). This statement is formalized in the following conjecture. For \( 2 \leq k \leq n \), let \( \Delta_k \) denote the “\( k \)-diagonal” in \( \text{Conf}_n(D^2) \) consisting of all configurations in which at least \( k \) of the \( n \) points are collinear, and let \( j_k \) denote the inclusion of \( \text{Conf}_n(D^2) \setminus \Delta_k \) into \( \text{Conf}_n(D^2) \). If \( r > \frac{1}{k} \), then \( \text{Conf}_{n,r}(D^2) \) is contained in \( \text{Conf}_n(D^2) \setminus \Delta_k \), so automatically we have

\[
\ker j^*_k \subseteq \ker i^*_n.
\]

**Conjecture 3.** Let \( r \leq \frac{3}{2n+3} \), and suppose that \( \frac{1}{k} < r \leq \frac{1}{k-1} \). Then we have

\[
\ker i^*_n = \ker j^*_k.
\]

**Question 4.** Let \( r \) and \( k \) be as above, so that \( \text{Conf}_{n,r}(D^2) \subseteq \text{Conf}_n(D^2) \setminus \Delta_k \). Is \( \text{Conf}_{n,r}(D^2) \) a deformation retract of \( \text{Conf}_n(D^2) \setminus \Delta_k \)?

**Question 5.** For general \( r \), is it possible to describe the homotopy type of \( \text{Conf}_{n,r}(D^2) \) only in terms of the balanced configurations of radius at most \( r \), as in Morse theory?

**Question 6.** How can Theorem 13 be generalized to a meaningful statement about all dimensions \( d \)?

5. **Naive conjecture and revised conjecture**

Larry Guth suggested the following conjecture in analogy with Theorem 7.3 of Gromov’s book [5]: that theorem is about homology of spaces of paths or loops with a given maximum length. We consider the degree-\( j \) part of the cohomology kernel \( \ker i^*_n \) for each \( j \). We define \( r_{\text{min}}(n,j) \) to be the infimal \( r \) such that \( i^*_n \) is not injective on \( H^j(\text{Conf}_n(D^2)) \), and define \( r_{\text{max}}(n,j) \) to be the supramal \( r \) such that \( i^*_n \) is not zero on \( H^j(\text{Conf}_n(D^2)) \). That is, the interval from \( r_{\text{min}}(n,j) \) to \( r_{\text{max}}(n,j) \) is where \( \ker i^*_n \) is changing in degree \( j \).

**Conjecture 7** (Naive conjecture). There exists a constant \( C > 1 \) such that for all \( n \) and \( j \), we have the bound

\[
\frac{r_{\text{max}}(n,j)}{r_{\text{min}}(n,j)} \leq C.
\]
In this section, we show in Lemmas 17 and 18 that this naive conjecture is false and replace it by a modified conjecture. Then we make some first observations about what would be needed to address the modified conjecture. In what follows, we use the notation $f \lesssim g$ to mean that $f$ is at most a constant times $g$, or $f = O(g)$. Likewise, $f \gtrsim g$ means that $f$ is at least a positive constant times $g$, or $f = \Omega(g)$, and $f \sim g$ means that the ratio between $f$ and $g$ is bounded between two positive constants, or $f = \Theta(g)$.

Lemma 17. We have

$$r_{\max}(2j,j) \gtrsim \frac{1}{\sqrt{j}}.$$ 

Proof. For $jr^2 \sim 1$, we can fit $j$ disjoint medium-sized disks of radius $2r$ inside the unit disk $D^2$. Then we can spin two small disks of radius $r$ inside each medium-sized disk, as depicted in Figure 7. That is, we define a map

$$f : (S^1)^j \to \text{Conf}_{2j,r}(D^2)$$

such that if $\vec{x} = f(\theta_1, \ldots, \theta_j)$, then for each $i = 1, 2, \ldots, j$ the points $x_{2i-1}$ and $x_{2i}$ are the centers of two tangent disks of radius $r$ inside the $i$th medium-sized disk, and the angle $\theta_i$ is equal to $x_{2i} - x_{2i-1}$.

Let $G$ be the ordered forest on $2j$ vertices with edges $(2i - 1) \to 2i$—in other words, a matching. Then the pairing of $G$ with the homology class corresponding to $f$ has value 1, so $i_{\ast,\gamma_1}(G) \neq 0$ and $r_{\max}(2j,j) \geq r$. \hfill \Box

Lemma 18. We have

$$r_{\min}(j + 1, j) = \frac{1}{j + 1}.$$ 

Proof. This is the content of Theorems 2 and 6 in the case of dimension 2. \hfill \Box

Because the ratio between $\frac{1}{j + 1}$ and $\frac{1}{\sqrt{j}}$ is unbounded as $j \to \infty$, this pair of lemmas implies that Conjecture 7 is false. We can express this argument informally by saying that the cohomology class corresponding to a matching requires much less empty space than the cohomology class corresponding to a path with
the same number of edges. The reason they require different amounts of space is that the matching has many small connected components, and the path has one large connected component. We modify the conjecture so that it only compares cohomology classes corresponding to ordered forests that all have the same list of sizes of connected components, as follows.

For each ordered forest $G$, we can list the number of vertices in each connected component of $G$, omitting the isolated vertices. The result is a multiset $m = \{m_1, \ldots, m_k\}$ of integers all at least 2. The number of edges in $G$ must be 

$$j(m) := \sum_{i=1}^{k} (m_i - 1).$$ 

We define $H^\infty(Conf_n(D^2))$ to be the subspace of $H^{j(m)}(Conf_n(D^2))$ spanned by the ordered forests for which the list of sizes of connected components is $m$—that is, the ordered forests shaped roughly like $G$. We define $r_{\text{min}}(n, m)$ to be the infimal $r$ such that $i_{n,r}^*$ is not injective on $H^m(Conf_n(D^2))$, and define $r_{\text{max}}(n, m)$ to be the supremal $r$ such that $i_{n,r}^*$ is not zero on $H^m(Conf_n(D^2))$.

**Conjecture 8** (Modified conjecture). There exists a constant $C > 1$ such that for all $n$ and all finite multisets $m$ of integers greater than 2, we have 

$$\frac{r_{\text{max}}(n, m)}{r_{\text{min}}(n, m)} \leq C.$$

In the remainder of this section, we make some observations about this modified conjecture. As a first step, we observe that it suffices to check only those cases in which there are no isolated vertices.

**Theorem 19.** For any $m$, we let $m = \sum_i m_i$. If we have 

$$\frac{r_{\text{max}}(m, m)}{r_{\text{min}}(m, m)} \leq C,$$

then for all $n \geq m$ we have 

$$\frac{r_{\text{max}}(n, m)}{r_{\text{min}}(n, m)} \leq 2C.$$

**Proof.** Let $r < \frac{1}{2r} \cdot r_{\text{max}}(n, m)$. We want to show $r < r_{\text{min}}(n, m)$—that is, that $i_{n,r}^*$ is injective on $H^m(Conf_n(D^2))$. Because $n \geq m$, we know $r_{\text{max}}(n, m) \leq r_{\text{max}}(m, m)$, and so $i_{n,2r}^*$ is injective on $H^m(Conf_m(D^2))$. We construct inclusions 

$$f_{S,r} : \text{Conf}_{m,2r}(D^2) \hookrightarrow \text{Conf}_{n,r}(D^2),$$

one for each subset $S \subset \{1, \ldots, n\}$ of size $m$, as in Figure 8. Specifically, we first fix an element $\tilde{y} \in \text{Conf}_{n-m,2r}(D^2)$; such an element exists because $2r < r_{\text{max}}(n, m)$ and so in particular $n$ disks of radius $2r$ must fit into the unit disk. Then for every $\tilde{x} \in \text{Conf}_{m,2r}(D^2)$, the configuration $f_{S,r}(\tilde{x}) \in \text{Conf}_{n,r}(D^2)$ is obtained by drawing two disjoint medium-sized disks of radius $\frac{1}{2}$ and putting a half-scaled copy of $\tilde{x}$ into one and a half-scaled copy of $\tilde{y}$ into the other. In the half-scaled copy of $\tilde{x}$, the labels of the disks are the elements of $S$ in order instead of 1, ..., $m$, and in the half-scaled copy of $\tilde{y}$, the labels are the $n - m$ numbers not in $S$.

We can extend each $f_{S,r}$ to a map 

$$f_S : \text{Conf}_m(D^2) \hookrightarrow \text{Conf}_n(D^2)$$
1
6
7
2
3
4
5
8
1
2
3
4
5
8

Figure 8. We can include $\text{Conf}_{m,2r}(D^2)$ into $\text{Conf}_{n,r}(D^2)$ by putting a half-scaled version in one medium-sized disk and putting the other $n - m$ disks in fixed locations in another medium-sized disk.

by the same construction, using the same fixed $\vec{y}$. The induced map on cohomology

$$f_S^*: H^*(\text{Conf}_n(D^2)) \to H^*(\text{Conf}_m(D^2))$$

is the projection to the span of the ordered forests for which all non-isolated vertices are in $S$. Therefore we have an isomorphism

$$\bigoplus_S f_S^*: H^m(\text{Conf}_n(D^2)) \to \bigoplus_S H^m(\text{Conf}_m(D^2)),$$

where the direct sum in the target consists of $\binom{n}{m}$ copies of the same space, one for each $S$.

In order to show that $i^*_n$ is injective on $H^m(\text{Conf}_n(D^2))$, we compare the two equal compositions

$$\left(\bigoplus_S f_S^*\right) \circ i^*_n : H^m(\text{Conf}_n) \to H^*(\text{Conf}_n,r) \to \bigoplus_S H^*(\text{Conf}_{m,2r})$$

and

$$i^*_{m,2r} \circ \left(\bigoplus_S f_S^*\right) : H^m(\text{Conf}_n) \to \bigoplus_S H^m(\text{Conf}_m) \to \bigoplus_S H^*(\text{Conf}_{m,2r}).$$

Both maps of the second composition are injections, so both maps of the first composition must be injections.

This Theorem 19 implies that it suffices to study $r_{\min}(m, m)$ and $r_{\max}(m, m)$. Thus we denote $r_{\min}(m, m)$ and $r_{\max}(m, m)$ by $r_{\min}(m)$ and $r_{\max}(m)$. The next theorem is a lower bound on $r_{\min}(m)$ which we prove using a similar argument of scaling configurations to fit into medium-sized disks.

**Theorem 20.** For any $m$ we have

$$r_{\min}(m) \gtrsim \frac{1}{\sqrt{\sum_i m_i^2}}.$$

That is, for $\sum_i (rm_i)^2 \sim 1$, the map $i^*_{m,r}$ is injective on $H^m(\text{Conf}_m(D^2))$, where $m = \sum_i m_i$ as before. The condition $\sum_i (rm_i)^2 \sim 1$ corresponds to the condition
that we can fit medium-sized disks of radii \( r_m \) inside the unit disk. For completeness we prove this fact in the following lemma.

**Lemma 21.** Given \( k \) disks of radii \( r_1 \geq \cdots \geq r_k \), they can be translated to fit inside a disk of radius \( R \) with

\[
R^2 \leq 36 \sum_{i=1}^{k} r_i^2.
\]

**Proof.** We use induction on \( k \). For \( k = 1 \) we can take \( R = r_1 \). For \( k > 1 \), we start with a configuration of the first \( k - 1 \) disks inside a disk of smallest possible radius \( R_{k-1} \). If there is an empty space big enough to fit the \( k \)th disk, then we are done because, by the inductive hypothesis, we have

\[
(R_{k-1})^2 \leq 36 \sum_{i=1}^{k-1} r_i^2 < 36 \sum_{i=1}^{k} r_i^2.
\]

Otherwise, as in the Vitali covering lemma, if we scale the first \( k - 1 \) disks by a factor of 3 they cover the disk of radius \( R_{k-1} \), so in fact we have

\[
(R_{k-1})^2 \leq 9 \sum_{i=1}^{k-1} r_i^2.
\]

Then because \( r_k \leq r_1 \leq R_{k-1} \), the disk of radius \( 2R_{k-1} \) is big enough to fit both the disk of radius \( R_{k-1} \) and the disk of radius \( r_k \) inside it, so we take \( R = 2R_{k-1} \) and we have

\[
R^2 = (2R_{k-1})^2 \leq 36 \sum_{i=1}^{k-1} r_i^2 < 36 \sum_{i=1}^{k} r_i^2.
\]

\( \Box \)

**Proof of Theorem 20.** Let \( r \) be such that we can fit disjoint medium-sized disks of radii \( r m_1, \ldots, r m_k \) inside the unit disk; we fix such a configuration of medium-sized disks. We want to show that \( i_m^* \) is injective on \( H^\bullet(\text{Conf}_m(D^2)) \).

Let \( S = \{ S_1, \ldots, S_k \} \) be a partition of \( \{ 1, \ldots, m \} \), such that each subset \( S_i \) has size \( m_i \). For each such \( S \), we construct an injection

\[
f_{S, r} : \prod_i \text{Conf}_{m_i, \frac{r}{m_i}}(D^2) \hookrightarrow \text{Conf}_{m, r}(D^2),
\]

as follows. Let \( (\vec{x}_1, \ldots, \vec{x}_k) \) be an arbitrary element of \( \prod_i \text{Conf}_{m_i, \frac{r}{m_i}}(D^2) \). Then \( f_{S, r}(\vec{x}_1, \ldots, \vec{x}_k) \) is a configuration in \( \text{Conf}_{m, r}(D^2) \) such that for each \( i \)th medium-sized disk, the configuration inside that medium-sized disk is a scaled copy of \( \vec{x}_i \), such that the disks are relabeled to have the values in \( S_i \). The scale factor on the \( i \)th configuration \( \vec{x}_i \) is \( r m_i \), so that the resulting disks all have radius \( r \).

The remainder of the proof is just like the proof of Theorem 19. The map \( f_{S, r} \) may be extended to a map

\[
f_{S} : \prod_i \text{Conf}_{m_i}(D^2) \to \text{Conf}_{m}(D^2),
\]

by the same formula. The induced cohomology map

\[
f_{S}^* : H^\bullet(\text{Conf}_{m}(D^2)) \to H^\bullet \left( \prod_i \text{Conf}_{m_i}(D^2) \right)
\]
corresponds to the projection of $H^*(\text{Conf}_m(D^2))$ onto its subspace generated by the ordered forests in which the connected components are contained in the sets $S_i$. Thus, there is an isomorphism

$$\bigoplus S \xrightarrow{f^*_S} H^m(\text{Conf}_m(D^2)) \rightarrow \bigoplus H^j(m) \left( \prod_i \text{Conf}_{m_i}(D^2) \right),$$

where the direct sum is taken over all partitions $S = \{S_1, \ldots, S_k\}$ such that $|S_i| = m_i$ for all $i$, and $j(m)$ denotes the number of edges, $\sum_i (m_i - 1)$.

In order to show that $i^*_m$ is injective on $H^m(\text{Conf}_m(D^2))$, we compare the two equal compositions

$$\left( \bigoplus S \xrightarrow{f^*_S} \right) \circ i^*_m : H^m(\text{Conf}_m) \rightarrow H^*(\text{Conf}_{m,r}) \rightarrow \bigoplus S \xrightarrow{H^*(\text{Conf}_{m_i})} \longrightarrow \bigoplus S \xrightarrow{H^j(m)} \left( \prod_i \text{Conf}_{m_i} \right).$$

Both maps of the second composition are injections, so both maps of the first composition must be injections. \hfill \Box

**Question 9.** Theorem 20 says that 

$$r_{\min}(m) \gtrsim \frac{1}{\sqrt{\sum_i m_i^2}}.$$

Is the reverse inequality also true?

We might even hope to prove the following conjecture, which would answer both this question and Conjecture 8.

**Conjecture 10 (Strong conjecture).** For any $m$ we have

$$r_{\max}(m) \lesssim \frac{1}{\sqrt{\sum_i m_i^2}},$$

which implies

$$r_{\min}(m) \sim r_{\max}(m) \sim \frac{1}{\sqrt{\sum_i m_i^2}}.$$

So far, we have only two upper bounds on $r_{\max}(m)$. One bound is

$$r_{\max}(m) \leq \frac{1}{\max_i m_i},$$

because if for some $m_i$ it is impossible to have $m_i$ collinear disks, then every ordered forest with a connected component of size $m_i$ becomes zero. The other bound is

$$(r_{\max}(m))^2 \leq \frac{1}{\sum_i m_i},$$

because if $r^2 \cdot \sum_i m_i$ is greater than 1 then the $m$ disks cannot fit into the unit disk at all.
**Question 11.** Can we prove any upper bound for $r_{\text{max}}(m)$ stronger than the bound

$$r_{\text{max}}(m) \lesssim \min \left( \frac{1}{\max_i m_i}, \frac{1}{\sqrt{\sum_i m_i}} \right)?$$

6. **Question about segments**

In the next two sections we consider a similar problem that ought to be easier than determining the asymptotic behavior of $r_{\text{min}}(m)$ and $r_{\text{max}}(m)$. Let $\text{Seg}_{n,r}(D^2)$ denote the space of configurations of $n$ disjoint labeled oriented line segments of length $r$ in the unit disk. Each configuration can be specified by the center of each segment, along with a unit vector indicating the direction of that segment. In this way every space $\text{Seg}_{n,r}(D^2)$ includes into the space $\text{Conf}_n(D^2)$, which we denote by $\text{Seg}_n(D^2)$. Just like on $\text{Conf}_n(D^2)$, there is a tautological function on $\text{Seg}_n(D^2)$ that indicates, for each configuration in $\text{Seg}_n(D^2)$, the supremal length $r$ such that the configuration is in $\text{Seg}_{n,r}(D^2)$.

There is a “torus map”

$$t : \text{Seg}_n(D^2) \to (S^1)^n$$

given by projection to the second factor. We denote the inclusion of $\text{Seg}_{n,r}(D^2)$ into $\text{Seg}_n(D^2)$ by

$$i_{n,r} : \text{Seg}_{n,r}(D^2) \hookrightarrow \text{Seg}_n(D^2).$$

Then we can examine the induced map on cohomology

$$i_{n,r}^* \circ t^* : H^*((S^1)^n) \to H^*(\text{Seg}_{n,r}(D^2)),$$

and in particular we can ask whether the fundamental cohomology class of $(S^1)^n$ pulls back to a nonzero class in $H^*(\text{Seg}_{n,r}(D^2))$. Let $r_{\text{crit}}(n)$ denote the threshold value of $r$, below which the pullback class $i_{n,r}^* t^* [(S^1)^n]$ is nonzero and above which it is zero. In some sense $r_{\text{crit}}(n)$ describes for which $r$ it is possible to spin $n$ line segments of length $r$ “independently” in the unit disk. The main question is to determine the asymptotic behavior of $r_{\text{crit}}(n)$.

**Proposition 22.** For all $n$ we have

$$r_{\text{crit}}(n) \gtrsim \frac{1}{\sqrt{n}}.$$

*Proof.* For $r^2 n \sim 1$ it is possible to fit $n$ medium-sized disks of diameter $r$ into the unit disk. For such $r$ there is a map

$$f : (S^1)^n \to \text{Seg}_{n,r}(D^2)$$

that spins each segment in its own medium-sized disk, such that the composition $t \circ i_{n,r} \circ f$ is the identity map on $(S^1)^n$. Then the homology class corresponding to $f$ has a nonzero pairing with the cohomology class $i_{n,r}^* t^* [(S^1)^n]$, so that cohomology class must be nonzero. Thus $r \leq r_{\text{crit}}(n)$. □

The following conjecture corresponds to the strong conjecture about disks, Conjecture 10.

**Conjecture 12.** For all $n$ we have

$$r_{\text{crit}}(n) \lesssim \frac{1}{\sqrt{n}}.$$
Figure 9. The $n$-segment configuration $\psi_n(\theta_1, \ldots, \theta_n)$ is constructed by placing the $n$th segment tangent to a scaled copy of the $(n-1)$-segment configuration $\psi_{n-1}(\theta_1, \ldots, \theta_{n-1})$. Each segment has length $\ell_n$.

which implies

$$r_{\text{crit}}(n) \sim \frac{1}{\sqrt{n}}.$$

In the remainder of this section, we construct a sequence of nice maps $\psi_n$ much like the sequence of disk maps $\varphi_n$. Each map $\psi_n$ corresponds to a nonzero homology class of dimension $n$, and even though the segments are quite long their lengths are still proportional to $\frac{1}{\sqrt{n}}$, giving some evidence for Conjecture 12 above.

Figure 9 depicts the construction of the maps $\psi_n$. The lengths of the segments appearing in the various $\psi_n$ form a sequence $\ell_n$ that we compute later. As with $\varphi_n$, we construct the map

$$\psi_n : (S^1)^n \to \text{Seg}_{n, \ell_n}(D^2)$$

recursively in $n$, as follows. The goal is for the composition $t \circ i_{n, \ell_n} \circ \psi_n$ to be the identity map on $(S^1)^n$.

For $n = 1$ we put $\ell_1 = 2$ so that $\psi_1(\theta_1)$ is the configuration with one segment equal to a diameter at angle $\theta_1$. For $n > 1$, we would like to construct the configuration $\psi_n(\theta_1, \ldots, \theta_n)$. Inside the unit disk, we draw a medium-sized disk of radius $\frac{\ell_n}{\ell_{n-1}}$ and draw the $n$th segment of length $\ell_n$ tangent to the medium-sized disk and with endpoints on the boundary of the unit disk. We also require that the vector along the $n$th segment points counterclockwise along the boundary of the medium-sized disk and has angle $\theta_n$. Then inside the medium-sized disk we draw an $\frac{\ell_n}{\ell_{n-1}}$-scaled copy of $\psi_{n-1}(\theta_1, \ldots, \theta_{n-1})$, which has $n-1$ segments of length $\ell_{n-1}$. The resulting picture is $\psi_n(\theta_1, \ldots, \theta_n)$.

This procedure recursively determines the sequence $\ell_n$ of the lengths of the segments in $\psi_n$. The construction implies a bound on $r_{\text{crit}}(n)$: the composition $t \circ i_{n, \ell_n} \circ \psi_n$ is the identity map on $(S^1)^n$, so the pullback class $i_{n, \ell_n}^* t^*[(S^1)^n]$ must be nonzero and we have the inequality

$$\ell_n \leq r_{\text{crit}}(n).$$
Figure 10. The sequence of diameters $d_n$ satisfies the recursion $d_n = d_{n-1} + \frac{1}{d_{n-1}}$. This picture is a rescaled version of the construction of $\psi_n$; the radius 1 and the segment length $\ell_n$ are replaced by the radius $\frac{1}{2}d_n$ and the segment length 2.

However, the next proposition shows that $\ell_n$ is asymptotically no larger than the bound already proven in Proposition 22.

**Proposition 23.** The sequence $\ell_n$ of lengths appearing in the map $\psi_n$ satisfies

$$\ell_n \sim \frac{1}{\sqrt{n}}.$$

**Proof.** The computation is a little easier if we rescale by a factor $\frac{2}{\ell_n}$. Instead of spinning segments of length $\ell_n$ in a disk of radius 1, we imagine spinning segments of length 2 in a disk of radius $\frac{2}{\ell_n}$, or of diameter $\frac{1}{\ell_n}$, which we denote as $d_n$. We compute $\ell_n$ by computing $d_n$.

Figure 10 shows how a computation with similar triangles implies the recursion

$$d_n = d_{n-1} + \frac{1}{d_{n-1}}.$$

To determine the asymptotic behavior, we prove the following claim: for all $n$, we have

$$2(n+1) \leq (d_n)^2 \leq 3(n+1).$$

The base case $n = 1$ has $d_1 = 2$. Then $d_1^2$ is equal to 4, which is indeed between $2(n+1) = 4$ and $3(n+1) = 6$. Suppose $n > 1$. Squaring both sides of the recursion we obtain

$$d_n^2 = d_{n-1}^2 + 2 + \frac{1}{d_{n-1}^2}.$$

Because $d_{n-1}^2 > 1$ for all $n$, the quantity $2 + \frac{1}{d_{n-1}^2}$ is strictly between 2 and 3, and so

$$d_{n-1}^2 + 2 < d_n^2 < d_{n-1}^2 + 3,$$

and we obtain the desired inequality by plugging in the inductive hypothesis.

Thus $d_n \sim \sqrt{n}$ and $\ell_n \sim \frac{1}{\sqrt{n}}$. \qed
It is reasonable to hope that the lengths $\ell_n$ are close to the true values of $r_{\text{crit}}(n)$. In the case $n = 2$ we do have $\ell_2 = r_{\text{crit}}(2)$, because if two segments are longer than $\ell_2$ it is impossible to arrange them at right angles. But for all $n > 2$ it is possible to deform the map $\psi_n$ slightly to make the segments longer.

7. Partial progress for segments

We would at least like to prove that
\[ \lim_{n \to \infty} r_{\text{crit}}(n) = 0, \]
but it is not clear how to do this. In Theorems 24 and 26 we exhibit two strategies that do not help; the first is to determine whether $t \circ i_{n,r}$ is surjective, and the second is to determine whether there is a connected component of $\text{Seg}_{n,r}(D^2)$ on which $t \circ i_{n,r}$ is surjective. Then in Theorem 29 we show a partial result:
\[ \lim_{n \to \infty} r_{\text{crit}}(n) < 1. \]

**Theorem 24.** There exists $\varepsilon > 0$ such that for all $n$, the map
\[ t \circ i_{n,1+\varepsilon} : \text{Seg}_{n,1+\varepsilon}(D^2) \to (S^1)^n \]
is surjective.

The main part of the proof is contained in the following lemma.

**Lemma 25.** For sufficiently small $\varepsilon > 0$, there exists a configuration of infinitely many segments of length $1 + \varepsilon$ in $D^2$ such that for each angle in $[0, \pi)$, one segment has that angle.

**Proof.** We start with the upper half of the disk of radius $1 + \varepsilon$ and make two radial cuts, at some angle $\theta$ and also at $\pi - \theta$, so that the resulting three pieces can be translated to fit disjointly inside $D^2$. We take the translated radii of the $(1 + \varepsilon)$-halfdisk as the infinitely many segments of our configuration.

Specifically, the construction goes as follows, depicted in Figure 11. We start by choosing two horizontal line segments $s_0$ and $s_\pi$ of length 1 in the lower half of $D^2$. Coordinatizing $D^2$ as a subset of $\mathbb{R}^2$, we label the rightmost and leftmost points of the boundary by $A = (1, 0)$ and $B = (-1, 0)$. We let $C$ be the point $(0, -\delta)$ for some small $\delta > 0$, just below the origin. Then we set $1 + \varepsilon$ to be the distance from $C$ to $A$ or $B$, and set $\theta$ to be the angle of the vector from $C$ to $A$. The sector with center $C$ and radius $1 + \varepsilon$, from angle $\theta$ to angle $\pi - \theta$, sweeps from $A$ to $B$ inside $D^2$.

As for the two other sectors, the sector with angles from 0 to $\theta$ fits into a small neighborhood of the segment $s_0$, and the sector with angles from $\pi - \theta$ to $\pi$ fits into a small neighborhood of segment $s_\pi$. If $\delta$ is chosen to be sufficiently small, then the three pieces all fit disjointly inside $D^2$. \qed

**Proof of Theorem 24.** We use the same $\varepsilon$ as in the lemma. Let $(\theta_1, \ldots, \theta_n)$ be an arbitrary element of $(S^1)^n$. We want to construct an element of $\text{Seg}_{n,1+\varepsilon}(D^2)$ that maps to $(\theta_1, \ldots, \theta_n)$ under $t \circ i_{n,1+\varepsilon}$—that is, its segments are at angles $\theta_1, \ldots, \theta_n$. If no two angles among $\theta_1, \ldots, \theta_n$ are equal or opposite, then it is easy to find a preimage: we use the configuration from the lemma, reversing the orientations of some segments as appropriate so that their angles are in $[\pi, 2\pi)$ instead of $[0, \pi)$.

Otherwise, we use induction on $n$. Suppose that $\theta_n$ is equal or opposite to $\theta_{n-1}$. By the inductive hypothesis, there is an element of $\text{Seg}_{n-1,1+\varepsilon}(D^2)$ with angles
For small enough $\varepsilon$, the upper half disk of radius $1 + \varepsilon$ can be cut into three pieces that can be translated to fit into the unit disk.

$\theta_1, \ldots, \theta_{n-1}$. In this configuration, some small neighborhood of segment $n - 1$ does not intersect any other segment, so we may insert segment $n$ very close to segment $n - 1$ and parallel, to make an element of $\text{Seg}_{n,1+\varepsilon}(D^2)$ with angles $\theta_1, \ldots, \theta_n$. □

Theorem 24 does not completely rule out the possibility of using surjectivity to find an upper bound on $r_{\text{crit}}(n)$. If $\text{Seg}_{n,r}(D^2)$ consists of several connected components, and the image of each component under $t \circ i_{n,r}$ does not cover $(S^1)^n$, then the pullback class $i_{n,r}^* t^* [(S^1)^n]$ is zero and so $r \geq r_{\text{crit}}(n)$. However, the next theorem shows that this argument does not work for any $r < 1$.

**Theorem 26.** For all $\varepsilon > 0$ and all $n$, there is a connected component of $\text{Seg}_{n,1-\varepsilon}(D^2)$ on which the restriction of $t \circ i_{n,1-\varepsilon}$ is surjective onto $(S^1)^n$.

**Proof.** First we select a preimage under $t \circ i_{n,1-\varepsilon}$ for most of the points of $(S^1)^n$. Let $\text{Conf}_n(S^1)$ (as in earlier sections of the paper) denote the subspace of $(S^1)^n$ consisting of all elements for which no two coordinates are the same. We construct a continuous map

$$s : \text{Conf}_n(S^1) \to \text{Seg}_{n,1-\varepsilon}(D^2)$$

that takes each point of $\text{Conf}_n(S^1)$ to one of its $(t \circ i_{n,1-\varepsilon})$-preimages, as follows. For each $(\theta_1, \ldots, \theta_n) \in \text{Conf}_n(S^1)$, we let $s(\theta_1, \ldots, \theta_n)$ be the configuration such that for each $i$ the $i$th segment points from $\varepsilon^2 \cdot \theta_i$ to $(1 - \varepsilon^2) \cdot \theta_i$, where we view $\theta_i \in S^1$ as a unit vector in $\mathbb{R}^2$. The $i$th segment points outward along a radius of $D^2$.

Lemmas 27 and 28, which follow, complete the proof by showing that there is a connected component of $\text{Seg}_{n,1-\varepsilon}(D^2)$ that contains the image of $s$ and maps surjectively onto $(S^1)^n$. □

**Lemma 27.** Let the map

$$s : \text{Conf}_n(S^1) \to \text{Seg}_{n,1-\varepsilon}(D^2)$$
be defined as above. The images of the different connected components of \( \text{Conf}_n(S^1) \) under \( s \) are all contained in the same connected component of \( \text{Seg}_{n,1-\varepsilon}(D^2) \).

**Proof.** Each connected component of \( \text{Conf}_n(S^1) \) corresponds to one of the ways to order \( n \) points around the circle, so it suffices to connect the \( s \)-images of two elements of \( \text{Conf}_n(S^1) \) that differ by one swap of consecutive angles. A path connecting these \( s \)-images is depicted in Figure 12. It is constructed by pushing the first segment across the disk, turning both segments so that their angles swap, and then pushing the first segment back. \( \square \)

**Lemma 28.** For every point of \((S^1)^n \setminus \text{Conf}_n(S^1)\), its preimage under \( t \circ i_{n,1-\varepsilon} \) contains a point in the same connected component of \( \text{Seg}_{n,1-\varepsilon}(D^2) \) that contains \( s(\text{Conf}_n(S^1)) \).

**Proof.** Let \((\phi_1, \ldots, \phi_n)\) be an arbitrary element of \((S^1)^n \setminus \text{Conf}_n(S^1)\). We perturb it slightly to find a nearby point \((\theta_1, \ldots, \theta_n) \in \text{Conf}_n(S^1)\). We construct a path in \( \text{Seg}_{n,1-\varepsilon}(D^2) \) from \( s(\theta_1, \ldots, \theta_n) \) to an element of the preimage of \((\phi_1, \ldots, \phi_n)\) under \( t \circ i_{n,1-\varepsilon} \) in the following way.

For the \( i \)th segment of \( s(\theta_1, \ldots, \theta_n) \), we keep its inner endpoint fixed near the center of \( D^2 \) and pivot the rest of the segment so that its angle changes at a constant rate from \( \theta_i \) to \( \phi_i \). This process is depicted in Figure 13. If \((\theta_1, \ldots, \theta_n)\) is close enough to \((\phi_1, \ldots, \phi_n)\), then the segments do not cross while pivoting, so the path stays in \( \text{Seg}_{n,1-\varepsilon}(D^2) \). \( \square \)

The two previous theorems show which strategies will not work to prove that \( r_{\text{crit}}(n) \) decreases to zero. The next theorem shows that \( r_{\text{crit}}(n) \) decreases below 1; its proof strategy may be useful for proving stronger upper bounds on \( r_{\text{crit}}(n) \). The rest of this section contains the proof of the theorem and ends with some questions about configurations of segments.

**Theorem 29.** We have

\[
\lim_{n \to \infty} r_{\text{crit}}(n) < 1.
\]

The proof relies on the Lusternik-Schnirelmann theorem, which is stated as follows. A proof can be found on pages 2 and 3 of the book [4].

**Theorem 30** (Lusternik-Schnirelmann theorem). Let \( X \) be a topological space. Suppose that \( \alpha_1 \) and \( \alpha_2 \) are cohomology classes in \( H^*(X) \), and suppose that \( A_1 \) and \( A_2 \) are closed subsets of \( X \). If the restrictions \( \alpha_1|_{A_1^c} \) and \( \alpha_2|_{A_2^c} \) to the complements
of $A_1$ and $A_2$ are zero, then the restriction $(\alpha_1 \sim \alpha_2)|_{(A_1 \cap A_2)^c}$ of the cup product to the complement of the intersection $A_1 \cap A_2$ is also zero.

The Lusternik-Schnirelmann theorem allows us to prove the following curious lemma, which implies that we can remove an arbitrarily large finite set of points from the unit disk $D^2$ without changing the limit of $r_{\text{crit}}(n)$. In what follows, for $i = 1, \ldots, n$ we let $d\theta_i$ denote the cohomology class in $H^1(\text{Seg}_n(R^2))$ corresponding to the angle of the $i$th segment. In this language, $r_{\text{crit}}(n)$ is the threshold at which the restriction $(d\theta_1 \wedge \cdots \wedge d\theta_n)|_{\text{Seg}_n,r(U)}$ changes between zero and something nonzero.

**Lemma 31.** Suppose that $r > 0$ and $U \subseteq R^2$ are such that for all $n$, we have

$$(d\theta_1 \wedge \cdots \wedge d\theta_n)|_{\text{Seg}_n,r(U)} \neq 0.$$ 

Then for any finite set of points $S \subseteq U$, we also have

$$(d\theta_1 \wedge \cdots \wedge d\theta_n)|_{\text{Seg}_n,r(U \setminus S)} \neq 0.$$

**Proof.** It suffices to prove the statement where $S$ is a single point $p \in U$. Suppose for contradiction that there is some large $n$ such that

$$(d\theta_1 \wedge \cdots \wedge d\theta_n)|_{\text{Seg}_n,r(U \setminus \{p\})} = 0.$$ 

We apply the Lusternik-Schnirelmann theorem (Theorem 30), taking $X$ to be $\text{Seg}_{2n,r}(U)$, taking $\alpha_1$ to be $d\theta_1 \wedge \cdots \wedge d\theta_n$, taking $\alpha_2$ to be $d\theta_{n+1} \wedge \cdots \wedge d\theta_{2n}$, taking $A_1$ to be the subset of $\text{Seg}_{2n,r}(U)$ where one of the segments $1, \ldots, n$ passes through $p$, and taking $A_2$ to be the subset where one of the segments $n + 1, \ldots, 2n$ passes through $p$. By the contradiction hypothesis we have

$$\alpha_1|_{A_1^c} = 0, \quad \alpha_2|_{A_2^c} = 0,$$

and because $A_1$ and $A_2$ do not intersect, this implies

$$(d\theta_1 \wedge \cdots \wedge d\theta_{2n})|_{\text{Seg}_{2n,r}(U)} = 0,$$

which is a contradiction. 

The next lemma shows that removing points as in the previous lemma can restrict the movement of long segments.
Lemma 32. Let $U$ be an infinite horizontal strip, and suppose that $r > 0$ is greater than half the height of the strip. Then for any $\delta > 0$, there is a discrete set of points $S \subseteq U$ such that any vertical segment of length $r$ confined to $U \setminus S$ must stay in the vertical strip of width $\delta$ centered on the segment.

Proof. Let $U$ be the strip defined in coordinates by

$$U = \{(x, y) \in \mathbb{R}^2 : |y| < 1\}.$$

First suppose that we start with a vertical segment along the $y$-axis, and remove from $U$ the four points with coordinates $(\pm a, \pm b)$. Then as long as the segment is more than a certain length—namely $\sqrt{(a + 1)^2 + (b + 1)^2}$—it is trapped in an hourglass-shaped region, shown in Figure 14.

Given $r > 1$ and $\delta > 0$, we choose $S$ in the following way. First we choose the ratio $\frac{a}{b}$ small enough that the width $2\frac{a}{b}$ of the corresponding hourglass is less than $\frac{\delta}{2}$ and its diagonal $2\sqrt{\left(\frac{a}{b}\right)^2 + 1}$ is less than $2r$. Then we choose $a$ and $b$ in that ratio, small enough that the length $\sqrt{(a + 1)^2 + (b + 1)^2}$ is less than $r$. We take $S$ to be the set of points $\{(2k + 1)a, \pm b\}_{k \in \mathbb{Z}}$, so that no matter where we place a vertical segment of length $r$, it is trapped by the two points of $S$ immediately to its left and the two points of $S$ immediately to its right. \[\square\]

The next lemma completes the construction needed for the proof of Theorem 29.

Lemma 33. For all $\varepsilon > 0$, there exist $r < 1$ and a finite set $S$ of points in the unit disk $D^2$ such that the following is true. Let $A$ be the set of configurations in $\text{Seg}_{1,r}(D^2)$ for which the segment is in the vertical strip $[-\varepsilon, \varepsilon] \times \mathbb{R}$ and the midpoint of the segment is in the box $[-\varepsilon, \varepsilon] \times [-\varepsilon, \varepsilon]$. Then we have

$$d\theta_1|_{\text{Seg}_{1,r}(D^2 \setminus S) \setminus A} = 0.$$

Proof. To show that the restriction $d\theta_1|_{\text{Seg}_{1,r}(D^2 \setminus S) \setminus A}$ is zero, it suffices to show

$$t_* \pi_1(\text{Seg}_{1,r}(D^2 \setminus S) \setminus A) = 0,$$

where

$$t: \text{Seg}_{1,r}(D^2 \setminus S) \setminus A \to S^1$$
denotes the map that records the angle of the segment. If this criterion on fundamental group is satisfied, then the map \( t \) lifts to the universal cover \( \mathbb{R} \) of \( S^1 \), and so \( t \) is null-homotopic.

Let us say for a given configuration in \( \text{Seg}_{1,r}(D^2 \setminus S) \setminus A \) that a segment is “trapped” if every loop in \( \text{Seg}_{1,r}(D^2 \setminus S) \setminus A \) containing that configuration maps to a loop in \( S^1 \) that is null-homotopic. Our goal is to construct \( S \) so that every vertical segment is trapped; this then implies that every segment is trapped.

We apply Lemma 32 three times. First we define a middle strip and remove points from it so that the vertical segments on either side of \([-\varepsilon, \varepsilon] \times \mathbb{R}\) are trapped. Then we define upper and lower strips and remove points so that the vertical segments in the upper and lower parts of \([-\varepsilon, \varepsilon] \times \mathbb{R}\) are trapped.

Specifically, we choose \( \varepsilon_1, \varepsilon_2 \) close to \( \varepsilon \) with \( 0 < \varepsilon_1 < \varepsilon_2 < \varepsilon \). We define the middle strip \( U_{\text{middle}} \) to be the smallest horizontal strip containing \( D^2 \setminus ([-\varepsilon_1, \varepsilon_1] \times \mathbb{R}) \), which has height less than 2. Then, applying Lemma 32 we find a length \( r_{\text{middle}} < 1 \) and a set of points \( S_{\text{middle}} \subseteq D^2 \) such that if a vertical segment of length \( r_{\text{middle}} \) starts in \( D^2 \setminus ([-\varepsilon_2, \varepsilon_2] \times \mathbb{R}) \) and moves in \( U_{\text{middle}} \setminus S_{\text{middle}} \), then it is trapped and stays in \( U_{\text{middle}} \setminus ([-\varepsilon_1, \varepsilon_1] \times \mathbb{R}) \).

We define the upper strip \( U_{\text{upper}} \) to be the horizontal strip with boundaries \( y = -\frac{1}{2} + \varepsilon \) and \( y = 1 \), and define the lower strip \( U_{\text{lower}} \) to have boundaries \( y = -1 \) and \( y = \frac{1}{2} - \varepsilon \), so that if a segment of length less than 1 has midpoint above \( y = \varepsilon \) it is contained in the upper strip, and if its midpoint is below \( y = -\varepsilon \) it is in the lower strip. We find a length \( r_{\text{upper}} < 1 \) and a set of points \( S_{\text{upper}} \subseteq D^2 \) such that if a vertical segment of length \( r_{\text{upper}} \) starts in the part of the upper strip in \([-\varepsilon_2, \varepsilon_2] \times \mathbb{R}\) and moves in \( U_{\text{upper}} \setminus S_{\text{upper}} \), then it is trapped and stays in \([-\varepsilon, \varepsilon] \times \mathbb{R}\). Similarly we find \( r_{\text{lower}} \) and \( S_{\text{lower}} \).

We set \( r = \max(r_{\text{middle}}, r_{\text{upper}}, r_{\text{lower}}) \) and \( S = S_{\text{middle}} \cup S_{\text{upper}} \cup S_{\text{lower}} \). Then every vertical segment is trapped: if it is outside \([-\varepsilon_2, \varepsilon_2] \times \mathbb{R}\) then it is trapped by the middle strip, if it is in \([-\varepsilon_2, \varepsilon_2] \times \mathbb{R}\) with midpoint above \( \varepsilon \) then it is trapped by the upper strip, and if it is in \([-\varepsilon_2, \varepsilon_2] \times \mathbb{R}\) with midpoint below \(-\varepsilon\) then it is trapped by the lower strip.

**Proof of Theorem 29.** By Lemma 31 it suffices to find \( r < 1 \) and a finite set of points \( S \subseteq D^2 \) such that we have

\[
(d\theta_1 \wedge d\theta_2)|_{\text{Seg}_{2,r}(D^2 \setminus S)} = 0.
\]

We apply Lemma 33 once to the first segment, and then again to the second segment with a right-angle rotation. That is, for small \( \varepsilon > 0 \) we let \( A_1 \) be the set of configurations in \( \text{Seg}_{2,r}(D^2) \) for which segment 1 is in the vertical strip \([-\varepsilon, \varepsilon] \times \mathbb{R}\) and its midpoint is in the box \([-\varepsilon, \varepsilon] \times [-\varepsilon, \varepsilon]\), and let \( A_2 \) be the configurations for which segment 2 is in the horizontal strip \( \mathbb{R} \times [-\varepsilon, \varepsilon]\) and its midpoint is in \([-\varepsilon, \varepsilon] \times [-\varepsilon, \varepsilon]\). If \( \varepsilon \) is small and \( r \) is not too small, then \( A_1 \) and \( A_2 \) are disjoint because on their intersection the two segments would have to cross.

Lemma 33 gives \( r < 1 \) and two finite sets of points \( S_1, S_2 \subseteq D^2 \) such that

\[
d\theta_1|_{\text{Seg}_{2,r}(D^2 \setminus S_1) \setminus A_1} = 0, \quad d\theta_2|_{\text{Seg}_{2,r}(D^2 \setminus S_2) \setminus A_2} = 0.
\]

Taking \( S = S_1 \cup S_2 \), we apply the Lusternik-Schnirelmann theorem (Theorem 30) to obtain

\[
(d\theta_1 \wedge d\theta_2)|_{\text{Seg}_{2,r}(D^2 \setminus S)} = 0.
\]

□
Question 13. How can the idea of balanced configurations be adapted to find the values of $r$ for which $\text{Seg}_{n,r}(D^2)$ changes homotopy type?

Question 14. How can we compute (perhaps numerically) $r_{\text{crit}}(n)$ for small values of $n$? For instance, what is $r_{\text{crit}}(3)$?

Question 15. Can we adapt the proof of Theorem 29 to show that

$$\lim_{n \to \infty} r_{\text{crit}}(n) = 0?$$

8. Conclusion

It appears that almost all possible questions about this subject are open. The most ambitious questions stated in this paper are Conjectures 10 and 12, which are the strongest possible upper bounds on the disk radius $r_{\text{max}}(m)$ and the segment length $r_{\text{crit}}(n)$. The questions that seem the most approachable are the following two: Question 2, which asks how to construct homology classes for $\text{Conf}_{n,r}(D^2)$ when $r$ is small relative to $n$; and Question 14, which asks how to define balanced configurations in $\text{Seg}_{n,r}(D^2)$.

REFERENCES


MIT, CAMBRIDGE, MA 02139 USA

E-mail address: hcalpert@math.mit.edu