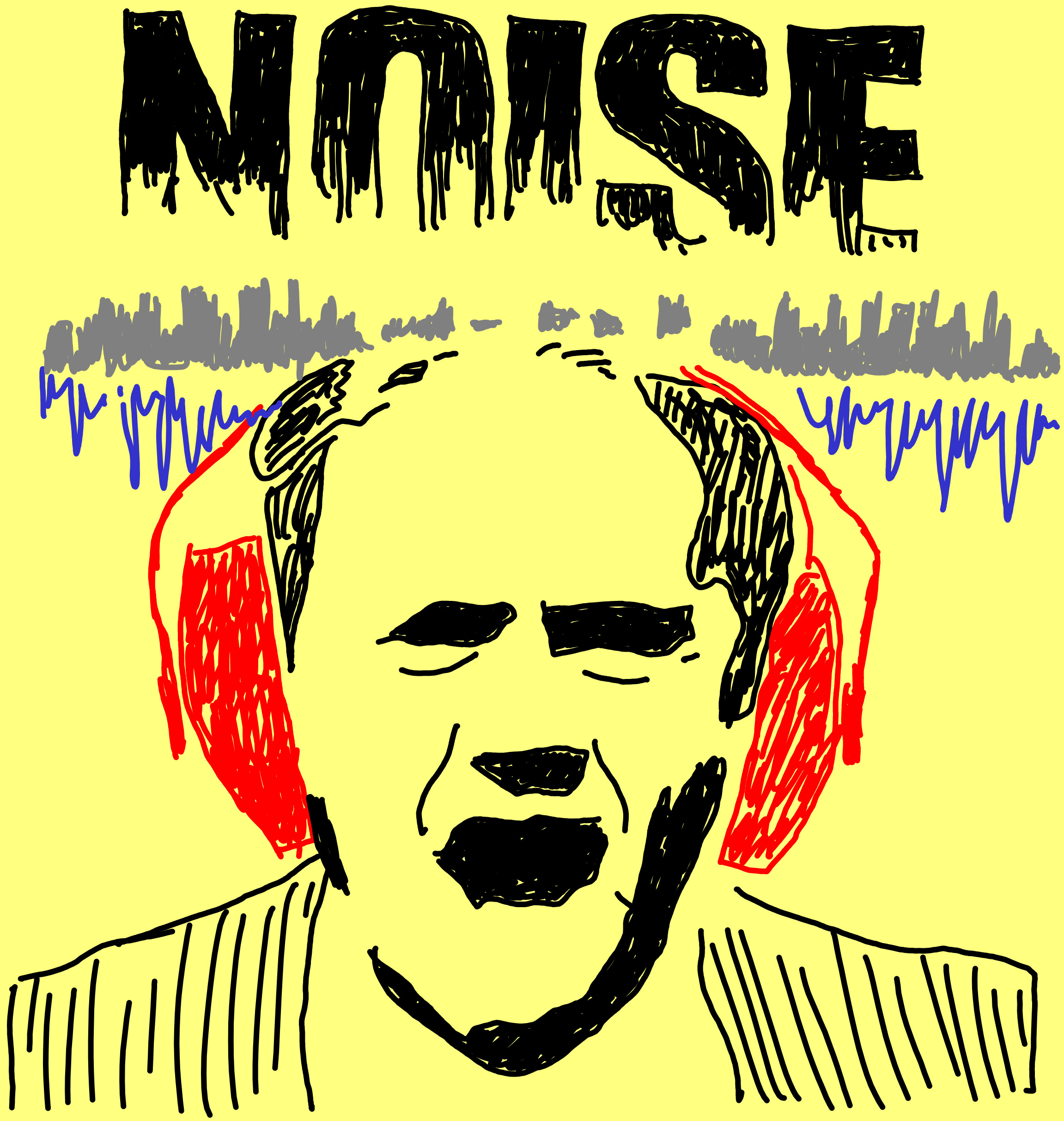
Measuring with Algebra

CTIC '09

Measuring with Algebra

Herbert Edelsbrunner

IST Austria,
Duke University & Geomagic


NOISE

NOISE

. . . IS ALWAYS AND EVERYWHERE.

DOGMA:

measure features and threshold noise

(without changing the data)

I Persistent Homology

II L_p-Stability and Somites

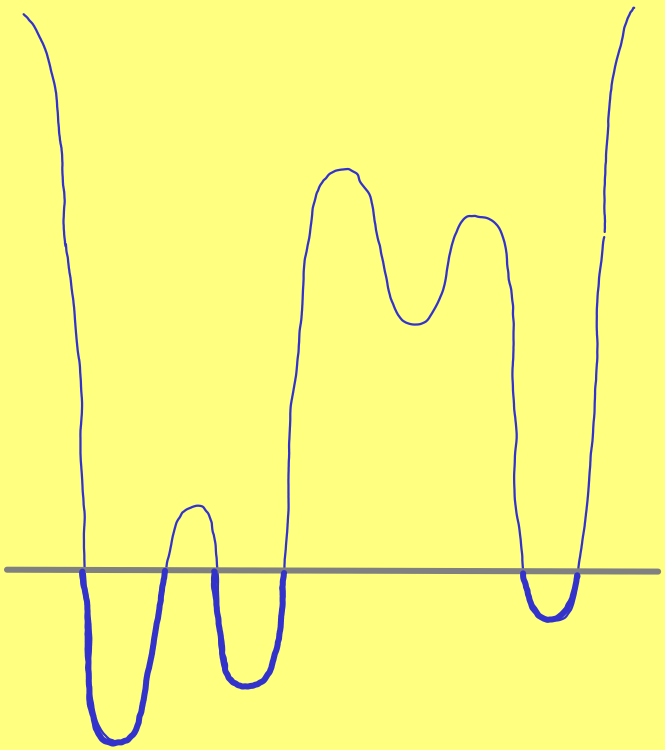
III Contour Stability

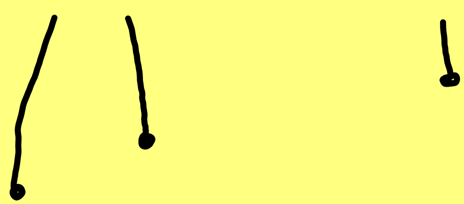
I.1 One-D Functions

function

$f : \mathbb{S}^1 \to \mathbb{R}$

I.1 One-D Functions

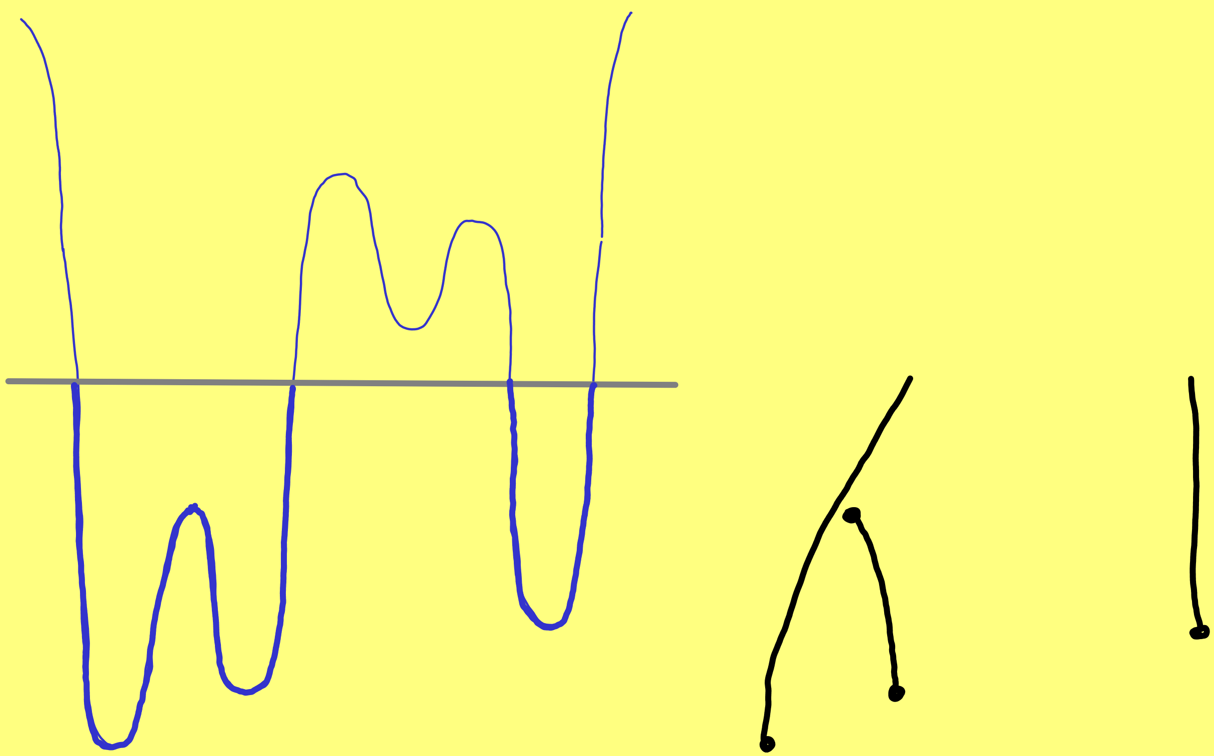




function

$f : \mathbb{S}^1 \to \mathbb{R}$

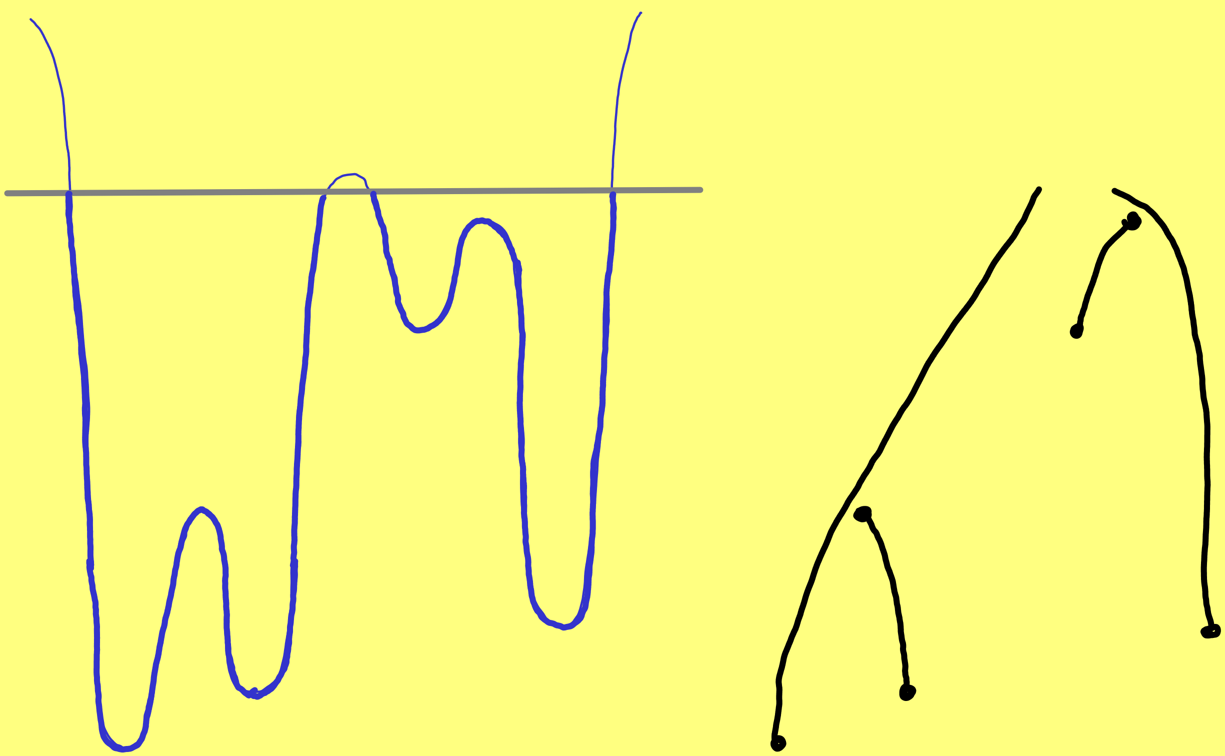
I.1 One-D Functions



function

$f : S^1 \to \mathbb{R}$

I.1 One-D Functions



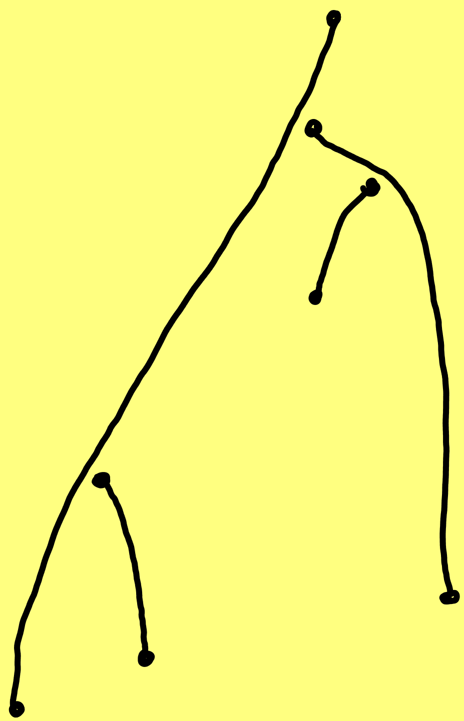
function

$f : S^1 \to \mathbb{R}$

I.1 One-D Functions

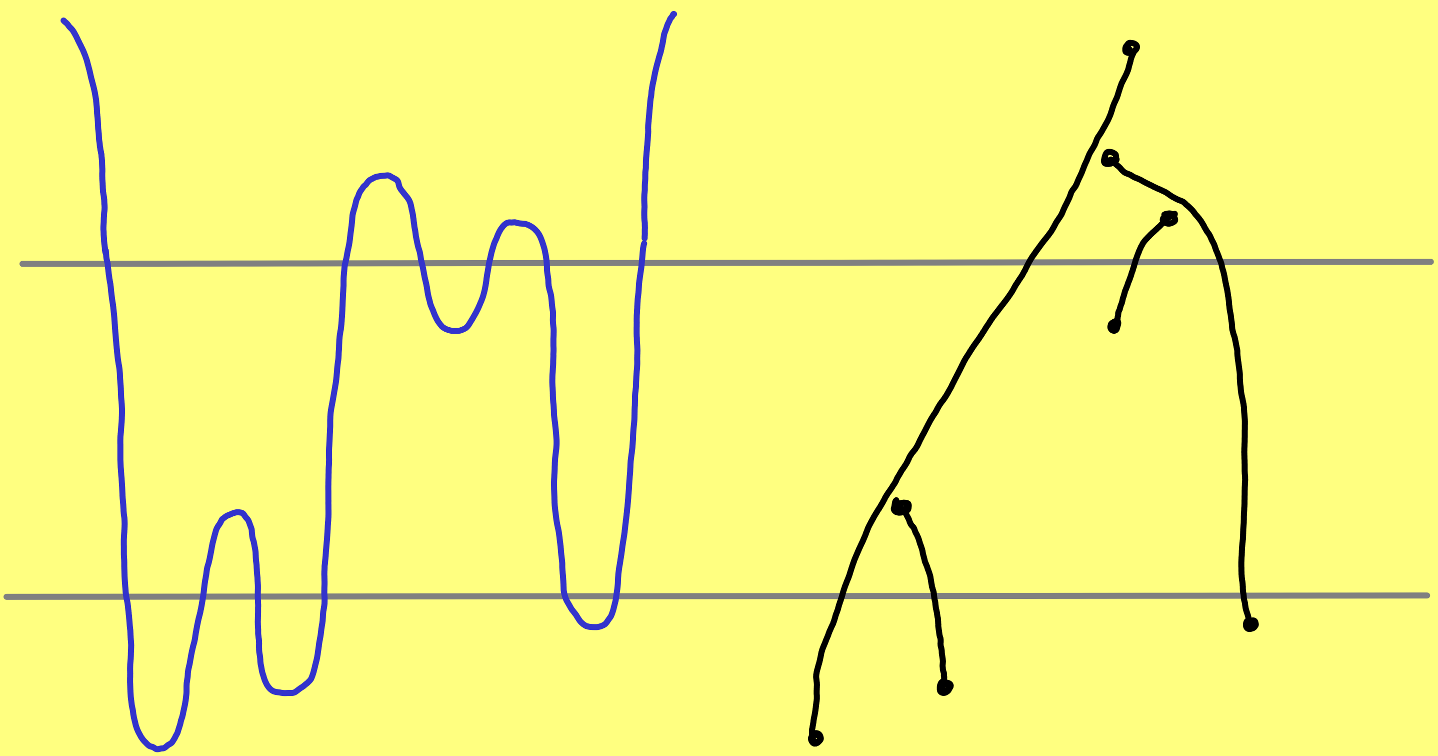
function

$f : \mathbb{S}^1 \to \mathbb{R}$



merge tree

I.1 One-D Functions

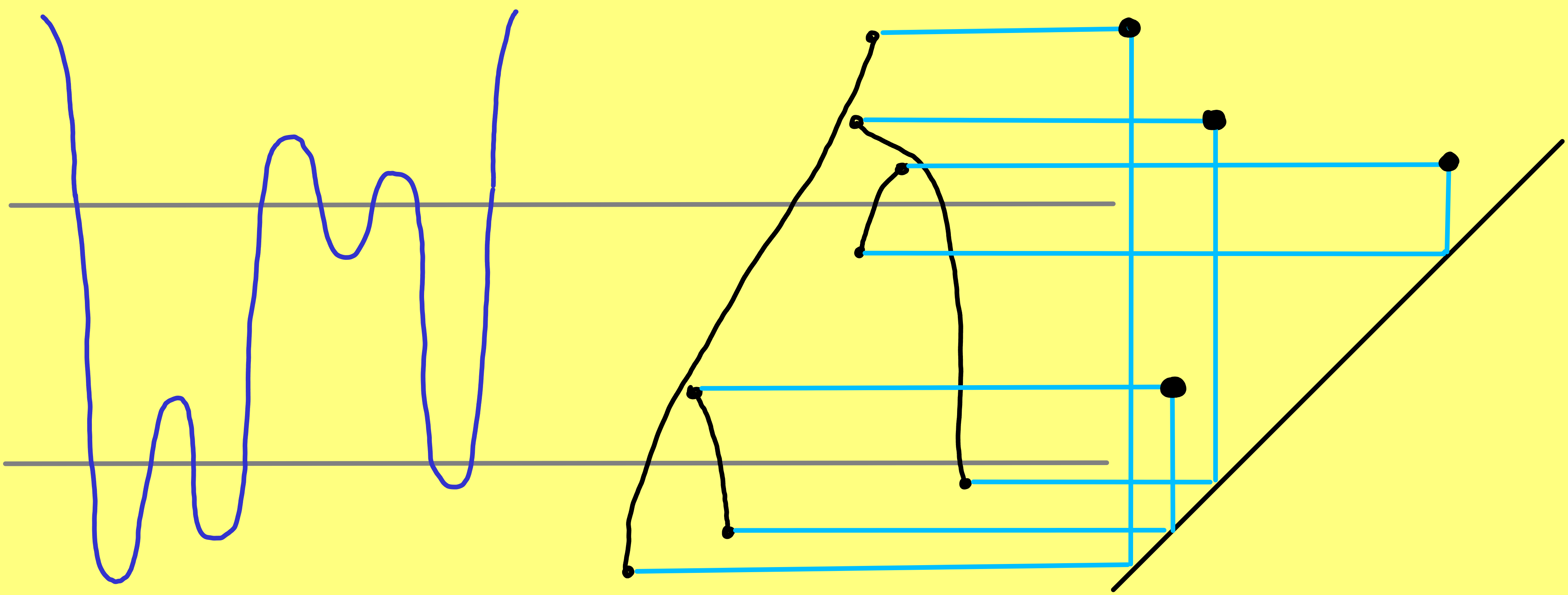


function

$f : \mathbb{S}^1 \to \mathbb{R}$

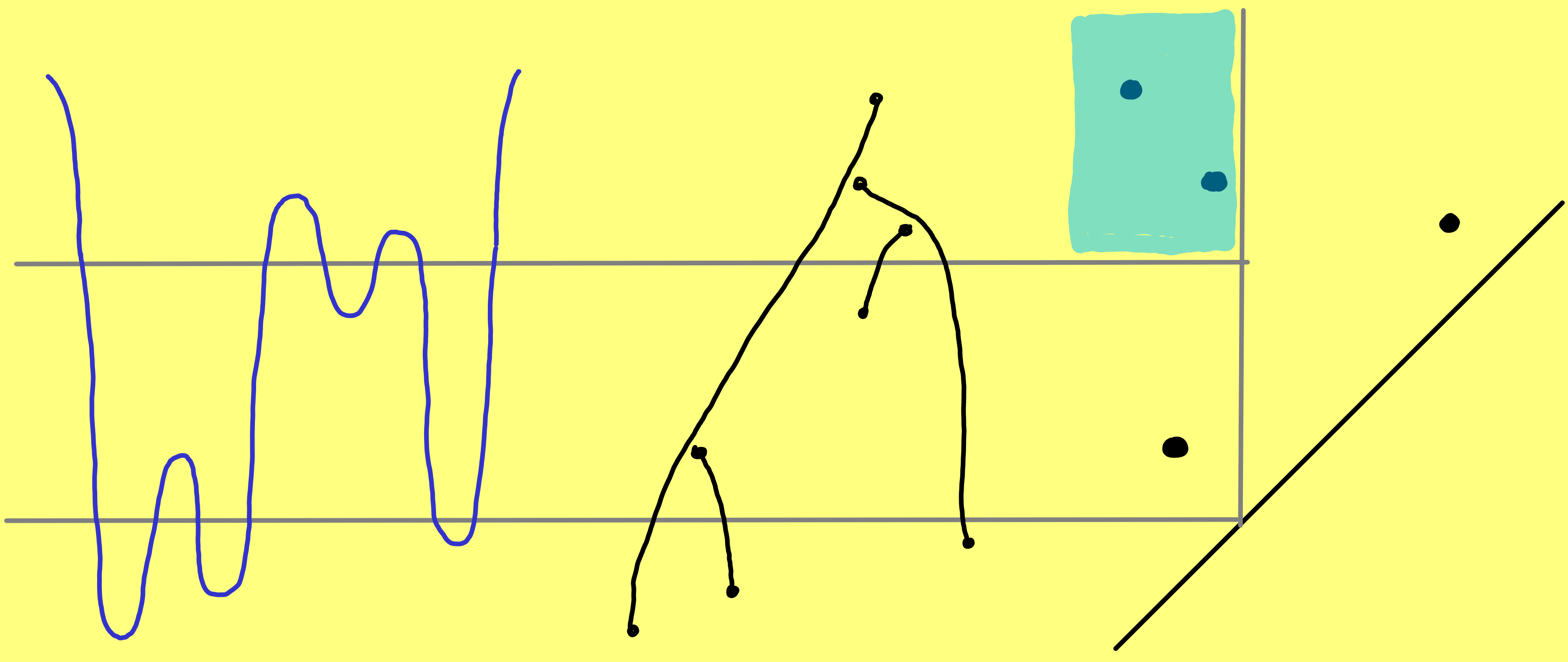
merge tree

I.1 One-D Functions



function $f : \mathbb{S}^1 \to \mathbb{R}$ | merge tree | persistence diagram $\mathrm{Dgm}_0(f)$

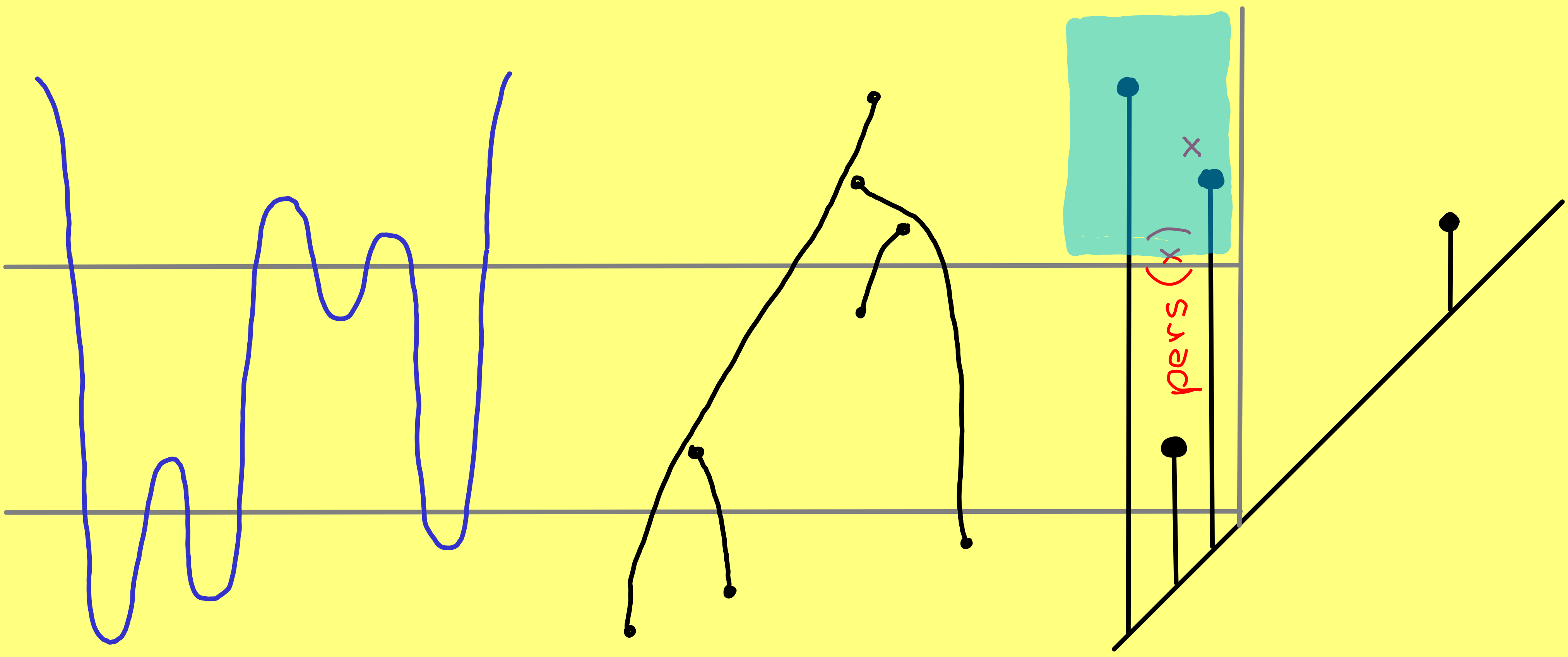
I.1 One-D Functions



function $\quad$ merge tree $\quad$ persistence diagram

$f : \mathbb{S}^1 \to \mathbb{R}$ $\quad$ $\mathrm{Dgm}_0(f)$

I.1 One-D Functions



function $f : \mathbb{S}^1 \to \mathbb{R}$

merge tree

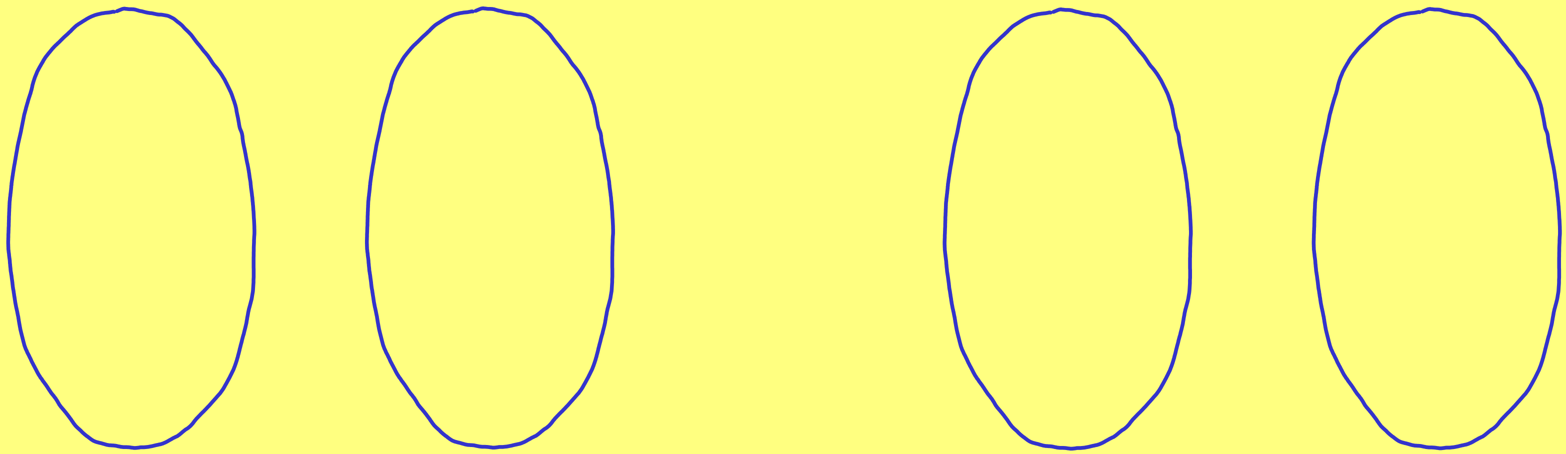
persistence diagram $Dgm_0(f)$

I.2 Filtrations

$$\ldots \subseteq K_{i-1} \subseteq K_i \subseteq \ldots \subseteq K_{j-1} \subseteq K_j \subseteq \ldots$$

I.2 Filtrations

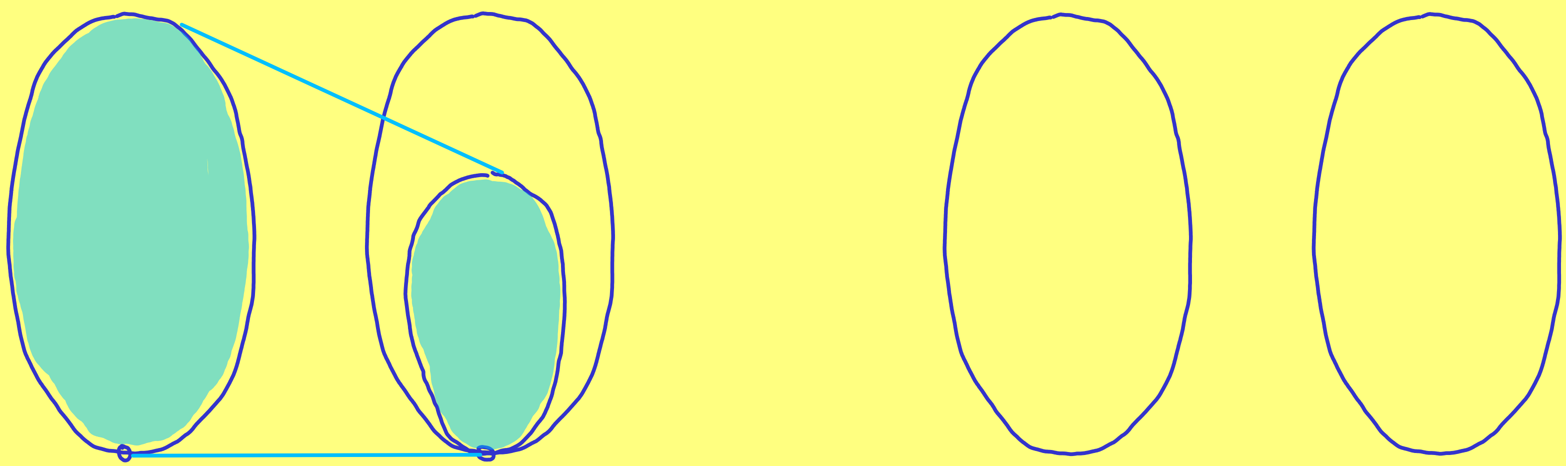
$$\ldots \subseteq K_{i-1} \subseteq K_i \subseteq \ldots \subseteq K_{j-1} \subseteq K_j \subseteq \ldots$$



$$\ldots \to H_p(K_{i-1}) \to H_p(K_i) \to \ldots \to H_p(K_{j-1}) \to H_p(K_j) \to \ldots$$

I.2 Filtrations

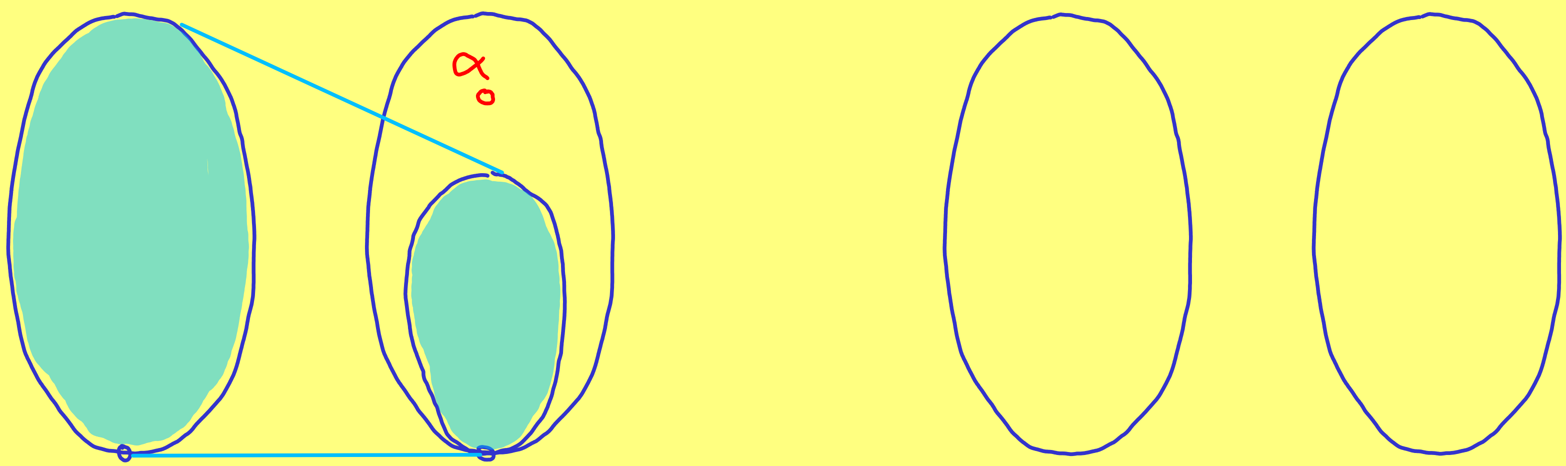
$$\dots \subseteq K_{i-1} \subseteq K_i \subseteq \dots \subseteq K_{j-1} \subseteq K_j \subseteq \dots$$



$$\dots \longrightarrow H_p(K_{i-1}) \longrightarrow H_p(K_i) \longrightarrow \dots \longrightarrow H_p(K_{j-1}) \longrightarrow H_p(K_j) \longrightarrow \dots$$

I.2 Filtrations

$$\ldots \subseteq K_{i-1} \subseteq K_i \subseteq \ldots \subseteq K_{j-1} \subseteq K_j \subseteq \ldots$$

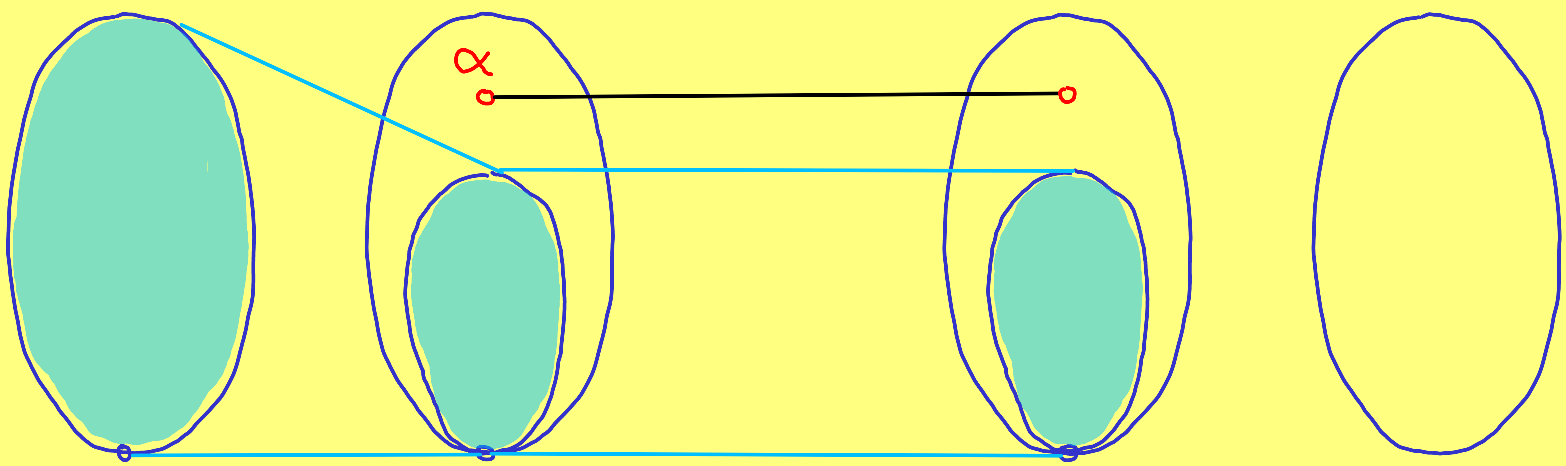


$$\ldots \longrightarrow H_p(K_{i-1}) \longrightarrow H_p(K_i) \longrightarrow \ldots \longrightarrow H_p(K_{j-1}) \longrightarrow H_p(K_j) \longrightarrow \ldots$$

α is born at K_i

I.2 Filtrations

$$\ldots \subseteq K_{i-1} \subseteq K_i \subseteq \ldots \subseteq K_{j-1} \subseteq K_j \subseteq \ldots$$

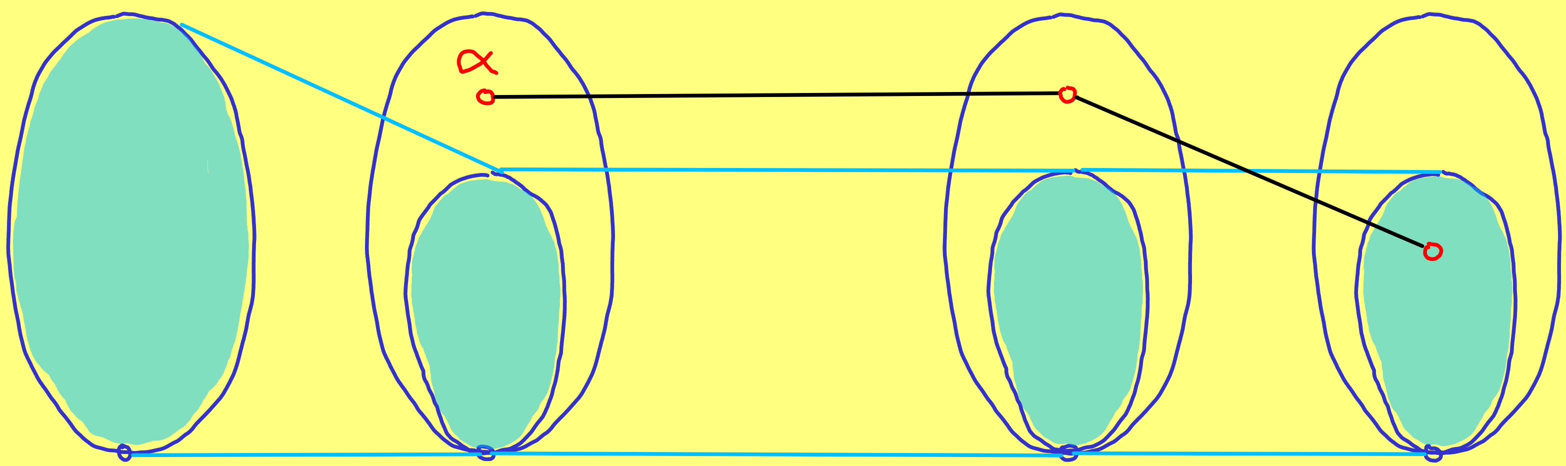


$$\ldots \rightarrow H_p(K_{i-1}) \rightarrow H_p(K_i) \rightarrow \ldots \rightarrow H_p(K_{j-1}) \rightarrow H_p(K_j) \rightarrow \ldots$$

α is born at K_i

I.2 FILTRATIONS

$$\dots \subseteq K_{i-1} \subseteq K_i \subseteq \dots \subseteq K_{j-1} \subseteq K_j \subseteq \dots$$

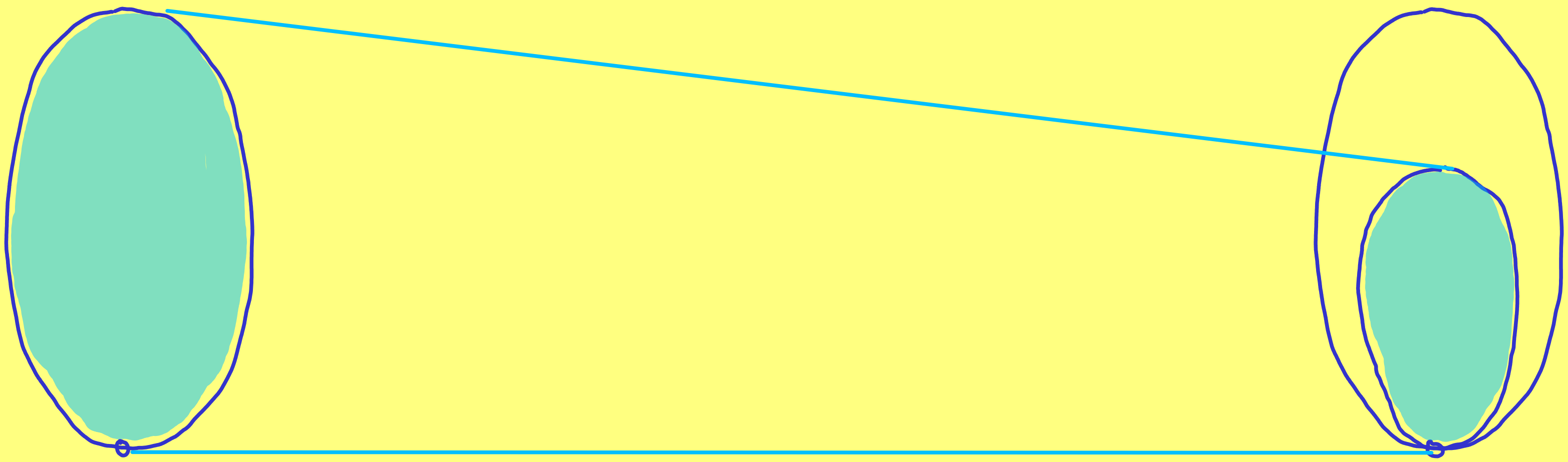


$$\dots \to H_p(K_{i-1}) \to H_p(K_i) \to \dots \to H_p(K_{j-1}) \to H_p(K_j) \to \dots$$

α is born at K_i and dies entering K_j

I.2 FILTRATIONS

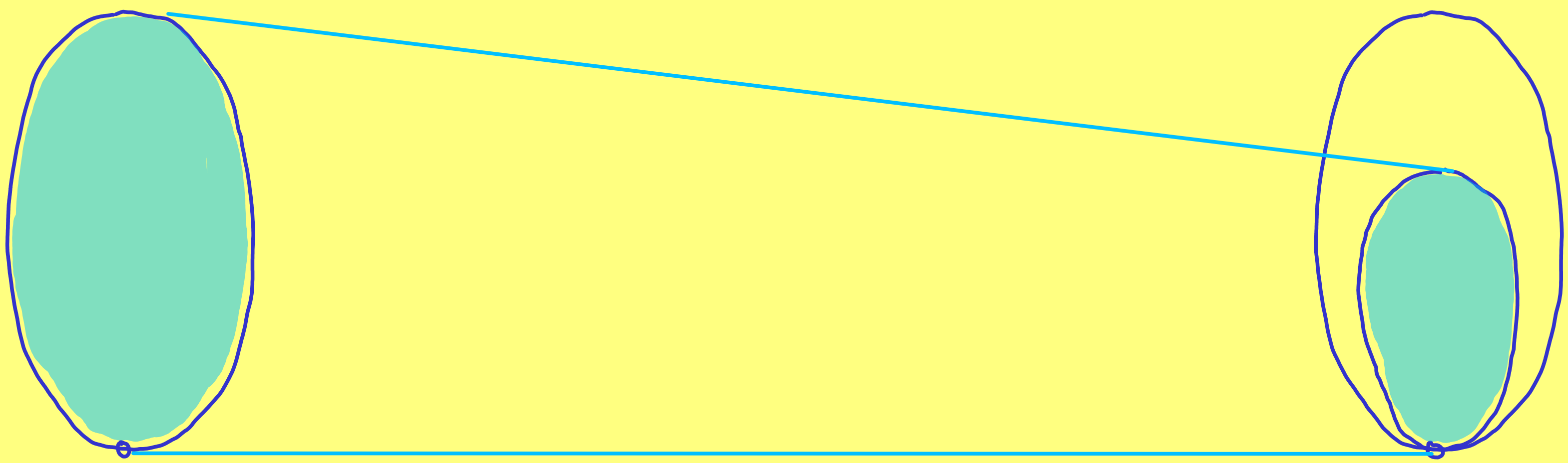
$$\ldots \subseteq K_{i-1} \subseteq K_i \subseteq \ldots \subseteq K_{j-1} \subseteq K_j \subseteq \ldots$$



$$H_p(K_{i-1}) \to \quad \ldots \quad \to H_p(K_j)$$

I.2 Filtrations

$$\ldots \subseteq K_{i-1} \subseteq K_i \subseteq \ldots \subseteq K_{j-1} \subseteq K_j \subseteq \ldots$$

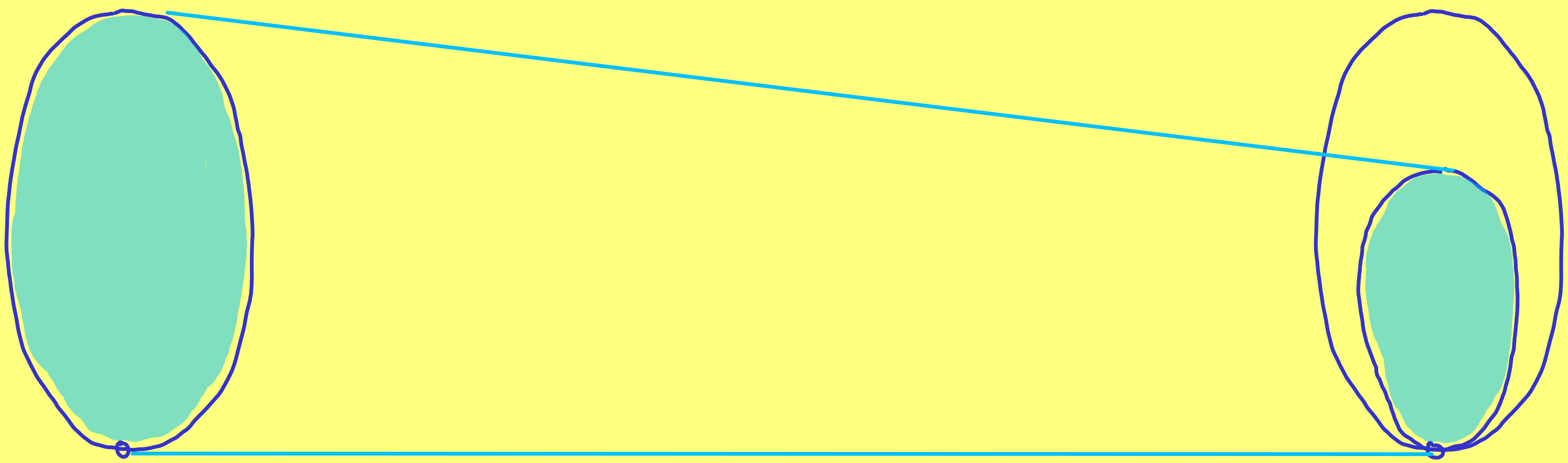


$$h_p^{i-1,j} : H_p(K_{i-1}) \rightarrow \quad \ldots \quad \rightarrow H_p(K_j)$$

induced homomorphism

I.2 Filtrations

$$\ldots \subseteq K_{i-1} \subseteq K_i \subseteq \ldots \subseteq K_{j-1} \subseteq K_j \subseteq \ldots$$



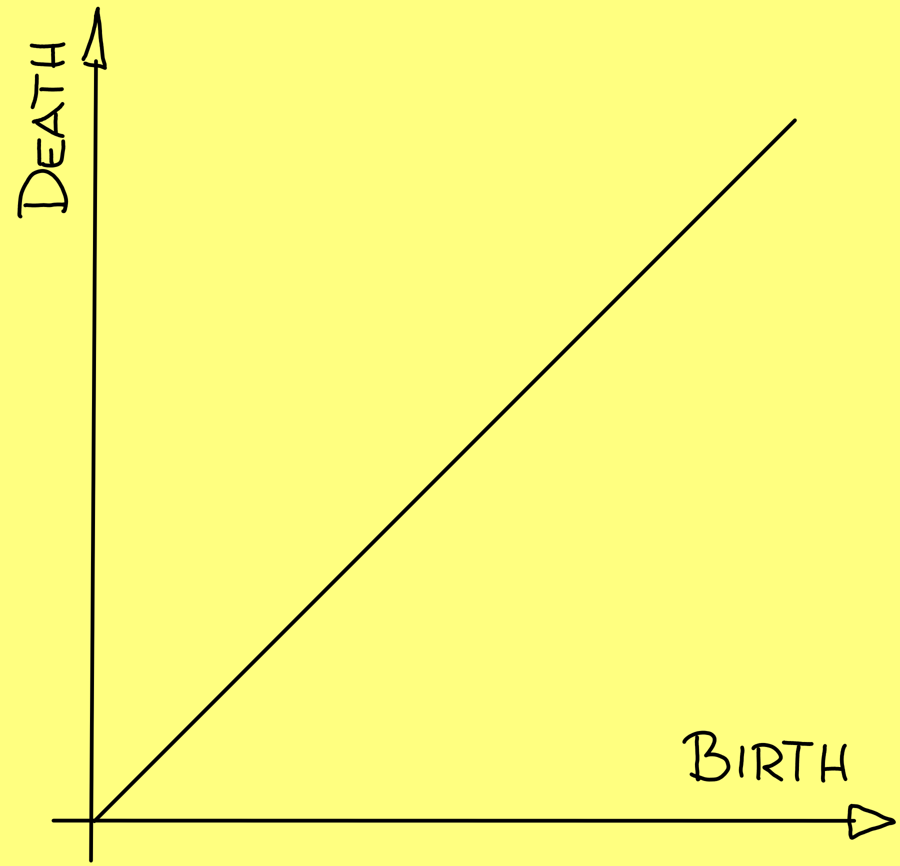
$$h_p^{i-1,j} : H_p(K_{i-1}) \to \quad \ldots \quad \to H_p(K_j)$$

induced homomorphism

$\operatorname{im} h_p^{i-1,j}$ is persistent homology group

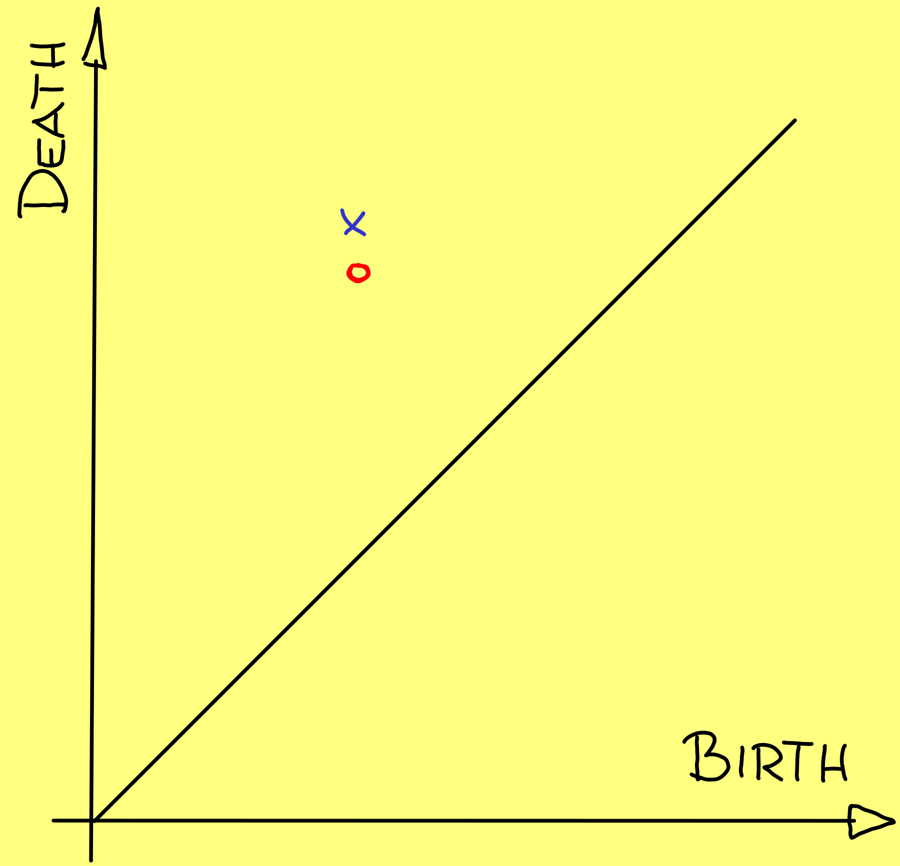
I.3 Diagrams

$\mathrm{Dgm}_p(f)$:



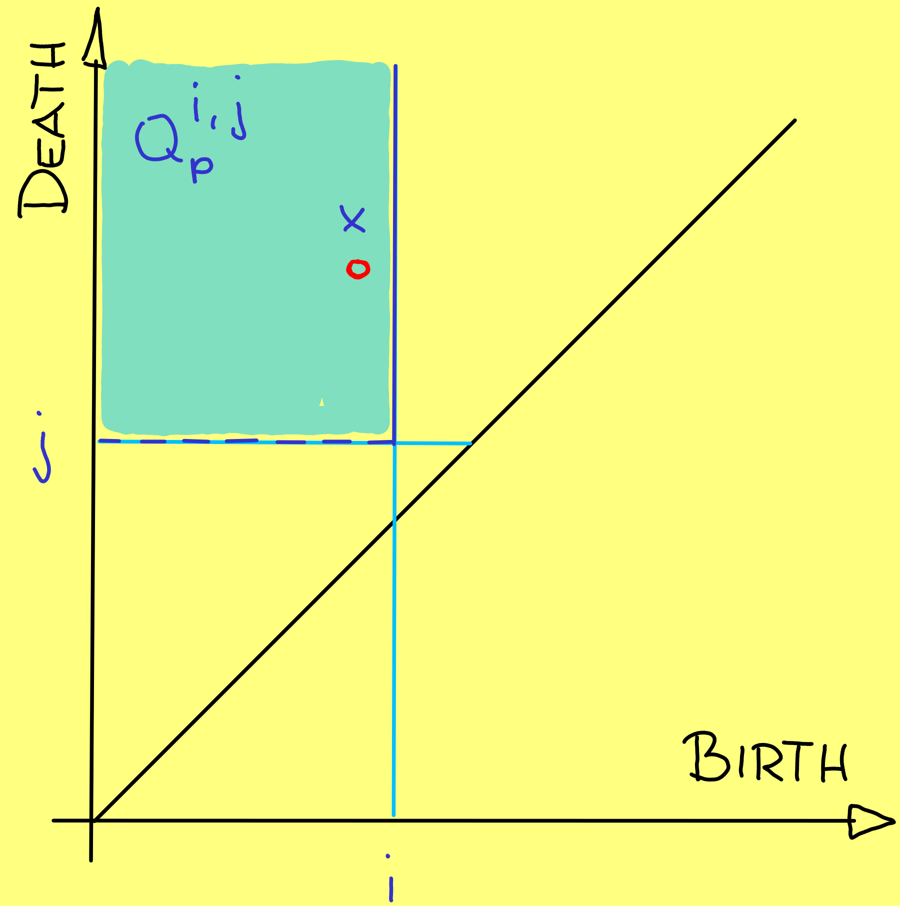
I.3 DIAGRAMS

$\mathrm{Dgm}_p(f)$:



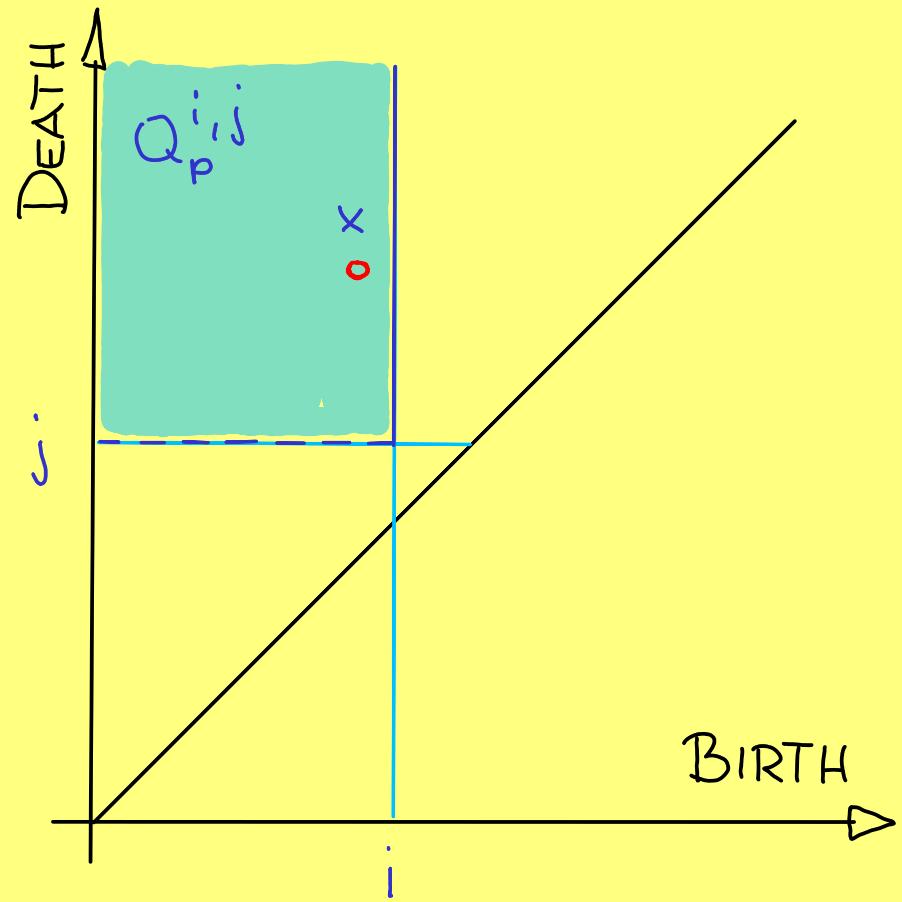
I.3 DIAGRAMS

$\mathrm{Dgm}_p(f)$:



I.3 DIAGRAMS

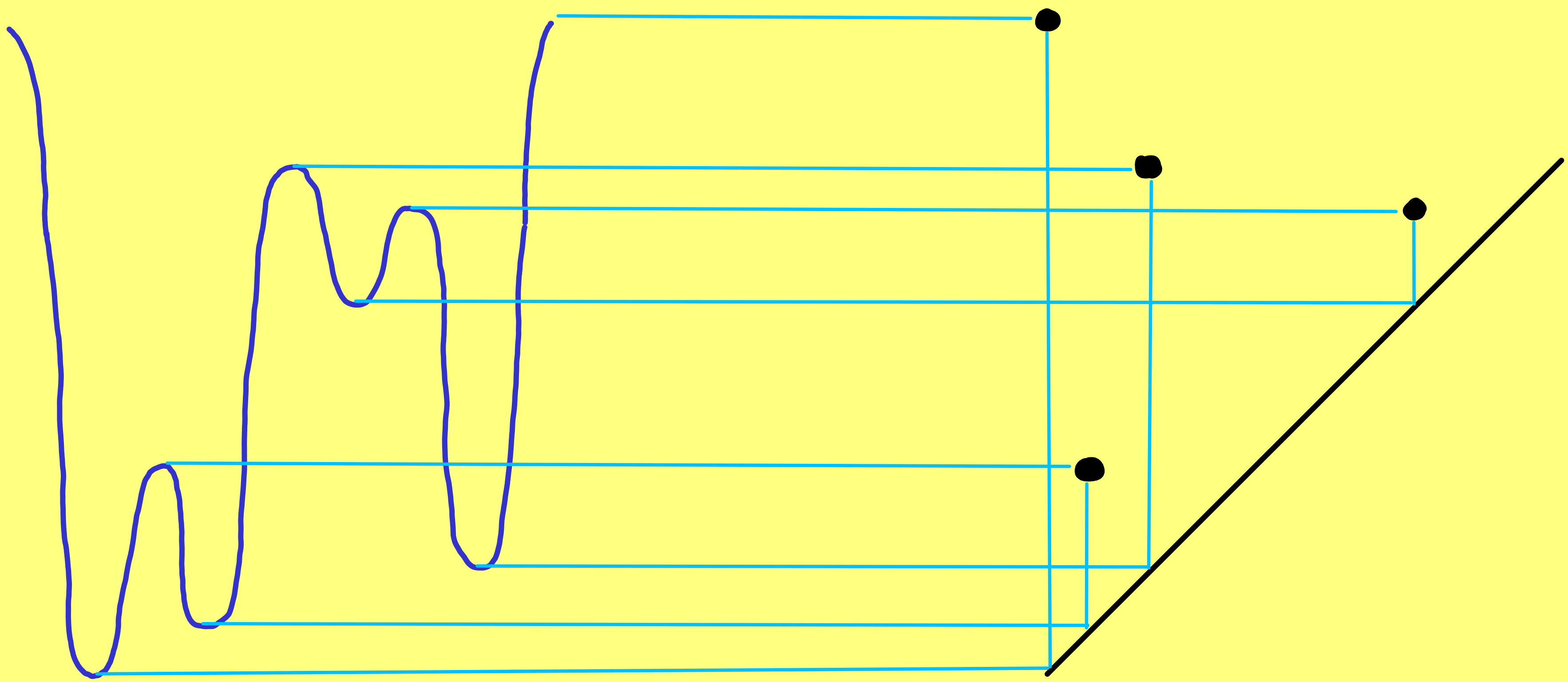
$\mathrm{Dgm}_p(f)$:



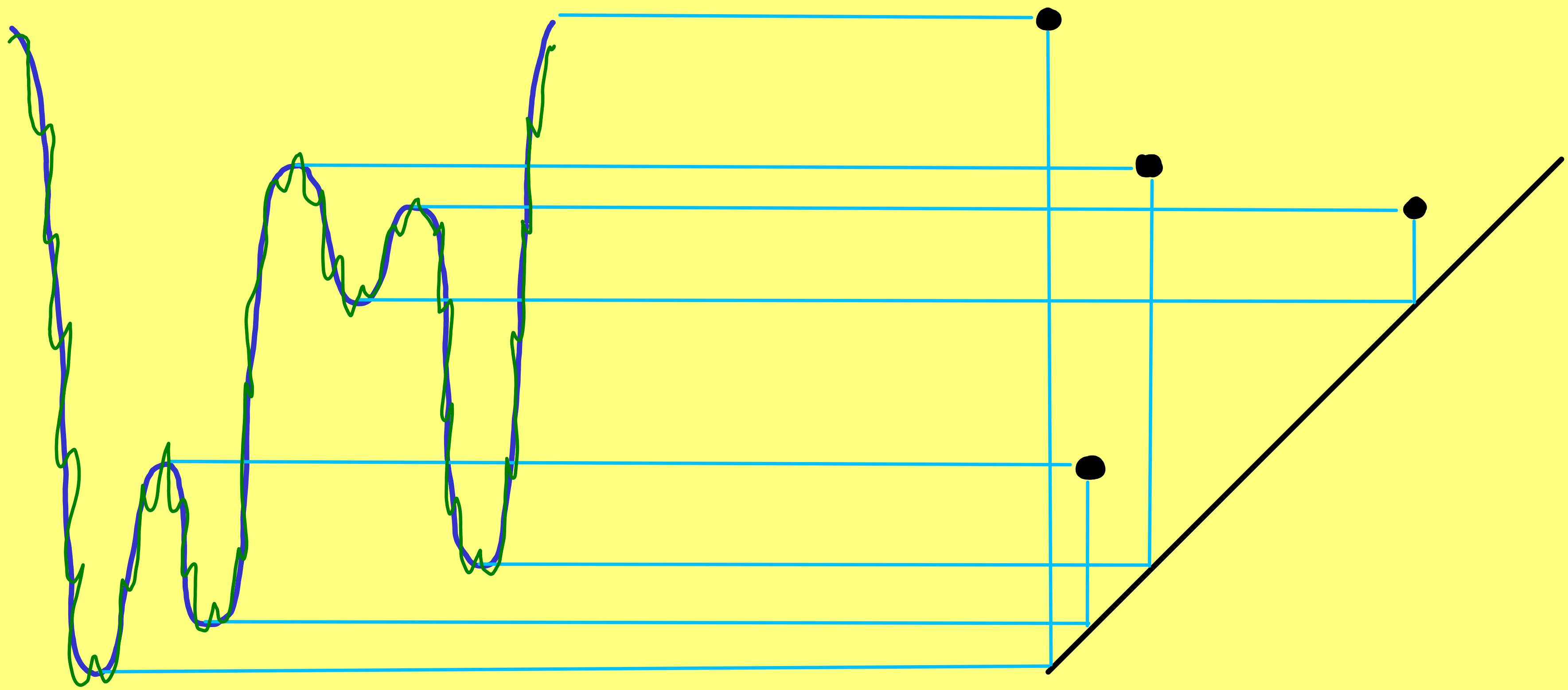
FUNDAMENTAL LEMMA OF PERSISTENT HOMOLOGY.

$\#\text{points in } Q_p^{i,j} = \operatorname{rank} \operatorname{im} h_p^{i,j}$.

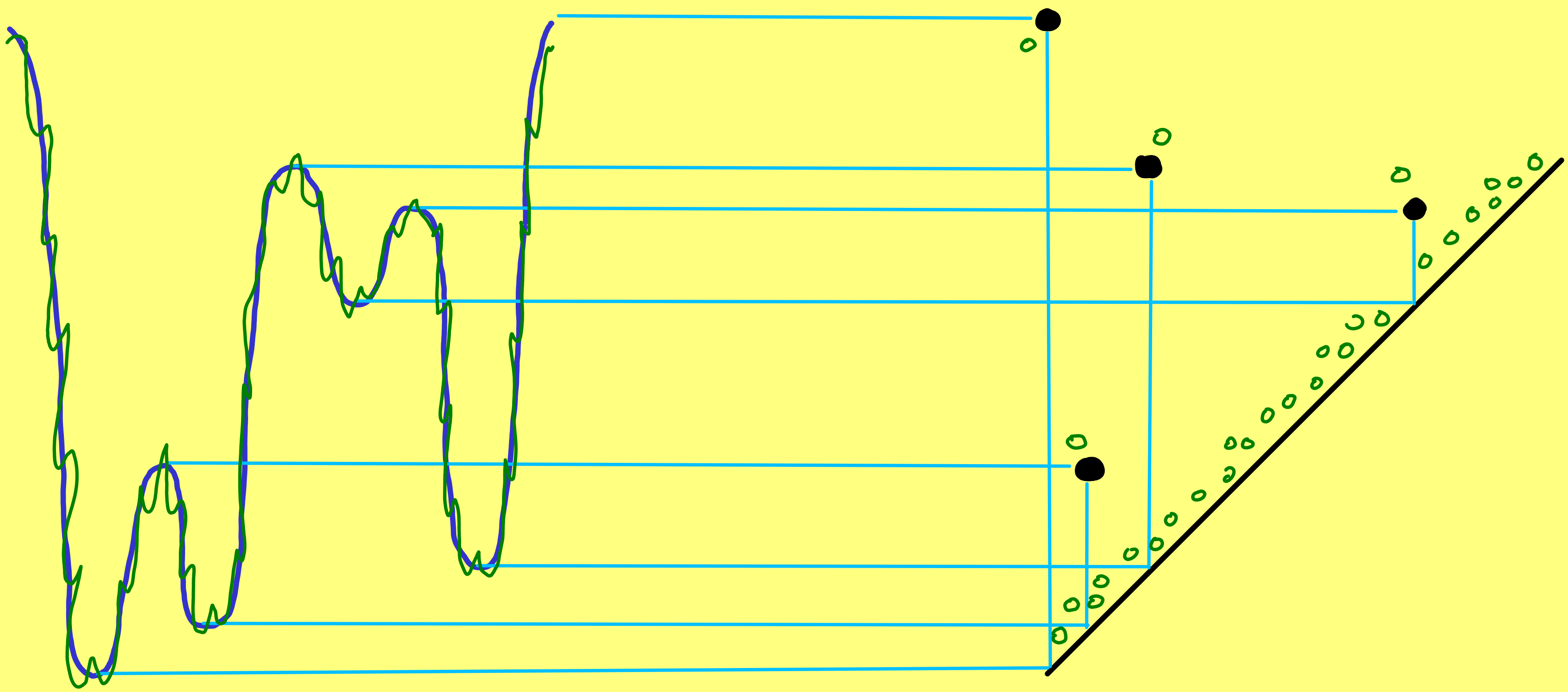
I.4 Bottleneck Distance



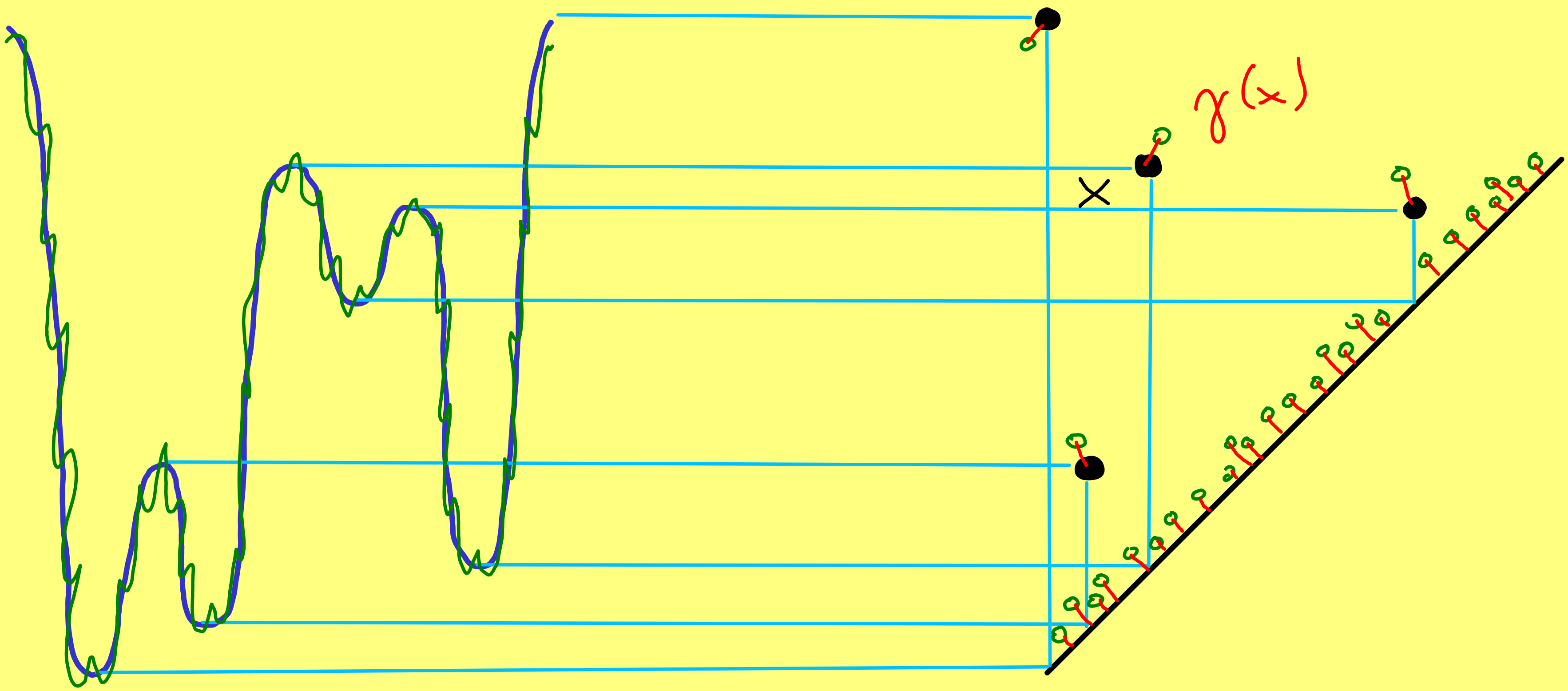
I.4 Bottleneck Distance



I.4 Bottleneck Distance



I.4 Bottleneck Distance



The bottleneck distance between the diagrams is

$$W_\infty(\mathrm{Dgm}_p(f), \mathrm{Dgm}_p(g)) = \inf_\gamma \sup_x \|x - \gamma(x)\|_\infty .$$

I.4 BOTTLENECK DISTANCE

L_∞-STAB. THM. For tame functions $f, g : \mathbb{X} \to \mathbb{R}$ the bottleneck distance between their diagrams is bounded by the max-difference between the functions,

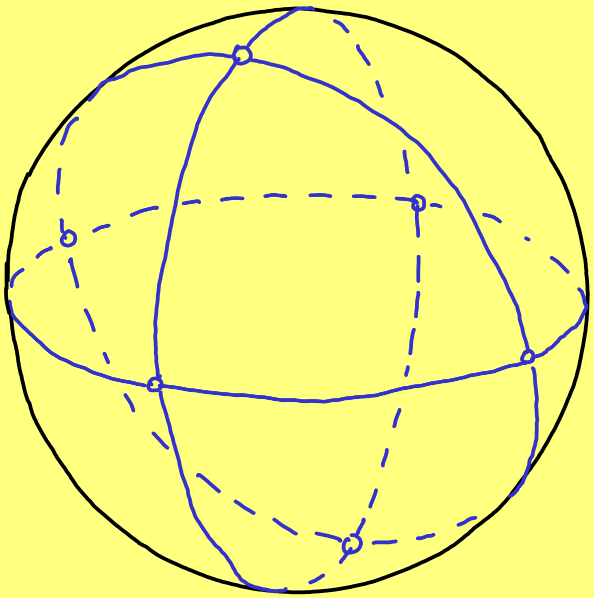
$$W_\infty(\mathrm{Dgm}_p(f), \mathrm{Dgm}_p(g)) \leq \|f - g\|_\infty .$$

I Persistent Homology

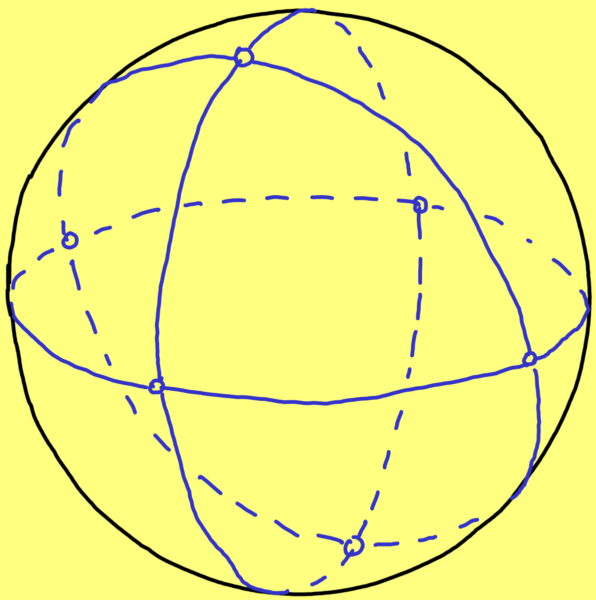
II L_p-stability and Somites

III Contour Stability

II.1 Size and Growth



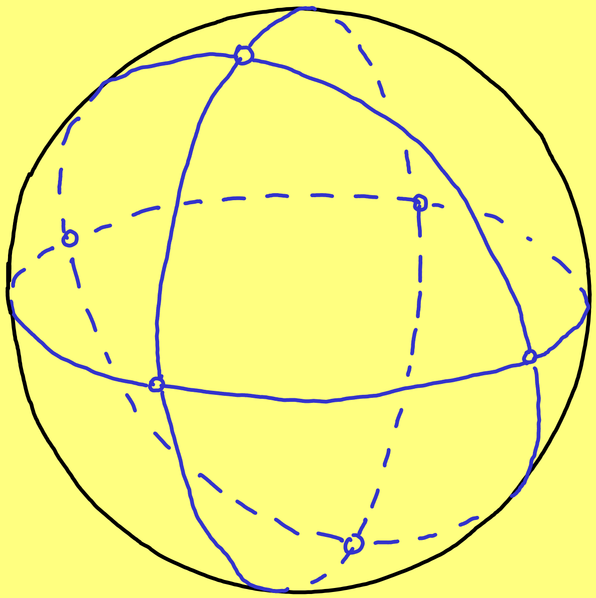
II.1 Size and Growth



mesh (K) = max. diameter of a simplex

size (K) = number of simplices

II.1 Size and Growth

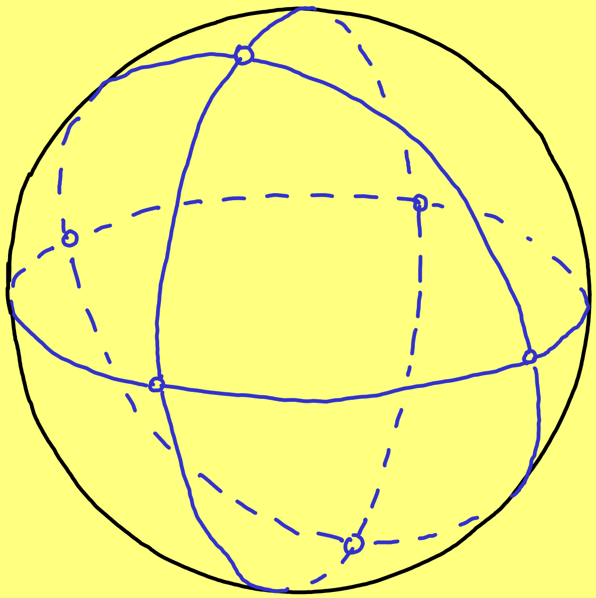


mesh (K) = max. diameter of a simplex

size (K) = number of simplices

$$N(r) = \min_{\text{mesh}(K) \le r} \text{size}(K)$$

II.1 Size and Growth



$\text{mesh}(K)$ = max. diameter of a simplex

$\text{size}(K)$ = number of simplices

$$N(r) = \min_{\text{mesh}(K) \leq r} \text{size}(K)$$

The triangulation grows polynomially if $\exists$ constants C, M s.t.

$$N(r) \leq C/r^M.$$

II.2 Total Persistence

The q-th total persistence of f is

$$\mathrm{Pers}_q(f) = \sum_p \sum_x \mathrm{pers}(x)^q$$

II.2 Total Persistence

The q-th total persistence of f is

$$\mathrm{Pers}_q(f) = \sum_p \sum_x \mathrm{pers}(x)^q$$

f Lipschitz and $N(r) \leq C / r^M$

$\Longrightarrow$ $\mathrm{Pers}_q(f) \leq \mathrm{const}$ for all $q > M$

II.2 Total Persistence

TP-STAB. THM. Let $\mathbb{X}$ be a compact metric space with polynomially growing triangulation and $f, g : \mathbb{X} \to \mathbb{R}$ Lipschitz functions. Then

$$\mathrm{Pers}_q(f) - \mathrm{Pers}_q(g) \leq \mathrm{const} \cdot \|f - g\|_\infty$$

for every $q > M + 1$.

II.2 SIMPLIFICATION

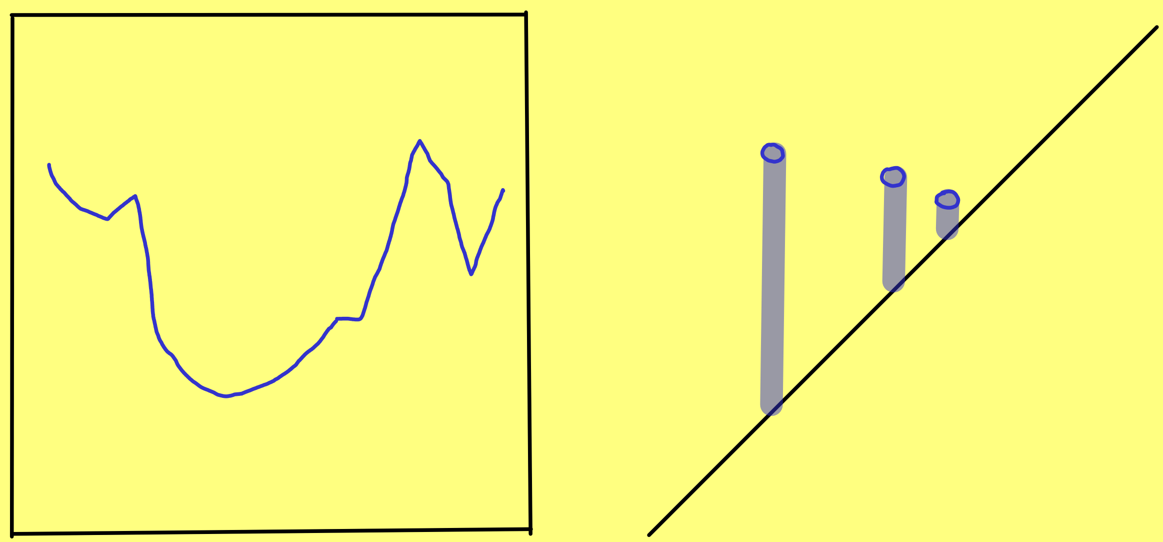
normalized to $\operatorname{amp}(f) = \max_x f(x) - \min_x f(x) = 1$

II.2 SIMPLIFICATION

An ε-simplification of f is a function $f_\varepsilon : \mathbb{X} \to \mathbb{R}$ with $\| f - f_\varepsilon \|_\infty \leq \varepsilon$ whose diagrams $\mathrm{Dgm}_p(f_\varepsilon)$ are same as $\mathrm{Dgm}_p(f)$ without points of persistence at most ε.

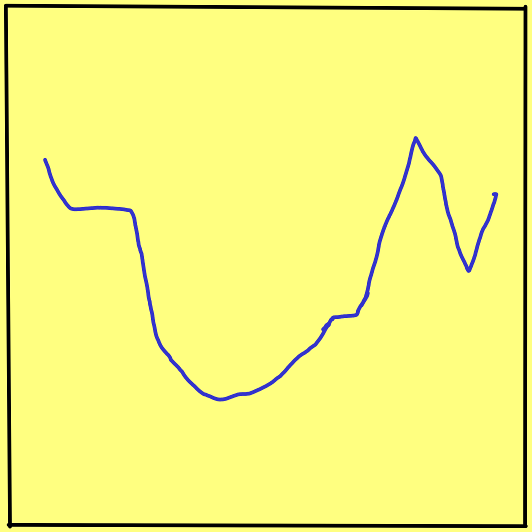
II.2 SIMPLIFICATION

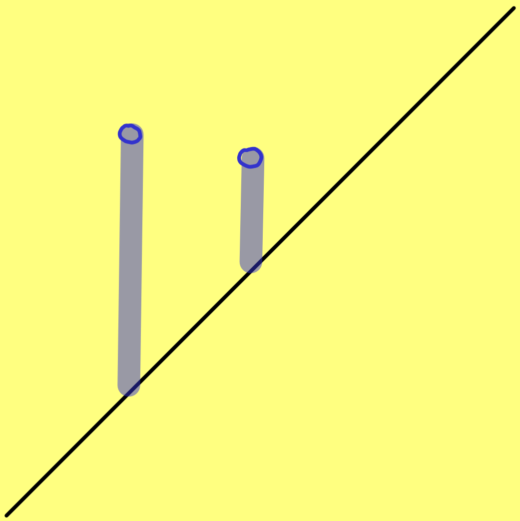
An ε-simplification of f is a function $f_\varepsilon : \mathbb{X} \to \mathbb{R}$ with $\|f - f_\varepsilon\|_\infty \leq \varepsilon$ whose diagrams $\mathrm{Dgm}_p(f_\varepsilon)$ are same as $\mathrm{Dgm}_p(f)$ without points of persistence at most ε.



II.2 SIMPLIFICATION

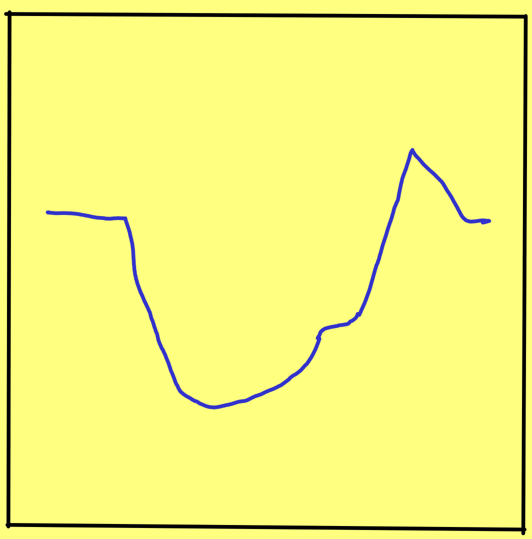
An ε-simplification of f is a function $f_\varepsilon : \mathbb{X} \to \mathbb{R}$ with $\|f - f_\varepsilon\|_\infty \leq \varepsilon$ whose diagrams $\mathrm{Dgm}_p(f_\varepsilon)$ are same as $\mathrm{Dgm}_p(f)$ without points of persistence at most ε.

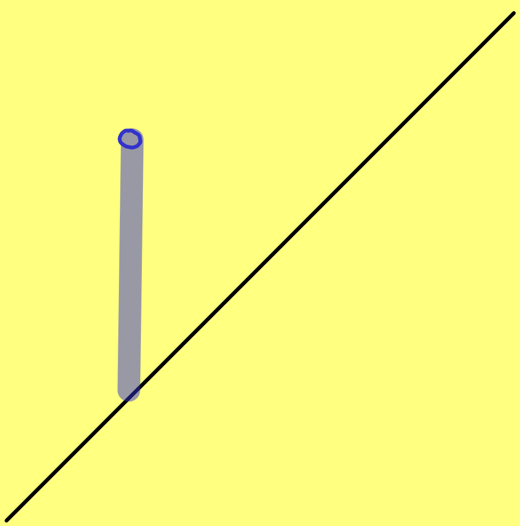




II.2 SIMPLIFICATION

An ε-simplification of f is a function $f_\varepsilon : \mathbb{X} \to \mathbb{R}$ with $\|f - f_\varepsilon\|_\infty \leq \varepsilon$ whose diagrams $\mathrm{Dgm}_p(f_\varepsilon)$ are same as $\mathrm{Dgm}_p(f)$ without points of persistence at most ε.





II.2 Simplification

$f : S^1 \to \mathbb{R}$

$M_0(f) = \frac{1}{2} \#\text{crit. pts.}$

II.2 SIMPLIFICATION

$f : S^1 \to \mathbb{R}$

$$M_0(f) = \frac{1}{2} \#\text{crit. pts.}$$

$$M_1(f) = \int_{\varepsilon=0}^{1} M_0(f_\varepsilon)\, d\varepsilon$$

II.2 SIMPLIFICATION

$f : \mathbb{S}^1 \to \mathbb{R}$

$$M_0(f) = \frac{1}{2} \#\text{crit. pts.}$$

$$M_1(f) = \int_{\varepsilon=0}^{1} M_0(f_\varepsilon)\, d\varepsilon$$

$$M_2(f) = \int_{\varepsilon=0}^{1} M_1(f_\varepsilon)\, d\varepsilon$$

etc.

II.2 SIMPLIFICATION

$f : \mathbb{S}^1 \to \mathbb{R}$

$$M_0(f) = \frac{1}{2} \,\#\text{crit. pts.} = \mathrm{Pers}_0(f)$$

$$M_1(f) = \int_{\varepsilon=0}^{1} M_0(f_\varepsilon)\, d\varepsilon = \mathrm{Pers}_1(f)$$

$$M_2(f) = \int_{\varepsilon=0}^{1} M_1(f_\varepsilon)\, d\varepsilon = \mathrm{Pers}_2(f)$$

etc.

II.2 Simplification

$$f : \mathbb{S}^1 \to \mathbb{R}$$

not stable

$$M_0(f) = \frac{1}{2}\,\#\text{crit. pts.} = \mathrm{Pers}_0(f)$$

$$M_1(f) = \int_{\varepsilon=0}^{1} M_0(f_\varepsilon)\,d\varepsilon = \mathrm{Pers}_1(f)$$

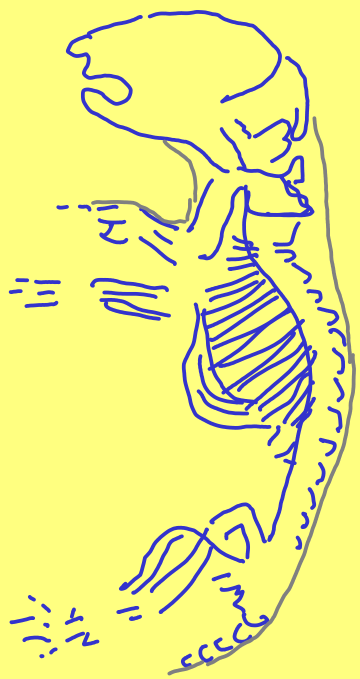
stable

$$M_2(f) = \int_{\varepsilon=0}^{1} M_1(f_\varepsilon)\,d\varepsilon = \mathrm{Pers}_2(f)$$

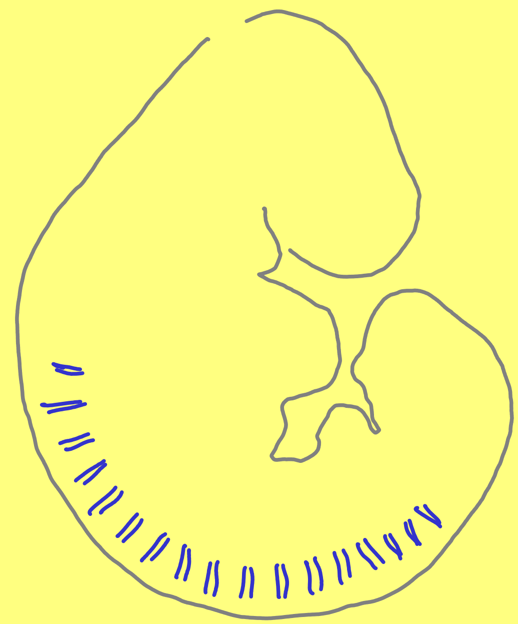
etc.

II.3 Rhythmic Gene Expression

adult mouse

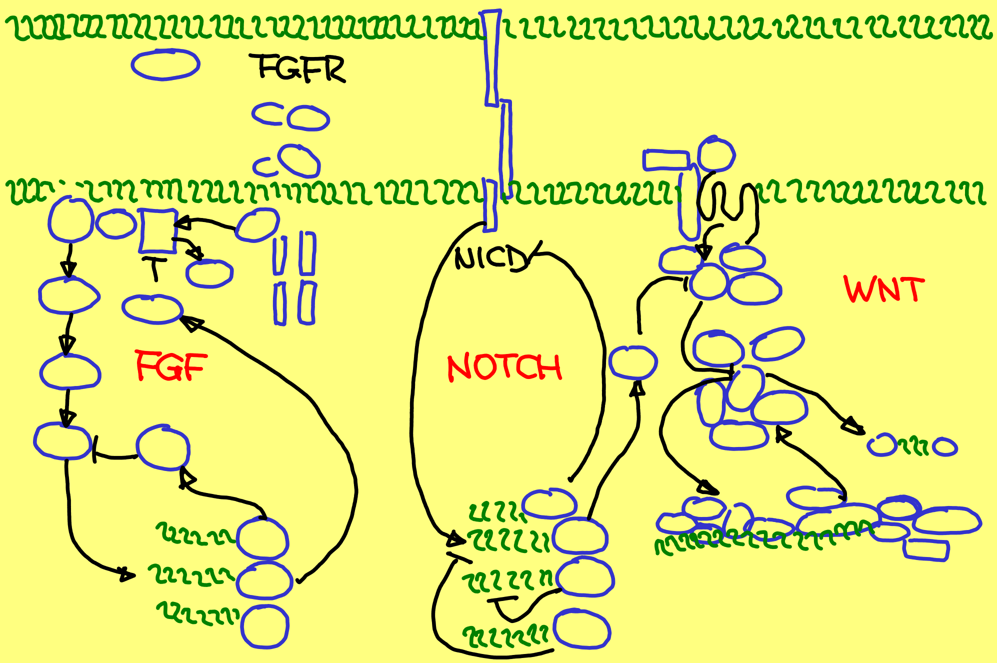


mouse embryo



somite development is a rythmic process

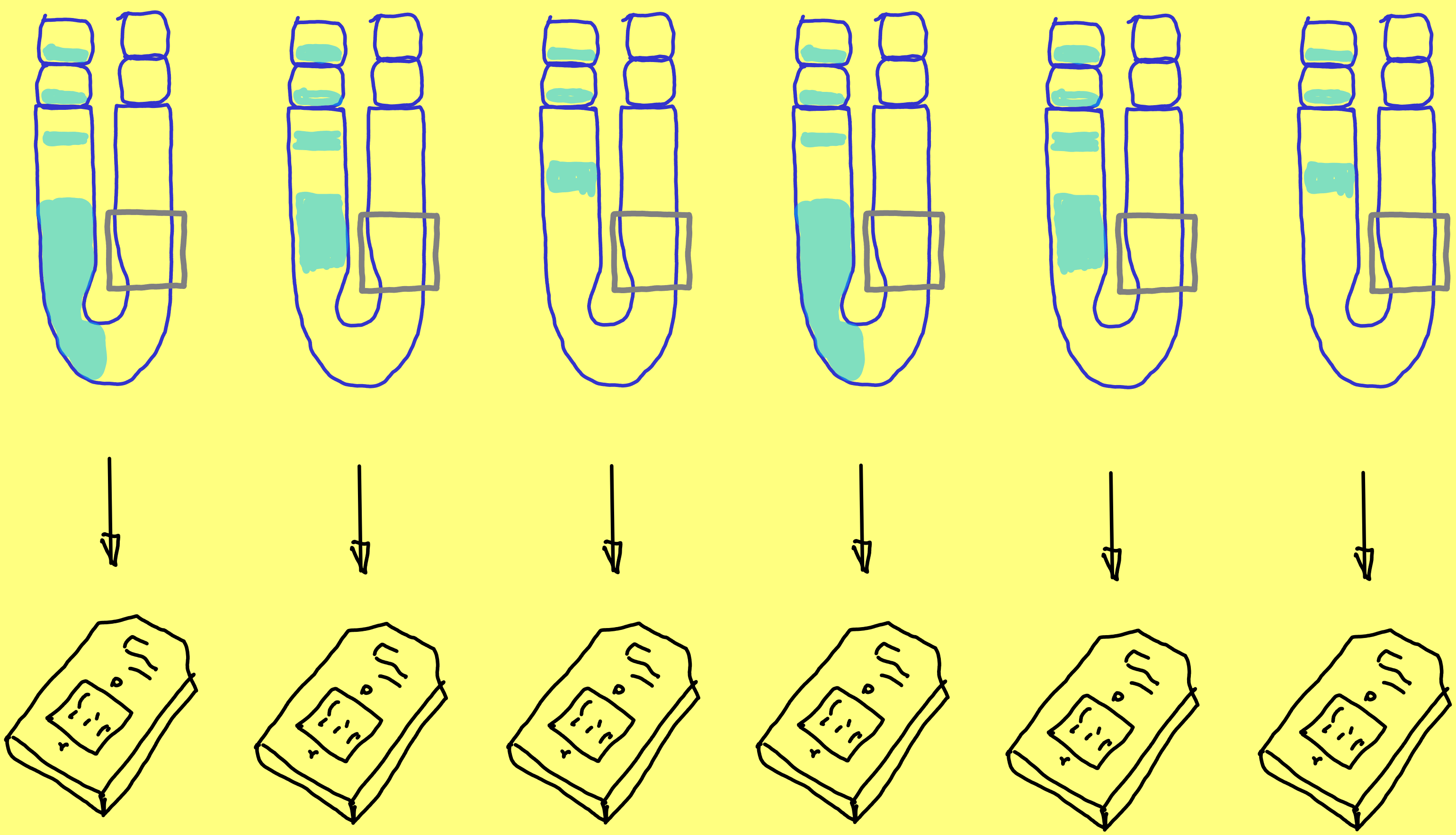
II.3 Rhythmic Gene Expression



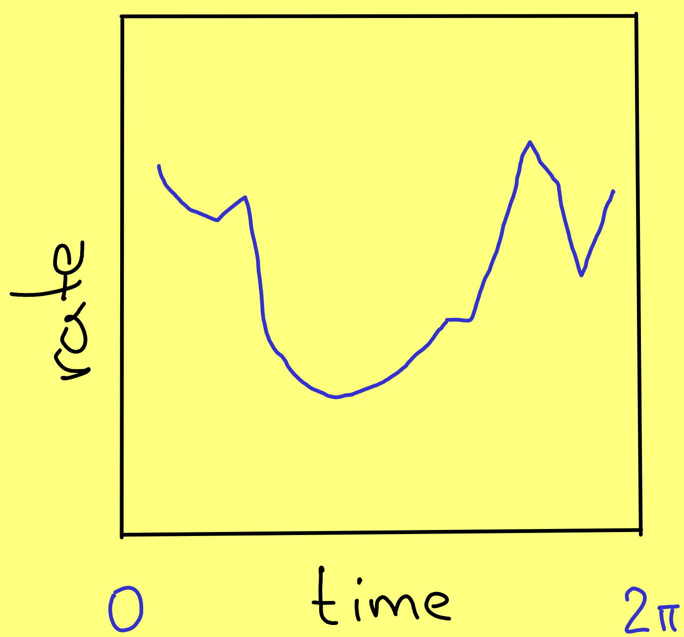
the proposed clock

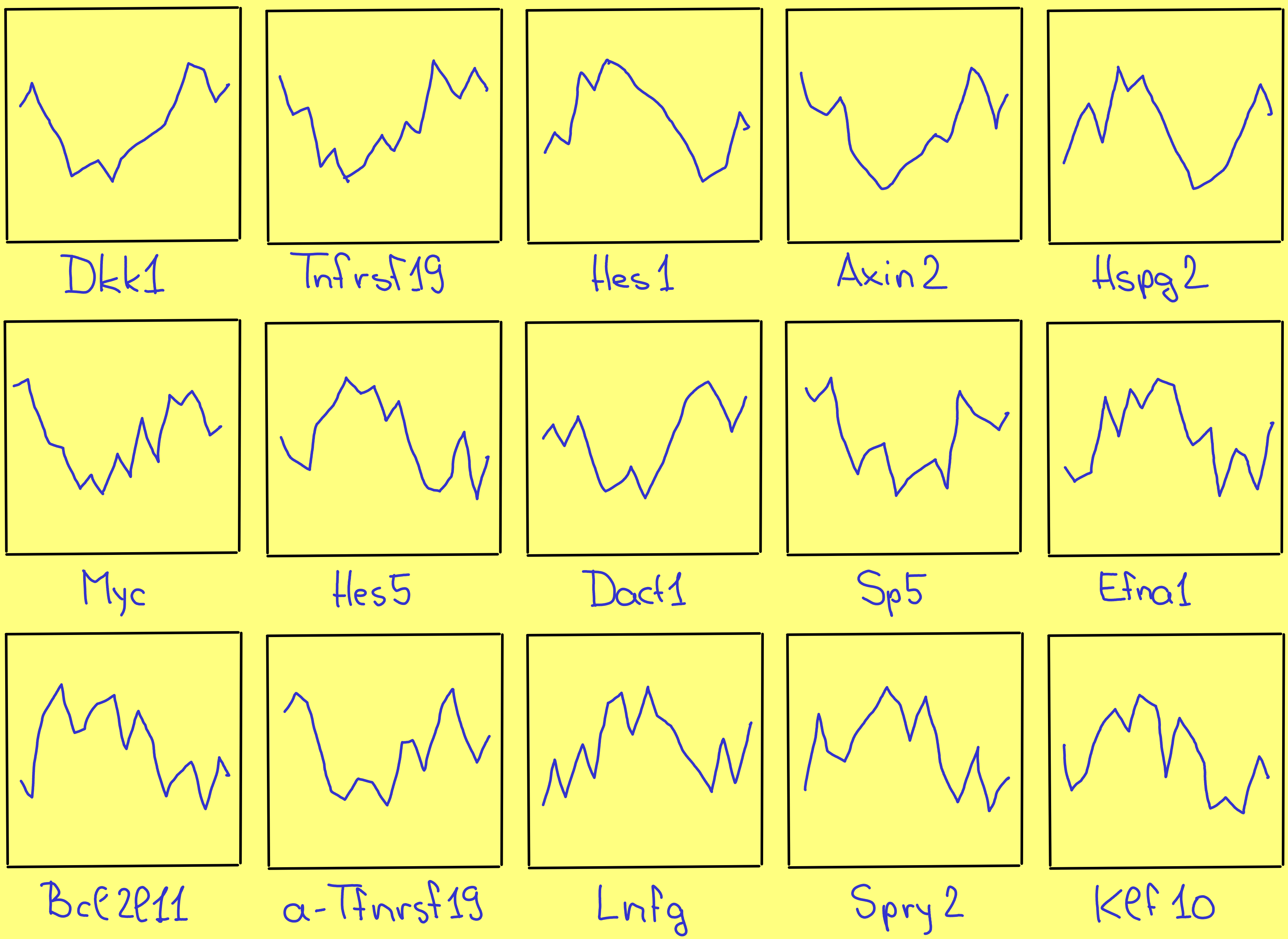
II.3 Rhythmic Gene Expression

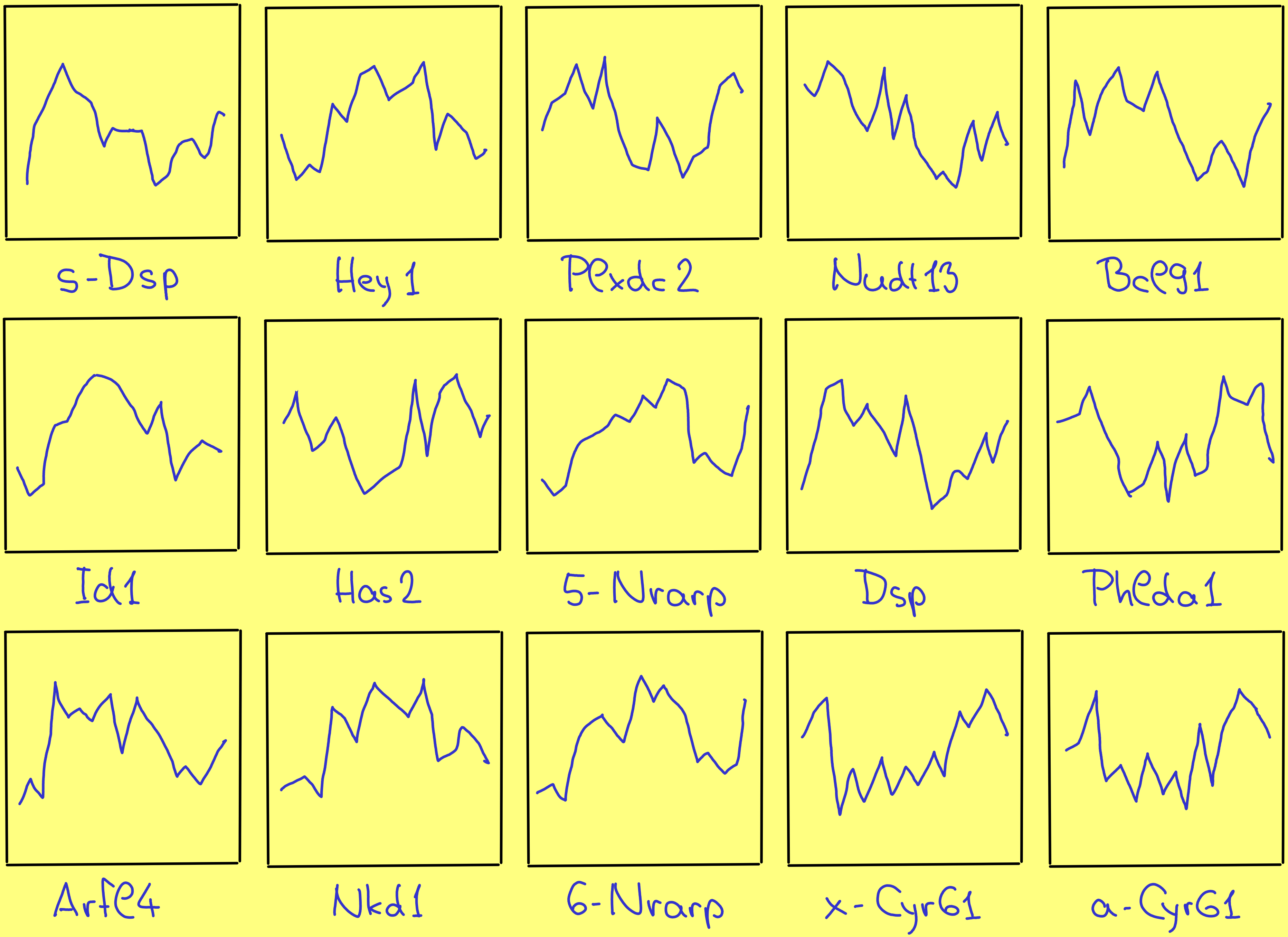
... time series microarray analysis



II.3 Rhythmic Gene Expression



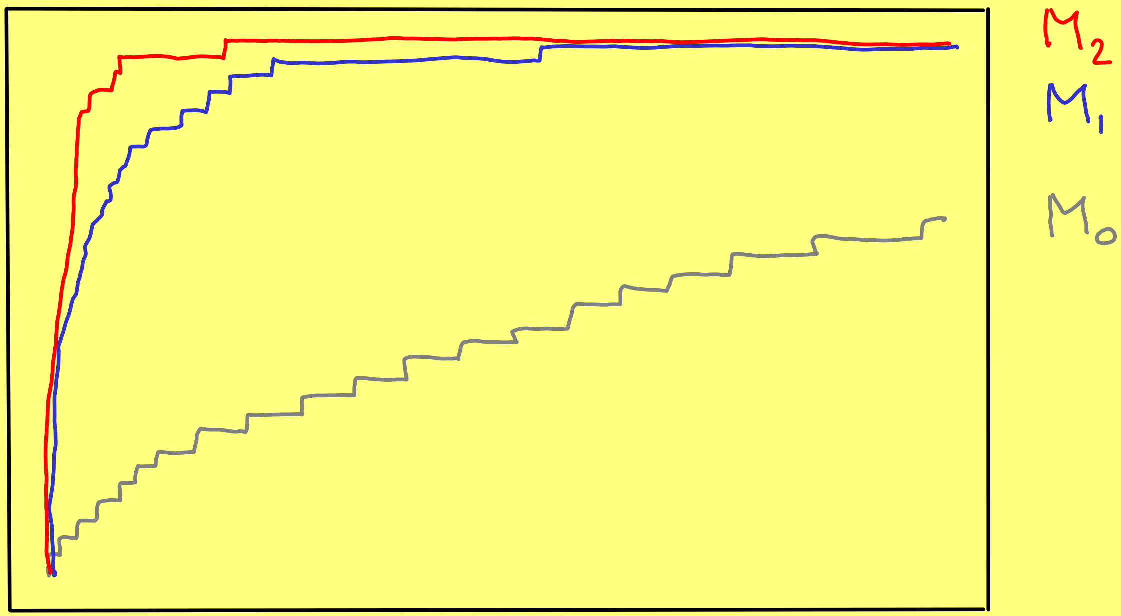

Dkk1
Tnfrsf19
Hes1
Axin2
Hspg2
Myc
Hes5
Dact1
Sp5
Efna1
Bcl2l11
a-Tfnrsf19
Lnfg
Spry2
Klf10


s-Dsp
Hey1
Plxdc2
Nudt13
Bcl91
Id1
Has2
5-Nrarp
Dsp
Phlda1
Arfl4
Nkd1
6-Nrarp
x-Cyr61
a-Cyr61

II.3 Rhythmic Gene Expression

yield of 30 biologically confirmed genes

ROC Curves



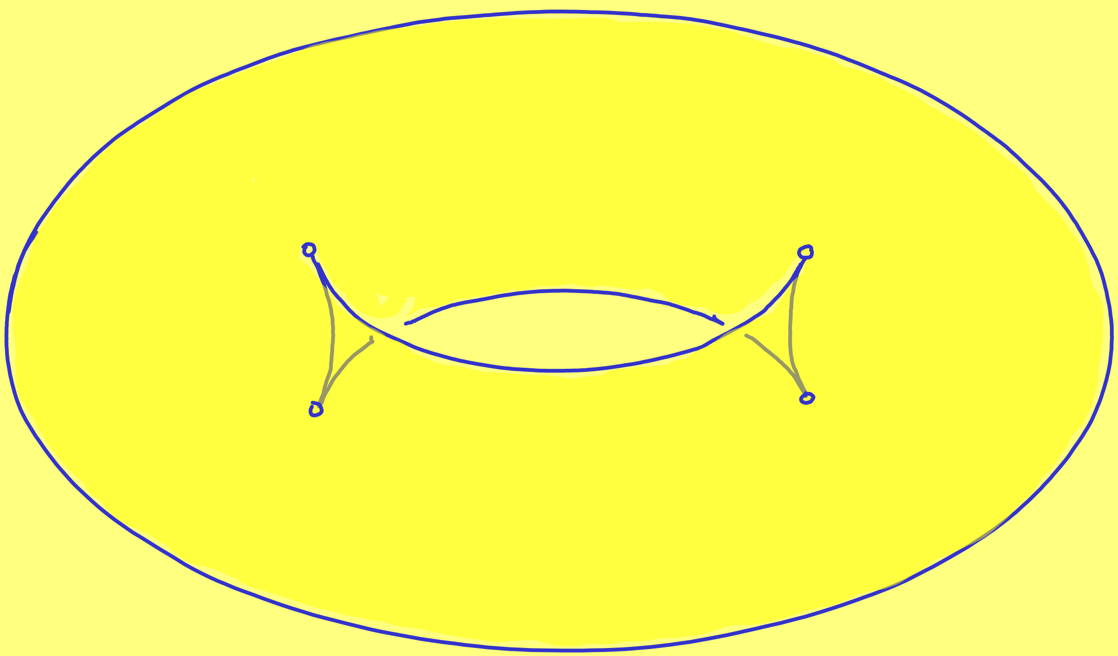
	M_0	M_1	M_2	M_3	M_4	max
Area	4.42	9.75	10.18	10.19	10.15	10.50

I Persistent Homology

II L_p-stability and Somites

III Contour Stability

III.1 The Contour

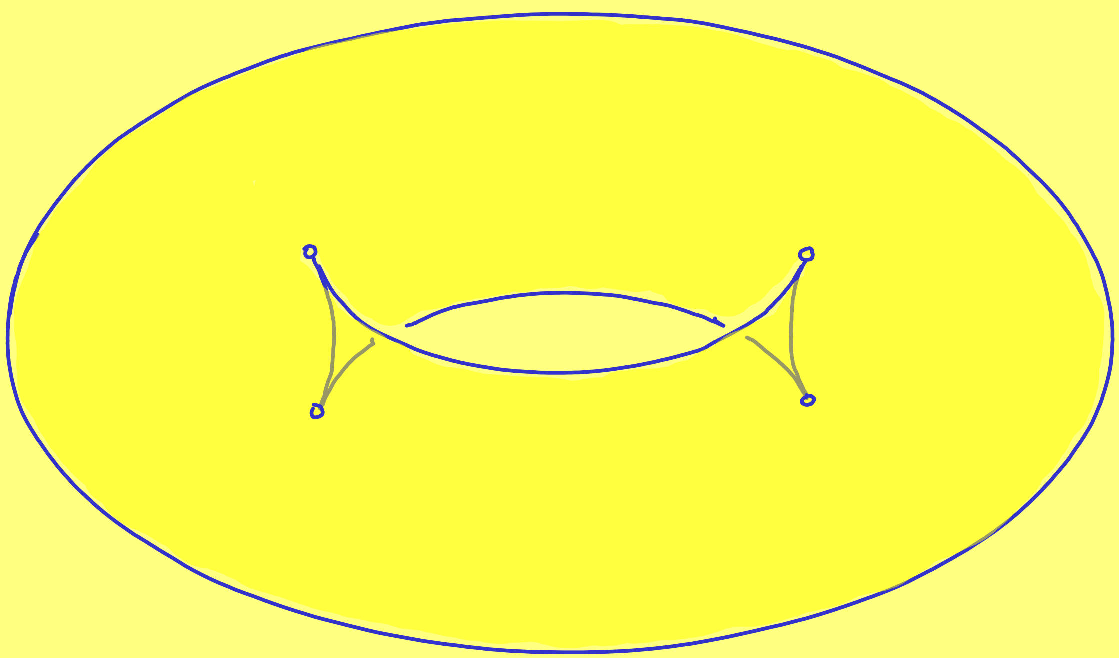


III.1 The Contour

$$f : \mathbb{M} \to \mathbb{R}^2$$

generic smooth

compact, orientable 2-mfld w/o boundary



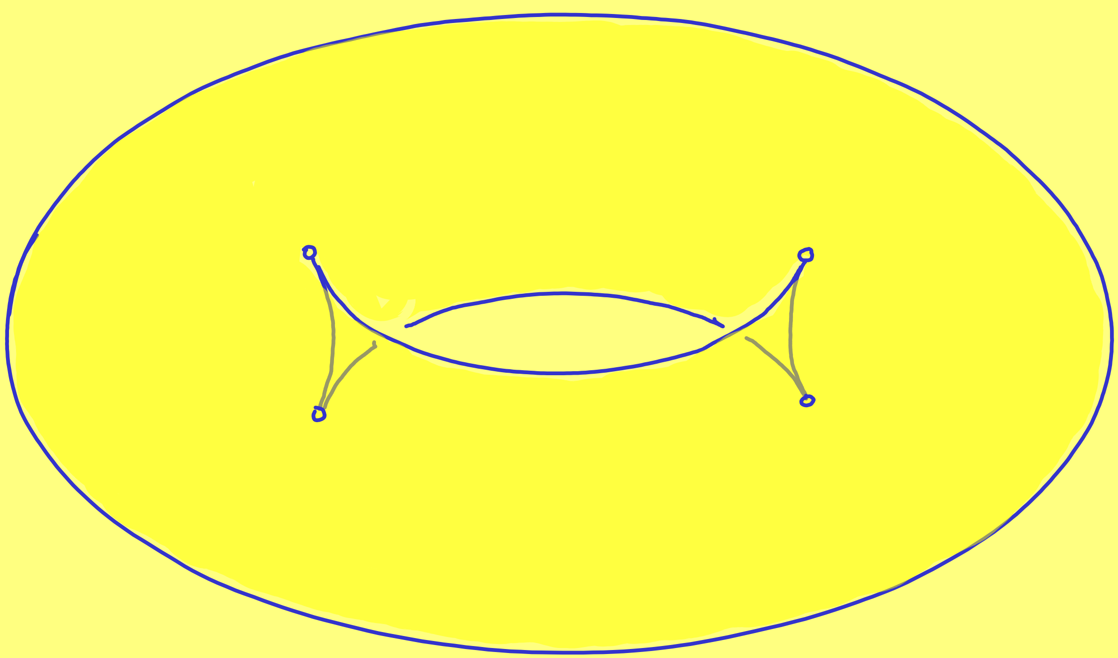
III.1 THE CONTOUR

$f : \mathbb{M} \to \mathbb{R}^2$

generic smooth

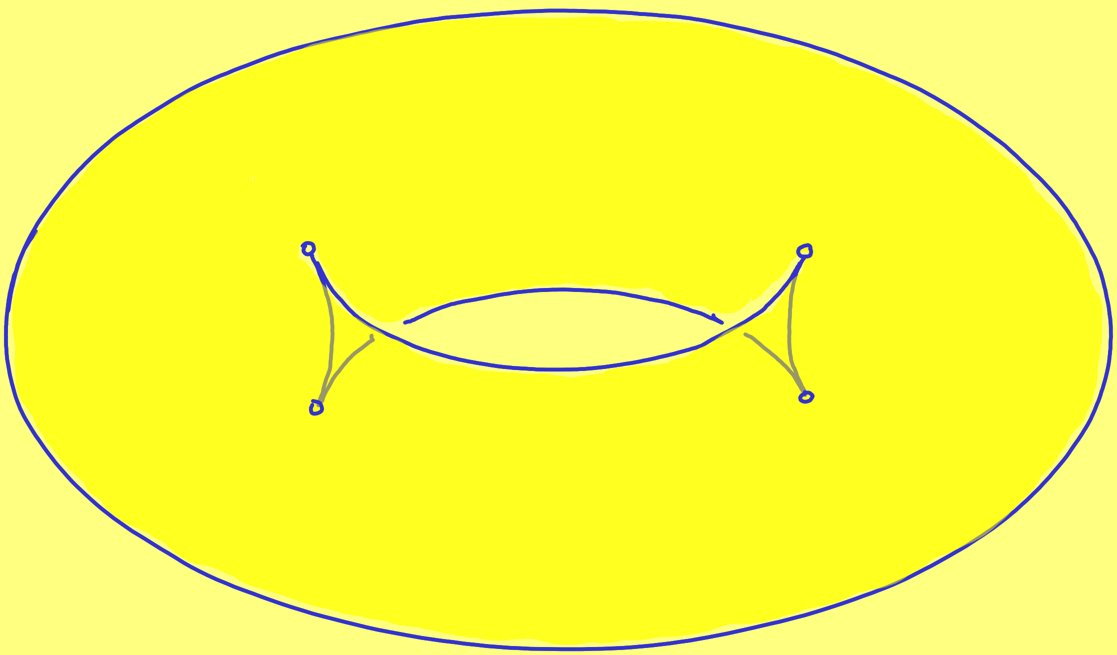
compact, orientable 2-mfld w/o boundary

Contour(f) =
set of critical values



III.2 A Perturbation

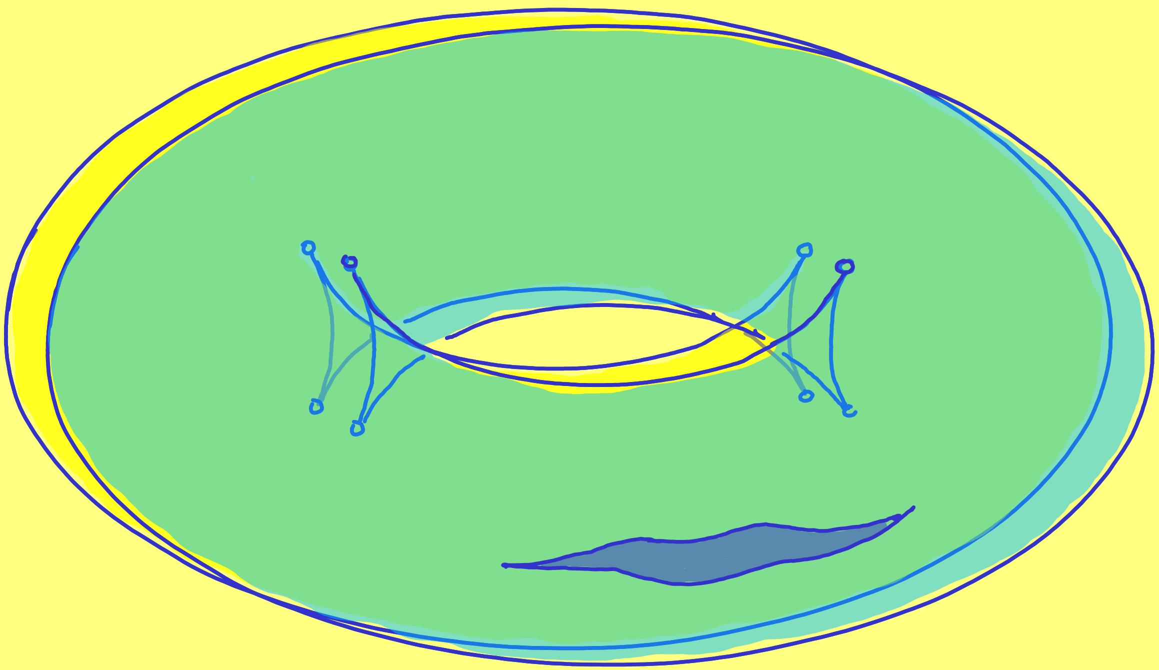
$f : \mathbb{M} \to \mathbb{R}^2$



III.2 A Perturbation

$$f, g : \mathbb{M} \to \mathbb{R}^2$$

$$\varepsilon = \max_{x \in \mathbb{M}} \| f(x) - g(x) \|_2$$

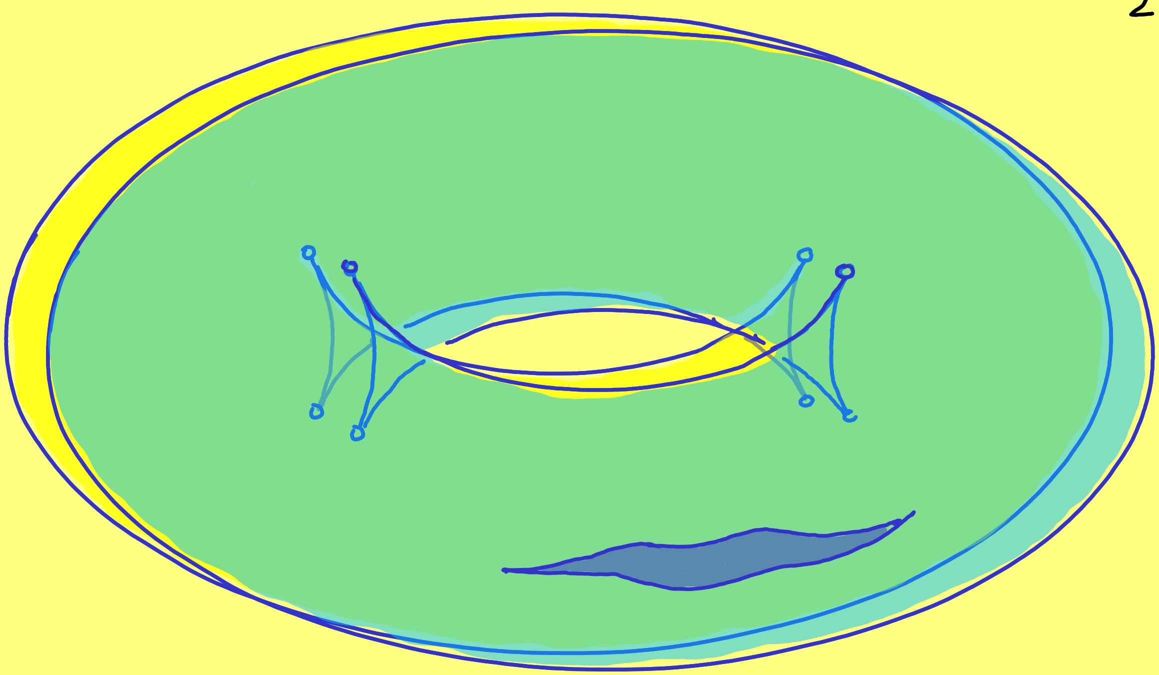


III.2 A Perturbation

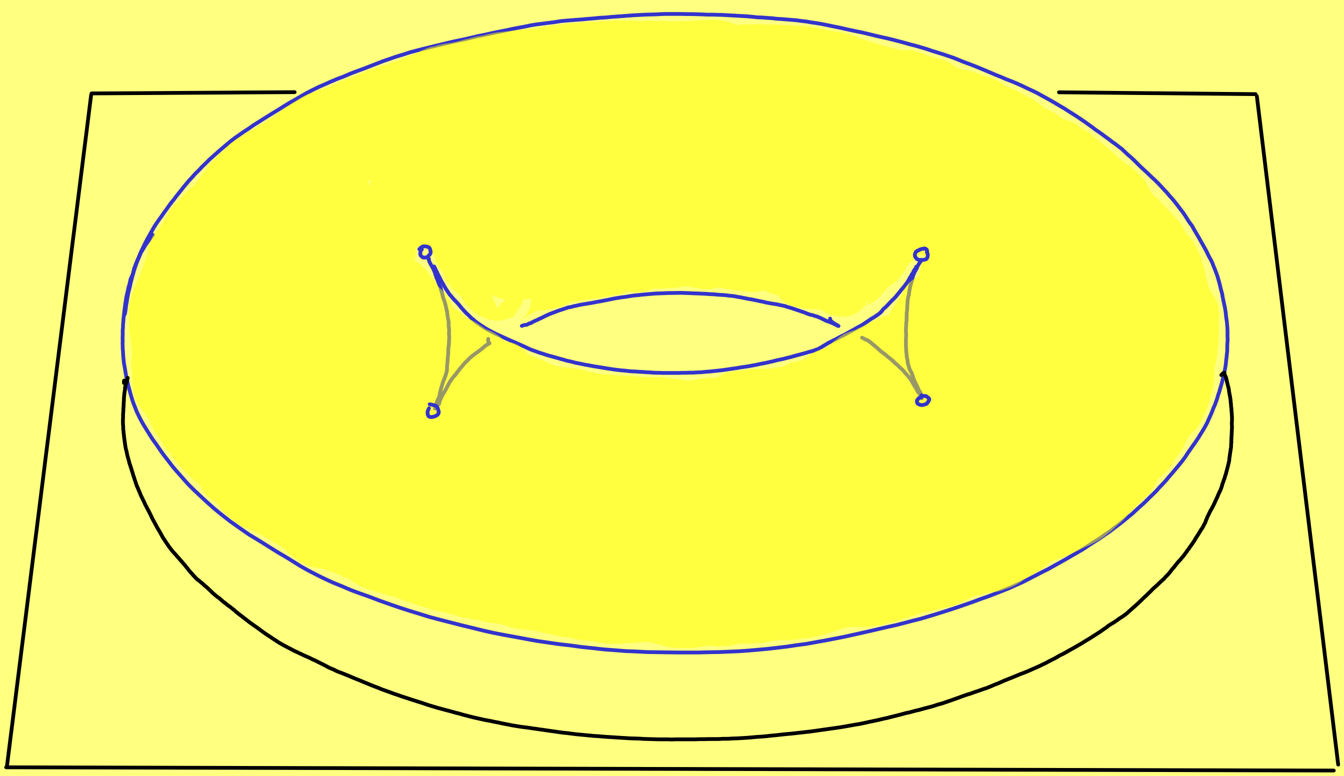
$$f, g : \mathbb{M} \to \mathbb{R}^2$$

$$\varepsilon = \max_{x \in \mathbb{M}} \| f(x) - g(x) \|_2$$

1. close contour lines;
2. thin creases.



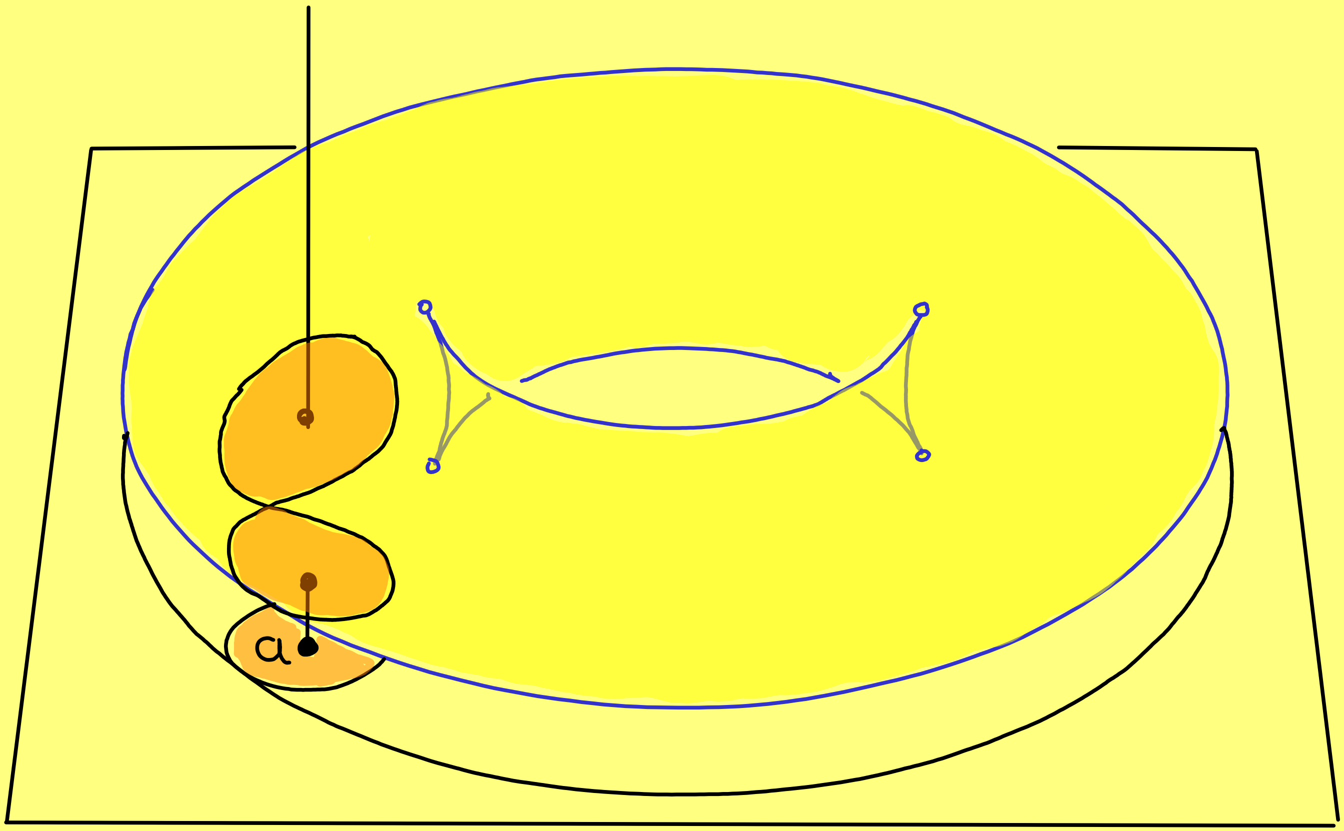
III.3 Distance Function



III.3 Distance Function

$f_a : \mathbb{M} \to \mathbb{R}$ defined by

$$f_a(x) = \| f(x) - a \|_2$$



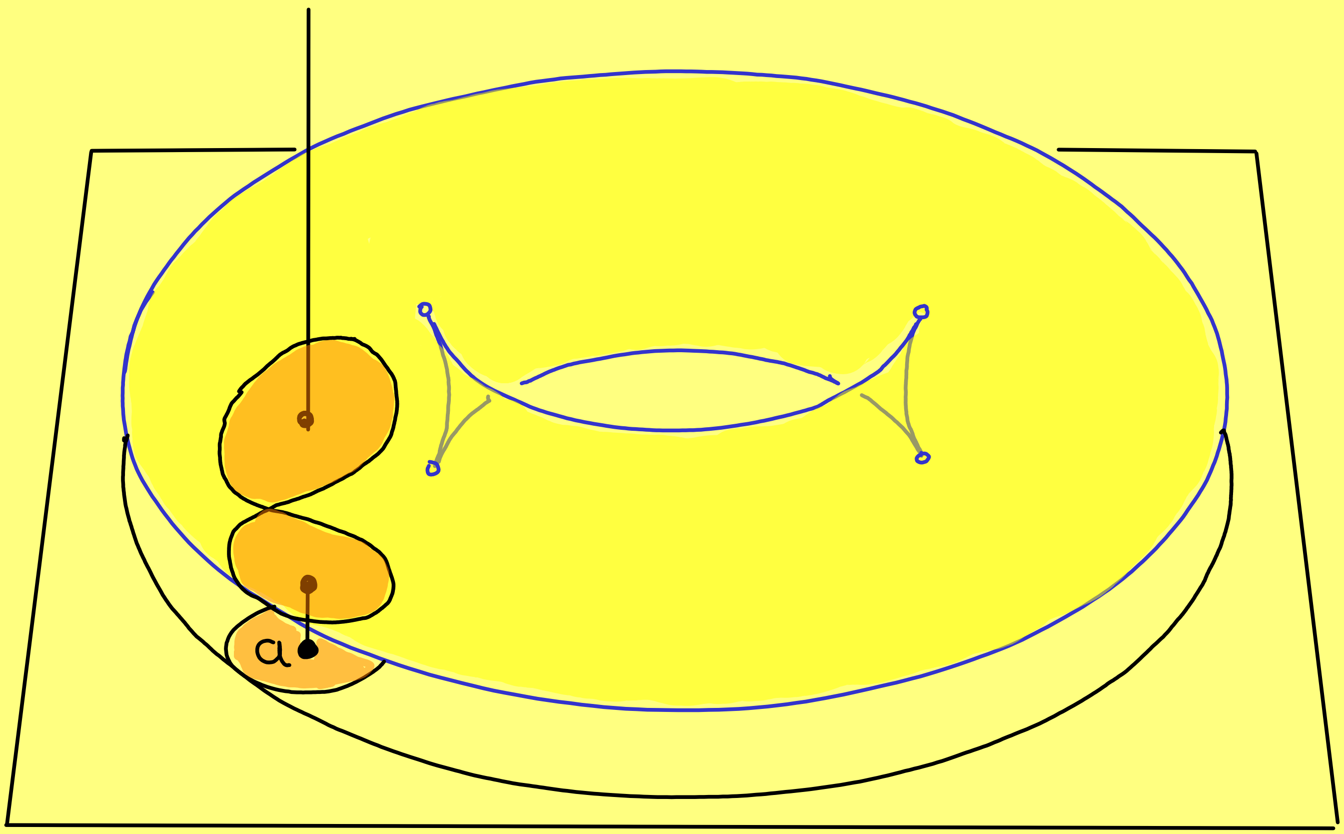
III.3 DISTANCE FUNCTION

$f_a : \mathbb{M} \to \mathbb{R}$ defined by

$$f_a(x) = \| f(x) - a \|_2$$

sublevel set is

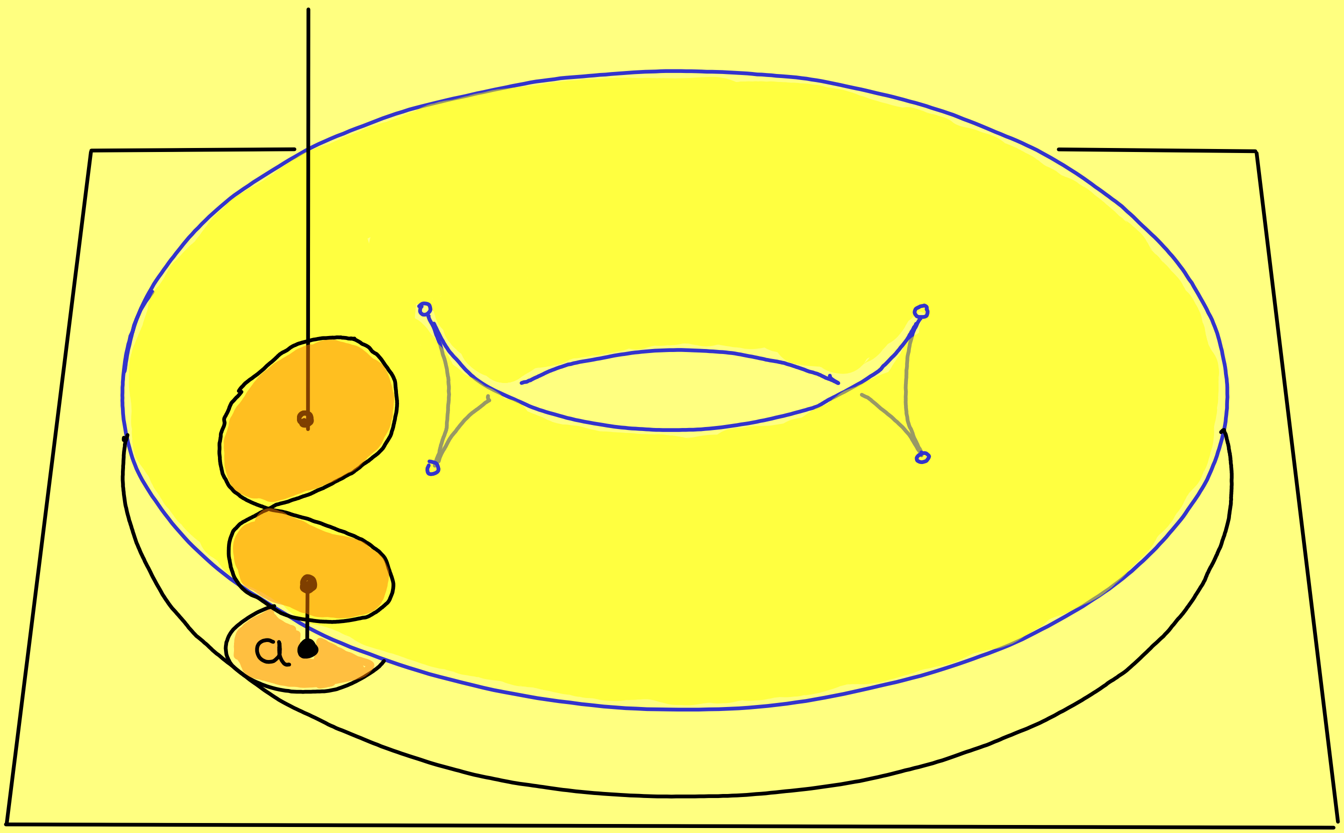
$$\mathbb{M}_r(a) = f_a^{-1}[0, r]$$



III.3 Distance Function

$f_a : \mathbb{M} \to \mathbb{R}$ defined by

$$f_a(x) = \| f(x) - a \|_2$$



sublevel set is

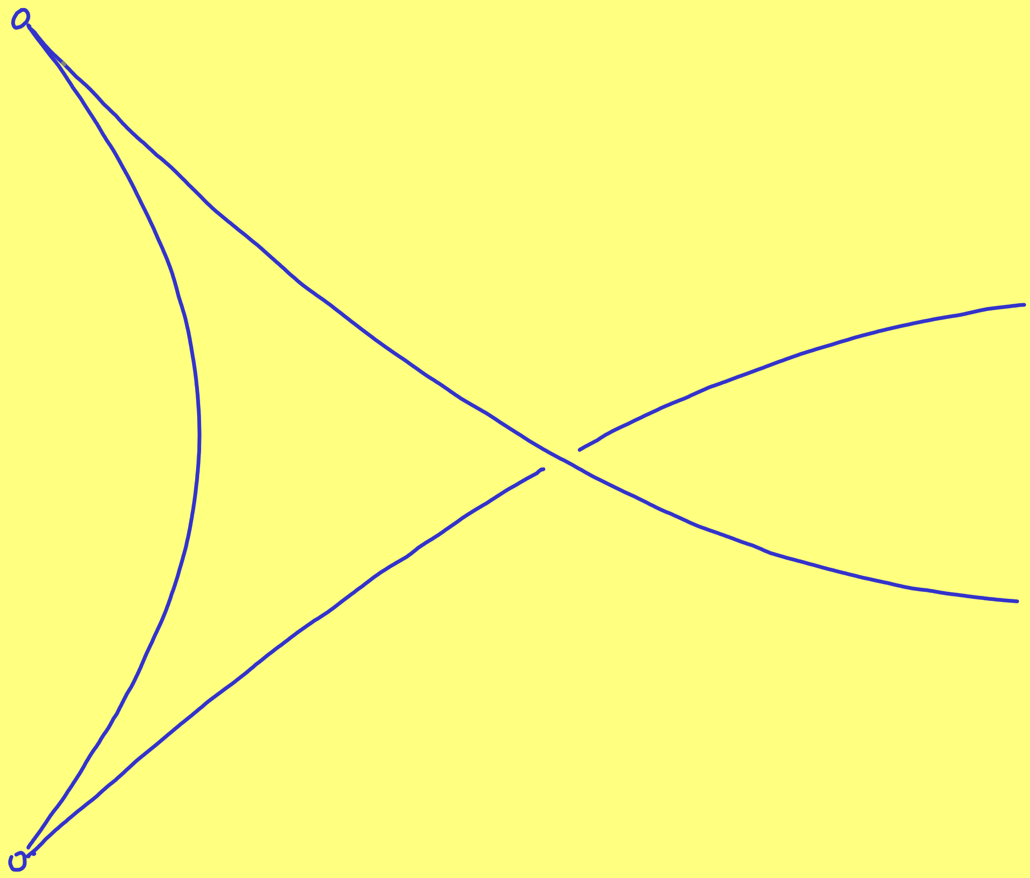
$$\mathbb{M}_r(a) = f_a^{-1}[0, r]$$

x crit. for f_a

$\Downarrow$

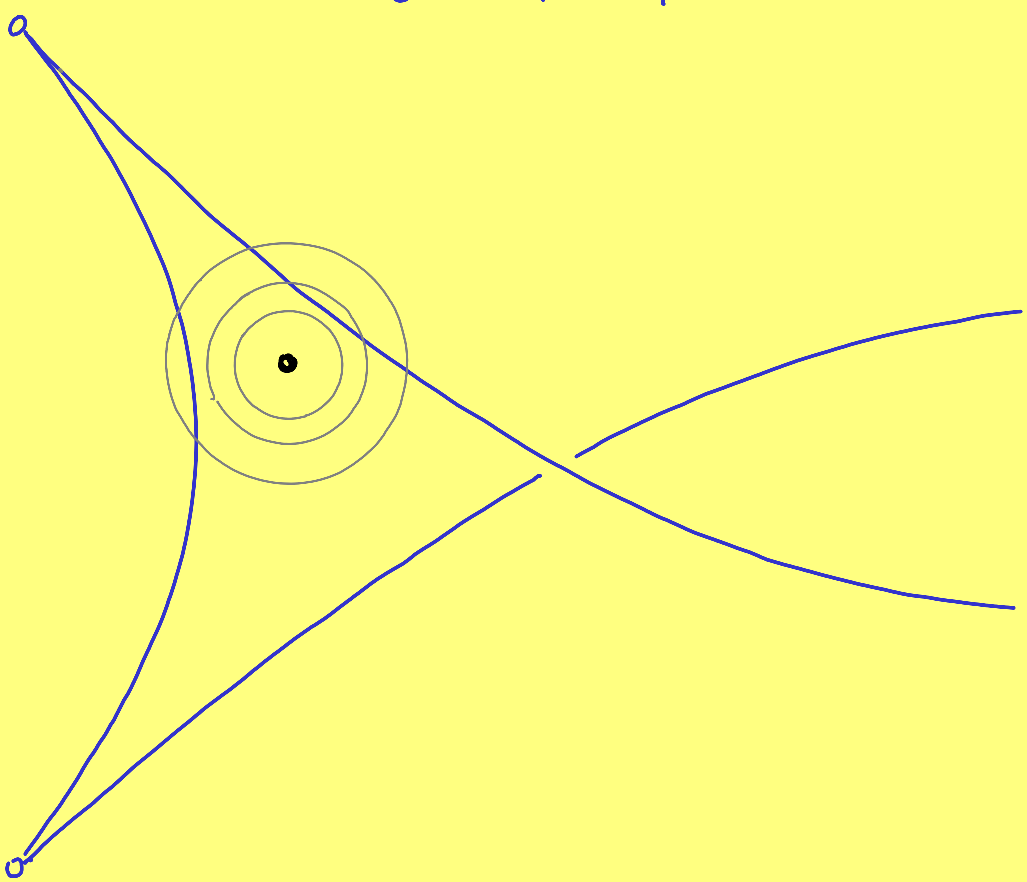
x crit. for f

III.4 Degree



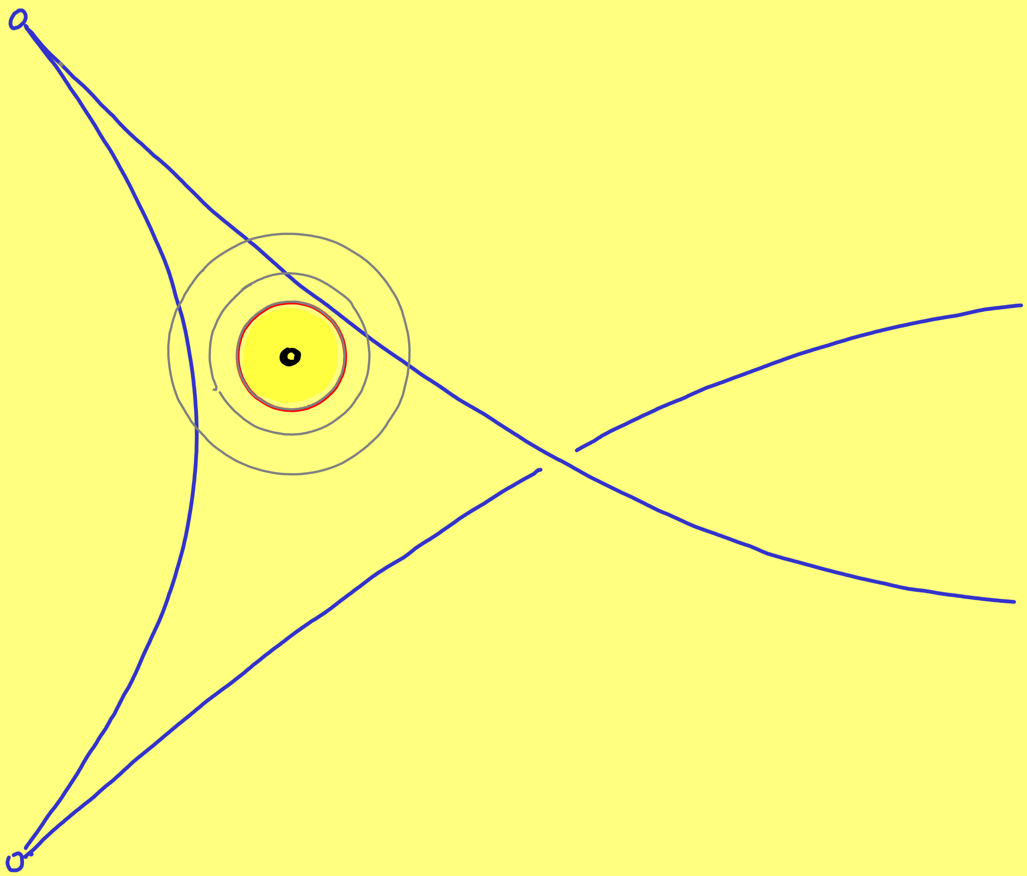
III.4 DEGREE

$\mathbb{F}_r = H_0(\mathbb{M}_r(a))$ indexed by radius



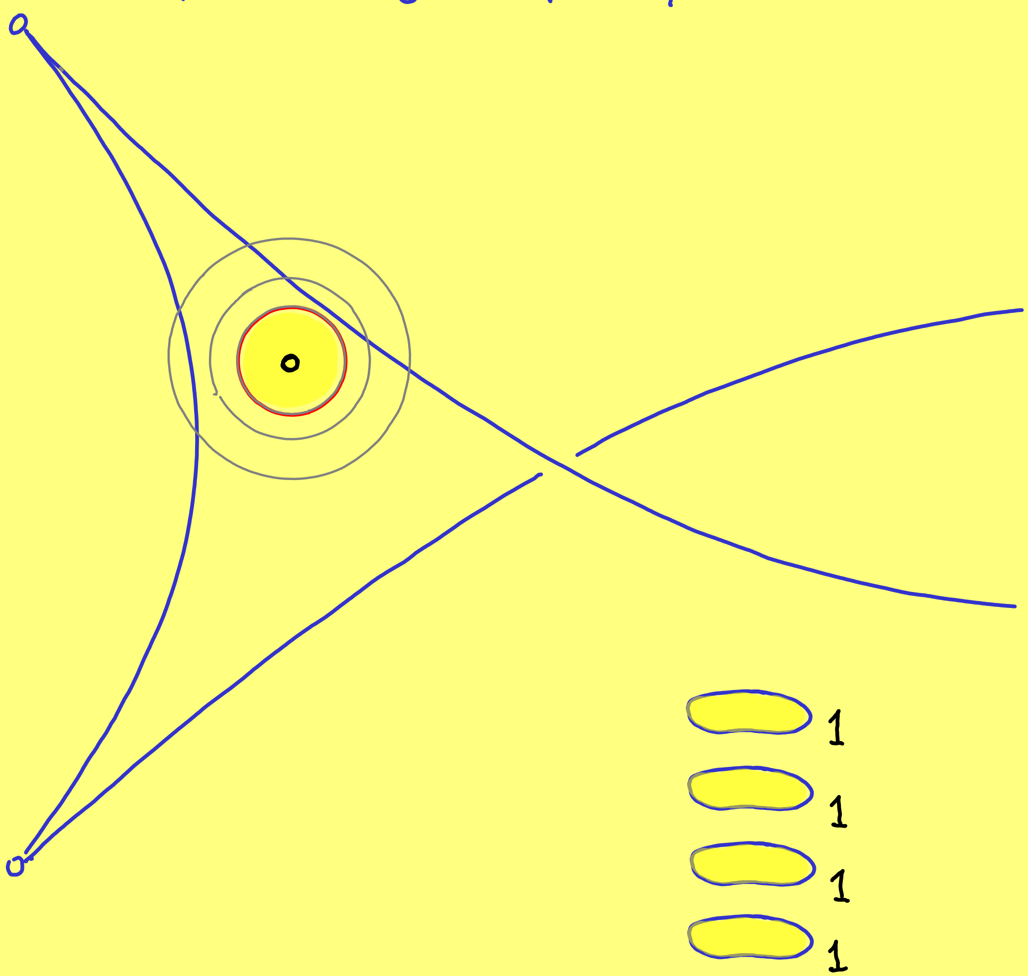
III.4 Degree

$F_r = H_0(\mathbb{M}_r(a))$ indexed by radius



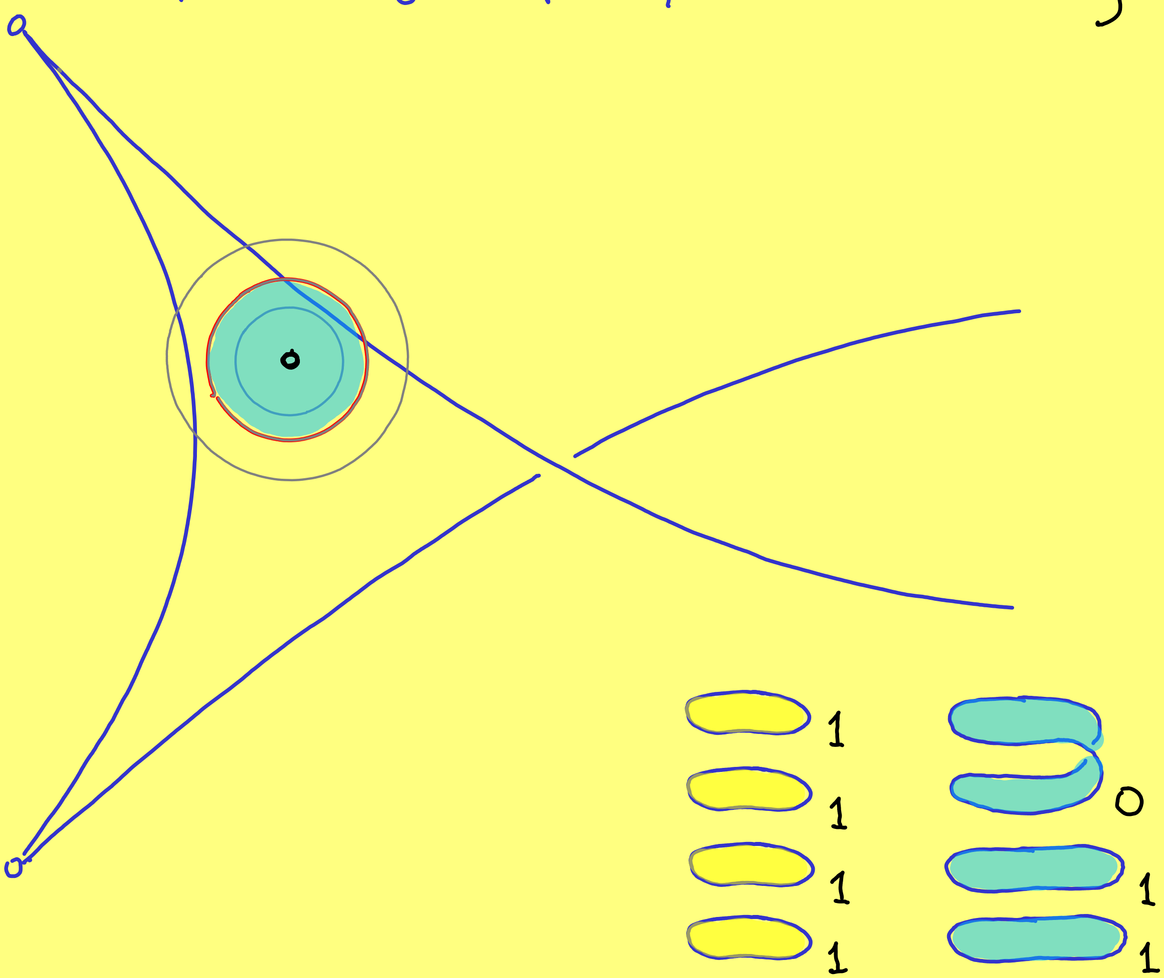
III.4 DEGREE

$F_r = H_0(\mathbb{M}_r(a))$ indexed by radius



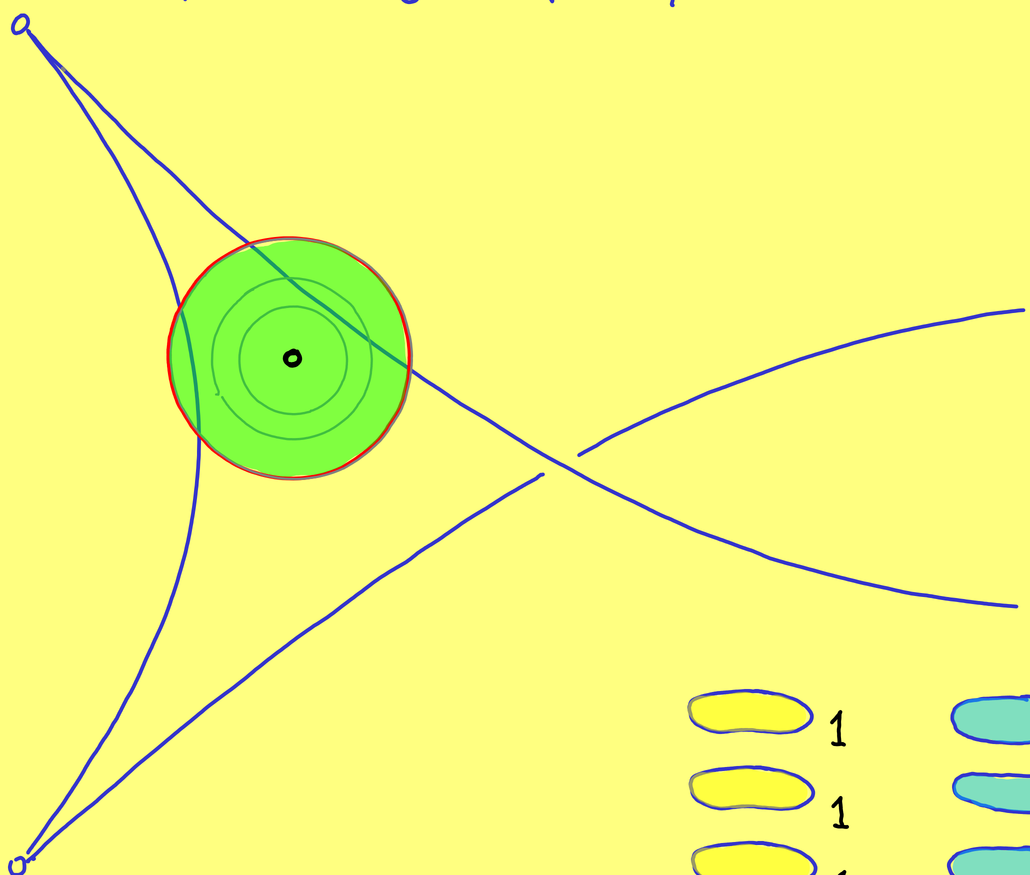
III.4 DEGREE

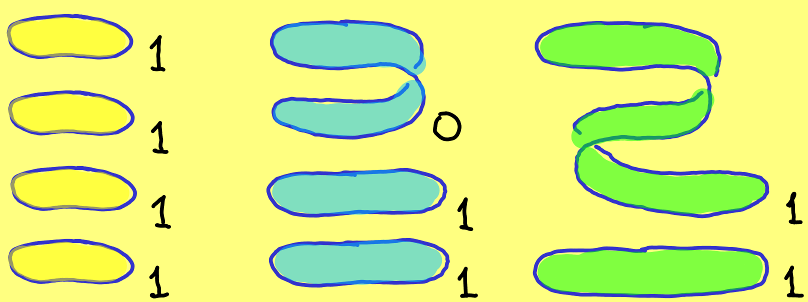
$F_r = H_0(\mathbb{M}_r(a))$ indexed by radius



III.4 DEGREE

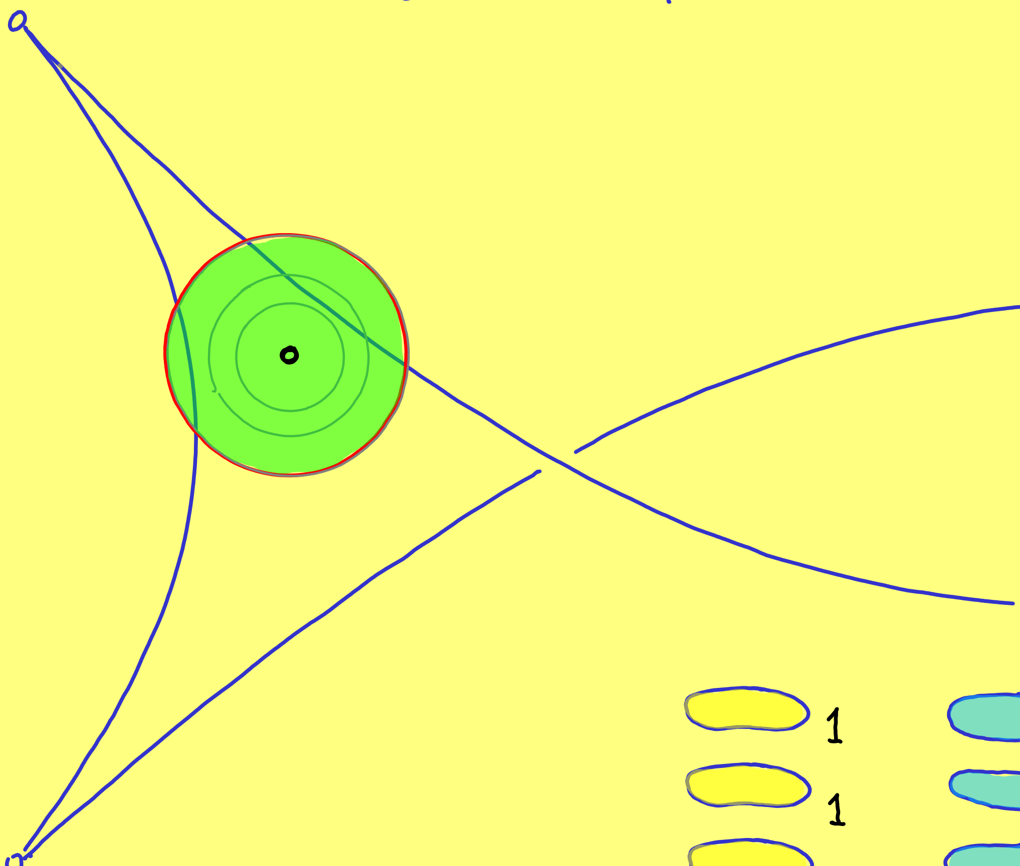
$F_r = H_0(\mathbb{M}_r(a))$ indexed by radius

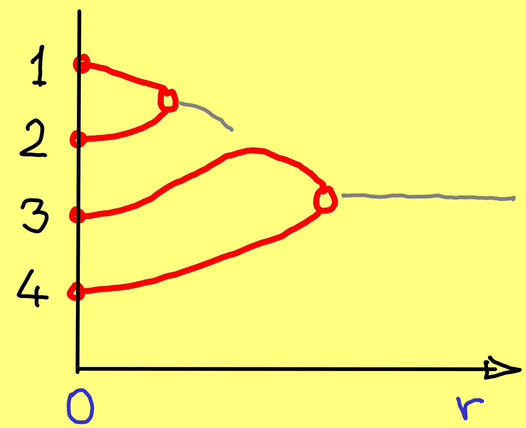


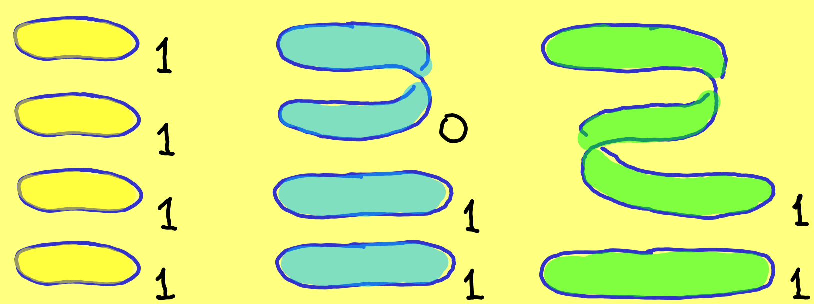


III.4 Degree

$F_r = H_0(\mathbb{M}_r(a))$ indexed by radius

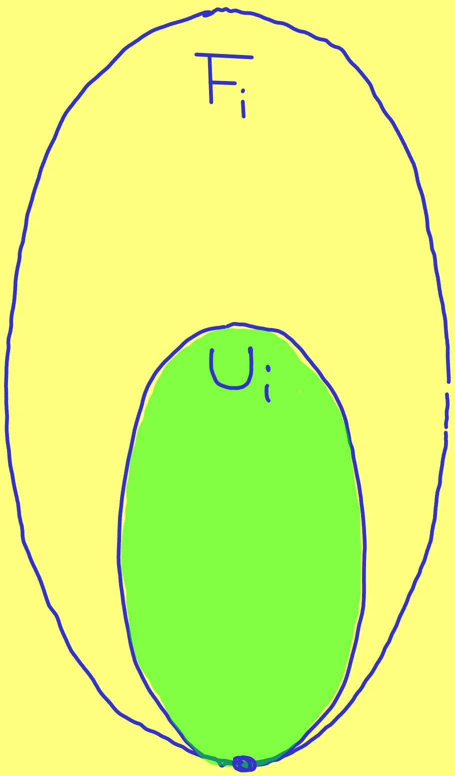


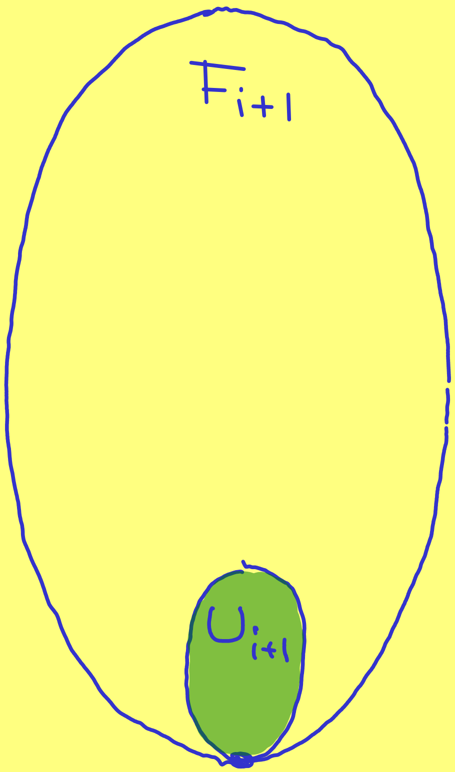




III.5 WELL GROUP

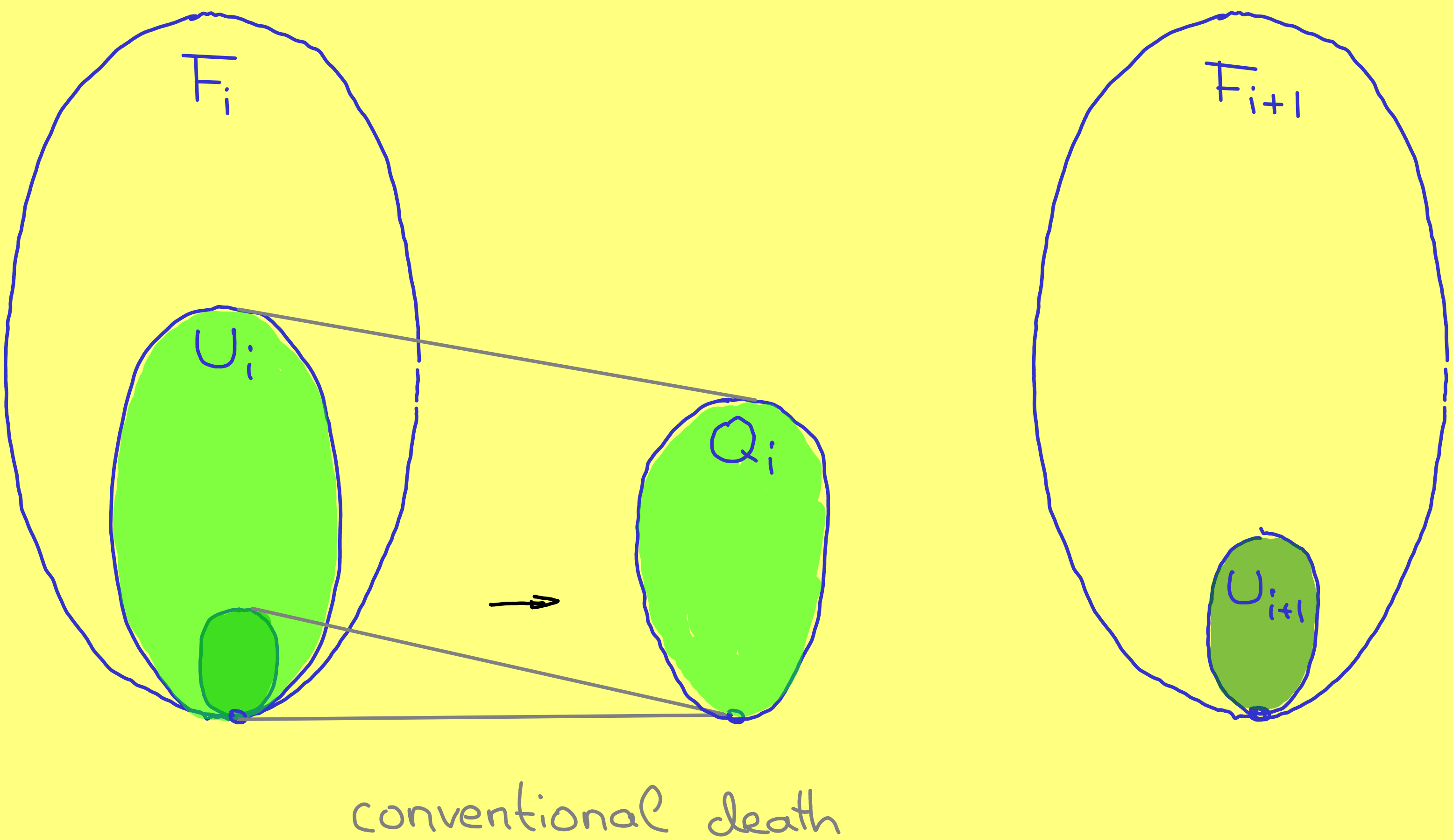
$U_i \subseteq F_i$ well subgroup of $H_0(\mathbb{M}_r(a))$





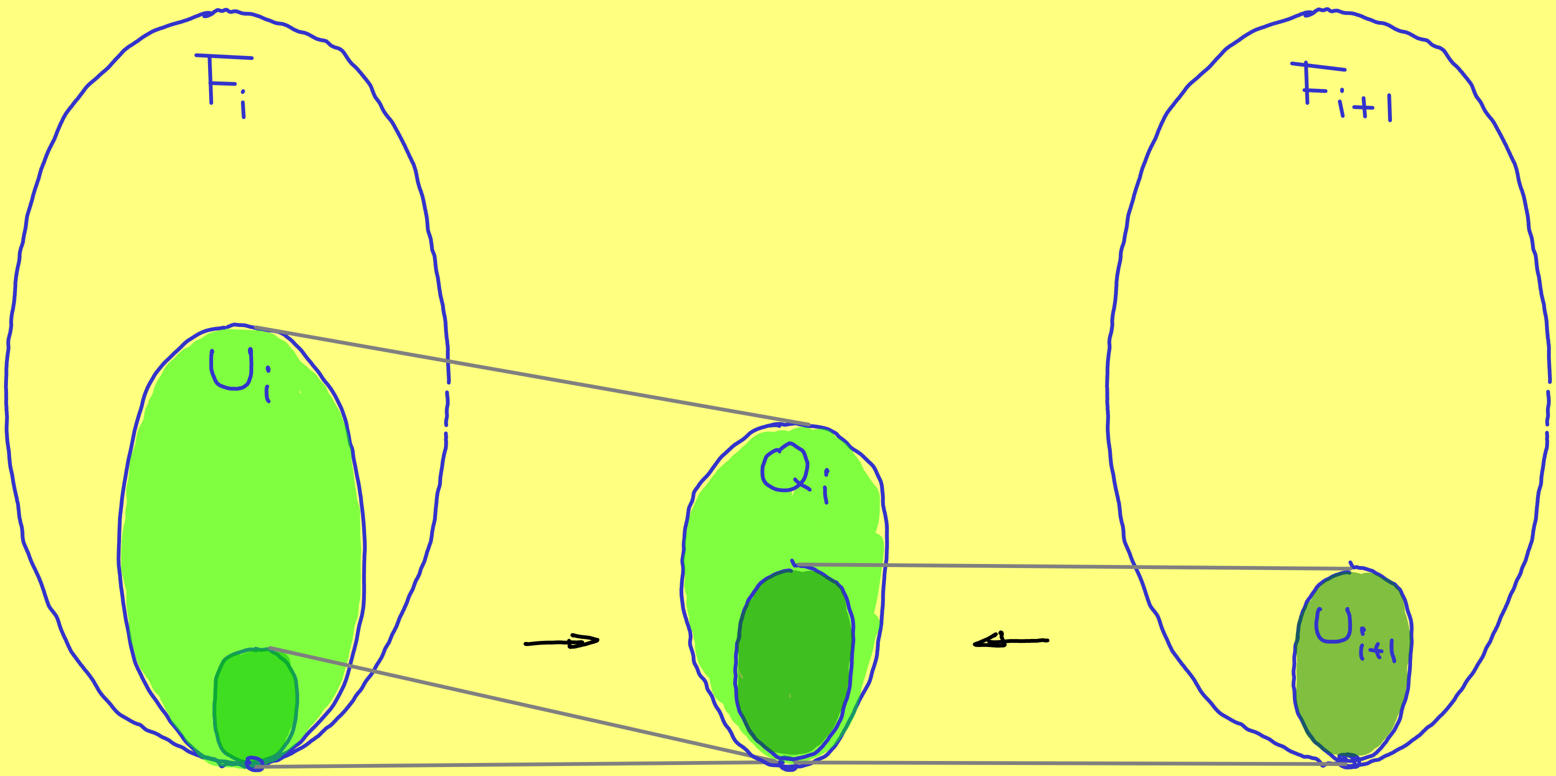
III.5 WELL GROUP

$U_i \subseteq F_i$ well subgroup of $H_0(\mathbb{M}_r(a))$



III.5 WELL GROUP

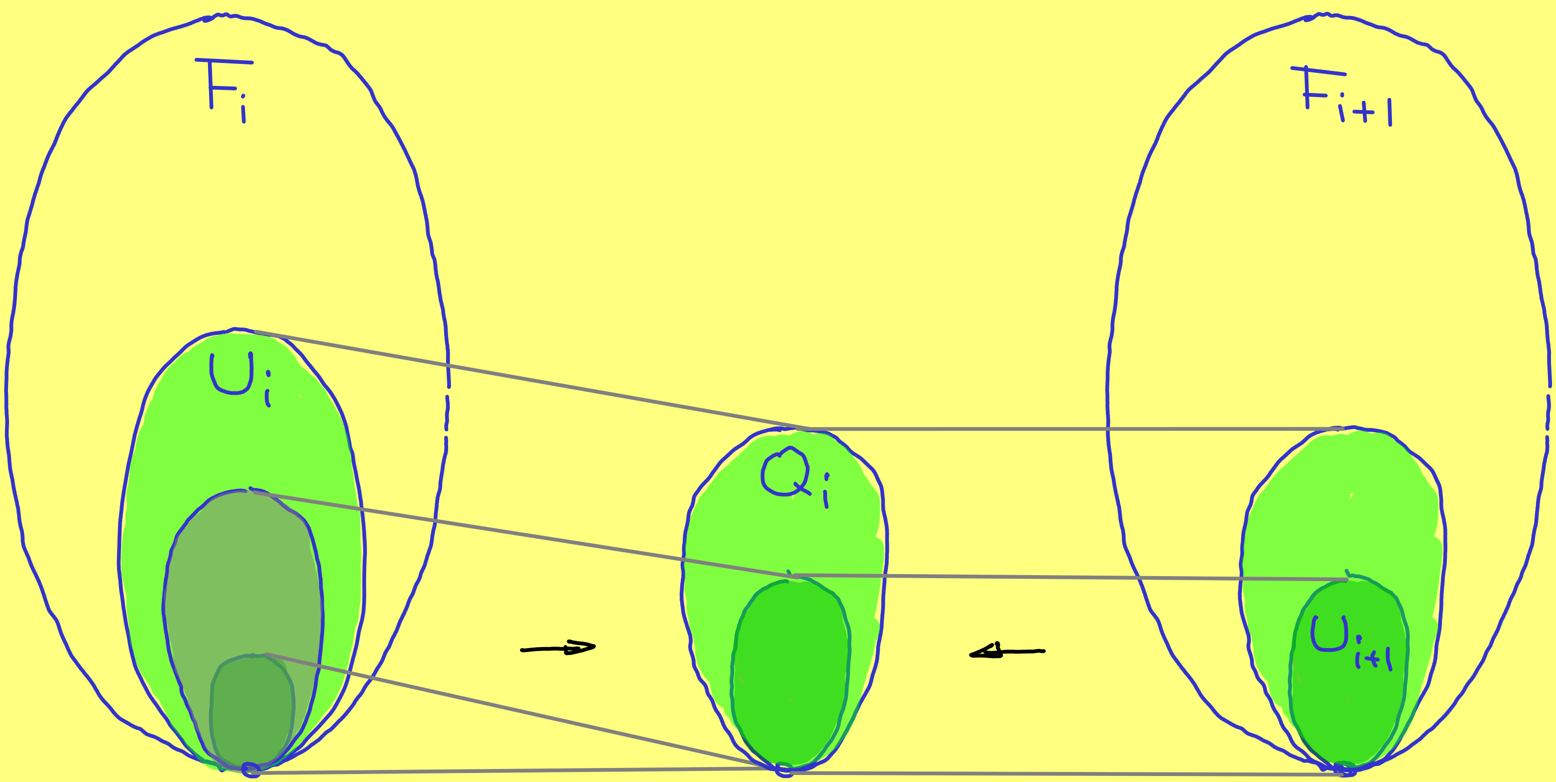
$U_i \subseteq F_i$ well subgroup of $H_0(\mathbb{M}_r(a))$



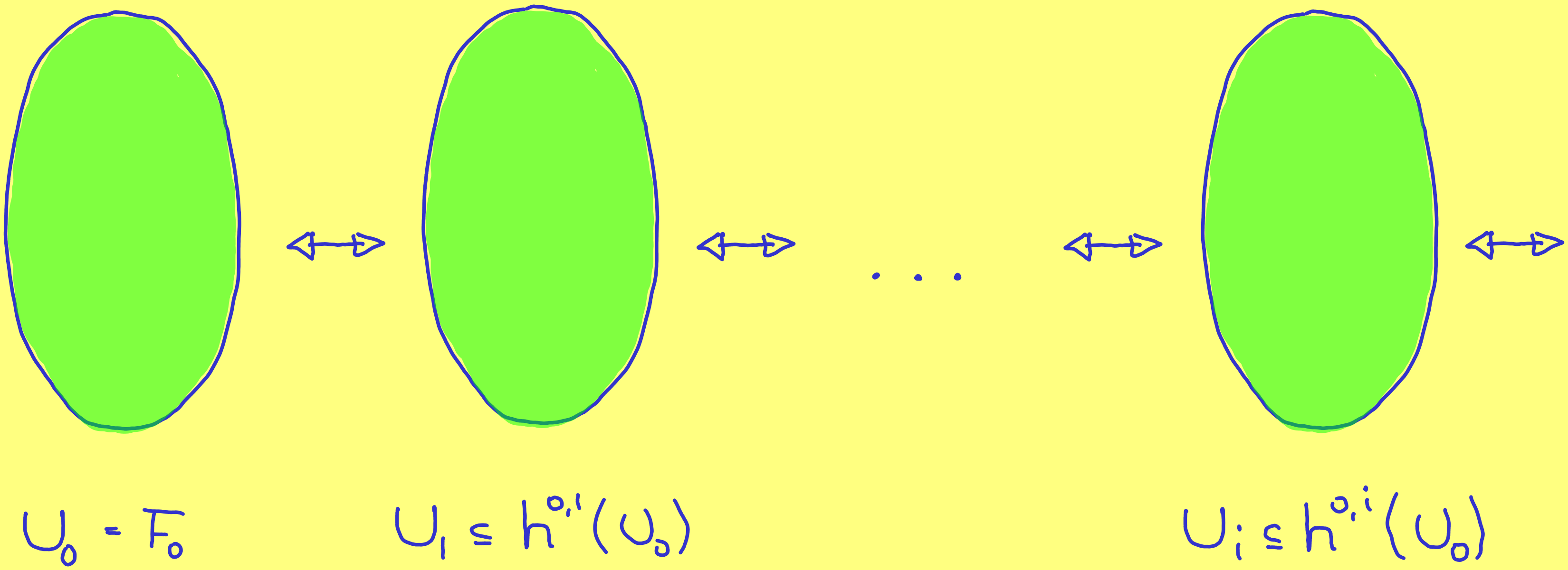
unconventional death

III.5 WELL GROUP

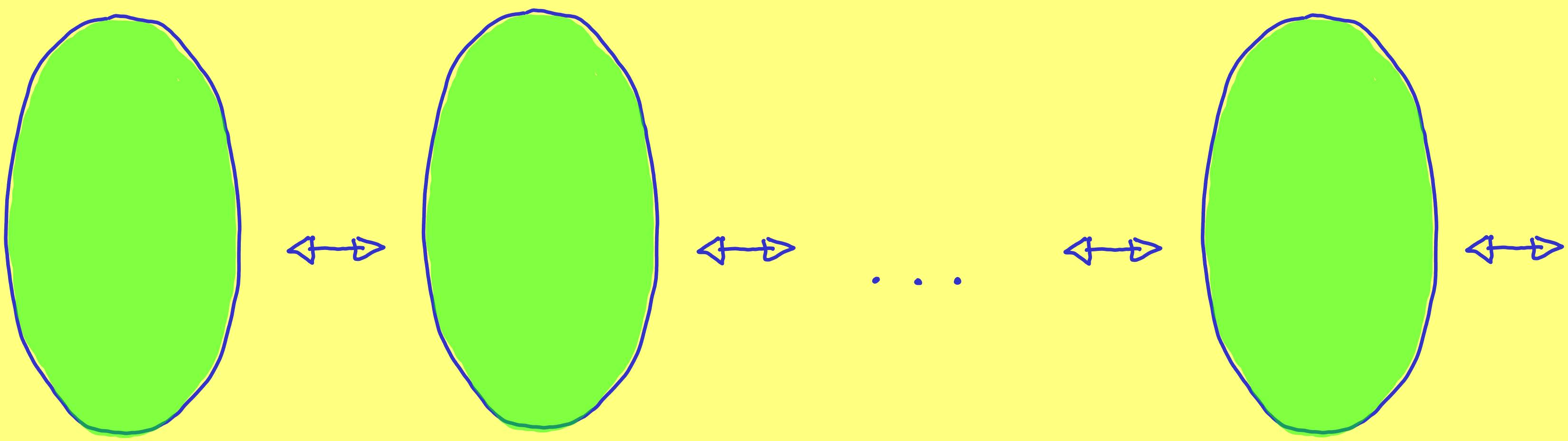
$U_i \subseteq F_i$ well subgroup of $H_0(\mathbb{M}_r(a))$



III.6 Zigzag Module



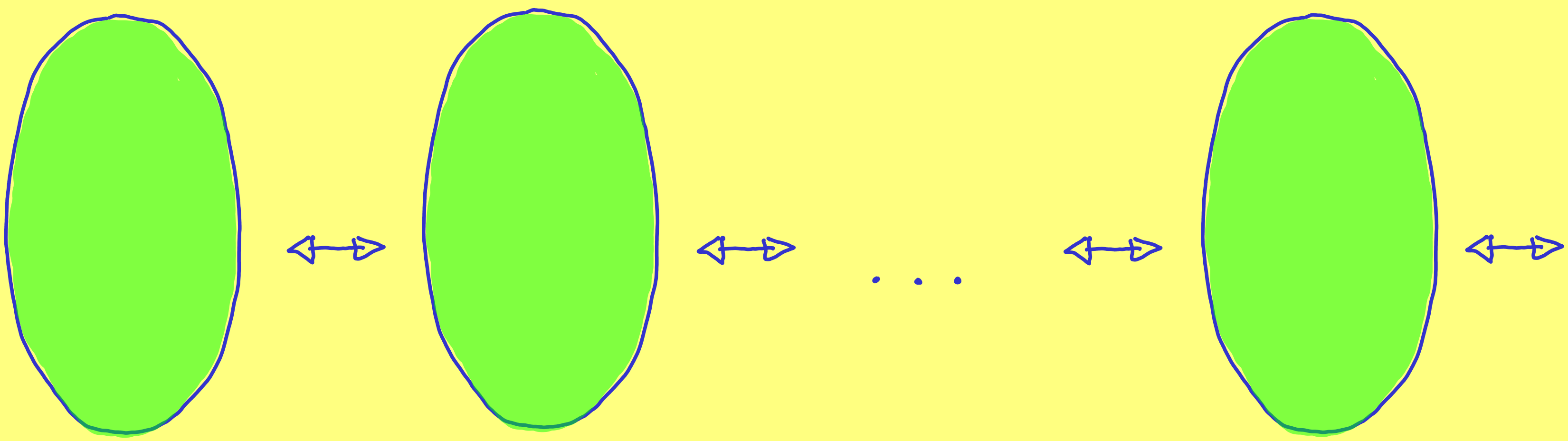
III.6 Zigzag Module



$U_0 = F_0$ $\qquad$ $U_1 \subseteq h^{0,1}(U_0)$ $\qquad$ $U_i \subseteq h^{0,i}(U_0)$

persistence diagram exists

III.6 Zigzag Module

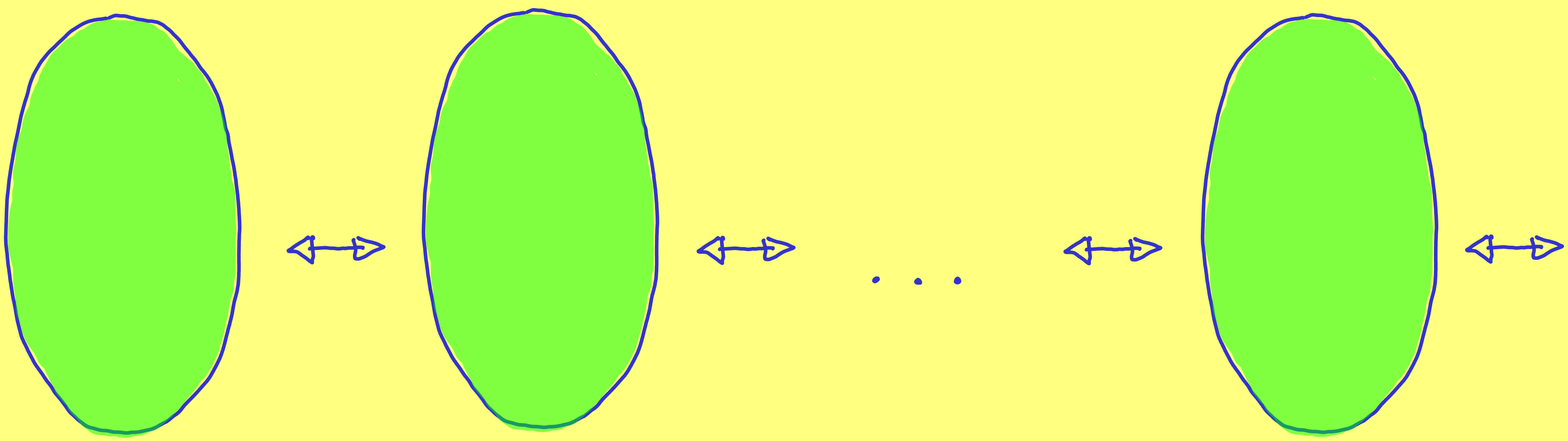


$U_0 = F_0$ $\quad U_1 \leq h^{0,1}(U_0)$ $\quad U_i \leq h^{0,i}(U_0)$

persistence diagram

exists
one-dimensional

III.6 Zigzag Module

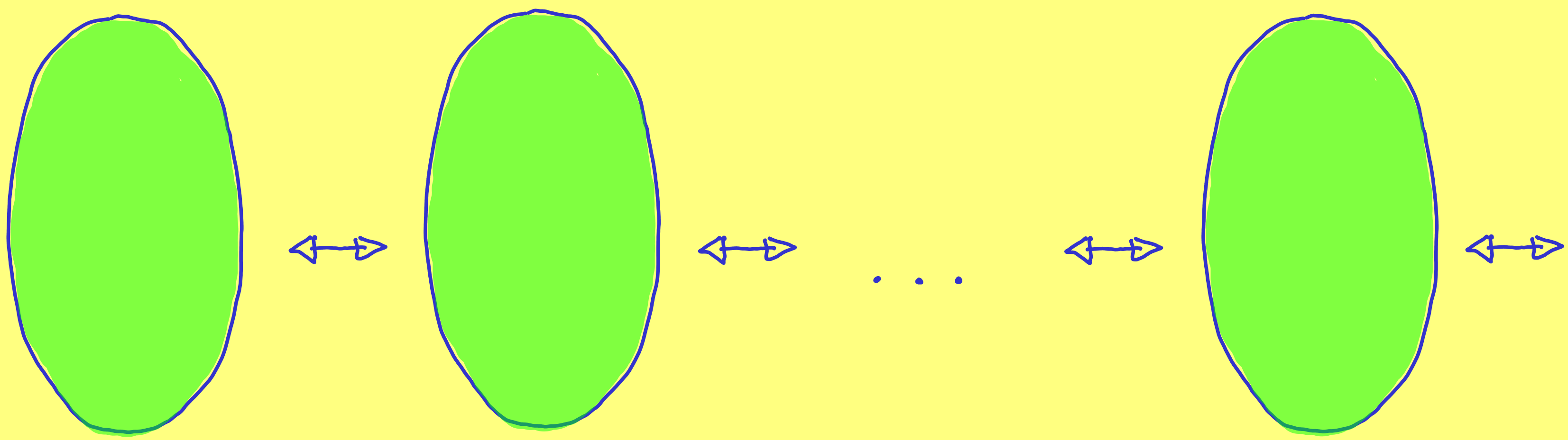


$U_0 = F_0$ $\qquad$ $U_1 \subseteq h^{0,1}(U_0)$ $\qquad$ $U_i \subseteq h^{0,i}(U_0)$

persistence diagram

exists
one-dimensional
stable

III.6 Zigzag Module



$U_0 = F_0$ $\quad$ $U_1 \subseteq h^{0,1}(U_0)$ $\quad \ldots \quad$ $U_i \subseteq h^{0,i}(U_0)$

persistence diagram

exists
one-dimensional
stable $\Longrightarrow$ contour stable

THANK YOU