Last Time

<u>Recall</u>: A persistence module is a collection of vector spaces and maps $\varphi_{\alpha,\beta} : H_k(X_\alpha) \to H_k(X_\beta)$ if $\alpha \leq \beta$.

A persistence module has a persistence diagram.
There is a matching distance between diagrams called the bottleneck distance.

<u>Stability Thm</u>: Let $f : \Delta \to \mathbb{R}$ & $g : \Delta \to \mathbb{R}$. Then

$$d_B(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq \|f - g\|_\infty$$

bottleneck distance

persistence diagrams

<u>Distances between modules</u>: Let F & G denote persistence modules (not necessarily over the same space). F & G are ε-interleaved if the following diagrams commute

$$\begin{tikzcd} F_\alpha \arrow[rr] \arrow[dr] & & F_{\alpha+2\varepsilon} \\ & G_{\alpha+\varepsilon} \arrow[ur] & \end{tikzcd} \qquad \begin{tikzcd} F_\alpha \arrow[r] \arrow[d] & F_{\alpha+\delta} \arrow[d] \\ G_{\alpha+\varepsilon} \arrow[r] & G_{\alpha+\varepsilon+\delta} \end{tikzcd}$$

$$\begin{tikzcd} & F_{\alpha+\varepsilon} \arrow[dr] & \\ G_\alpha \arrow[ur] \arrow[rr] & & G_{\alpha+2\varepsilon} \end{tikzcd} \qquad \begin{tikzcd} F_{\alpha+\varepsilon} \arrow[r] & F_{\alpha+\varepsilon+\delta} \\ G_\alpha \arrow[u] \arrow[r] & G_{\alpha+\delta} \arrow[u] \end{tikzcd}$$

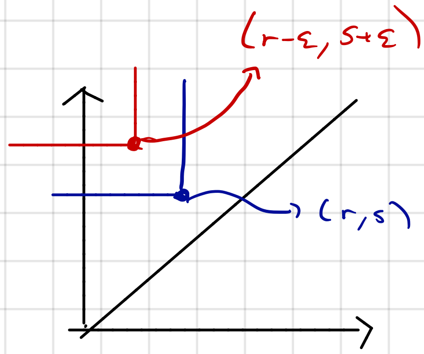
Observation: $\|f-g\|_\infty$ implies F & G are ε-interleaved

Proof: $f^{-1}(-\infty,\alpha] \subseteq g(-\infty,\alpha+\varepsilon] \subseteq f^{-1}(-\infty,\alpha+2\varepsilon]$

+ (induced maps from inclusions commute)

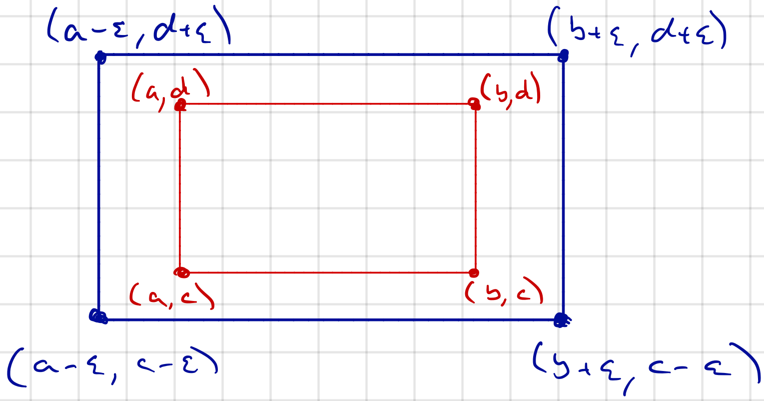
Outline:

1) Quadrant Lemma



$$F \sim_\varepsilon G \implies \boxed{\beta_k^{r,s}(F) \geq \beta_k^{r-\varepsilon,s+\varepsilon}(G)}$$

2) Box Lemma



$\square$: # points in blue rectangle for F

$\square$: # points in red rectangle for G

$$\boxed{\square \leq \square}$$

3) Easy Bijection Lemma & Intepolation

if ε is small enough

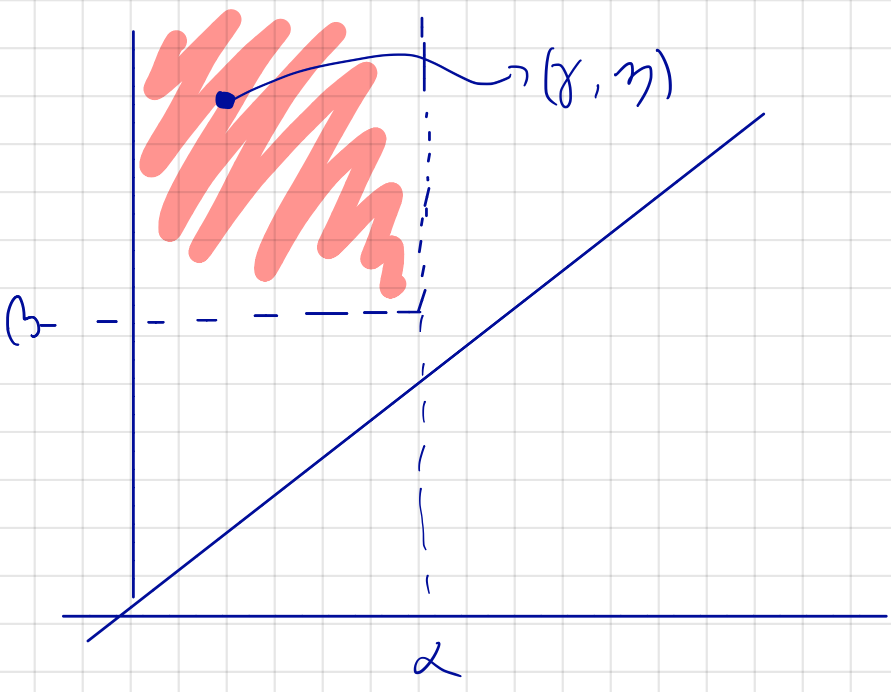
$$\boxed{\square = \square}$$

intepolate between f & g in small enough steps

Quadrant Lemma

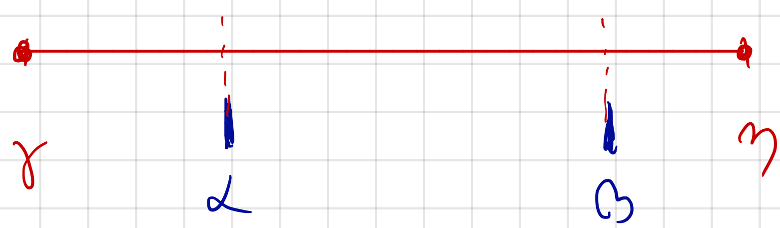
Recall: A point (α, β) in $Dgm(f)$ implies an element in $im\,(H_k(X_\alpha) \to H_k(X_\beta))$, ie a class is born at α & dies a β.

Quadrant: Assume there are no topological changes at α & β. The number of points in $Dgm(f)$ that are in the upper-left quadrant of (α, β) is equal to $rk\,(im\, H_k(X_\alpha) \to H_k(X_\beta))$



Note: this point represents an element born before α (since $\gamma < \alpha$) & dies after β (since $\eta > \beta$), so it contributes to the image.

Barcode picture



Limit: Quandrant of (α, α) is the rank of the vector space at α.

Denote the # of points in the upper-left quadrant of (r,s) (or rk im $H_k(X_\alpha) \to H_k(X_\beta)$) by $\beta_k^{r,s}$ or the r,s persistent Betti number

Quadrant Lemma

$$\beta_k^{(r,s)}(F) \geq \beta_k^{(r-\varepsilon, s+\varepsilon)}(G)$$

Proof: Assume $F \sim_\varepsilon G$, there exists a diagram

$$\begin{tikzcd}
H_k(G_{r-\varepsilon}) \arrow[rrr, "g"] \arrow[dr, "\varphi"'] & & & H_k(G_{s+\varepsilon}) \\
& H_k(F_r) \arrow[r, "f"'] & H_k(F_s) \arrow[ur, "\psi"'] &
\end{tikzcd}$$

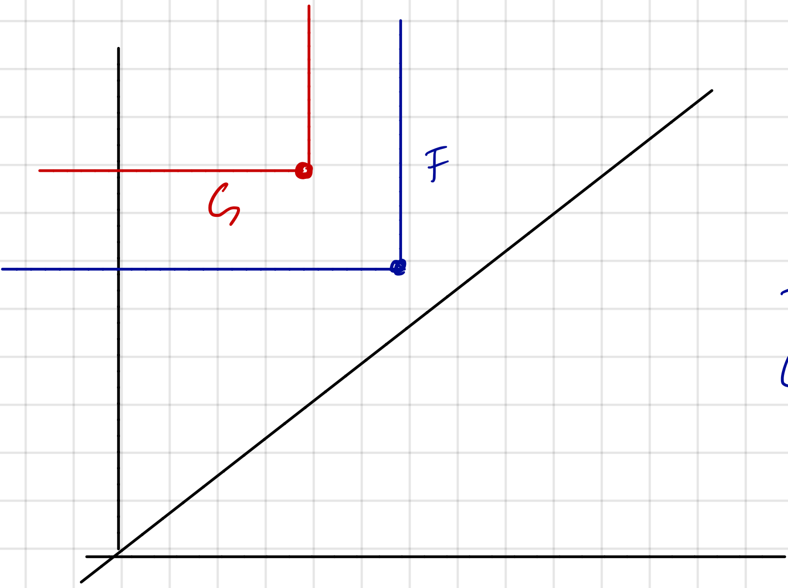
By the definition of ε-interleaving, $g = \psi \circ f \circ \varphi$

Hence $rk(\text{im } f) \geq rk(\text{im } \psi \circ f \circ \varphi) = rk(\text{im } g)$

This implies the result.

Remark: If you are comfortable with diagram chasing, the above is obvious otherwise try to prove the above diagram commutes using the definitions of ε-interleaving.

What does this imply?

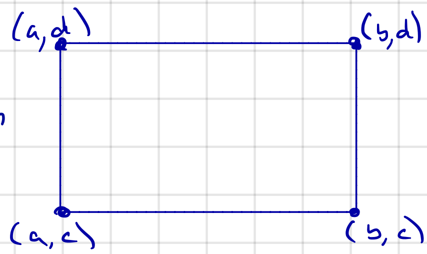


We can upper bound the number of points in quadrant (r,s) in $\mathrm{Dgm}(G)$ by $\#$ of points in quadrant $(r+\varepsilon, s-\varepsilon)$ in $\mathrm{Dgm}(F)$

Generalization: Box Lemma

Box Lemma

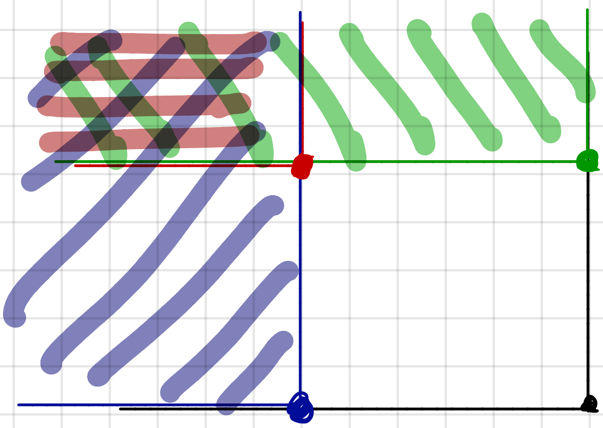
Obs.: The number of points in



is

$$\boxed{rk(\square) = \beta_k^{b,c} - \beta_k^{b,d} - \beta_k^{a,c} + \beta_k^{a,d}}$$

Proof:



inclusion-exclusion

__Assumption__: all chosen values do not correspond to topological changes. r is a homological regular value if there exists a neighborhood $(r-\delta, r+\delta)$ such that for all $\alpha \leq \beta \in (r-\delta, r+\delta)$

$$H_k(X_\alpha) \xrightarrow{\cong} H_k(X_\beta)$$

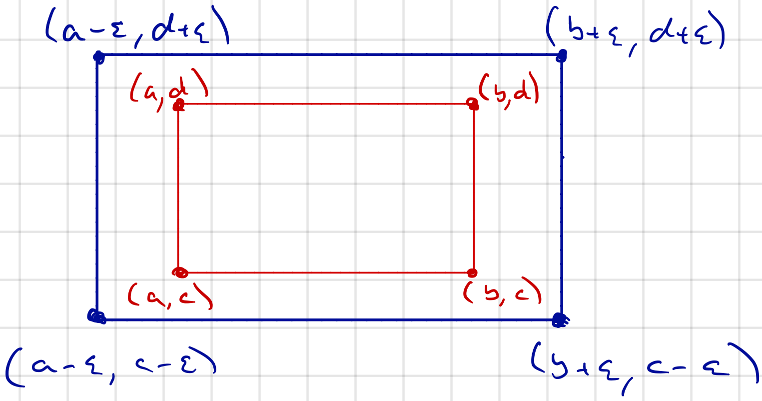
__Q1__ Why is this ok for finite simplicial complexes

__Q2__ In the original paper, there is a small mistake in how homological regular value is defined can you figure out the problem? __Hint__: it does not appear when considering simplicial complexes.

__Box Lemma__: Let F & G be ε-interleaved.

$rk(\square)$: # points in blue rectangle for F

$rk(\square)$: # points in red rectangle for G



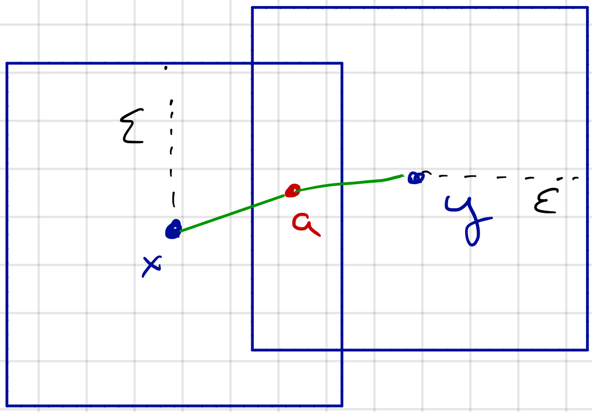
$$\boxed{rk(\square) \leq rk(\square)}$$

__Remark__: # of points in a box in $Dgm(G)$ can be upperbounded by # of points in ε-bigger box in $Dgm(F)$.

Proof at the end

Easy Bijection Lemma

First, observe that if $F \sim_\varepsilon G$, then for every point in $\mathrm{Dgm}(F)$ which is farther than ε away from the diagonal, ie the bar is of length at least 2ε, there is a corresponding point in $\mathrm{Dgm}(G)$. This is called the Hausdorff distance



$x, y \in \mathrm{Dgm}(F)$
$a \in \mathrm{Dgm}(G)$

Say we are lucky, F & G are ε-interleaved, where all points in $\mathrm{Dgm}(F)$ are 2ε-seperated from each other & the diagonal. Then the box lemma implies

$$d_B(\mathrm{Dgm}(F), \mathrm{Dgm}(G)) \leq \varepsilon$$

<u>Proof</u>: The box lemma around each point in $\mathrm{Dgm}(F)$ implies there is an injective set map $\mathrm{Dgm}(F) \to \mathrm{Dgm}(G)$ with longest edge of length at most ε. Likewise the box lemma around each point in $\mathrm{Dgm}(G)$ implies the map is a bijection.

Restatement

Observe: For every point p in the persistence diagram of F ($Dgm(f)$), there must be a point in the ε-box around p ($\square_p^{\varepsilon}$) in the persistence diagram of G ($Dgm(g)$)

Proof: Quadrant Lemma

Easy case: If ε is small enough such that all points are at least 2ε apart, then the bottleneck distance is bounded from above by ε.

Proof: Construct the following matching: for each point $p \in Dgm(f)$ there exists a point $q \in Dgm(g)$, by the Quadrant Lemma. However by the assumption on separation of points, there is no other point in $Dgm(f)$ which is within ε of q. Hence it is a bijection.

Final Step: Reduce general case to easy case via interpolation

Interpolation

Lemma: Let $\|f_0 - f_1\|_\infty \leq \varepsilon$ & $f_t = (1-t)f_0 + tf_1$ for $t \in [0,1]$

then, $\|f_0 - f_t\|_\infty \leq t\varepsilon$ & $\|f_t - f_1\|_\infty \leq (1-t)\varepsilon$

Corollary: If we have a sequence of functions $f_1, f_2, \ldots f_n$ such that $\|f_i - f_{i-1}\|_\infty \leq \varepsilon_i$, then

$$\|f_1 - f_n\|_\infty \leq \sum_{i=1}^{n-1} \varepsilon_i$$

Proof: Triangle inequality.

Theorem: If $f, g : K \to \mathbb{R}$ & $\|f - g\|_\infty \leq \varepsilon$ then

$$d_B(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq \varepsilon.$$

Proof: Let $h_t = tf - (1-t)g$. Let δ denote the minimum separation of points in $\mathrm{Dgm}(f)$ & $\mathrm{Dgm}(g)$. Sample t_i such that $|t_i - t_{i-1}| \leq \frac{\delta}{2}$

At each step apply the easy case (easy bijection lemma) to see that

$$d_B(\mathrm{Dgm}(h_{t_i}), \mathrm{Dgm}(h_{t_{i-1}})) \leq |t_i - t_{i-1}|$$

Summing up yields the result.

Remarks

* This proof relies on interpolation of persistence modules. In the case we covered, we used function interpolation. (so the two functions had to be defined on the same space)

* It is possible to do algebraic interpolation. Either

  - constructing an explicit interp.
  - via categorical arguments (either general or via universal objects)

Outline of alternative proof

<u>Proof</u>: In the simplicial complex case, each simplex creates or kills a cycle. One can "track" points in the diagram through the interpolation by tracking how simplicies function values change through the interpolation.

* The alternate proof makes use of the fact that if $f(\sigma)$ is unique for all $\sigma \in \Delta$, then the points in the diagram can be viewed as $(f(\sigma), f(\tau))$ $\Rightarrow$ so we get a map Ω.

$(b,d) \in \mathrm{Dgm}(f)$ $\quad \Omega(b,d) = (\sigma,\tau) \in \Delta \times \Delta$ st $f(\sigma) = b$ & $f(\tau) = d$

Studying this map has other applications (last lecture)

* The most natural explicit construction of algebraic interleaving is via universal constructions.

Proof of Box Lemma

The following proof follows the original proof

Recall, we have two functions $f, g : K \to \mathbb{R}$

Defs: $F_\alpha := H_k(f^{-1}(-\infty, \alpha])$

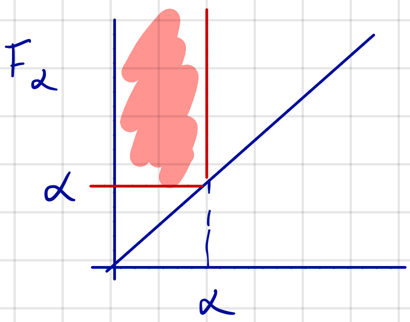
$$F_\alpha^\beta := \operatorname{im}(F_\alpha \longrightarrow F_\beta)$$

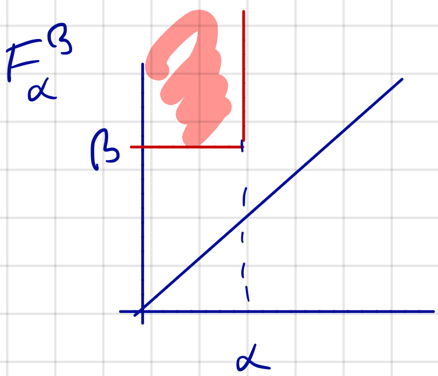
$$F_\alpha^{\beta,\gamma} := \ker(F_\alpha^\beta \longrightarrow F_\alpha^\gamma)$$

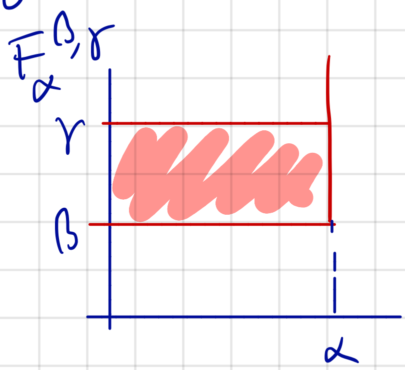
$$\begin{tikzcd} {[x]} \arrow[r] & {[x]} \\ F_\alpha \arrow[r] \arrow[d, "\cong"'] & F_\beta \arrow[d] \\ F_\alpha \arrow[r] & F_\gamma \\ {[x]} \arrow[r] & 0 \end{tikzcd}$$

$F_\alpha^{\beta,\gamma}$

What do these mean in terms of diagrams.



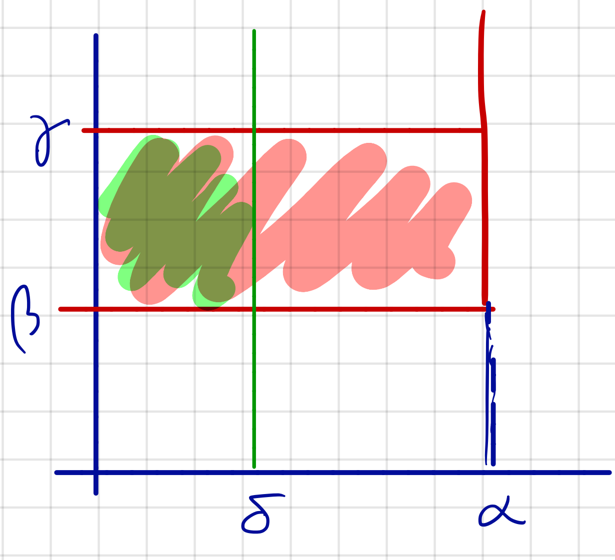




A <u>box</u> is $F_\alpha^{\beta,\gamma} \big/ F_\delta^{\beta,\gamma}$

or

$\operatorname{rank} F_\alpha^{\beta,\gamma} - \operatorname{rank} F_\delta^{\beta,\gamma}$



Relating F & G

Define $\varphi_\alpha : F_\alpha \to G_{\alpha+\varepsilon}$

$\psi_\alpha : G_\alpha \to F_{\alpha+\varepsilon}$

Consider

$$\begin{tikzcd} F_{a+\varepsilon}^{d-\varepsilon} \arrow[r, hook, "r_2"] & F_{b-\varepsilon}^{d-\varepsilon} \\ F_{a+\varepsilon}^{c+\varepsilon} \arrow[r, hook, "r_1"'] \arrow[u, two heads, "u_2"] & F_{b-\varepsilon}^{c+\varepsilon} \arrow[u, two heads, "u_3"'] \end{tikzcd}$$

Recall $\hookrightarrow$ means an injective (monic) morphism, & $\twoheadrightarrow$ is a surjective (epic) morphism.

This assumes

$$a+\varepsilon < b-\varepsilon < c+\varepsilon < d-\varepsilon$$

Additionally,

$$\begin{tikzcd} G_a^d \arrow[r] & G_b^d \\ G_a^c \arrow[r, hook] \arrow[u, two heads] & G_b^c \arrow[u, two heads] \end{tikzcd}$$

$E_b^c = \psi^{-1}(\underbrace{F_{b-\varepsilon}^{c+\varepsilon, d-\varepsilon}}_{\ker u_3}) \cap G_b^c$

To relate the two, take the preimage of $\psi : G \to F$ and consider the classes which are in G_b^c (persistent from b to c)

$$E_b^c = \psi^{-1}(F_{b-\varepsilon}^{c+\varepsilon, d-\varepsilon}) \cap G_b^c$$

$$= \psi^{-1}(\ker u_3) \cap G_b^c$$

There is the following diagram commutes by construction of E_b^c

$$\begin{tikzcd}
 & F_{b-\varepsilon}^{c+\varepsilon} \arrow[r] & F_{b-\varepsilon}^{d-\varepsilon} \arrow[dr] & \\
E_b^c \arrow[ur] \arrow[rrr, "u_4"'] & & & G_b^d
\end{tikzcd} \qquad (1)$$

u_4 is just the restriction of $G_b^c \to G_b^d$ to E_b^c.

Define: $E_a^d = E_b^c \cap G_a^c$: this is well-defined since $G_a^c \hookrightarrow G_b^c$.

We can combine them.

$$\begin{tikzcd}
G_a^d \arrow[rrr, hook, "r_3"] & & & G_b^d \\
 & F_{a+\varepsilon}^{d-\varepsilon} \arrow[r, hook, "r_2"] & F_{b-\varepsilon}^{d-\varepsilon} \arrow[ur, "4"] & \\
 & F_{a+\varepsilon}^{c+\varepsilon} \arrow[u, "u_2"] \arrow[r, hook, "r_1"'] & F_{b-\varepsilon}^{c+\varepsilon} \arrow[u, "u_3"'] & \\
E_a^c \arrow[uuu, "u_1"] \arrow[ur, "\ell"] \arrow[rrr, hook] & & & E_b^d \arrow[uuu, "u_4"'] \arrow[ul]
\end{tikzcd}$$

From (1) we can see $\ker u_3 = \ker u_4$. From the above since r_3 is injective we can conclude that $\ker u_2 = \ker u_1$. Since we consider the subspace E_b^d, the inequality follows.

Machine Learning & Statistics with Persistence Diagrams

Example: We are given multiple black & white images, say coming from two different materials. Using pixel intensity, we construct a filtration for each image and compute the corresponding persistence diagrams.

Question: Can we classify based on the diagrams?

Observation: We have distances between diagrams so it should be possible. However the space of persistence diagrams is not particularly nice, e.g. no unique geodesics.
Most Machine Learning / Statistics requires something nicer.

General Approach: Define functionals on diagrams & work with those.

Kernel trick: Given a distance matrix, one can lift this to a Hilbert space using a (Mercer) kernel

$$\text{e.g. } f(x,y) = e^{-d(x,y)^2}$$

Summary of Functionals

General idea: Extend diagrams to function over $\mathbb{R}^2$ & compare functions.

1) Gaussian-weighted kernel (Bauer, Kerber, Reininghaus)

   * convolve gaussian $(0, \sigma)$ with the diagram & compare function

2) Rank function or weighted variants (Robbins & Turner)

3) Persistent Images (Adams et al)

4) Landscapes (Bubenik)

Many more...

Note 1: Average of functional does not necessarily correspond to a persistence diagram

   e.g. the average of two rank functions is not necessarily a rank function

Note 2: Once we have a functional, we can treat it as a vector and use standard techniques:

- regression, SVM, PCA, neural networks, clustering, etc.