Applied topology in static and dynamic sensor networks

Jennifer Gamble
Electrical and Computer Engineering
North Carolina State University
jpgamble@ncsu.edu

Harish Chintakunta
Electrical and Computer Engineering
North Carolina State University
hkchinta@ncsu.edu

Hamid Krim
Electrical and Computer Engineering
North Carolina State University
ahk@ncsu.edu

***Abstract*—In the study of sensor networks, many applications require topological analysis, and for some problems topological information is even sufficient. Here, we review how algebraic topology (and specifically simplicial homology theory) can be used as a general framework for detection of coverage holes in a coordinate-free sensor network. Extensions to distributed processing and localization algorithms are also reviewed, before progressing into discussion of a new way to apply algebraic topological methods to the analysis of coverage properties in dynamic sensor networks. Zigzag persistent homology is a recently developed method to track homological features (such as holes) over a sequence of spaces. This paper demonstrates the promise of this method for the identification of coverage holes in a time-varying coordinate-free sensor network, as well as the designation of coverage holes as significant or not, based on the length of time they are present in the sequence.**

I. Introduction

Wireless sensor networks have attracted much attention in the past two decades, with a host of applications [8]. In many of these scenarios the sensors are reasonably simple devices, and one aims to extract potentially complicated information about the sensing environment by intelligently combining together many local measurements. In particular, to avoid the necessity of including GPS functionality on each sensor, algorithms which are executable in a coordinate-free setting are particularly important. Analysis of topological features is a necessary step in most problems in this domain, and in some cases, all that is needed. Given a topological space, (ex., subset of $\mathbb{R}^2$, a graph, etc), the features of interest may be the number of connected components, cycles or "holes", or some higher dimensional equivalents. Such features may be studied using tools from an area of mathematics called algebraic topology.

In brief, algebraic topology is a field associated with assigning algebraic objects to topological spaces, and then studying how the assigned algebraic objects change when we change the topological space in a specific way. A good introduction may be found in [6]. One problem in sensor networks which has shown to be particularly well-suited [4] to an algebraic topological formulation is that of detecting (and even localizing) coverage holes. These tools give a way to take local, coordinate-free information about which sensors are "neighbors" (i.e. lie within some communication radius of each other), and determine global properties about the presence or absence of holes in the coverage area of the network.

We will first review some key concepts from algebraic topology, especially simplicial homology theory, and discuss how the topology of sensor networks can be analyzed within this general framework. In Section III-A, we will illustrate some existing results on coverage hole detection, including extensions to hole localization, and distributed computations.

Thus far, such methods have been applied to static sensor networks with great success, but are not able to track coverage holes which persist over time in dynamic networks. In Section III-B we show how a recently developed tool from computational algebraic topology (called zigzag persistent homology) can be leveraged for these purposes. A brief outline of the theory behind this method is given in Section II-B, with full exposition available in [1].

II. Algebraic Topology

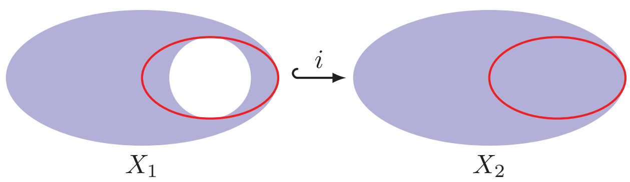
A. Simplicial Homology

Algebraic topology works with general topological spaces, but for computational purposes, discrete objects called simplicial complexes are often used to approximate these spaces and make calculations more tractable. Figure 1(a) shows a topological space (left). A simplicial complex representation of it is shown in Figure 1(b) (left). The simplicial complex has the same topology as the space (left) in Figure 1(a), in the sense that both are surfaces with a hole. The simplicial complex has the following convenient features: 1) It is made of simple pieces (simplices): points, edges and triangles, 2) It is completely described by specifying how to glue these pieces together (edges with vertices, triangles with edges, etc) and 3) The "gluing" may be expressed algebraically with linear operators.

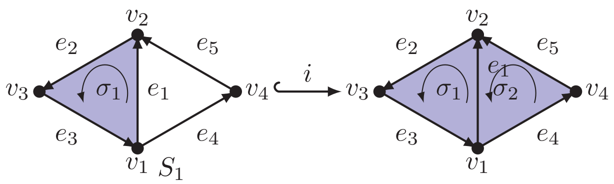
The simplices are named by their dimension, nodes are 0-simplices, edges are 1-simplices, triangles 2-simplices and so on. An edge is glued to two distinct vertices (to avoid self loops), a triangle is glued to 3 distinct edges (pairwise adjacent), and in general, a k-simplex is glued to $k+1$ $(k-1)-$simplices. To express this construction algebraically, we build a sequence $\{C_k\}$, of vector spaces called chain spaces. [1] The space C_k is abstractly generated by considering each $k-$simplex as a basis vector, and an element in this

[1] chain spaces may be defined as groups in general. We restrict ourselves to vector spaces, as they are sufficient for the applications.

978-1-4673-2014-6/12/$31.00 ©2012 IEEE



(a) inclusion of a topological space



(b) inclusion of the corresponding simplicial complexes

Fig. 1. Mappings between topological spaces can be represented as mapping between corresponding simplicial complexes.

vector space is a linear combination of these simplices (a chain). Each simplex is also endowed with an orientation. The orientation for the simplices is shown by arrows and arcs in Figure 1(b). A $k-$simplex is glued to the $(k-1)-$simplices along its canonical boundary. This gluing is represented by linear operators $\partial_k : C_k \to C_{k-1}$, called boundary operators, whose action is dictated by the orientation on the simplices. For example, $\partial_1(e_1) = v_2 - v_1$, $\partial_2(\sigma_1) = e_1 + e_2 + e_3$ and $\partial_2(\sigma_2) = e_4 + e_5 - e_1$.

A $k-$cycle is defined as a vector in the null space of ∂_k (i.e. chains whose boundary is zero). When $k = 1$, the resulting cycles are cycles in the usual sense (i.e. closed loops). In Figure 1(b), the edges e_1, e_2, e_3 form a cycle. As a simple demonstration, the action of ∂_1 on these edges is given as $\partial_1(e_1 + e_2 + e_3) = v_2 - v_1 + v_3 - v_2 + v_1 - v_3 = 0$. Holes on a surface are cycles which are "empty" inside. In other words, a 1-cycle is not a hole if it can be expressed as a boundary of some 2-simplices. The range of ∂_2 is exactly the space containing the boundaries of 2-dimensional simplices. This idea of "cycles which are not boundaries" is captured algebraically by the *homology spaces* which are the quotient spaces $H_k = ker(\partial_k)/Img(\partial_{k+1})$. Any element in $ker(\partial_k)$ which is also in $Img(\partial_{k+1})$ is considered to be equivalent to zero (these are cycles which *are* boundaries, thus do not encircle a hole), and any two vectors whose difference is in $Img(\partial_{k+1})$ are equivalent (i.e. homologous) to each other.

In Figure 1(b), for the simplicial complex S_1 (on the left), the space of cycles ($ker(\partial_1)$) is spanned by the vectors $e_1 + e_2 + e_3$ and $e_4 + e_5 - e_1$, and the space of boundaries of 2-simplices ($Img(\partial_2)$) is spanned by $e_1 + e_2 + e_3$. Therefore, H_1 is spanned by a single basis vector, $e_4 + e_5 - e_1$. It is not a coincidence that the dimension of H_1, also called the first Betti number [2], is equal to number of holes in the surface. It is true in general, that for a given surface, the first Betti number is equal to the number of holes [3]. For simplicial complexes, because the chain spaces and boundary operators have linear algebraic representation, topological computations (such as computing Betti numbers) reduce to matrix calculations.

In addition to analyzing topological features of spaces as described above, algebraic topology also provides powerful tools to understand the changes in these features under some changes to the space. Figure 1(a) shows one such situation where a hole in X_1 is filled in to form X_2. When there are multiple holes present, it is important to identify which particular hole was filled in, a problem which can be non-trivial in the absence of coordinate information. However, given inclusion maps between the spaces, such information can be easily obtained from the maps between the homology spaces (induced by the inclusion maps). Consider the inclusion map $i : X_1 \hookrightarrow X_2$ where $i(x) = x$ and the induced map $i_* : H_1(X_1) \to H_1(X_2)$. The cycle c (shown in red) in X_1 surrounding the hole represents a non-zero homology class in X_1 (this homology class, denoted $[c]$, represents the equivalence class of all homologous cycles to c). However, the cycle $i(c)$ in X_2 is contractible and zero in homology. In general, given any cycle c in X_1 for which $[c] \neq 0$, we say c persists upon inclusion into X_2 if $i_*([c]) \neq 0$. Such ideas are the basis for the theory of persistent homology.

B. *Zigzag persistent homology*

Persistent homology [9] is a method which takes a nested sequence of spaces

$$\emptyset = X_0 \subseteq X_1 \subseteq X_2 \subseteq \ldots, \subseteq X_n \tag{1}$$

and tracks the associated homology spaces

$$0 \to H_k(X_1) \to H_k(X_2) \to \ldots \to H_k(X_n)$$

(where the maps between homology spaces are induced by inclusion maps between the X_i's). Here, the homology can be tracked for all $k = 0, \ldots, \dim(X_n)$, but for the upcoming applications in sensor networks H_1 will be of predominant concern. A given homology class is said to be "born" at time i if it is present in X_i but was not present in X_{i-1} (i.e. is not in the image of $i_* : H_1(X_{i-1}) \to H_1(X_i)$). Similarly, a homology class $[c]$ present in X_i "dies" at time $i+1$ if it is no longer present in X_{i+1} (i.e. if $i_* : H_1(X_i) \to H_1(X_{i+1})$ gives $i_*([c]) = 0$). Therefore, each homology class is present over some interval in the nested sequence, which begins at its birth time and ends at its death time.. Persistent homology is a method devised to calculate these intervals, and thereby identify important homological features in the sequence, which "persist" over a long interval.

The original idea behind this method was to estimate the homology of a manifold, when only observing a discrete sample of points from it. For example, when given a discrete sample of 100 points from an annulus (shown in Figure 2, top left), one would like to form a simplicial complex using

[2]The dimension of H_k is called the k^{th} Betti number.

[3]In general, this is equal to the number of non-homologous, non-contractible cycles

these points as vertices in such as way as to capture the true topology of the underlying space. One way to form a simplicial complex from a point set is to fix some distance, ϵ, and build a k-simplex out of any $(k+1)$ points that are all pairwise within ϵ from each other. This is called the Vietoris-Rips complex, and will be discussed again in Section III. The resulting complex will depend on the ϵ chosen, and a nested sequence of simplicial complexes, as in Equation 1 can be formed by choosing increasing values of ϵ. For the annulus example, simplicial complexes with $\epsilon = 0, 0.2, 0.4, 0.6$ and 0.8 are shown (where the annulus has inner radius of 1 and outer radius of 2). The main hole in the annulus is present as a homology class for ϵ values ranging from about 0.45 to 1.9. Smaller holes can be seen in both complexes in the second row of Figure 2, and it seems that these are due to the nature of the discrete sampling. The set of intervals representing the birth time and death time of the homology features during the sequence can be plotted in a barcode (as in Figure 2, bottom right), with longer intervals representing features that are considered to be inherent or "significant", and short intervals corresponding to features appearing due to sampling "noise".

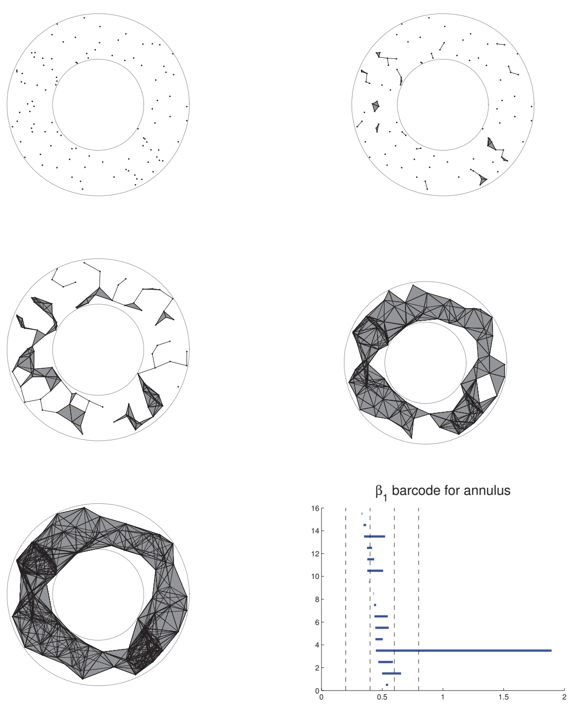


Fig. 2. 100 points sampled from an annulus in $\mathbb{R}^2$ (top left), and the Vietoris-Rips complex formed using $\epsilon = 0.2$ (top right), $\epsilon = 0.4$ (middle left), $\epsilon = 0.6$ (middle right), and $\epsilon = 0.8$ (bottom left). The barcode of the intervals representing degree-1 homology classes (calculated using persistent homology) is shown in the bottom right.

Zigzag persistent homology [1] extends this notion to situations where the sequence of spaces $X_1, \ldots, X_n$ are not necessarily nested (as they were in Equation 1). One is still interested in obtaining intervals indicating the birth and death times of homology classes in the sequence, but in this case the index is not a parameter such as ϵ, but simply the index $i = 1, \ldots, n$. Since the sequence of spaces is not nested, there are not inclusion maps from one space into the next from which to induce the corresponding homology spaces. Zigzag persistent homology addresses this problem by mapping two consecutive spaces into a common intermediate space (such as the union of the two spaces), and then computes the homology persisting through this "zigzag" sequence. This is illustrated in Figure 3. In fact, the methods behind zigzag persistence work for any general sequence of "forward" and "backward" inclusion maps between spaces (not only for alternating forward and backward maps into union spaces).

An example using continuous spaces is shown in Figure 4. Observing the homology of X_1, X_2, and X_3 individually, one sees that each of the three spaces has exactly two holes. By computing zigzag persistent homology using the inclusion maps from X_1 and X_2 into $X_1 \bigcup X_2$ and from X_2 and X_3 into $X_2 \bigcup X_3$, one is able to tell that there is in fact one hole which is present across all spaces in the sequence, while the other holes are only temporary (and present only in their individual spaces). The barcode corresponding to this would have one long interval from X_1 through X_3 representing the persistent hole, and three short intervals, each starting at X_i for $i = 1, 2, 3$, and ending immediately afterward.

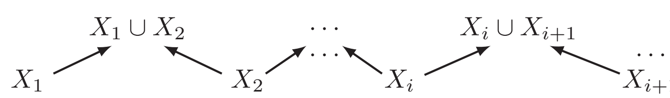


Fig. 3. A sequence of spaces, with "forward" and "backward" inclusion maps into the unions of each consecutive pair of spaces

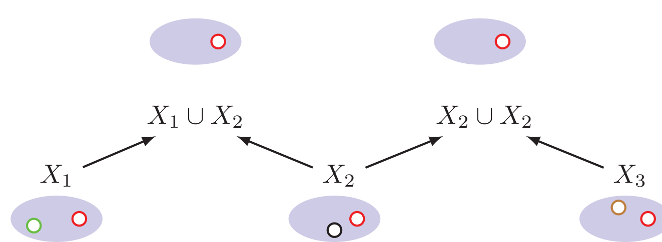


Fig. 4. A sequence of spaces with one hole persisting throughout.

The details of the persistent homology and zigzag persistent homology algorithms may be found in [9] and [1] respectively, but it is worth noting that in practice the spaces are simplicial complexes, and all the computations are performed using matrix calculations on vector spaces representing the chain spaces (as in Section II-A). In fact, the forward and backward inclusion maps can be broken down into a series of single-simplex additions and deletions, from which the matrix representations are updated, and the birth and death times of the homology classes are obtained.

III. Applications in Sensor Networks

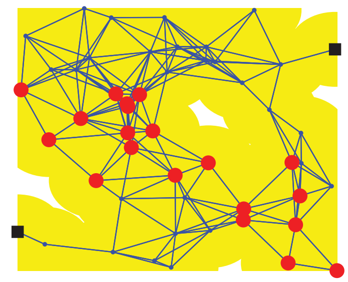
A. Stationary Networks

Researchers who study sensor networks are accustomed to representing them using a graph, with links between nodes indicating, for example, communication ability. Representing a sensor network using a simplicial complex (which includes not only nodes and edges but also triangles and potentially higher-dimensional simplices) allows the network to be subjected to topological analysis. Some widely known problems with straight-forward topological implications include coverage holes and worm hole attacks.

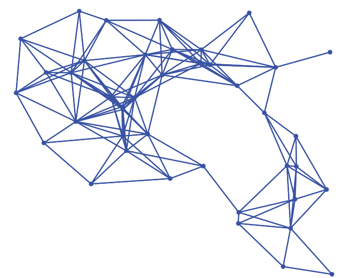
Consider a given sensor network deployed in some region of $\mathbb{R}^2$, and denote R_c^i as the coverage area of the i^{th} node, and $R_c = \cup_i R_c^i$ the total coverage area. A coverage hole is a hole in the surface R_c, which corresponds to a region which is not within the scope of any of the sensors. The coverage area for each node is defined as the set of points within some "coverage radius" of that node (i.e. a disc-shaped region centered at the node). For example, Figure 5(a) shows a network with the total coverage area in yellow (shaded).

For the same network, one can build its communication graph, by connecting together by an edge any two nodes which lie within communication range of each other (a distance defined by some fixed "communication radius"). A simplicial representation of the space may be obtained directly from the communication graph [5], by defining a $k-$clique in the communication graph (a set of k nodes which are all pairwise connected by an edge) as a $(k-1)-$simplex. This forms the Vietoris-Rips complex, often simply called the Rips complex, mentioned in Section II-B. Simplicial homology computations (as in Section II-A) on this complex can be performed to determine topological features. For example, a first Betti number of zero indicates no holes in the complex, and given information about the relationship between the coverage radius and communication radius, one can make some conclusions about the topology of R_c (including criteria for guaranteed coverage). Many different tools from algebraic topology have been applied to analyze the topology of R_c using the Rips complex [4], [3]. An important point to note is that the communication graph is all that is required as input for these algorithms. This consists of binary information about which nodes are connected by an edge, and does not contain any coordinates or edge lengths. This information can be stored as local lists at each node of which other nodes are its "neighbors" (i.e. which it is connected to by an edge).

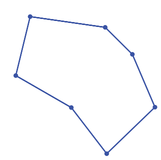
These tools have been combined with gossip algorithms to lead to efficient distributed algorithms for analyzing topology [2], [7]. Additionally, a divide and conquer algorithm for localizing coverage holes (after detection) was developed in [2], which involves an iterative partitioning procedure which hones in on the edges surrounding a hole, while preserving the topology within each partition (see reference for details). Figure 5(a) shows the partition boundary (red dots) obtained by flooding from two diameter nodes (black squares), and Figures 5(b) and 5(c) give the graph after 1 and 7 iterations (respectively) of the partitioning procedure.



(a) Communication Graph superimposed on the coverage area



(b) after partition 1



(c) after partition 7

Fig. 5. Figure showing the sequence of surviving subgraph with each partition

B. Dynamic Networks

The applications of the previous section are applicable to static sensor networks, and are able to achieve good results in that setting, but there are many real-world situations where the static assumption is unrealistic or undesirable. The sensors could be purposely mobile, programmed to move about randomly over the region of deployment, or they could be spatially affected by environmental factors. Given similar constraints as above (the communication graph at each time point is the only input to the algorithms, and is coordinate-free), it is desirable to not only detect coverage holes at each time point, but to track them, and determine whether any given hole is persisting over time.

The tools of zigzag persistent homology (outlined in Section II-B, and detailed in [1]) give a way to take a series of discrete-time snapshots of the network, and map the associated homology groups through this sequence of spaces to detect long-lasting coverage holes. In a setting where the sensors are mobile, one would expect that coverage holes that appear and then disappear again quickly are due to random movement of the nodes, while coverage holes that are continually present over longer time periods are due to some underlying problem (such as an obstruction, interference, or node/communication failures).

Here a time-varying sensor network is modeled by randomly deploying 100 sensors uniformly over the $[0, 1]$ unit square, and with an obstruction or failure region is defined by inhibiting communication for all sensors within the box 0.35 to 0.65 in both the x- and y-coordinates. The coverage region of one such network, and its corresponding simplicial complex are shown in Figure 6 (failure region shown as dotted square). The simplicial complex is formed using a Rips complex with radius 0.2 on the nodes outside of the failure region.

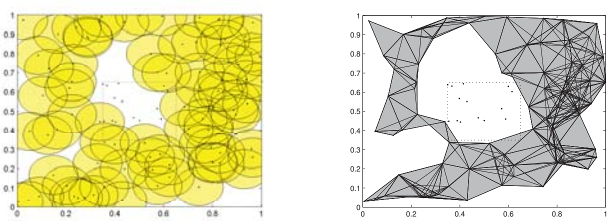


Fig. 6. Coverage area of a sensor network with an obstruction in the center (left), and its corresponding simplicial complex representation (right).

The time-varying nature of the network is modeled by allowing each sensor to move a small random distance (using a bivariate normal distribution with mean 0 and standard deviation 0.04) at each time step. At any given time point in the evolution of this complex, there may be a number of holes present, with some of them due to the random placement of the sensors at that moment, and with the "main" hole due to a true underlying problem (a region where communication is impossible). Zigzag persistent homology can be used to distinguish between these transient features and the persistent obstruction, while using only coordinate-free input at each time point (about which sensors connect to each other by an edge). For example, Figure 7 shows the simplicial complexes for a network at two consecutive time steps (X_1 and X_2), which each mapping into the union $X_1 \bigcup X_2$. Although both X_1 and X_2 have spurious holes due to the node placement, the only hole which persists through all three spaces (X_1, $X_1 \bigcup X_2$, and X_2) is the one corresponding to the failure region. This can be seen by observing the two holes on the left in X_i persist until $X_i \bigcup X_{i+1}$, but aren't present in X_{i+1}, and the two holes on the right aren't present in $X_i \bigcup X_{i+1}$. So the only hole that persists all the way through is the one surrounding the failure region.

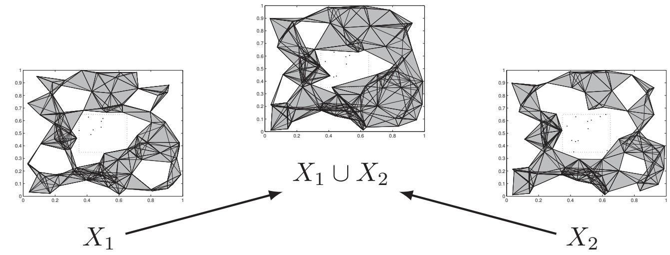


Fig. 7. Simplicial complexes representing a sensor network at time 1 and time 2, with the maps from the network at each of the the two time points into the union of the two simplicial complexes.

IV. CONCLUSION

Algebraic topology provides a useful framework for the analysis of topological properties of sensor networks. One area where this proves particularly useful is in the study of coverage properties where only local neighborhood information is available (without coordinates). Additionally, representing a sensor network as a simplicial complex gives access to the tools of simplicial homology theory, where homological features can be calculated explicitly using only linear algebra.

Zigzag persistent homology allows these properties to be observed in time-varying coordinate-free networks as well, which allows for additional judgements to be made not only about the presence or absence of coverage holes, but their relative importance, as judged by temporal persistence.

REFERENCES

[1] G. Carlsson, V. de Silva, and D. Morozov. Zigzag persistent homology and real-valued functions. *Proc. 25th Annual Symposium on Computational Geometry (SoCG)*, pages 247–256, June 2009.

[2] Harish Chintakunta and Hamid Krim. Divide and Conquer: Localizing Coverage Holes in Sensor Networks. *Proceedings of IEEE SECON*, pages 359–366, 2010.

[3] V. De Silva and R. Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7:339–358, 2007.

[4] Vin de Silva and Robert Ghrist. Coordinate-free coverage in sensor networks with controlled boundaries via homology. *The International Journal of Robotics Research*, 25(12):1205–1222, 2006.

[5] Robert Ghrist and Abubakr Muhammad. Coverage and hole-detection in sensor networks via homology. *Proceedings of the 4th international symposium on Information processing in sensor networks*, 2005.

[6] Allen hatcher. *Algebraic Topology*. Cambridge University Pr, 2002.

[7] Abubakr Muhammad and Ali Jadbabaie. Decentralized Computation of Homology Groups in Networks by Gossip. *American Control Conference*, 2007.

[8] M. Saleem, G.A. Di Caro, and M. Farooq. Swarm intelligence based routing protocol for wireless sensor networks: Survey and future directions. *Information Sciences*, 181(20):4597–4624, 2011.

[9] A. Zomorodian and G. Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33:249–274, 2005.