Classifying Field Lines with Topological Data Analysis

Nicholas Bohlsen, Vanessa Robins, and Matthew Hole

Applied Topology and Geometry Group

Fusion Plasma Theory and Modelling Group

Australian National University

December 13, 2023

Outline

Objective

Stellarator physics is inherently geometric and chaotic.

We can produce more simulations of plasma configurations under different configurations than we can analyse by hand.

Recent work (past 25 years) in computational topology (Robins, Edelsbrunner, Kaczynski, et. al [1–3]) has produced a series of tools for the analysis of the shape of datasets and demonstrated its use in the study of non-linear dynamical systems (Ghrist, Kaczynski, Tempelton, Kramar et. al [4–6]).

Aim to develop applications of these tools for the automated extraction of topological features from plasma physics simulations.

Field line classification as a first proof of concept

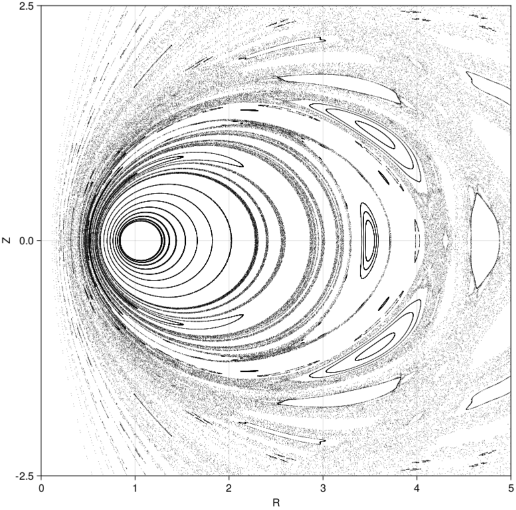


Figure 3: Diagram of many field line orbits of a toy field.

In Poincare sections of fields we usually see chains of islands of order and KAM toruses embedded in a stochastic (chaotic) see.

Problem Statement:

Can we automatically determine the *class* of a field line only from its orbit under a Poincare map?

Aim to capture the intuitive segmentation a person would construct on such a field.

The Shape of Data

Intuitively some data has "shape" or "topology".

Such as the hole in $\longrightarrow$

TDA seeks to make this notion mathematically rigorous.

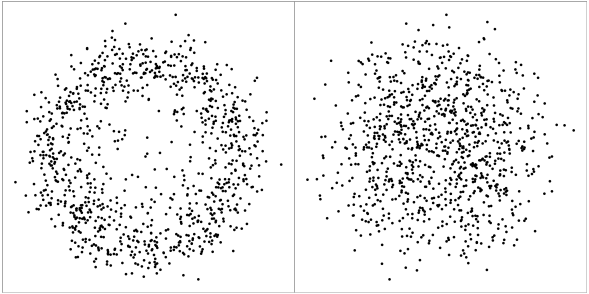


Figure 4: Two example point clouds.

- Convert data into a sequence of *simplicial complexes*. (Study how the data changes with scale)
- Study how the *topology* (really *homology*) changes over the sequence.
- Long-lived topological features represent geometric features in data.

Vietoris-Rips (VR) Complex

Definition

Let $\epsilon \in [0, \infty)$ *and* $X = \{x_0, \ldots x_n\}$ *be a point cloud then* $VR_\epsilon(X)$ *is a simplicial complex defined by*

$$\langle x_{k_1} \ldots x_{k_m} \rangle \in VR_\epsilon(X) \Leftrightarrow d(x_i, x_j) \leq \epsilon \text{ for all } i, j \in k_1, \ldots k_m.. \quad (1)$$

As ϵ varies from 0 to ∞ this defines a sequence of nested simplicial complexes $K_0 \hookrightarrow K_1 \hookrightarrow \ldots \hookrightarrow K_N = VR_\infty(X)$.

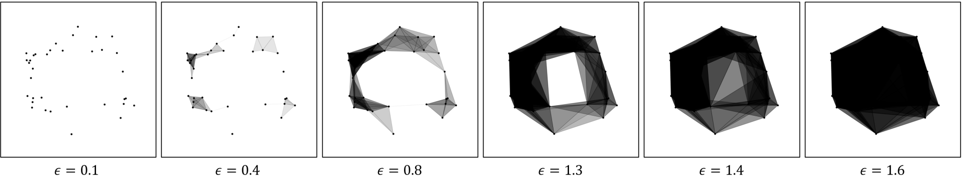


Figure 6: A subset of VR complexes from a filtration

Homology

Homology is a topological invariant counting the n-d holes in a space.

A hole (class) is a loop which is "closed but not a boundary".

Assigns a family of abelian groups H_n to the space with generator which represent topological features

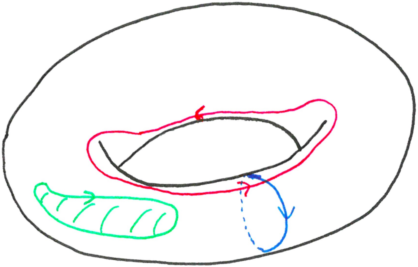
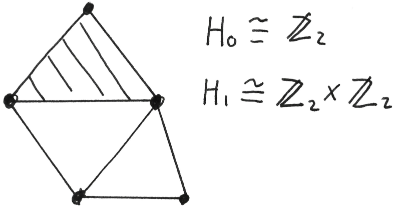


Figure 7: Diagrammatic introduction to Homology



Simplicial Homology Review

Suppose K is a simplicial complex we can construct abelian groups $C_n(K, \mathbb{Z}_2)$ as formal linear combinations of $n-$simplexes in K. We can these chain groups and there elements are

$$c \in C_n(K, \mathbb{Z}_2) \implies c = \sum_{\sigma_i^n \in K^n} a_i \sigma_i^n \,. \tag{2}$$

Homology theory tells us that we can construct a chain complex

$$\cdots \to C_n(K, \mathbb{Z}_2) \xrightarrow{\partial_n} C_{n-1}(K, \mathbb{Z}_2) \xrightarrow{\partial_{n-1}} \cdots \,, \tag{3}$$

where the boundary maps are defined by summing the co-dimension 1 faces of a $n-$simplex σ^n [7, 8].

Some examples

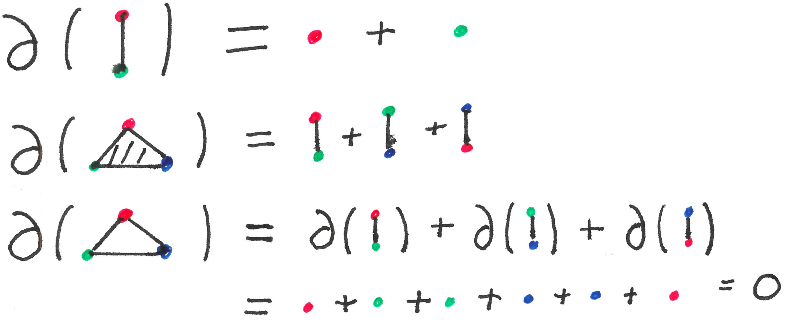


Figure 8: Diagrammatic representation of the action of the boundary operator on simplicial complexes.

Observe core fact that "boundaries are always closed"

$$\partial_{n-1}\partial_n = 0\,. \tag{4}$$

Fundamental Idea of Homology

Holes in spaces are related to chains which are "closed" but not "boundaries". Sometimes referred to as "cycles" which are not boundaries.

We formulate this in terms of the quotient group

$$H_n(K, \mathbb{Z}_2) = \frac{\ker \partial_n}{\operatorname{Im} \partial_{n+1}}, \tag{5}$$

which we refer to as the n-th homology groups.

2 cycles $c, c' \in C_n(K, \mathbb{Z}_2)$ belong to the same equivalence class (are homologous) in H_n if they differ by a boundary, that is if

$$c' = c + \partial b, \tag{6}$$

where $b \in C_{n+1}(K, \mathbb{Z}_2)$, then $[c] = [c']$ and they are associated the same hole.

Persistent Homology

Throughout our filtration the homology of the complex will change as we are analysing the dataset on different scales. Persistent homology is the process of tracking this change and recording topological features (homology classes) which are persistent, that is, exist for many different radii.

We convert our sequence of complexes to a sequence of simplicial homology groups by appealing to functoriality

$$H_n(K_0) \to H_n(K_1) \to \ldots \to H_n(K_N)\,. \tag{7}$$

A class in $H_n(K_i)$ is defined to *persist* if its image in $H_n(K_{i+1})$ is non-zero, *die* if its image is zero, and be *born* if it is not in the image of $H_n(K_{i-1})$.

The formal mathematics of this is left for another time.

Example barcode

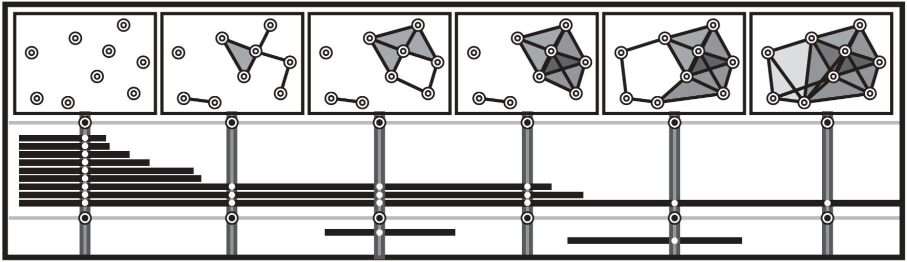


Figure: Barcode for an example point cloud. Reproduced from [9].

For the following section we will mostly work with the persistence diagram (scatter plot of birth and death times) instead of the barcodes.

Persistent Homology

The homology of the VR complex will change ϵ varies [3].

Classes will be born at some diameter b and may die at a diameter d.

The set of (b, d) pairs is the *Persistent Homology PH*.

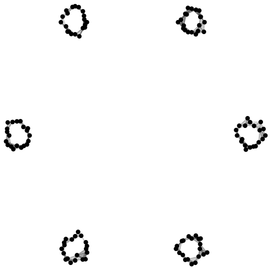


Figure 8a: $VR_{0.1}$ for a point cloud.

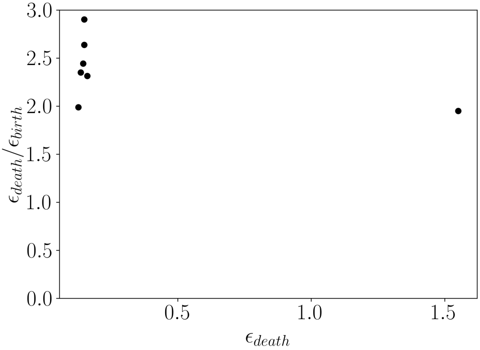


Figure 8b: Relative persistence of point cloud.

Concept for automated orbit classification

The finite time $T \in \mathbb{N}$ orbit of a point $x \in \Sigma$ under the Poincare map is a point cloud $X_T(x) \subset \Sigma$.

The persistent homology of this point cloud encodes information about the geometry (and topology) of the orbit.

Aim to develop a list of criteria using $PH(X_T(x))$ which can specify to which class (island chain, KAM torus, stochastic layer) of Hamiltonian orbit x belongs.

Note that this is currently very non-rigorous

Toy model for a tokamak device

To develop a computational procedure we need an example of the field line geometry.

We make a toy model for a perturbed tokamak device from an infinite line current and a circular loop.

Now, lets analyse some data.

Poincare map orbits of field lines specifically.

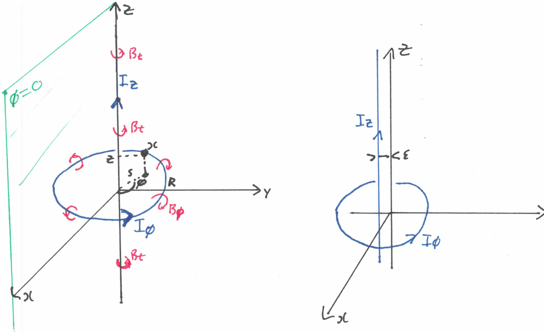


Figure 9: Toy tokamak currents

Orbit on a KAM torus

Rips persistence diagram and H_1 class relative persistence is as you would expect for a KAM torus

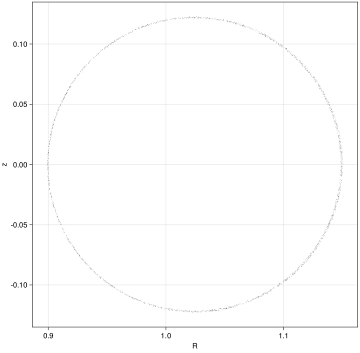
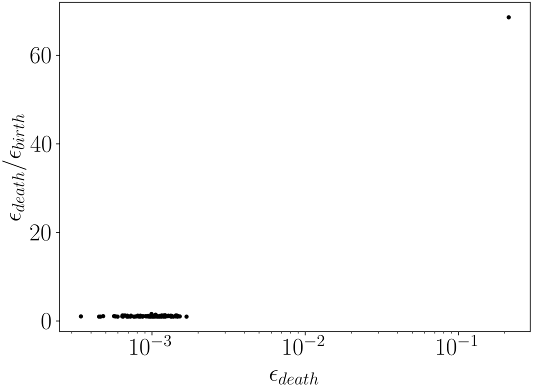


Figure 10: Poincare map orbit of KAM torus



Island chain

For an island chain, geometrically largest hole has low relative persistence.

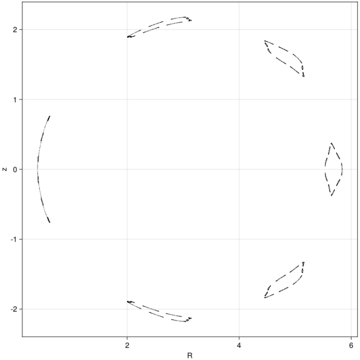
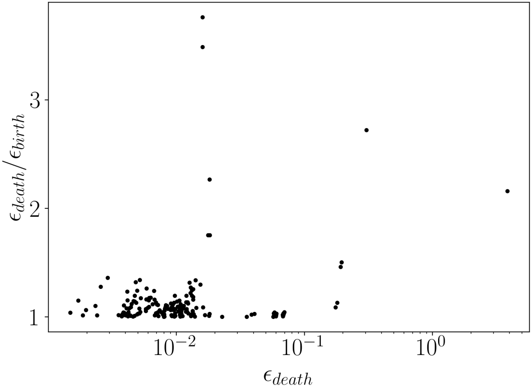
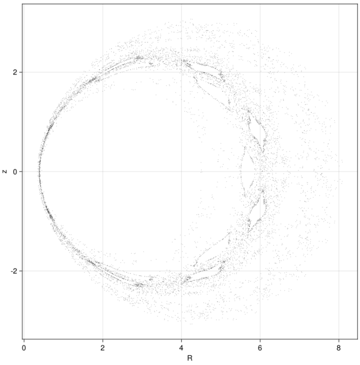


Figure 11: Poincare map orbit of an island chain.



Stochastic layer

Persistence diagram contains many short lived classes. Relative persistence of last H_1 class is high again.



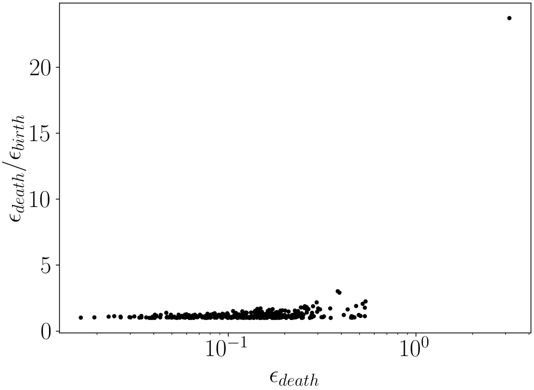


Figure 12: Poincare map orbit of a stochastic layer.

Separating island chains from other orbits

We can distinguish the island chains from other types of orbits by looking at the d/b for the last H_1 class.

Define c_l mapping point clouds $X \subset \Sigma$ to the class in $PH_1(X)$ with the largest death time. That is we define

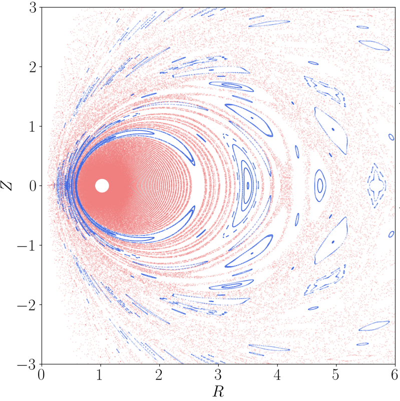
$$c_l(X) = \underset{c \in PH_1(X)}{\operatorname{argmax}} \, \epsilon_{\text{death}}(c) \,. \tag{8}$$

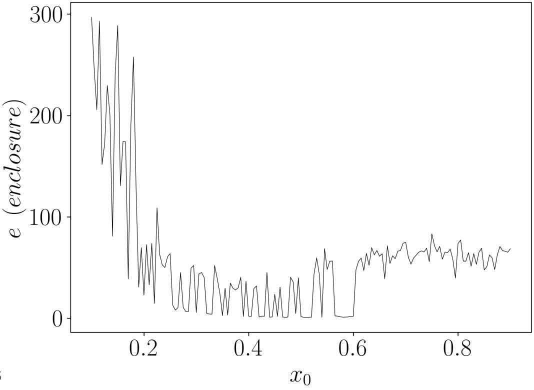
The relative persistence of this class defines a statistic we refer to as the *enclosure*, e, of the point cloud.

$$e(X) = \frac{\epsilon_{\text{death}}(c_l(X))}{\epsilon_{\text{birth}}(c_l(X))} \,. \tag{9}$$

Enclosure well discriminates the islands

Colouring point clouds of single orbit with $e < e_{\text{thesh}}$ gives the map of field lines below.





Map of islands with $e_{\text{thresh}} = 20$

Checking the island locations

Recall that surfaces should appear only on surfaces of rational q.

The number of islands in a chain is encoded in the PH_0 information (count number of persistent connected components).

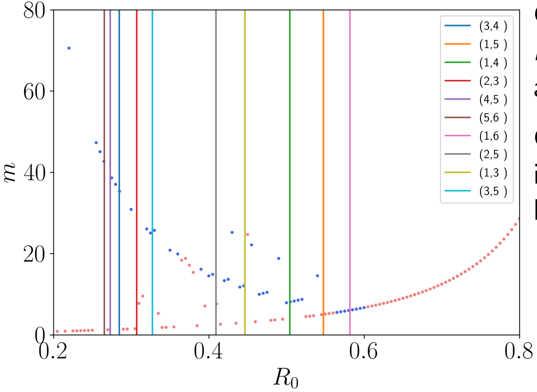


Figure 14: Placeholder

Can pull the number of islands n for a detected island chain and infer m from $q = n/m$.

Can then check if the detected islands occur at radius with low order rational q.

Towards separating KAM toruses

Exact KAM toruses should contain only 1 large H_1 class.

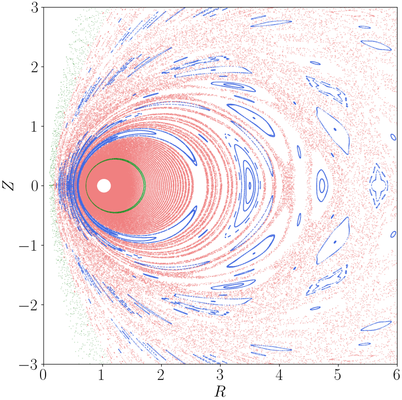


Figure 14: Map of islands with $e_{\text{thresh}} = 20, \epsilon_{\text{death}} < 10^{-3}$

Let *ihc* be the number of H_1 classes with death time above a threshold ϵ_{death}.

Then a non-island orbit is classed as a KAM torus if $ihc = 1$.

Some problems

There are problems with this procedure.

- Currently we need to fit the different thresholds largely by hand.
- Have to run a lot of field lines (between 1000-4000 iterations for my tests, but the actual requirement is unknown).
- Computing the persistent homology with RIPSER is slow but can be improved using recent work by Koyama.

Conclusion

With this work, we have shown that using TDA we can automatically detect the class of magnetic field line orbits in a chaotic perturbed tokamak field.

- Requires only the orbit of the field line on a Poincare section.
- No information on order of points are required.
- Currently only tested for “dense” point clouds. Minimum number of points required for accurate classification is unknown.
- Has not yet been tested for the case of a real tokamak field or a stellarator equilibrium.

This is a first practical application of TDA in a plasma physics context, more may follow.

References

1. Robins, V. *Computational topology at multiple resolutions: foundations and applications to fractals and dynamics.* PhD thesis (University of Colorado at Boulder, 2000).
2. Kaczynski, T., Mischaikow, K. M. & Mrozek, M. *Computational homology.* **7** (Springer, 2004).
3. Edelsbrunner, H. & Harer, J. *Computational Topology: An Introduction.* ISBN: 978-0-8218-4925-5 (Jan. 2010).
4. Ghrist, R. *Elementary Applied Topology.* 1.0. ISBN: 978-1502880857 (Createspace, July 2014).
5. Tempelman, J. R. & Khasawneh, F. A. A look into chaos detection through topological data analysis. *Physica D: Nonlinear Phenomena* **406,** 132446 (2020).
6. Kramár, M. *et al.* Analysis of Kolmogorov flow and Rayleigh–Bénard convection using persistent homology. *Physica D: Nonlinear Phenomena* **334.** Topology in Dynamics, Differential Equations, and Data, 82–98. ISSN: 0167-2789. https://www.sciencedirect.com/science/article/pii/S0167278916000270 (2016).
7. Nakahara, M. *Geometry, Topology, and Phyics.* (Taylor and Francis, 2003).
8. Hatcher, A. *Algebraic Topology.* (Cambridge University Press, 2000).

Future work

- Implement the automated detection scheme using the field of a real MHD solution.
- Incorporate information about the nearest neighbor distances obtained from the PH_0 data. May allow to detailed identification of the scales of structure.
- Attempt to reproduce results similar to the above with a more geometric filtration (alpha complexes for example)
- Investigate the VR persistent homology of particle trajectories (this is a high dimensional problem).
- Extract explicit representatives for the persistent cycles for phase space segmentation.

Other Applications

- The distribution of the size of islands in the phase space can be estimated using the sub-level set persistent homology on images.
- The renormalisation group transforms preserving phase space topology (which allow us to make exact predictions about dynamics and statistics) can be observed in the data and new symmetries may be found with the same methods.
- Location of maxima along field lines can be extracted using the sub-level set persistent homology.
- Adapt tools built to study topological states in condensed matter to plasma waves in spaces with non-trivial topology (plasma waves have topological phases too).