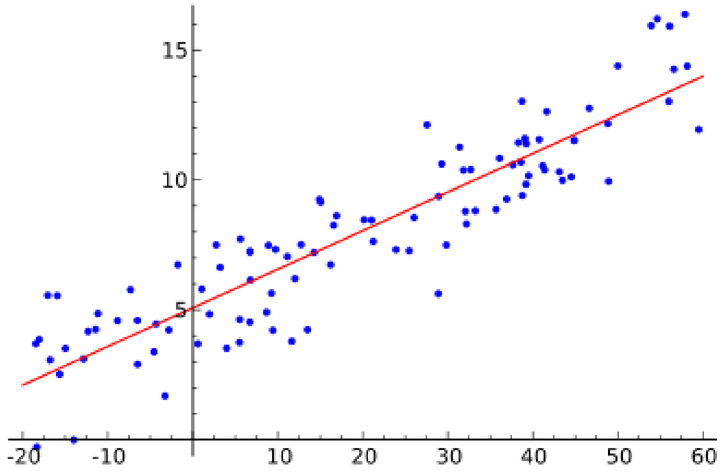
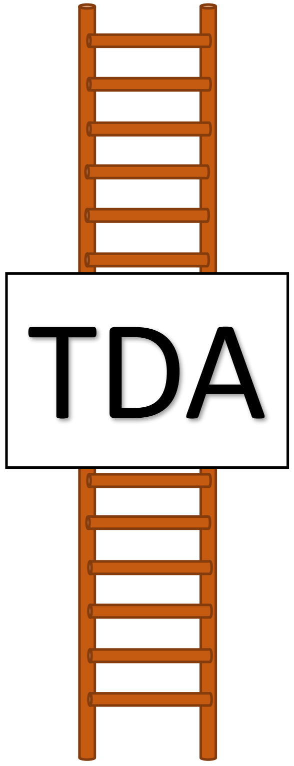
Topological Time Series Analysis

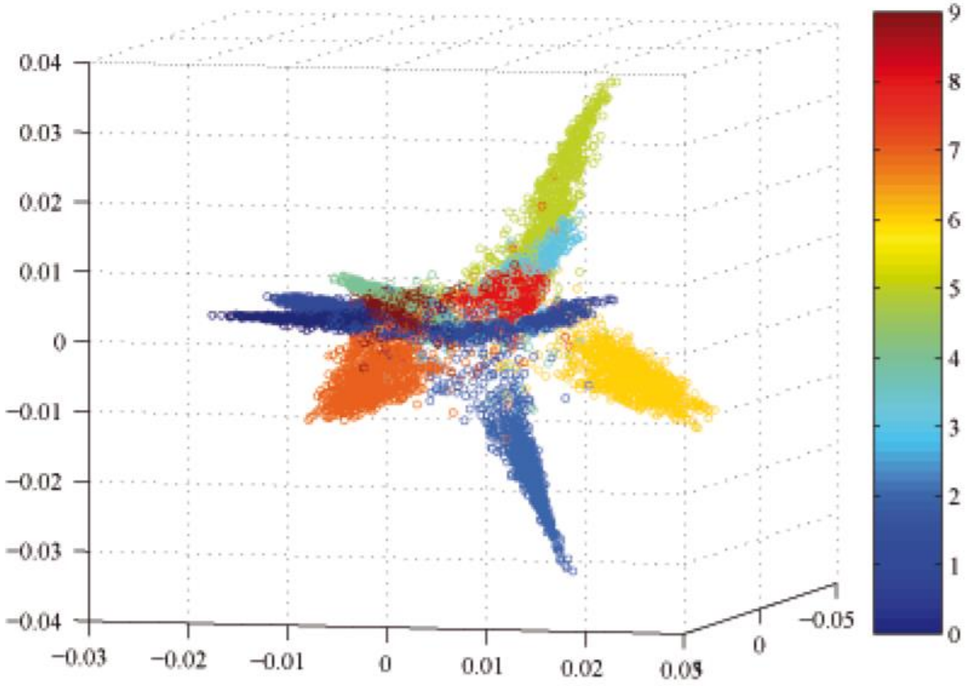
Day 1: Geometry of sliding window embeddings

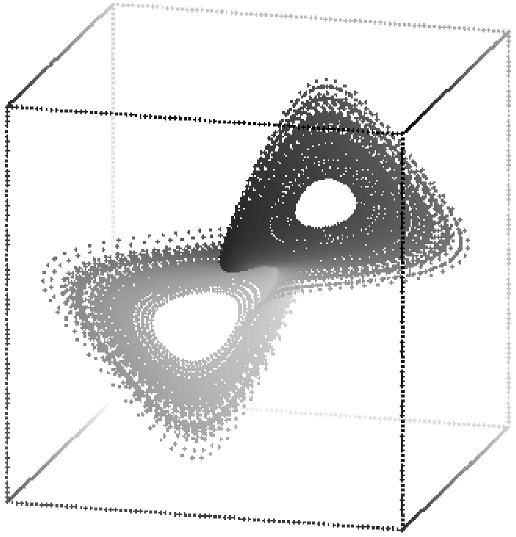
Jose Perea

Data has shape

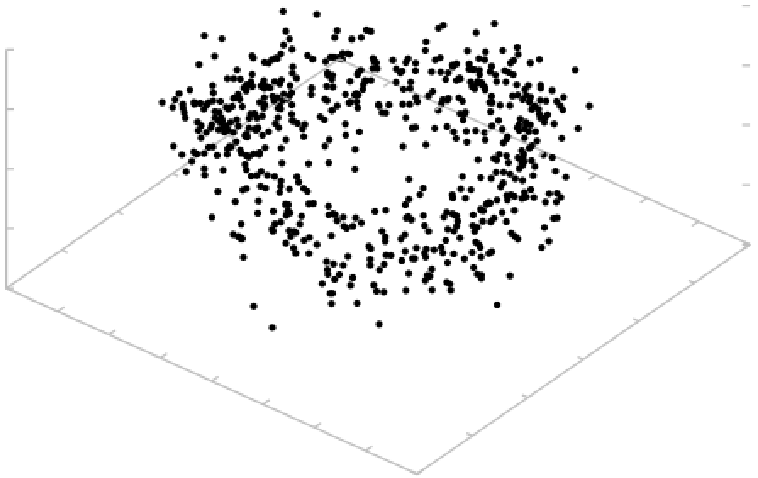








And shape matters



Topological Time Series Analysis

Set up:

$$\underbrace{(\mathbb{M}, \mathbf{d})}_{\text{Metric space}} \qquad \underbrace{f : S \subset \mathbb{N} \longrightarrow \mathbb{M}}_{\text{Time series}}$$

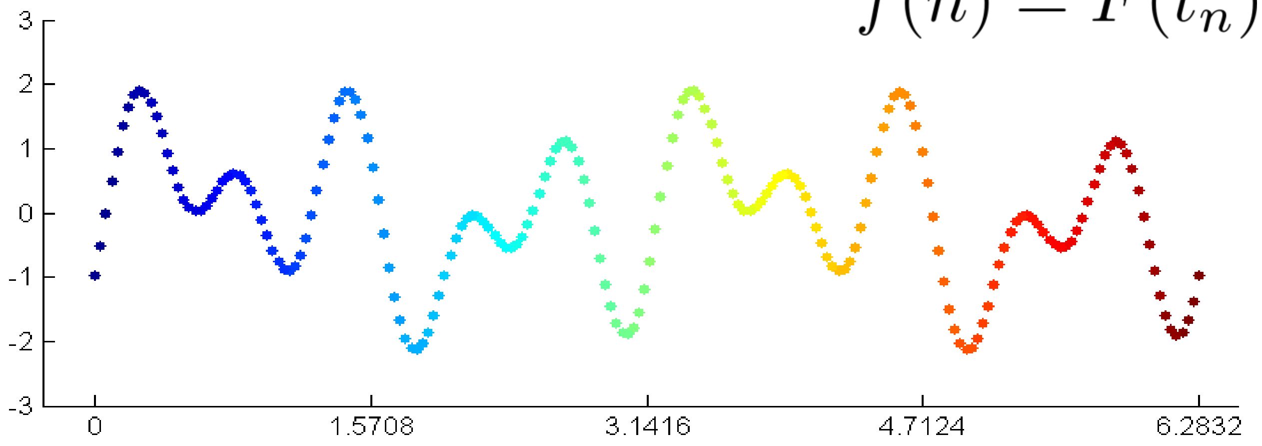
Example:

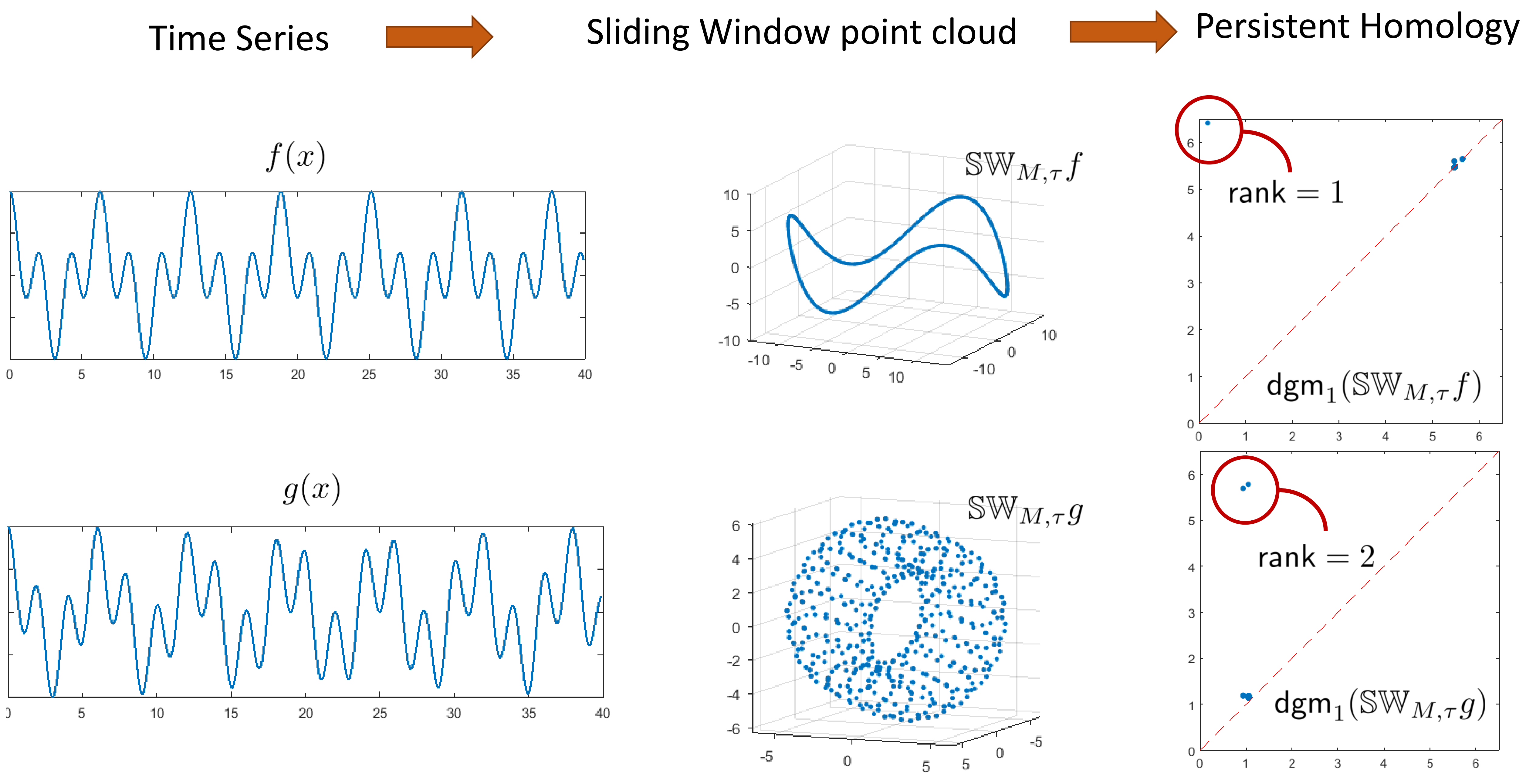
$$F : [a, b] \subset \mathbb{R} \longrightarrow \mathbb{R}$$

evaluated at

$$a \leq t_1 < t_2 < \cdots < t_N \leq b$$

$$f(n) = F(t_n)$$

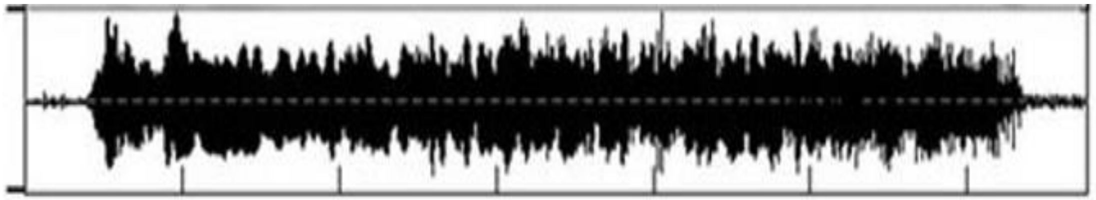




Time Series
Sliding Window point cloud
Persistent Homology
$f(x)$
$g(x)$
$\mathbb{SW}_{M,\tau} f$
$\mathbb{SW}_{M,\tau} g$
rank $= 1$
rank $= 2$
$\mathrm{dgm}_1(\mathbb{SW}_{M,\tau} f)$
$\mathrm{dgm}_1(\mathbb{SW}_{M,\tau} g)$

This Week...

Biphonation in mammals

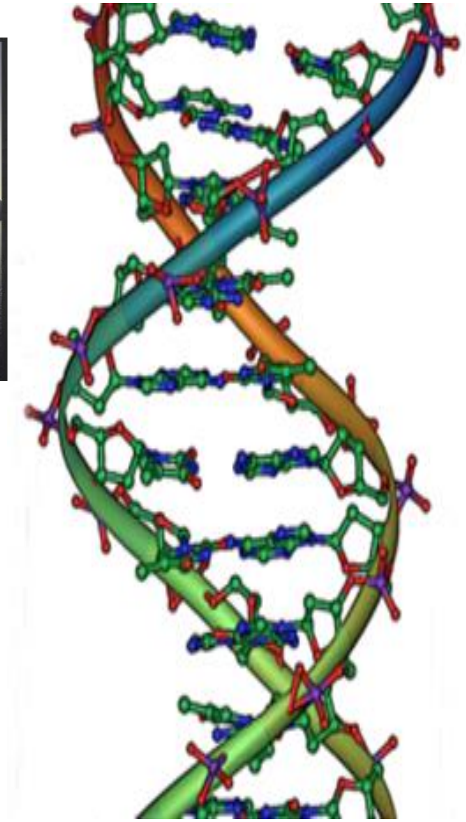


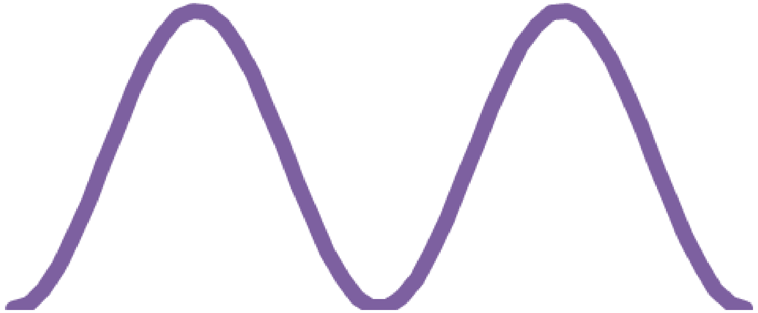
Periodicity in Video Data

Music analysis: rhythm

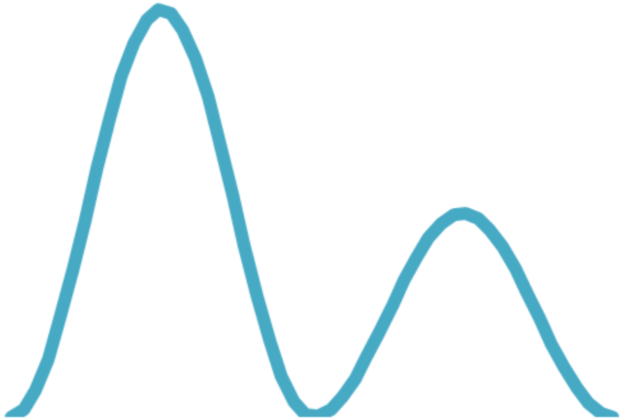
Sliding Window Embeddings

Biological Clocks

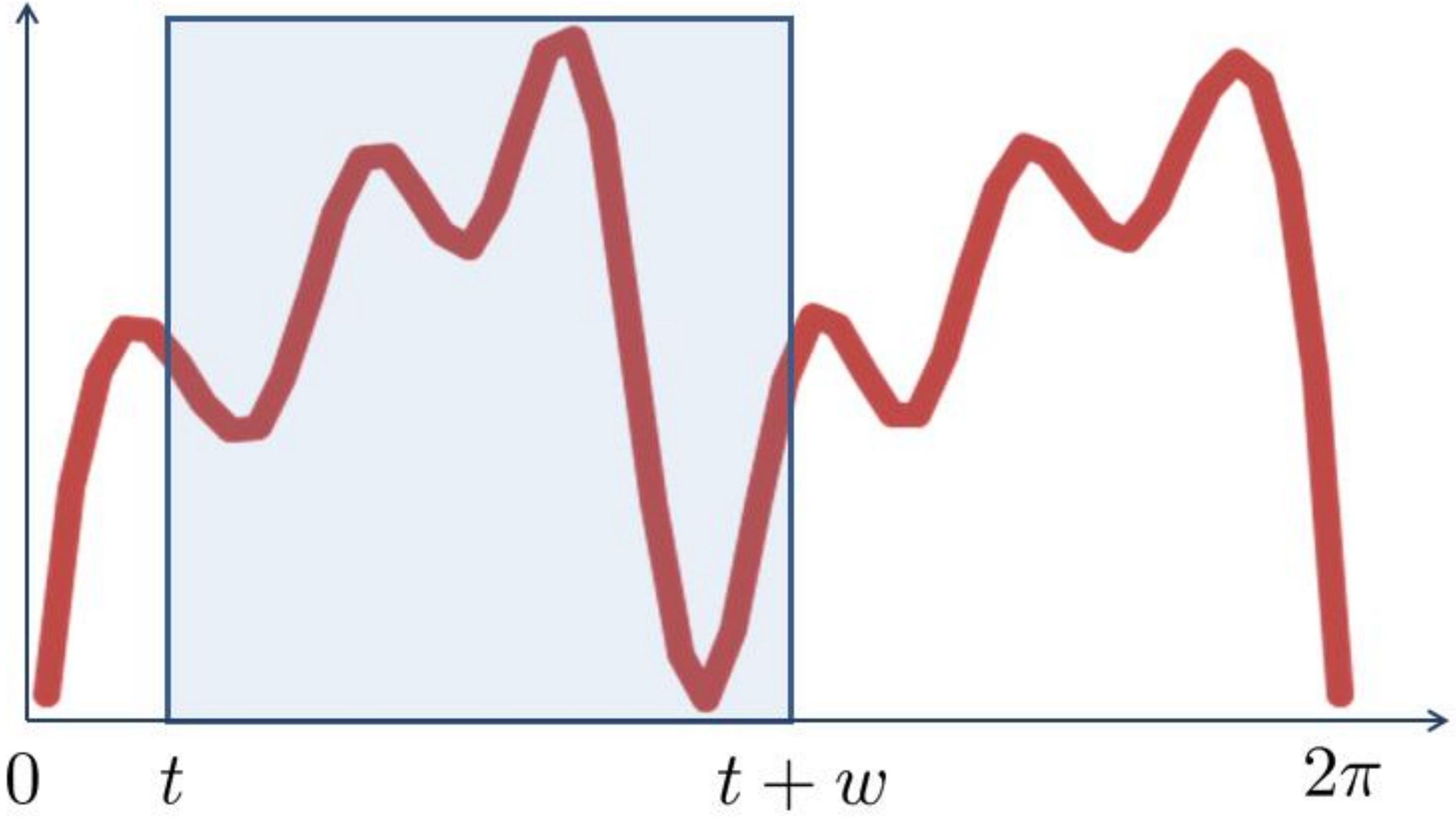




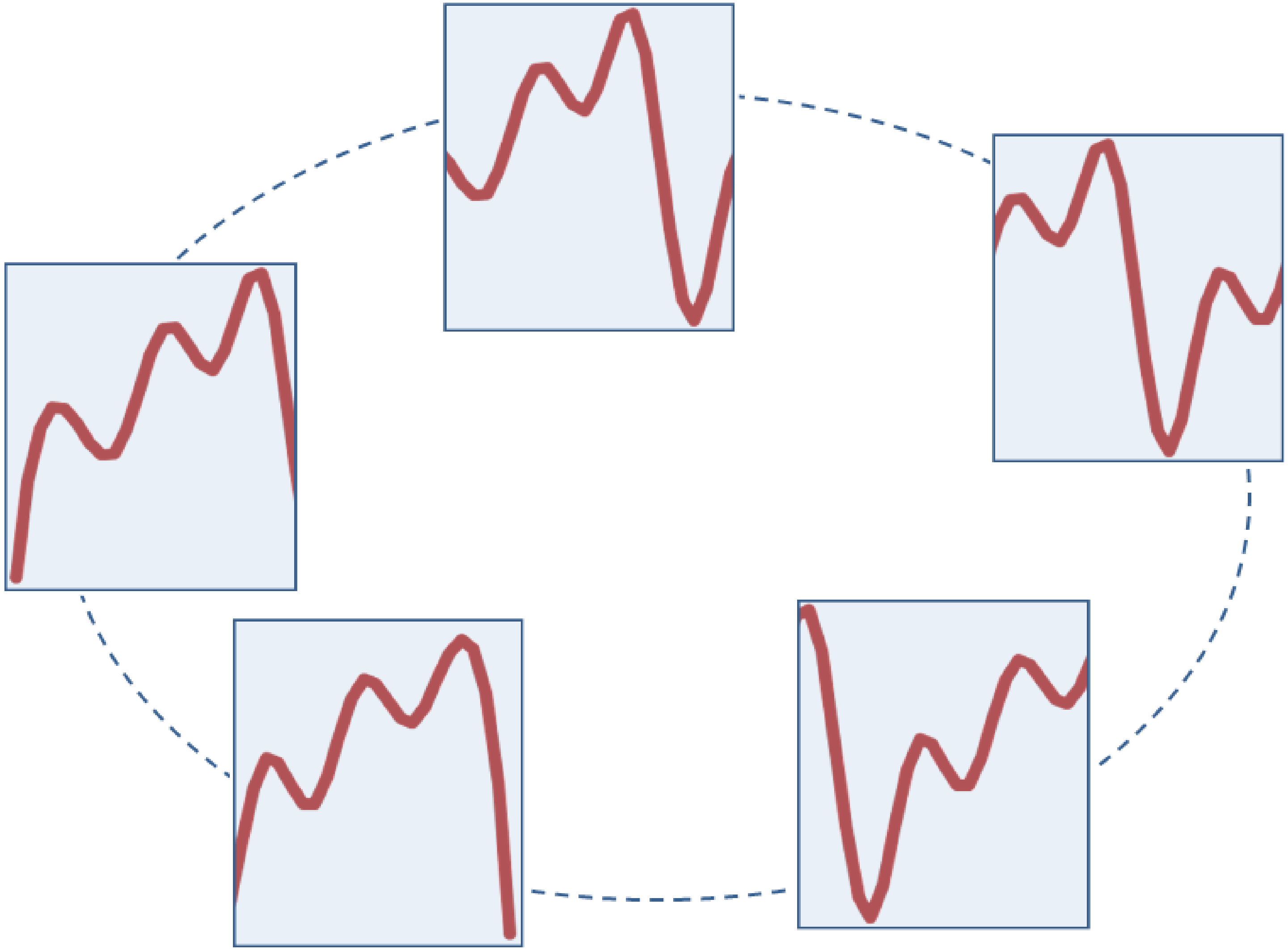
What is periodicity, and
how do we quantify it?



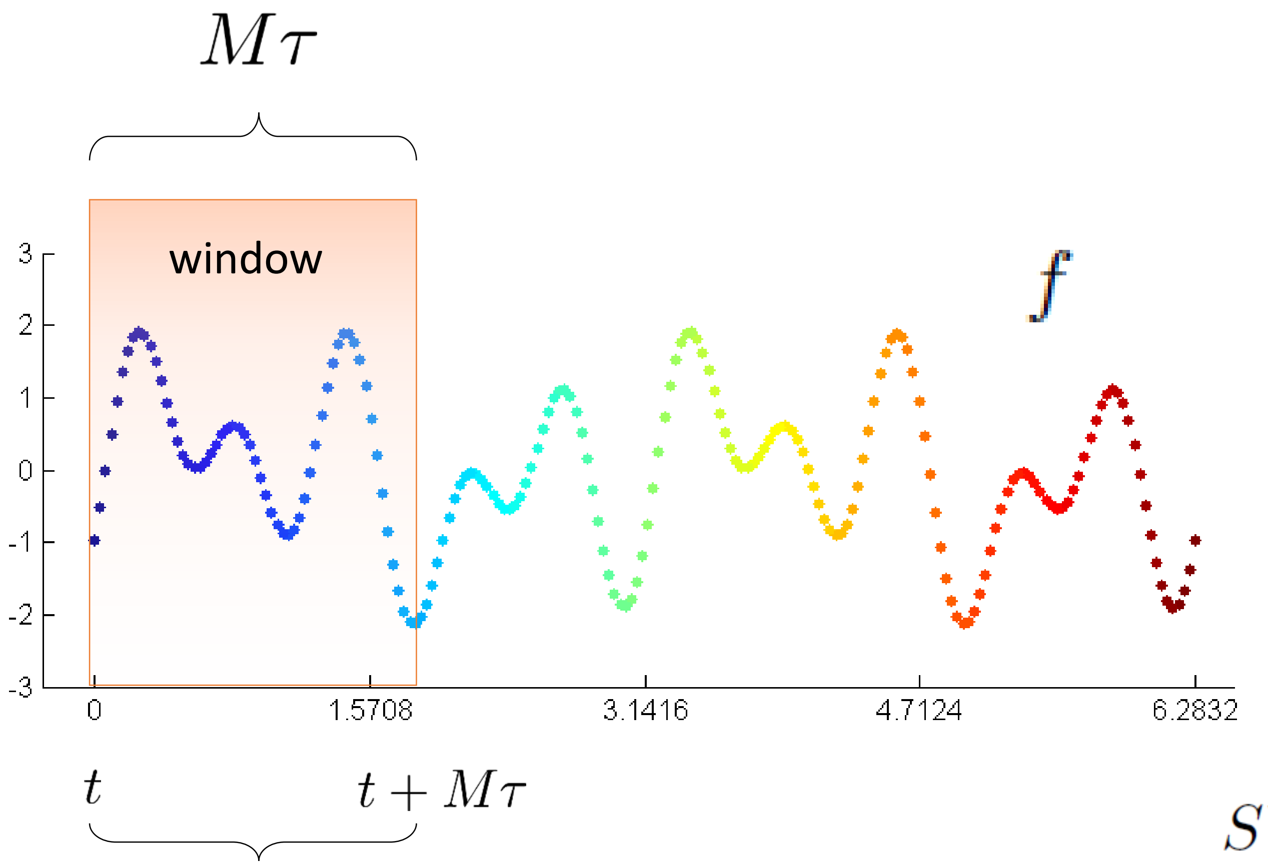
Sliding Windows



Sliding Windows



Sliding window embedding



$$f(t), f(t+\tau), \ldots, f(t+M\tau)$$

Fix

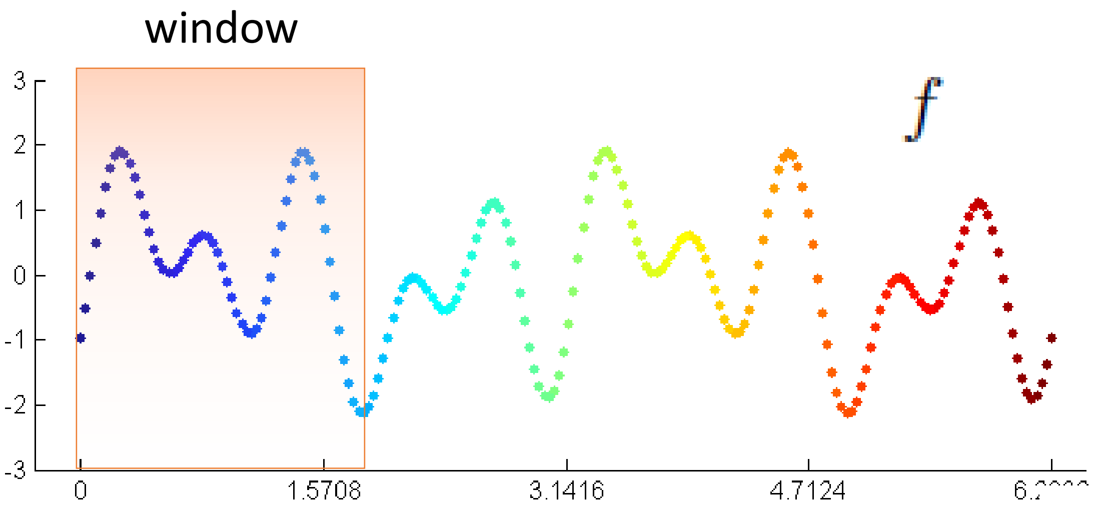
τ → Step/delay

$M+1$ → Dimension

$$SW_{M,\tau} f(t) = \begin{bmatrix} f(t) \\ f(t+\tau) \\ \vdots \\ f(t+M\tau) \end{bmatrix} \in \mathbb{R}^{M+1}$$

Sliding Windows and Persistence: An application of topology to signal analysis, J. Perea and J. Harer, 2015

Sliding window embedding

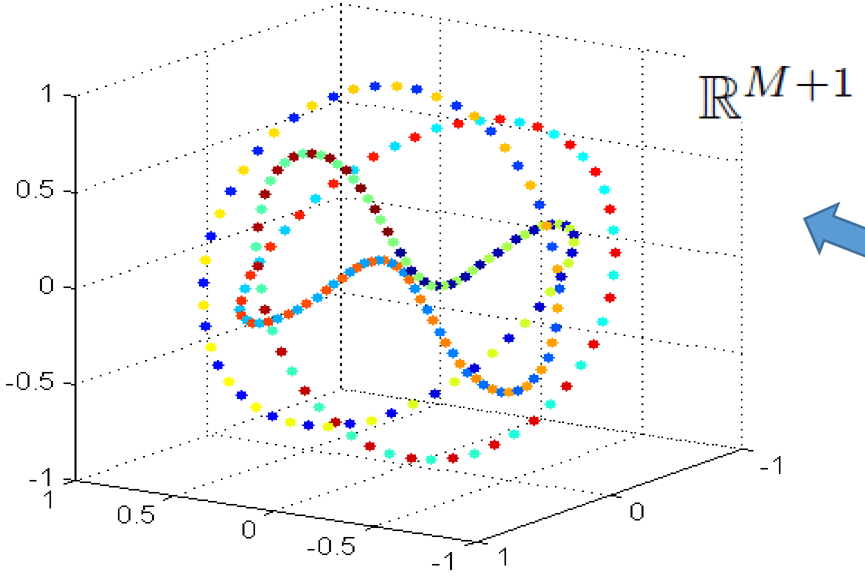


$\tau \rightarrow$ Step/delay

$M\tau \rightarrow$ Window size

$M+1 \rightarrow$ Dimension

$$SW_{M,\tau}f(t) = \begin{bmatrix} f(t) \\ f(t+\tau) \\ \vdots \\ f(t+M\tau) \end{bmatrix}$$



Sliding window point-cloud

$SW_{M,\tau}f(T)$

$T \subset \mathbb{R}$

Sliding Windows and Persistence: An application of topology to signal analysis, J. Perea and J. Harer, 2015

Takens' Embedding Theorem

Let $\mathcal{M}$ me be a smooth m-dimensional Riemannian manifold.

It is a generic property of $\phi \in \text{Diff}^{\,2}(\mathcal{M})$ and $f \in C^2(\mathcal{M}, \mathbb{R})$ that

$$
\begin{array}{ccc}
\mathcal{M} & \longrightarrow & \mathbb{R}^{2m+1} \\
x & \mapsto & \big(f(x), f \circ \phi(x), \ldots, f \circ \phi^{2m}(x)\big)
\end{array}
$$

is an embedding.

Activity 1: The effect of Window Size

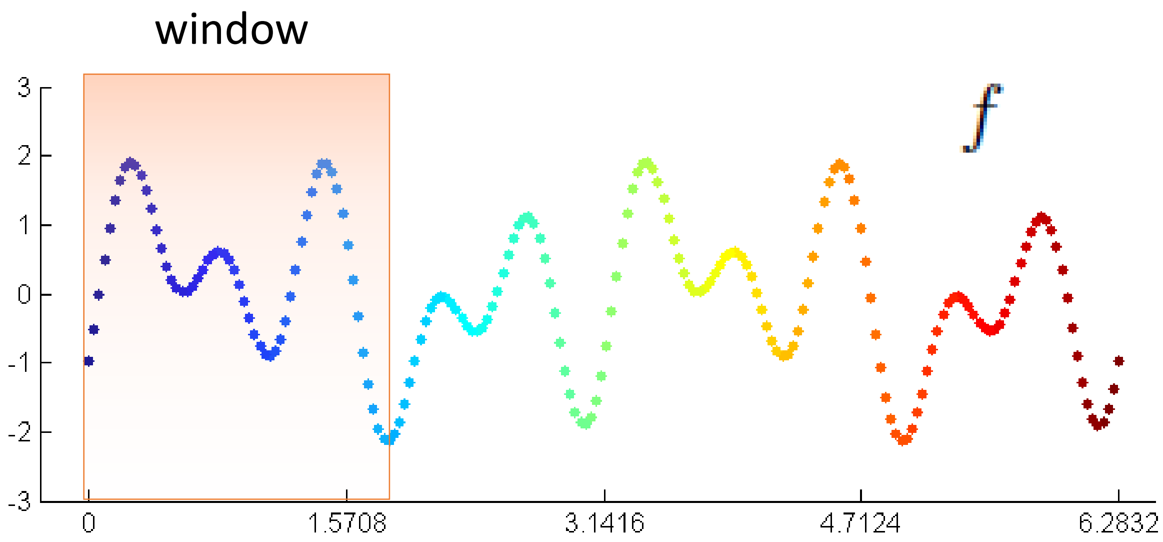
Activity 1

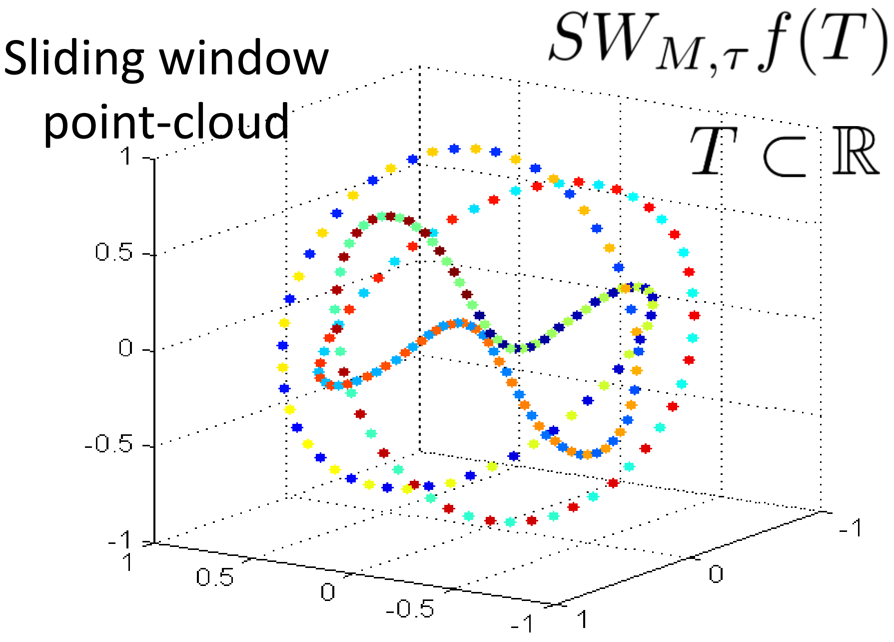
$$SW_{M,\tau}f(t) = \begin{bmatrix} f(t) \\ f(t+\tau) \\ \vdots \\ f(t+M\tau) \end{bmatrix}$$

$\tau \Rightarrow$ Step/delay

$M\tau \Rightarrow$ Window size

$M+1 \Rightarrow$ Dimension



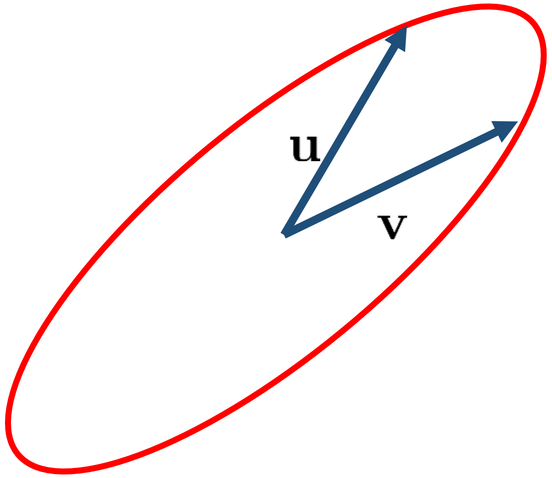


- Launch first python notebook: 1-SlidingWindowBasics
- Question: How does window size, $M\tau$, affect the geometry of $SW_{M,\tau}f(T)$?

Example. Let $L \in \mathbb{N}$ and $f(t) = \cos(Lt)$, then

$$SW_{M,\tau} f(t) = \begin{bmatrix} \cos(Lt) \\ \cos(Lt + L\tau) \\ \vdots \\ \cos(Lt + LM\tau) \end{bmatrix}$$

$$= \cos(Lt) \begin{bmatrix} 1 \\ \cos(L\tau) \\ \vdots \\ \cos(LM\tau) \end{bmatrix} - \sin(Lt) \begin{bmatrix} 0 \\ \sin(L\tau) \\ \vdots \\ \sin(LM\tau) \end{bmatrix}$$



$$= \quad \cos(Lt)\mathbf{u} - \sin(Lt)\mathbf{v}$$ ⟵ An ellipse

Example. Let $L \in \mathbb{N}$ and $f(t) = \cos(Lt)$.

$$SW_{M,\tau}f(t) = \cos(Lt)\mathbf{u} - \sin(Lt)\mathbf{v} \xrightarrow{\text{roundest}} \|\mathbf{u}\| - \|\mathbf{v}\| = \langle \mathbf{u}, \mathbf{v} \rangle = 0$$

$$4\langle \mathbf{u}, \mathbf{v} \rangle^2 + \left(\|\mathbf{u}\|^2 - \|\mathbf{v}\|^2\right)^2 = \frac{\sin^2(L(M+1)\tau)}{\sin^2(L\tau)} = 0$$

$$L(M+1)\tau \equiv 0 \mod \pi \xrightarrow{\text{In particular when}} M\tau = \left(\frac{M}{M+1}\right)\frac{2\pi}{L}$$

Window size

Period length

Generalizing to

$$f \in L^2(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{R})$$

Strategy

- Replace $f(t)$ by its N-truncated Fourier Series

$$S_N f(t) = \sum_{n=0}^{N} a_n \cos(nt) + b_n \sin(nt)$$

- Understand the geometry of $SW_{M,\tau} S_N f(t)$

Sliding Windows and Persistence: An application of topology to signal analysis, J. Perea and J. Harer, *2015*

Theorem 1 (P. and Harer)

Let $f \in C^1(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{R})$ be s.t. $f\left(t + \frac{2\pi}{L}\right) = f(t)$ for all t ($L \in \mathbb{N}$)
and so that $\|f\|_2 = 1$ and $\int f(t)\, dt = 0$.

1. $t \mapsto SW_{M,\tau} S_N f(t)$ is non-degenerate for $M \geq 2N$

2. $SW_{M,\tau} S_N f$ is roundest when $L(M+1)\tau = 2\pi$

$$M\tau = \frac{M}{M+1} \cdot \frac{2\pi}{L}$$

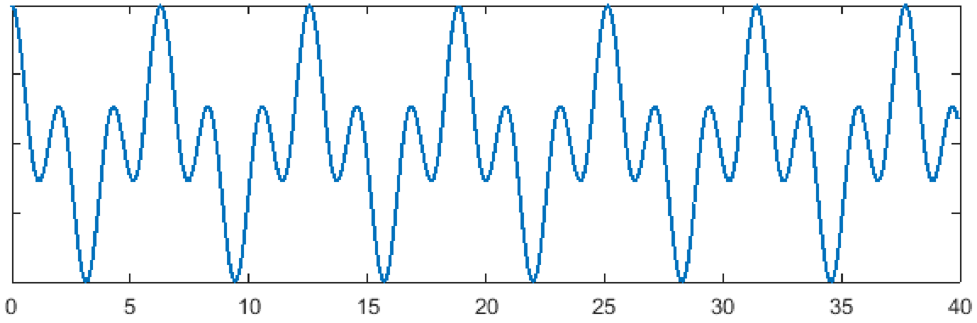
Sliding Windows and Persistence: An application of topology to signal analysis, J. Perea and J. Harer, *2015*

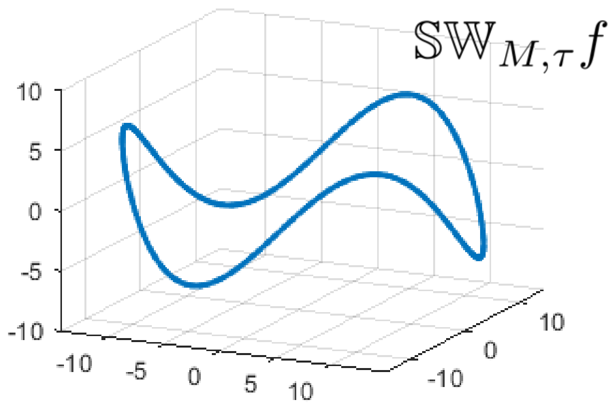
Activity 2: The Geometry in the Spectrum

- Check that the previous theorem is true: the sliding window point cloud is roundest if window ~ period ; the number of dimensions should be greater than twice the number of harmonics

- Question: are the results the same for sums of non-commensurate frequencies?

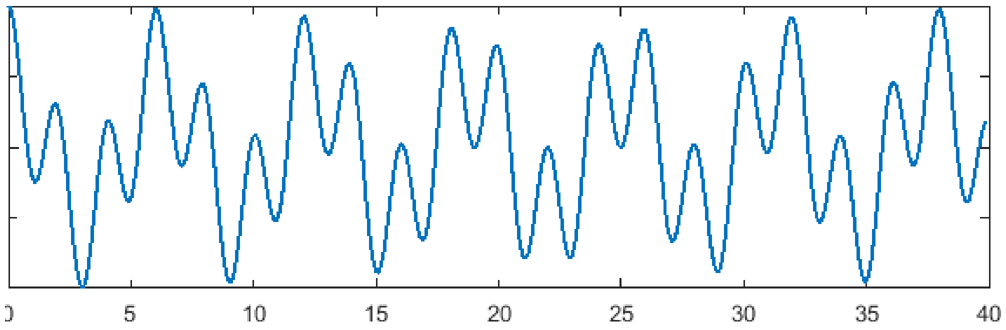
$$f(x) = \cos(t) + \cos(3t)$$

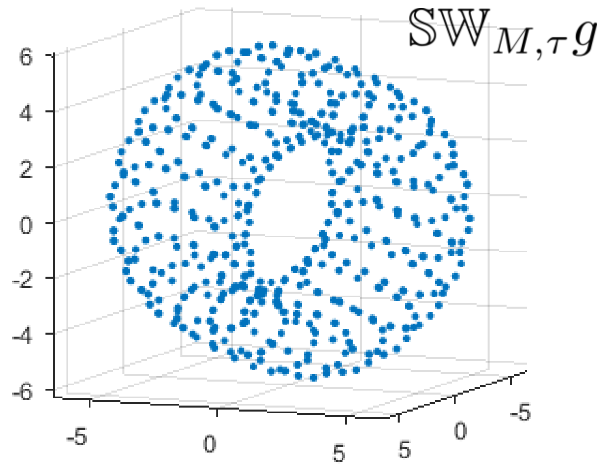




Commensurate

$$g(x) = \cos(t) + \cos(\pi t)$$





Non-Commensurate

Time Series

Sliding Window Embedding

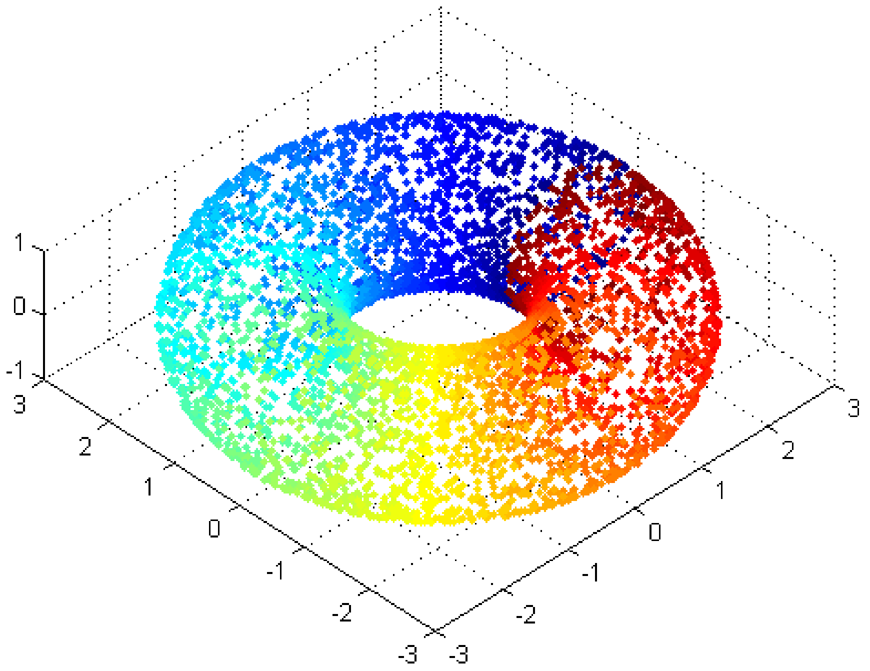
Set up (non-commensurate):

$f : \mathbb{R} \longrightarrow \mathbb{C}$ of the form $f(t) = \sum_{n=0}^{N} c_n e^{i\omega_n t}$

where $\begin{cases} c_0, \ldots, c_N \in \mathbb{C} \setminus \{0\} \\ \omega_0, \ldots, \omega_N \in \mathbb{R}^+ \end{cases}$

$1, \omega_0, \ldots, \omega_N$ are linearly independent over $\mathbb{Q}$

Why is $\mathbb{SW}_{M,\tau} f$ expected to be a hyper torus?



And of what dimension?

$$SW_{M,\tau}f(t) = \begin{bmatrix} f(t) \\ f(t+\tau) \\ \vdots \\ f(t+M\tau) \end{bmatrix}$$

$$f(t) = \sum_{n=0}^{N} c_n e^{i\omega_n t}$$

$$M \in \mathbb{N}, \tau > 0$$

$$= \sum_{n=0}^{N} c_n e^{i\omega_n t} \cdot \begin{bmatrix} 1 \\ e^{i\omega_n \tau} \\ \vdots \\ e^{i\omega_n M\tau} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{i\omega_0 \tau} & e^{i\omega_1 \tau} & \cdots & e^{i\omega_N \tau} \\ \vdots & \vdots & \ddots & \vdots \\ e^{i\omega_0 M\tau} & e^{i\omega_1 M\tau} & \cdots & e^{i\omega_N M\tau} \end{bmatrix} \cdot \begin{bmatrix} c_0 e^{i\omega_0 t} \\ c_1 e^{i\omega_1 t} \\ \vdots \\ c_N e^{i\omega_N t} \end{bmatrix}$$

$$SW_{M,\tau}f(t) = \Omega_f \cdot x_f(t)$$

$$M+1\left\{\underbrace{\begin{bmatrix} & \vdots & \\ \cdots & e^{i\omega_n m\tau} & \cdots \\ & \vdots & \end{bmatrix}}_{N+1}\right.$$

(Transpose of) Vandermonde matrix on

$$e^{i\omega_0\tau}, \ldots, e^{i\omega_N\tau}$$

$$\begin{bmatrix} c_0 e^{i\omega_0 t} \\ \vdots \\ c_N e^{i\omega_N t} \end{bmatrix}$$

Curve on a torus of dimension $N+1$

<u>Theorem</u> (Kronecker):

For $c \in \mathbb{C}$ let $S^1_c = \{z \in \mathbb{C} : |z| = |c|\}$ and let

$$x_f(t) = \begin{bmatrix} c_0 e^{i\omega_0 t} \\ \vdots \\ c_N e^{i\omega_N t} \end{bmatrix}$$

If $1, \omega_0, \ldots, \omega_N$ are linearly independent over $\mathbb{Q}$, then $\{x_f(k) : k \in \mathbb{Z}\}$ is dense on $\mathbb{T}^{N+1} = S^1_{c_0} \times \cdots \times S^1_{c_N}$

$$SW_{M,\tau}f(k) = \Omega_f \cdot x_f(k) \qquad k \in \mathbb{Z}$$

$$M+1\left\{\underbrace{\begin{bmatrix} & \vdots & \\ \cdots & e^{i\omega_n m\tau} & \cdots \\ & \vdots & \end{bmatrix}}_{N+1}\right.$$

(Transpose of) Vandermonde matrix on

$e^{i\omega_0\tau}, \ldots, e^{i\omega_N\tau}$

$$\begin{bmatrix} c_0 e^{i\omega_0 k} \\ \vdots \\ c_N e^{i\omega_N k} \end{bmatrix}$$

Dense subset of

the $(N+1)$ - torus

$\mathbb{T}^{N+1} = S^1_{c_0} \times \cdots \times S^1_{c_N}$

$$SW_{M,\tau}f(k) = \Omega_f \cdot x_f(k)$$

$$M+1 \left\{ \underbrace{\begin{bmatrix} & \vdots & \\ \cdots & e^{i\omega_n m\tau} & \cdots \\ & \vdots & \end{bmatrix}}_{N+1} \right.$$

(Transpose of) Vandermonde matrix on

$e^{i\omega_0\tau}, \ldots, e^{i\omega_N\tau}$

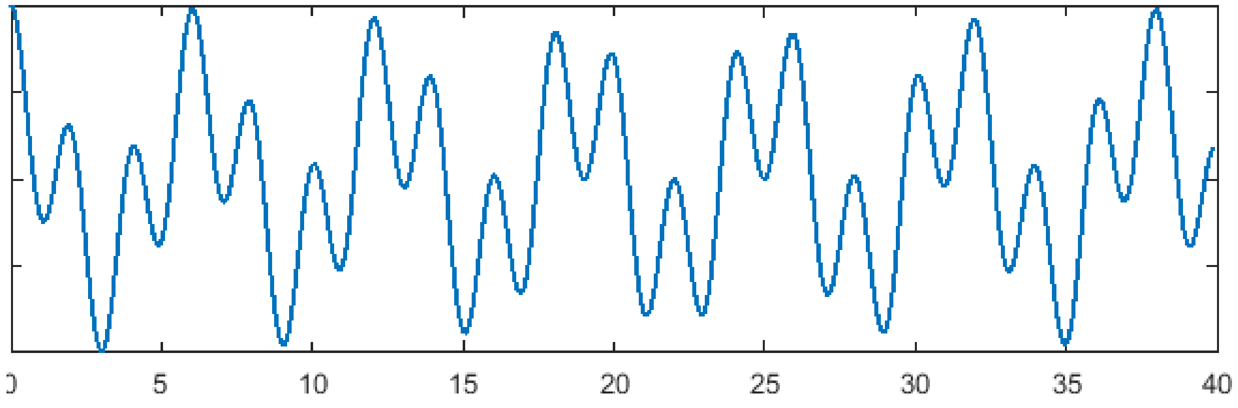
<u>Theorem:</u>

If $0 < \tau \cdot \max\{\omega_n\} < 2\pi$ then Ω_f is full-rank. Moreover, if in addition $M \geq N$ then $\mathbb{SW}_{M,\tau}f = SW_{M,\tau}f(\mathbb{Z})$ is dense in an $(N+1)$- torus.

Activity 3: The Spectrum in the Spectrum

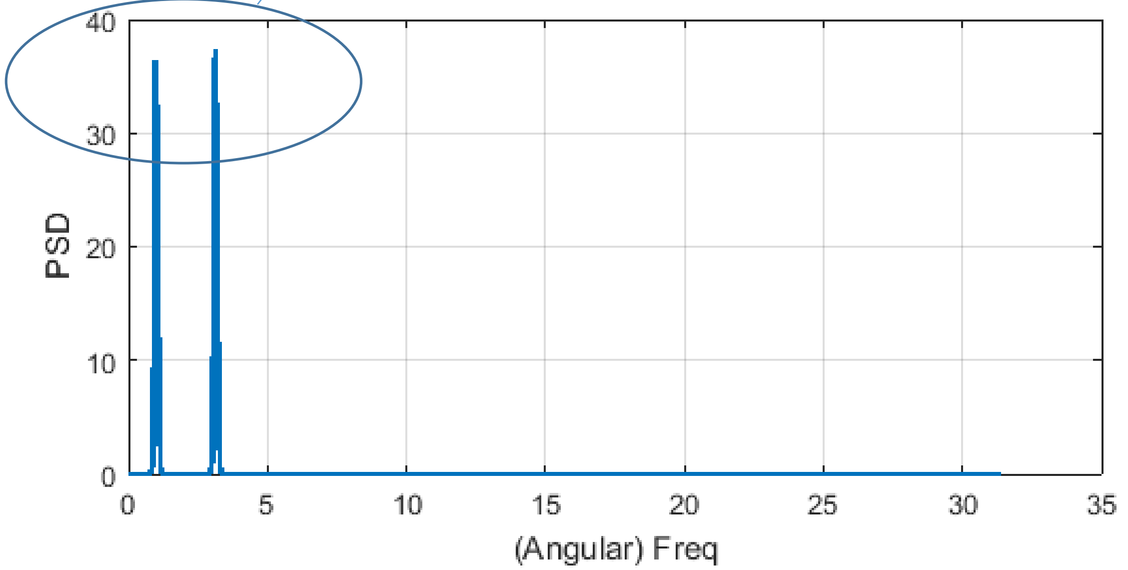
- Question: how would you distinguish between harmonic and harmonic functions using the power spectrum?

$$g(x) = \cos(t) + \cos(\pi t)$$



Time Series

$$\omega_0 \sim 1, \omega_1 \sim \pi$$



Power Spectral Density

Thanks!!

J. Perea and J. Harer, *Sliding Windows and Persistence: An Application of Topological Methods to Signal Analysis,* Foundations of Computational Mathematics, 2015.

J. Perea, A. Deckard, S. Haase and J. Harer, *SW1PerS: Sliding Windows and 1-Persistence Socring; Discovering Periodicity in Time Series Data,* BMC Bioinformatics, 2015.

J. Perea, *Persistent Homology of Toroidal Sliding Window Embeddings,* ICASSP, 2016.