Point-cloud topology via harmonic forms

Vin de Silva, Pomona College
vin.desilva@pomona.edu
Today’s goal

- Explore the use of **discrete Laplacian operators**...
- ...as applied to the topology of **point-cloud data**
- Discuss “**qualitative**” vs “**quantitative**”
- Discuss “**discrete**” vs “**continuous**”
- Run one or two **demos**
Thanks to my former colleagues at Stanford:

- Gunnar Carlsson
- Patrick Perry
- Afra Zomorodian
- Anne Collins
- Peter Lee
Discrete vs Continuous
Standard Pipeline (first attempt)

hidden/unknown space $X$

finite sample $Y \subseteq X$

simplicial complex $S = S(Y)$

homology invariants of $S$

$\beta_0 = 1$

$\beta_1 = 1$

$\beta_2 = 0$
Betti numbers ↔ features

For an object in 2D space
- $b_0$ is the number of components
- $b_1$ is the number of holes

For an object in 3D space
- $b_0$ is the number of components
- $b_1$ is the number of tunnels or handles
- $b_2$ is the number of voids

(and so on, in higher dimensions)
Reconstruction theorems

Various constructions for $S(Y)$

- Čech complex (folklore)
- Rips–Vietoris complex (folklore)
- $\alpha$-shape complex (Edelsbrunner, Mücke)
- strong/weak witness complexes (Carlsson, dS)

Desire theorems of the form:

If $Y$ is well-sampled from $X$ then $S(Y) \approx X$

e.g. Niyogi–Smale–Weinberger (2004), Čech complex
Discrete vs continuous

Betti numbers are discrete

Topological spaces
- topological spaces are continuous
- the space of topological spaces is discrete

Finite point-clouds
- point-clouds are discrete
- the space of point-clouds is continuous

Therefore, raw Betti numbers are
- ✔ very handy for topological spaces
- ❌ a bit dangerous for point-clouds
One lump or two?

At which parameter value does the number of components change?
One lump or two?

At which parameter value does the number of components change?
One lump or two?

At which parameter value does the number of components change?
One lump or two?

At which parameter value does the number of components change?
One lump or two?

At which parameter value does the number of components change?
One lump or two?

At which parameter value does the number of components change?
Standard Pipeline (second attempt)

- hidden/unknown space $X$
- finite sample $Y \subset X$
- labelled complex $S(r) = S(Y, r)$
- quantitative topology

WORKSHOP ON MODERN MASSIVE DATA SETS
Stanford University & Yahoo! Research

POINT-CLOUD TOPOLOGY VIA HARMONIC FORMS
Vin de Silva 2006–June–23
Example: Persistence
Persistent homology

- Edelsbrunner, Letscher, Zomorodian (2000)
  effective algorithm for persistence in 3-space

- Carlsson, Zomorodian (2005)
  general theory of persistent homology

- Cells of $S(Y)$ labelled by “time of birth”

- Bar-codes indicate feature lifetimes

**Continuous** measurements (interval length) coupled to **discrete** information (number of intervals)
Persistence pipeline

- hidden/unknown space $X$
- finite sample $Y \subset X$
- filtered complex $S(r) = S(Y, r)$
- persistent homology of $S(r)$
Discrete Laplacians
\[ \Delta_k \]

\( C_k = \{ \text{real-valued functions on } k\text{-simplices of } S(Y) \} \)
\( \text{floating point rather than exact arithmetic} \)

\( \text{Define discrete Laplacian operators } \Delta_k : C_k \rightarrow C_k \)

\( \text{Consider the harmonic spaces } H_k = \text{Ker}(\Delta_k) \)
\( H_k \text{ is isomorphic to standard homology of } X \)

\( \text{Consider eigenspaces } \{ f : \Delta_k f = \lambda f \} \text{ for } \lambda \text{ small} \)
\( \text{“almost homology” or “}\varepsilon\text{-homology”} \)

\( \text{Information derived from the ranks of these spaces (Betti numbers) and the eigenfunctions themselves} \)
Constructing $\Delta_k$

Given a chain complex over the real numbers...

$$\cdots \xrightarrow{\partial_k} C_k \xleftarrow{\partial_k} C_{k-1} \xrightarrow{\partial_k} \cdots$$

...and an inner product on each $C_k$, we can form the dual cochain complex:

$$\cdots \xrightarrow{\partial^*_k} C_k \xleftarrow{\partial^*_k} C_{k-1} \xrightarrow{\partial^*_k} \cdots$$

The discrete Laplacian is defined...

$$\Delta_k = \partial_k^* \partial_k + \partial^*_{k+1} \partial_{k+1}$$

...and one can easily prove (in the finite dimensional case):

$$\mathcal{H}_k := \text{Ker}(\Delta_k) \cong \frac{\text{Ker}(\partial_k)}{\text{Im}(\partial_{k+1})} =: H_k$$

homology is defined using a chain complex

cohomology is defined using a cochain complex
Aside: Hodge theory

For a 3-dimensional domain:

\[
\begin{align*}
\Omega^0 & \rightarrow \nabla \cdot \Omega^1 \rightarrow \nabla \times \Omega^2 \rightarrow \nabla \cdot \Omega^3 \\
\Omega^0 & \leftarrow \nabla \cdot \Omega^1 \leftarrow \nabla \times \Omega^2 \leftarrow \nabla \cdot \Omega^3
\end{align*}
\]

For example:

\[
\begin{align*}
\Delta_0 f & := - \nabla \cdot (\nabla f) = - \sum_{i=1}^{3} \frac{\partial^2 f}{\partial x_i^2} \\
\Delta_1 \vec{f} & := \nabla \times (\nabla \times \vec{f}) - \nabla (\nabla \cdot \vec{f}) = - \sum_{i=1}^{3} \frac{\partial^2 f}{\partial x_i^2}
\end{align*}
\]

Proof that \( \text{Ker}(\Delta_k) = H_k \) is much more difficult.
ε-Betti numbers

Structure theorem for homology and ε-homology

For every nonnegative integer k, and ε > 0:

Integers $b_k$ “Betti numbers”
Integers $b_{k+\frac{1}{2}}(\epsilon)$ “ε-Betti numbers”

such that:

$$\dim(\text{Ker}(\Delta_k)) = b_k$$
$$\dim(\text{Eig}(\Delta_k, \epsilon)) = b_{k-\frac{1}{2}}(\epsilon) + b_k + b_{k+\frac{1}{2}}(\epsilon)$$

space spanned by eigenfunctions with eigenvalue less than $\epsilon$
Laplacian pipeline

hidden/unknown space $X$

finite sample $Y \subset X$

weighted complex $S = S(Y), f: S \rightarrow \mathbb{R}$

$\varepsilon$-harmonic forms

$\beta_0 = 1$

$\beta_{0.5}(\varepsilon) = ?$

$\beta_1 = 1$

$\beta_{1.5}(\varepsilon) = ?$

$\beta_2 = 0$
Pros and cons

✔ Several ways to incorporate continuous parameters
  ✔ meaning of "λ is close to zero" — how close?
  ✔ simplices can be weighted prior to construction of $Δ_k$

✔ Harmonic cycles have global optimality properties
  ✔ localising features/minimal cycle problem

✔ Non-zero eigenfunctions encode subtle relationships between cells of adjacent dimensions

✘ More expensive than persistent homology

✘ Theory somewhat underdeveloped
  ✔ (except graph Laplacians, see "Spectral Graph Theory" by Chung)
Entropy
Local vs global features

Homological features can be local or global to varying degrees:

This example has a 2-dimensional space of harmonic 1-forms. Can we pick out 1-forms representing the two features?

persistent homology can do this very easily
Concentration

Heuristic arguments suggest that harmonic cycles concentrate energy...

- weakly along global features
- strongly along local features
Entropy & $L^p$ comparison

How to detect whether a cycle is highly concentrated in some region?

Some measure of entropy is called for
- high entropy $\leftrightarrow$ flat distribution $\leftrightarrow$ global feature
- low entropy $\leftrightarrow$ peaked distribution $\leftrightarrow$ local feature

Simple estimate: compare $L^1$ and $L^2$ norms
- $E[f] := \|f\|_1 / \|f\|_2$
- $E[f]$ large $\leftrightarrow$ global feature
- $E[f]$ small $\leftrightarrow$ local feature
Betti numbers: examples
## Examples

<table>
<thead>
<tr>
<th>$b_0$</th>
<th>$b_{0.5}(\varepsilon)$</th>
<th>$b_1$</th>
<th>$b_{1.5}(\varepsilon)$</th>
<th>$b_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Examples

hot spot for 1-chain $j$, where $\Delta_1j = \lambda j$

<table>
<thead>
<tr>
<th>$b_0$</th>
<th>$b_{0.5}(\varepsilon)$</th>
<th>$b_1$</th>
<th>$b_{1.5}(\varepsilon)$</th>
<th>$b_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
**Examples**

<table>
<thead>
<tr>
<th>$b_0$</th>
<th>$b_{0.5}(\varepsilon)$</th>
<th>$b_1$</th>
<th>$b_{1.5}(\varepsilon)$</th>
<th>$b_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

hot spot for 1-cycle $j$, where $\Delta_1 j = 0$
Examples

### Annulus

<table>
<thead>
<tr>
<th>$b_0$</th>
<th>$b_{0.5}(\epsilon)$</th>
<th>$b_1$</th>
<th>$b_{1.5}(\epsilon)$</th>
<th>$b_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Hot spot for 1-cycle $j$, where $\Delta_1 j = 0$.
Examples

hot spot for 1-cycle $j$, where $\Delta_1 j = \lambda j$

hot spot for 2-chain $k$, where $\Delta_2 k = \lambda k$

<table>
<thead>
<tr>
<th>$b_0$</th>
<th>$b_{0.5}(\epsilon)$</th>
<th>$b_1$</th>
<th>$b_{1.5}(\epsilon)$</th>
<th>$b_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
### Examples

#### Sphere

<table>
<thead>
<tr>
<th>$b_0$</th>
<th>$b_{0.5}(\epsilon)$</th>
<th>$b_1$</th>
<th>$b_{1.5}(\epsilon)$</th>
<th>$b_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Take-home message
What is a (1.5)-D feature?

punctured sphere
What is a (1.5)-D feature?

A 1-D cycle which is a boundary (but only just)

punctured sphere
What is a (1.5)-D feature?

A punctured sphere

A 1-D cycle which is a boundary (but only just)

A 2-D chain which is almost (but not quite) closed
What is a (1.5)-D feature?

A punctured sphere

A 1-D cycle which is a boundary (but only just)
A 2-D chain which is almost (but not quite) closed
Thank you