Persistent cohomology and period reconstruction

Mikael Vejdemo-Johansson Vin de Silva Primoz Skraba

October 25, 2010

Outline

Coordinatization

Essentially

It's all about finding *coordinate function* on a dataset $X \subseteq \mathbb{R}^d$. Preferably few coordinates - cognitive tools.

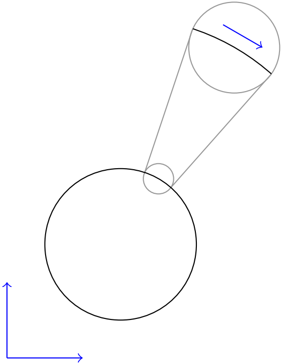
Classically

- Linear coordinatization: Find maps $X \to \mathbb{R}$, concentrating information.
- Principal Component Analysis
- Projection pursuit

Recently

Problematic cases

Some shapes take up too many coordinates.



Locally 1-dimensional.
Globally 2 coordinates needed to describe all points.
The shape doesn't fit in $\mathbb{R}$.

Fixes

How can we fix this?

Circle-valued coordinates

- Use $S^1 = [0, 1]/(0 \sim 1)$ as additional coordinate space
- Fixes the circle
- Fixes the torus
- Occurs naturally:
  - Phase coordinates for waves
  - Angle coordinates for directions

Circle-valued coordinates and cohomology

Problem remains: how do we find circle-valued coordinates?

It is enough to find maps $X \to S^1$.

We recall that $S^1 = K(\mathbb{Z}; 1)$ and therefore $H^1(X; \mathbb{Z}) = [X, S^1]$.

How do we compute cohomology for a point cloud?

Just computing cohomology: not useful. Discrete spaces are boring.

Instead, for a point cloud $X \subset \mathbb{R}^d$:

1. Pick a global, small ϵ
2. Cover each point x with $B_\epsilon(x)$
3. Take the nerve of the covering $C^v(X) = \mathcal{N}(\bigcup_x B_\epsilon(x)$
4. As ϵ grows, these *Čech complexes* include. Inclusion maps induce maps in cohomology.
5. At the addition of a single new simplex, either a cohomology class is born, or two classes merge. At a merge, we record the interval from the *most recent* birth to this death.
6. Long lives correspond to significant features.

From cohomology to circle-valued coordinatizations

Issues

- Easy to compute: Modular cohomology, coefficients in $\mathbb{F}_p$ for small primes p.
  Need for the isomorphism: Integer-valued cohomology.
  Smoothness: Integer cohomology gives constant values on all vertices, and wraps edges in the complex around the target circle.
- Numerical stability of cohomology computation and of the smoothing operations.

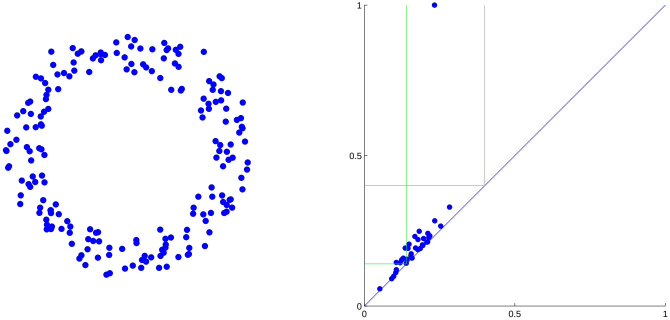
Smoothing

- Integral 1-cocycle: integer weighted edge graph.
- Coordinates found by edge traversal, increasing by edge weights.
- Each cohomologous cocycle guaranteed by cocycle condition to give compatible values, mod 1.0, to each vertex.
- Application straight on integral cocycle yields value 0 at each vertex.
- Given ζ integral cocycle, we wish to find cohomologous cocycle z such that the edges are small.
- Hence, we wish to find x such that $\zeta + \partial x$ has minimal L_2-norm.
- This is a well-known optimization problem. We use the LSQR algorithm.

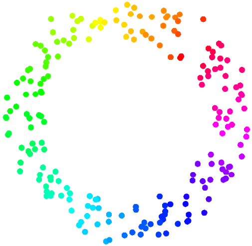
Reference

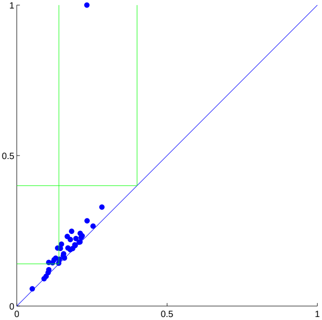
Morozov, de Silva, Vejdemo-Johansson: *Persistent Cohomology and Circular coordinates*, Proceedings of SoCG 2009; to appear: Discrete and Computational Geometry.

Examples

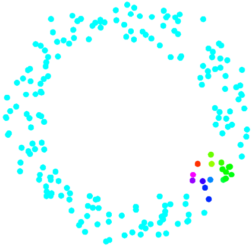


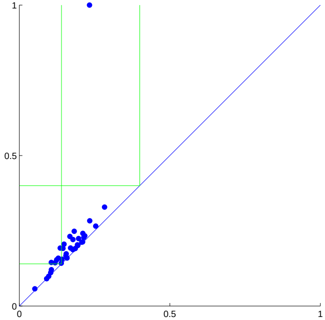
Examples



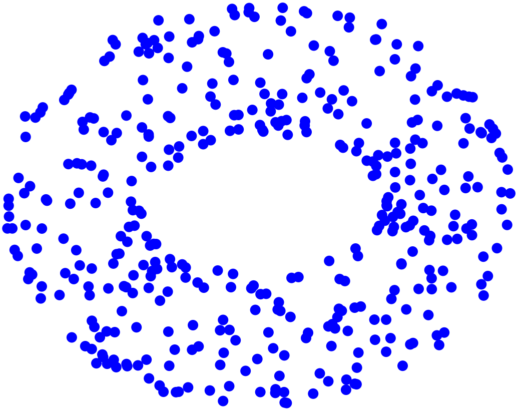


Examples

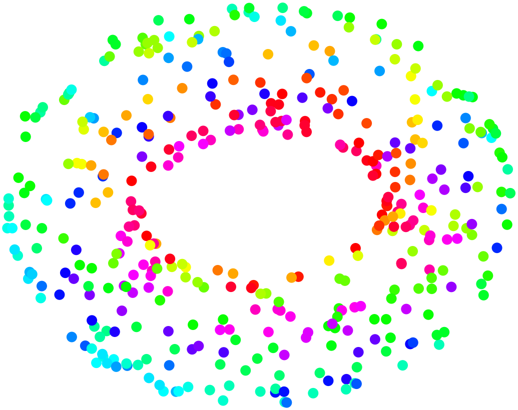




Examples

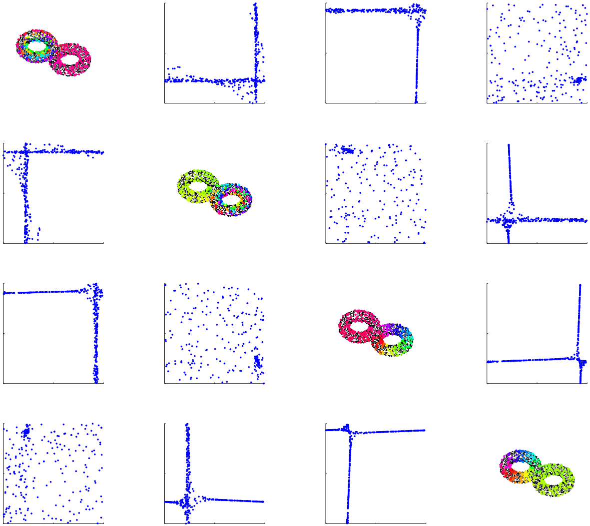


Examples



Examples

Examples



Outline

Periodic signals form circles

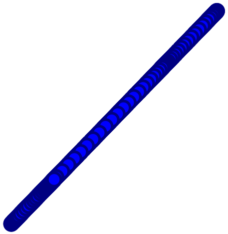
Sufficiently well described periodic or even recurrent signals will describe circles in the observation space.

Example

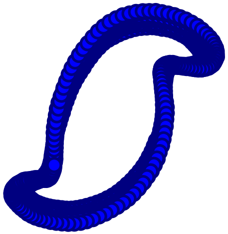
Take a clarinet tone as a 44 100 Hz sample of microphone membrane extensions rated in $[-32\,767, 32\,767]$.

We take the 1-dimensional signal and embed it in $\mathbb{R}^2$ by $f_\epsilon : a_t \mapsto (a_t, a_{t+\epsilon})$.

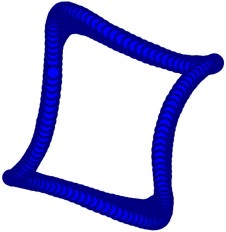
Clarinet delay embedding



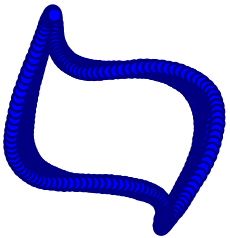
Clarinet delay embedding



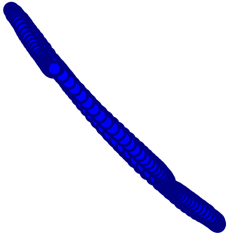
Clarinet delay embedding



Clarinet delay embedding



Clarinet delay embedding



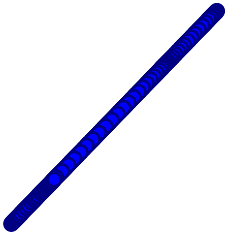
Coordinatizing a periodic system

Since the (good enough) image of a periodic system is an embedded circle, we can capture it with a circle-valued coordinate.

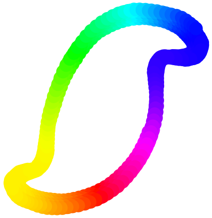
We can do this with our cohomological techniques.

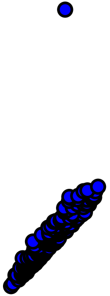
The quality of the capture can be determined by the length of the longest persistence interval.

Delay embeddings coordinatized

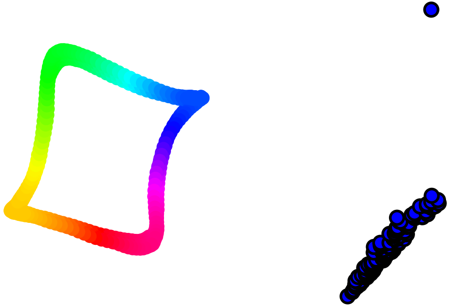


Delay embeddings coordinatized

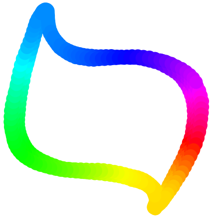


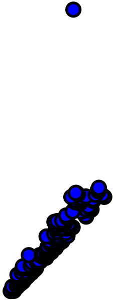


Delay embeddings coordinatized

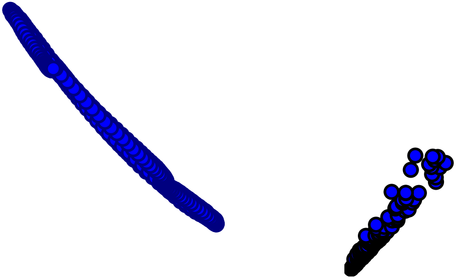


Delay embeddings coordinatized

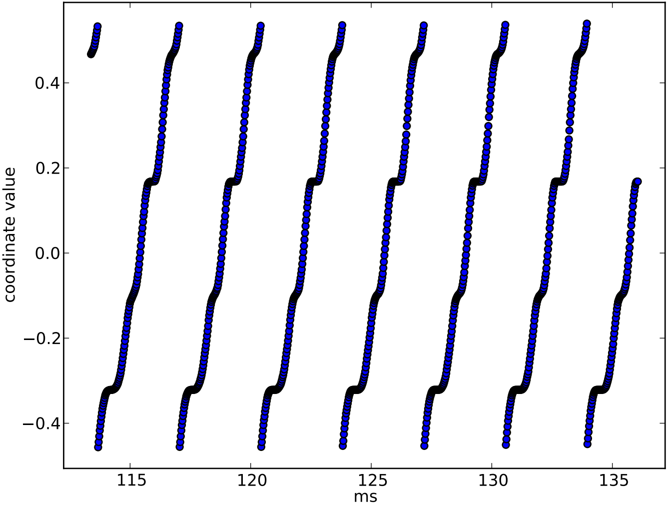




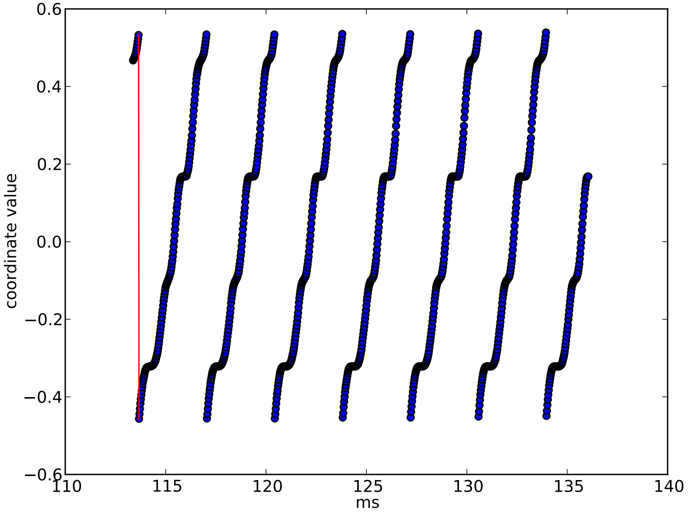
Delay embeddings coordinatized



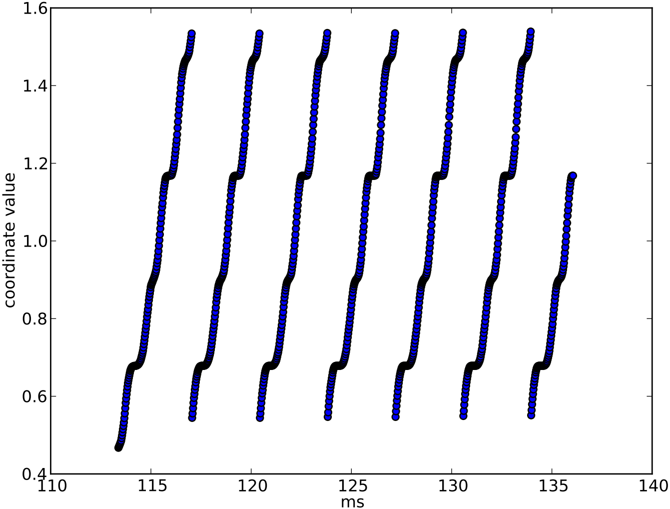
Unrolling the coordinates



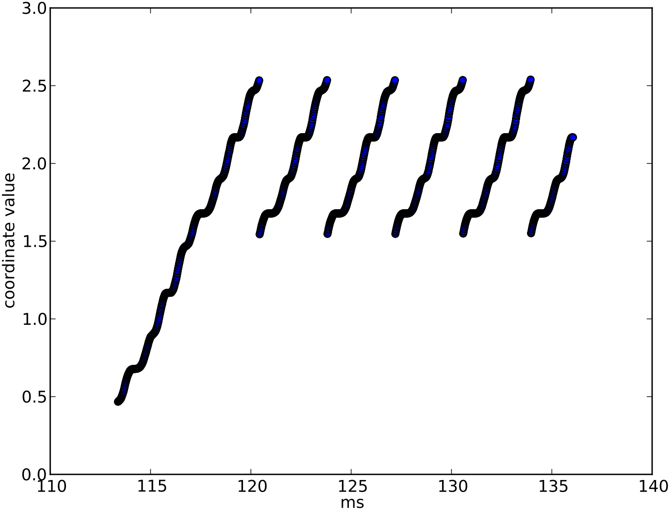
Unrolling the coordinates



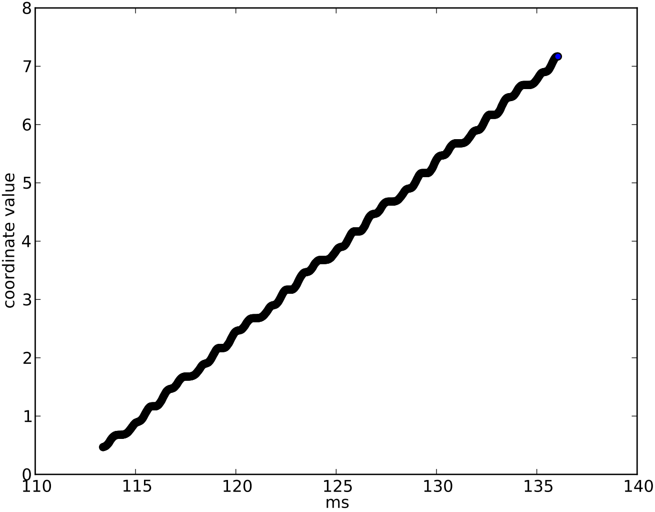
Unrolling the coordinates



Unrolling the coordinates



Unrolling the coordinates



Recovering period length

- Linear regression on unrolled coordinate: $y = 2.96x - 33.1$
- Slope: portion of full period / ms
- 1000/slope = 337.8 periods / s