Computational topology - group project

Cocycles and circular coordinates

Introduction: Sometimes (or often) it is useful to represent high-dimensional nonlinear data in terms of low-dimensional coordinates which somehow represent the intrinsic structure of the data. Sometimes the data will contain 1-dimensional features which cannot be faithfully represented with linear coordinates ($\mathbb{R}$), but can be nicely represented using circular coordinates (S^1). The task is to find examples of data containing such features and then also represent those features in an efficient way.

Goal: Use the techniques suggested below to find circular (and maybe even spherical) features of some data set.

Detailed description: Linear (real) coordinates on a topological space X are nothing else than functions $X \to \mathbb{R}$. (For example, the x-coordinate of a point $(x, y) \in \mathbb{R}^2$ is the function $(x, y) \mapsto x$.)
In an analogous sense *circular coordinates* on a space X are functions $X \to S^1$. A nice example is the torus $S^1 \times S^1$, for which the most apparent choice of circular coordinates would be to pick the projection onto either S^1 factor. These two choices correspond, at least geometrically, to the most apparent cocycles which generate $H^1(S^1 \times S^1; \mathbb{Z})$.
For us the space X will be given as a point cloud (a high dimensional dataset).

Let's assume that a point cloud X is given. Your computations should follow these steps:

1. Compute persistent H^1 of X with coefficients in $\mathbb{Z}_p$ for some prime p. Pick some of the most persistent 1-cocycles. Let's say $\sigma' \in H^1(\text{Rips}(X, r); \mathbb{Z}_p)$ is one of these cocycles.

2. Lift the cocycle σ' to a cocycle $\sigma \in H^1(\text{Rips}(X, r); \mathbb{Z})$ along the homomorphism

$$H^1(\text{Rips}(X, r); \mathbb{Z}) \to H^1(\text{Rips}(X, r); \mathbb{Z}_p)$$

   induced by the mod p reduction $\mathbb{Z} \to \mathbb{Z}_p$. If this is not possible, pick a new prime p and return to the previous step.

3. The inclusion homomorphism $\mathbb{Z} \hookrightarrow \mathbb{R}$ induces a homomorphism

$$H^1(\text{Rips}(X, r); \mathbb{Z}) \to H^1(\text{Rips}(X, r); \mathbb{R})$$

   and in turn determines a cocycle $\sigma_{\mathbb{R}} \in H^1(\text{Rips}(X, r); \mathbb{R})$. Specifically, the cocycle σ is determined by values $\sigma(AB) \in \mathbb{Z}$ for every 1-simplex $AB \in \text{Rips}(X, r)$, and now we simply view it as a real–valued function. Recall that σ is cohomologous to $\sigma + d_0 f$, where $d_0 \colon C^0 \to C^1$ is the coboundary map and $f \in C^0(\mathrm{Rips}(X,r);\mathbb{Z})$ is an arbitrary 0-cochain. The task now is to find f, which minimizes

$$\|\sigma_{\mathbb{R}} + d_0 f\|,$$

where $\|\alpha\|^2 := \sum_{AB} |\alpha(AB)|^2$ (this sum is taken over all 1-simplices $AB \in \mathrm{Rips}(X,r)$). We now replace $\sigma_{\mathbb{R}}$ with $\overline{\sigma}_{\mathbb{R}} = \sigma_{\mathbb{R}} + d_0 f$.

4. Use the cocycle $\overline{\sigma}_{\mathbb{R}}$ (or rather, f, the solution of the minimization problem above) to construct a map $s \colon X \to S^1$. View $S^1 \subseteq \mathbb{C}$ as the unit complex numbers and define $s(A) = e^{2\pi f(A)}$ for each $A \in X = \mathrm{Rips}(X,r)^{(0)}$. Explain why this s can be extended to the 1-skeleton of $\mathrm{Rips}(X,r)$. (The existence of this extension assures that this choice of s makes sense.) This s now defines our circular coordinates on X.

For a detailed description see `https://arxiv.org/abs/0905.4887`.
Steps 2 and 4 should be considered relatively straightforward. You can use already implemented algorithms for steps 1 and 3.
Test this procedure on a few interesting data sets X (artificial and/or real–world).

Results: The report should include a description of methodology and tasks undertaken, a pseudocode, methods of computation, results of experiments, and division of work.

Students are encouraged to take the initiative and possibly implement their own ideas on the theme of the project: perhaps thinking of their own topological optimization setting, etc.

Note that the generalization to *spherical coordinates*, i.e., functions $X \to S^n$ is not as straightforward. See, e.g., `https://arxiv.org/abs/2209.02791` and `https://arxiv.org/abs/2509.16102`.
The core of the algorithm that can be used to construct spherical coordinates relies on the following three theorems from algebraic topology:

- **Brown representability**: For a simplicial complex X we have

$$[X, K(\mathbb{Z},n)] \cong H^n(X;\mathbb{Z}),$$

where $[X,Y]$ denotes the set of (basepoint-preserving) homotopy classes of maps $X \to Y$ and $K(\mathbb{Z},n)$ is some specific space (called the *Eilenberg-MacLane space*). Note that $H^n(K(\mathbb{Z},n);\mathbb{Z}) \cong \mathbb{Z}$, so it has a single generator which we denote by $[S^n]$. The isomorphism between $[X, K(\mathbb{Z},n)]$ and $H^n(X;\mathbb{Z})$ can be explicitly described: Given a (homotopy class of a) map $f \colon X \to K(\mathbb{Z},n)$ the corresponding cocycle in $H^n(X;\mathbb{Z})$ is given as $f^*([S^n])$. (Also: $K(\mathbb{Z},1) = S^1$.)

- **Cellular approximation**: For simplicial complexes K and L, of dimensions k and l, respectively, every continuous map $|K| \to |L|$ is homotopic to a map $|K| \to |L^{(k)}|$, where $L^{(k)}$ is the k-skeleton of L.

- **Bockstein exact sequence**: There is a long exact sequence of cohomology groups

$$\cdots \to H^n(X;\mathbb{Z}) \xrightarrow{p} H^n(X;\mathbb{Z}) \xrightarrow{\text{mod } p} H^n(X;\mathbb{Z}_p) \xrightarrow{\beta} H^{n+1}(X;\mathbb{Z}) \to \cdots$$

  induced via the short exact sequence of coefficient groups

$$0 \to \mathbb{Z} \xrightarrow{p} \mathbb{Z} \xrightarrow{\text{mod } p} \mathbb{Z}_p \to 0.$$

  The resulting *connecting* homomorphism $\beta\colon H^n(X;\mathbb{Z}_p) \to H^{n+1}(X;\mathbb{Z})$ is called the *Bockstein homomorphism*.

All 4 steps that were suggested earlier to find circular coordinates can be generralized (or modified) to find spherical coordinates. See the linked papers for details if you'd like to take on that challenge.