JUNIOR TOPOLOGY SEMINAR — ANDREW GILLETTE
APRIL 9 2008

Biology → huge datasets / multiscale
- cells 10^{-4}
- ⋮
- viruses 10^{-7} m
- water mol. 10^{-9} m

→ geometric information

→ unknown shape / features

<u>Goal:</u> A way to measure <u>topological</u> features & their relative <u>geometric</u> significance

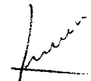
<u>Ex:</u> Gramicidin A (protein)

<u>Ex:</u> Protein docking (ID features)

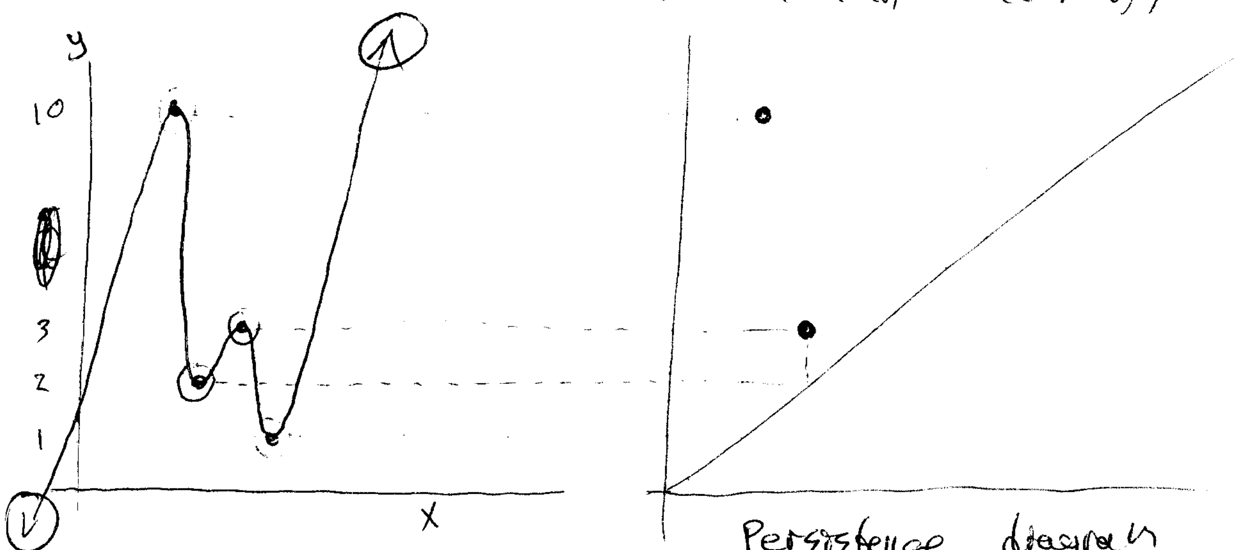
<u>Technique:</u> Persistent Homology

<u>Outline:</u>
- $\boxed{\text{I}}$ Basic Idea
- $\boxed{\text{II}}$ The Math
- $\boxed{\text{III}}$ Recent Theoretical Extensions

I Basic Idea

(c.f.  (no feat) or  (2 feat))

Ex:



Persistence diagram

- Let $f: \mathbb{R} \to \mathbb{R}$ smooth w/ only non-degenerate crit. pts w/ distinct values
- Consider the <u>SUBLEVEL SET</u> $f^{-1}(-\infty, t]$ as t increases
- min $\Longrightarrow$ create component
  max $\Longrightarrow$ merge components; match to higher (younger) of "mergees"
- The <u>PERSISTENCE</u> of a matched pair of critical points (x_1, x_2) is $|f(x_2) - f(x_1)|$
- This = vert. distance of pt to diag in persis. dgm
- "Noise" = points near diagonal (c.f. top)

II The Math

<u>Def:</u> A <u>FILTRATION</u> is a sequence of nested complexes inside a simplicial cplx K:

$$\emptyset = K^0 \subset K^1 \subset \cdots \subset K^m = K$$

<u>Def:</u> $Z_j^\ell :=$ j^{th} cycle group in K^ℓ

$B_j^\ell :=$ j^{th} boundary group in K^ℓ

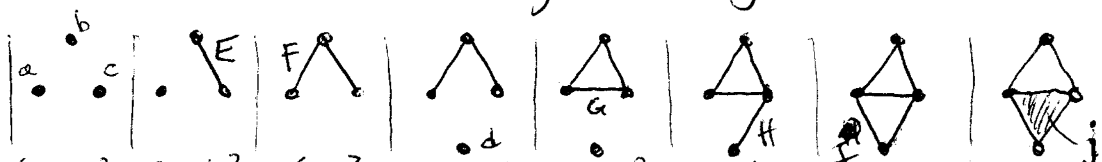
$$H_j^{\ell,p} := Z_j^\ell \big/ (B_j^{\ell+p} \cap Z_j^\ell)$$

= <u>p-PERSISTENT j^{th} hom. gp.</u> of K^ℓ

= {cycles in K^ℓ mod ∂ys in $K^{\ell+p}$}

<u>Rmk:</u> If $H_j^{\ell,p} \neq \{0\}$ for large p, $\exists$ "stable" homological j-features.

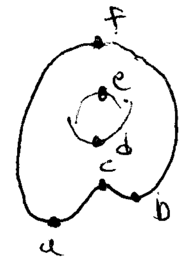
<u>Ex:</u>



H_0	{a,b,c}	{a,b}	{a}	{a,d}	{a,d}	{a}	{a}	{a}
H_1					{EFG} ({G})	{EFG} ({G})	{EFG, EFHI} ({G, I})	{EFG} ({G})

(<u>Def:</u> A <u>POSITIVE</u> simplex creates homology while a <u>NEGATIVE</u> one destroys it. $\exists$ bijection between <u>positive</u> simplices & homology bases)

III Recent Theoretical Extensions

Ex: 

a, b – create 0-hom
c – merge a & b (kills 0-hom)
d – creates 1-hom
e – " 1-hom
f – " 2-hom

Consider Morse function f on a compact surface.

Problems: ① Not all classes are matched
② Highly dependent on "direction" of height fn.

Solution to ①: Extended Persistence (Cohen-Steiner, Edelsbrunner, Harer } ≥2007)

Let $t_0 < t_1 < \cdots < t_m \ (= \infty)$ be regular values bracketing the critical values of f.

Def: $M_K := f^{-1}(-\infty, t_K]$

$M_K \overset{i}{\hookrightarrow} M_{K+1}$ induces $\begin{cases} i_* : H_r(M_K) \longrightarrow H_r(M_{K+1}) \\ i^* : H^r(M_{K+1}) \longrightarrow H^r(M_K) \end{cases}$

(pullback by i)

Thm: POINCARÉ DUALITY: M compact, oriented n-mfld then $H^K(M) \cong (H^{n-K}(M))^* \cong H_{n-K}(M)$

So we get the sequence

$$H_r(M_0) \xrightarrow[i_*]{} \cdots \longrightarrow H_r(M_m) \xrightarrow[\text{P.D.}]{\cong} H^{n-r}(M_m) \xrightarrow[i^*]{} \cdots \xrightarrow[i^*]{} H^{n-r}(M_0)$$

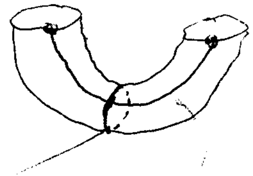
<u>Thm</u>: LEFSCHETZ DUALITY: $\exists$ non-degenerate pairing

$$\langle \cdot , \cdot \rangle : H_r(M, \partial M) \times H_{n-r}(M) \longrightarrow \mathbb{Z}/2\mathbb{Z}$$

given by counting $\cap$s of cycles in M.

$\longrightarrow$ Poincaré Duality is just Lef. Duality w/ $\partial M = \emptyset$

<u>Ex</u>: M



any generator of $H_1(M)$ $\cap$s any generator of $H_1(M, \partial M)$ exactly once

rep. of $H_1 M$ rep. of $H_1(M, \partial M)$

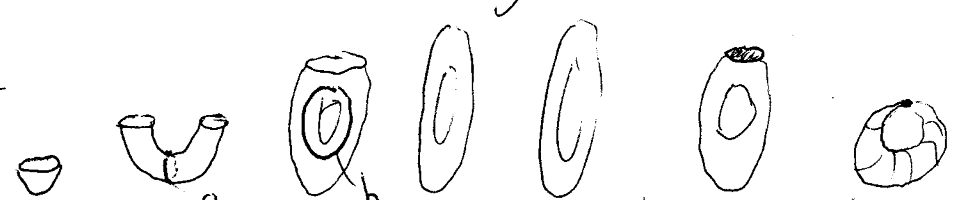
So the sequence becomes

$$H_r(M_0) \longrightarrow \cdots \longrightarrow H_r(M_n) \underset{(\partial M_n = \emptyset)}{\overset{\cong}{\longrightarrow}} H_r(M_n, \partial M_n) \longrightarrow \cdots \longrightarrow H_r(M_0, \partial M_0)$$

factor thru Lefschetz duality

Now all classes get matched!

<u>Ex</u>:



M_0 M_1 M_2 M_3 $(M_3, \partial M_2)$ $(M_2, \partial M_2)$ $(M_1, \partial M_1)$ $(M_0, \partial M_0)$

$\underbrace{M_1 \quad M_2 \quad M_3 \quad (M_3, \partial M_2) \quad (M_2, \partial M_2)}_{\text{life of a}}$

$\underbrace{M_2 \quad M_3 \quad (M_3, \partial M_2) \quad (M_2, \partial M_2) \quad (M_1, \partial M_1)}_{\text{life of b}}$

We have the decomposition:

REGULAR (ORDINARY) persistence: classes born & killed by $H_r(M_u)$

EXTENDED persistence: " born before & killed after $H_r(M_u)$

↳ represents essential classes

RELATIVE persistence: " born & killed after $H_r(M_u)$

(Rmk: Extended pairs are classes of complementary dimension)

What about choice of direction?

Obs: With extended persistence, all critical points get matched

$\therefore$ we gain information on geometrical features from distance between matched pts.

More formally:

Def: The HEIGHT OF $x \in M$ IN DIRECTION $\vec{u} \in S^2$ is

$$h: M \times S^2 \longrightarrow \mathbb{R}$$

where $h(x, \vec{u}) :=$ signed distance from $\vec{x}$ to the plane through the origin w/ normal $\vec{u}$

Rmk: h is Morse for a.e. $\vec{u} \in S^2$

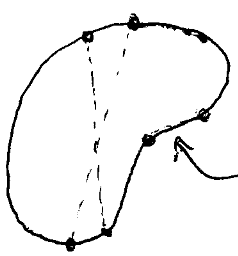
Not: Let $\vec{n}_x :=$ normal vector to $T_x M$.

Then x is a crit. point for $h(\cdot, \vec{n}_x)$ & $h(\cdot, -\vec{n}_x)$ & no other $\vec{u} \in S^2$.

$\therefore$ $\exists!$ $y \in M$ s.t. x paired to y by $h(\cdot, \pm\vec{n}_x)$.

Def: The ELEVATION AT $x \in M$ where y paired to x is $E: M \longrightarrow \mathbb{R}_{\geq 0}$ w/ $E(x) = |h(x, \vec{n}_x) - h(y, \vec{n}_x)|$

Ex:



at inflection points $E = 0$

Claim: Geometric "features" occur in regions of large E values

Claim: To "dock" two surfaces, features should be aligned.