Measuring Scale Before Simplification

Herbert Edelsbrunner

BMS

Measuring Scale Before Simplification

I. Functions

II. Curves

III. Somites

IV. Motion

I. Functions

I.1 PERSISTENCE

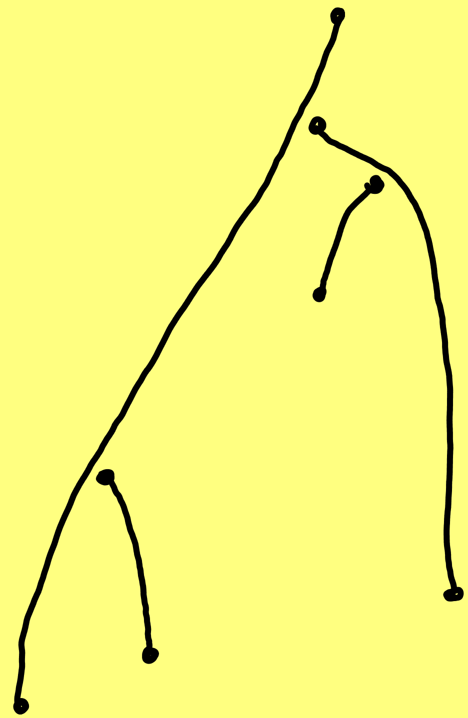
function

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

I.1 Persistence

function

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



merge tree

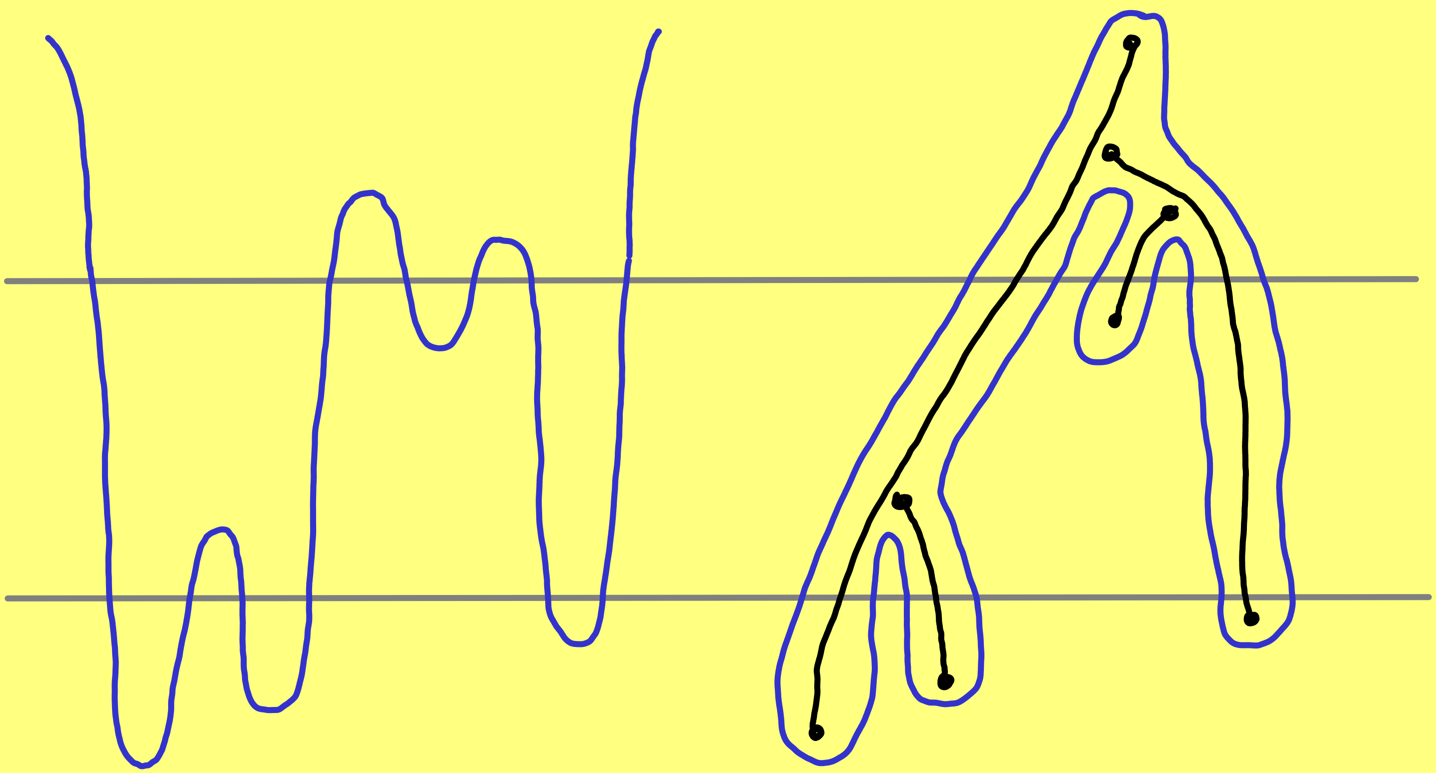
I.1 Persistence

function

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

merge tree

I.1 PERSISTENCE

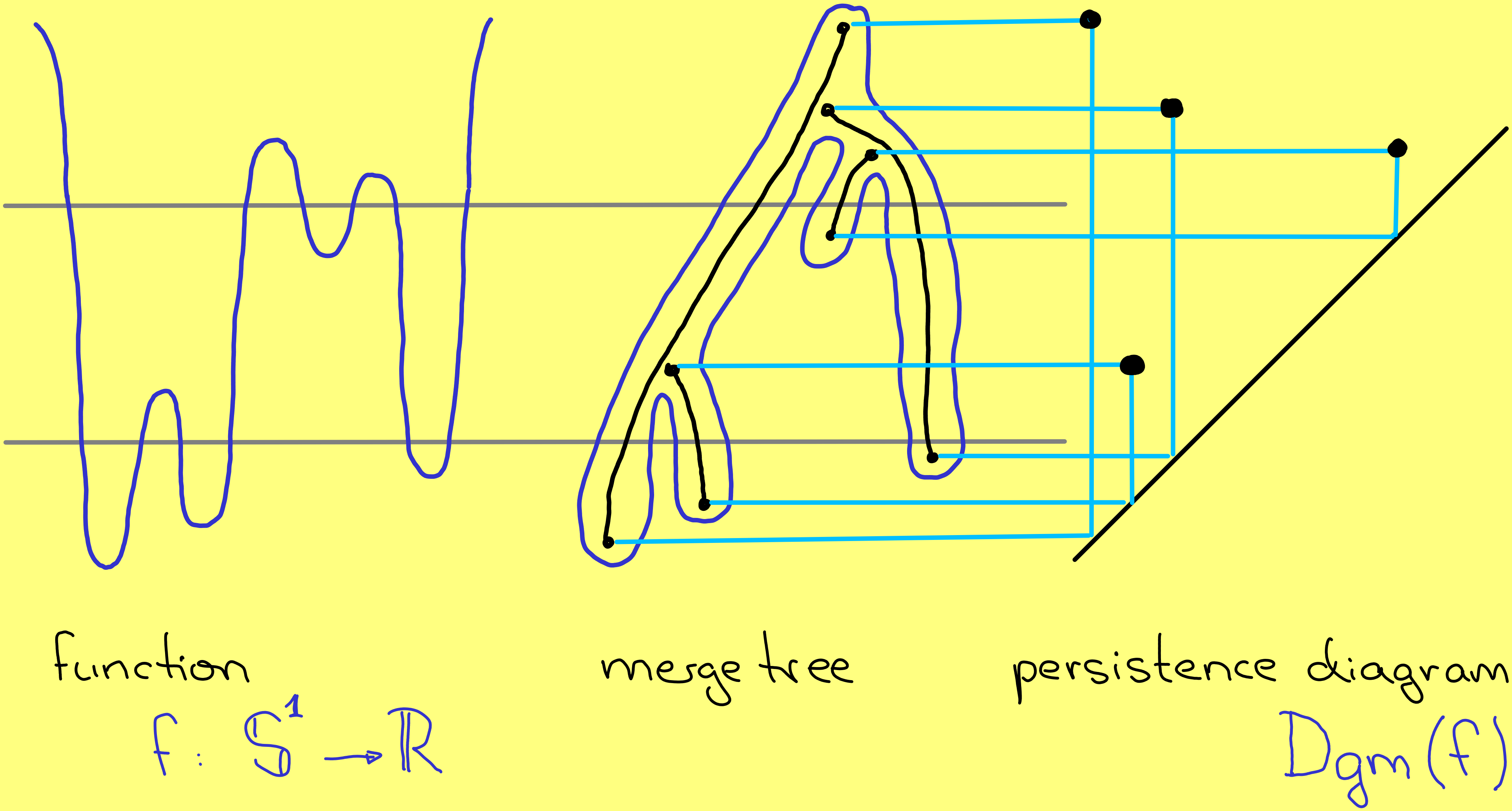


function

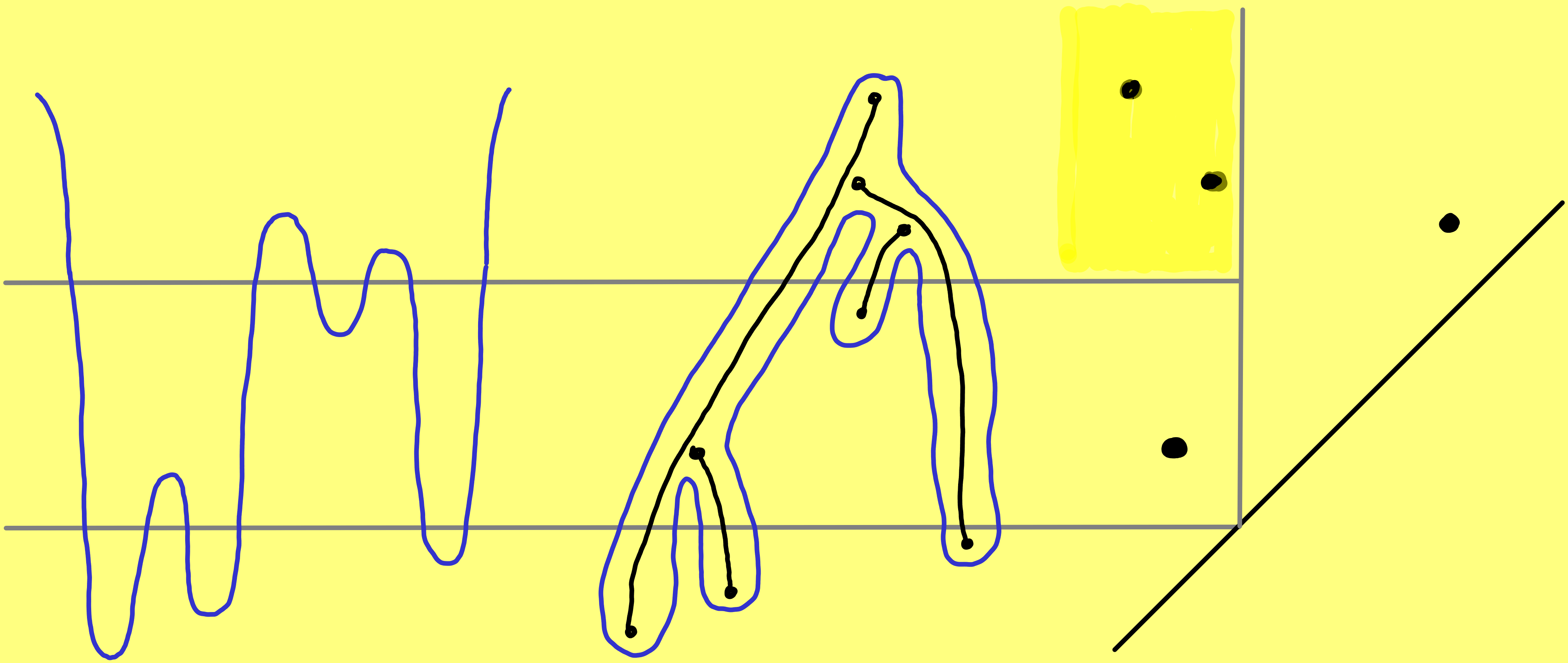
$f : \mathbb{S}^1 \to \mathbb{R}$

merge tree

I.1 Persistence



I.1 Persistence

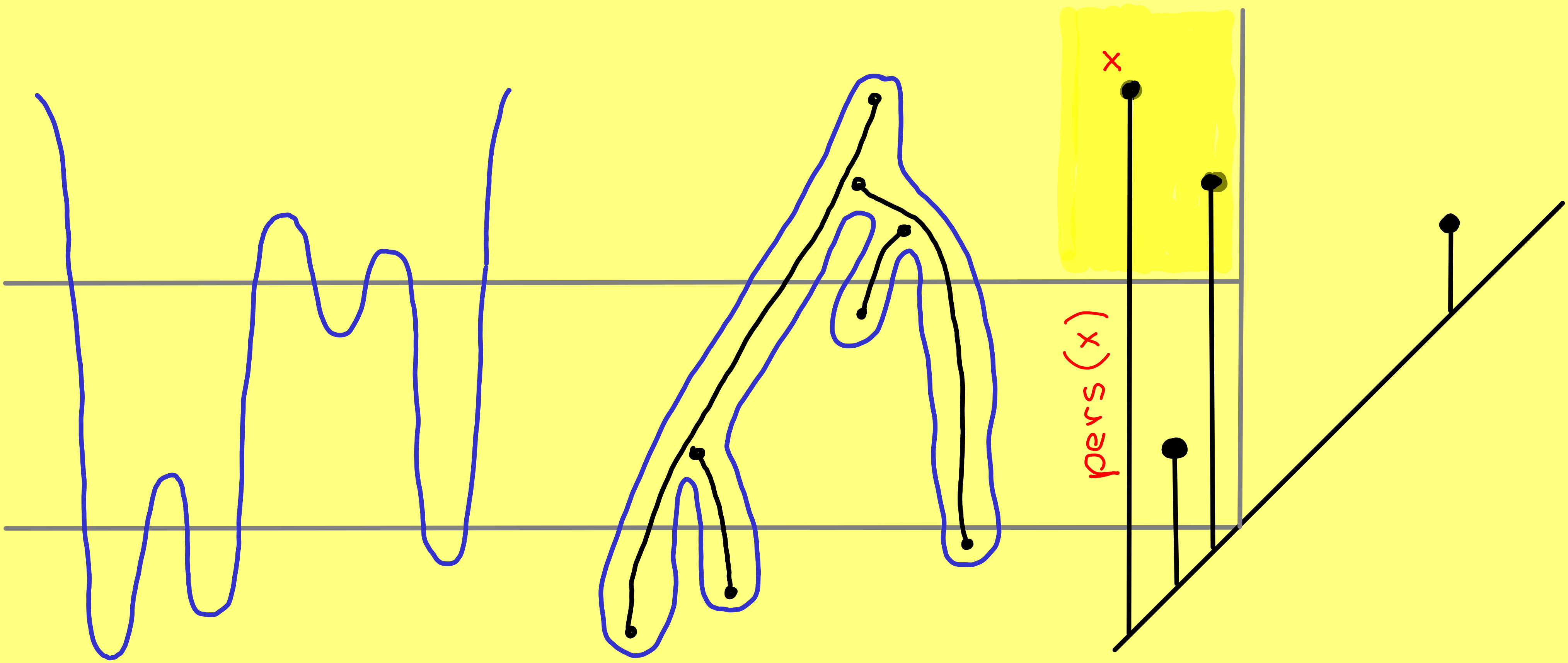


function $f: \mathbb{S}^1 \to \mathbb{R}$

merge tree

persistence diagram $\mathrm{Dgm}(f)$

I.1 PERSISTENCE



function

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

merge tree

persistence diagram

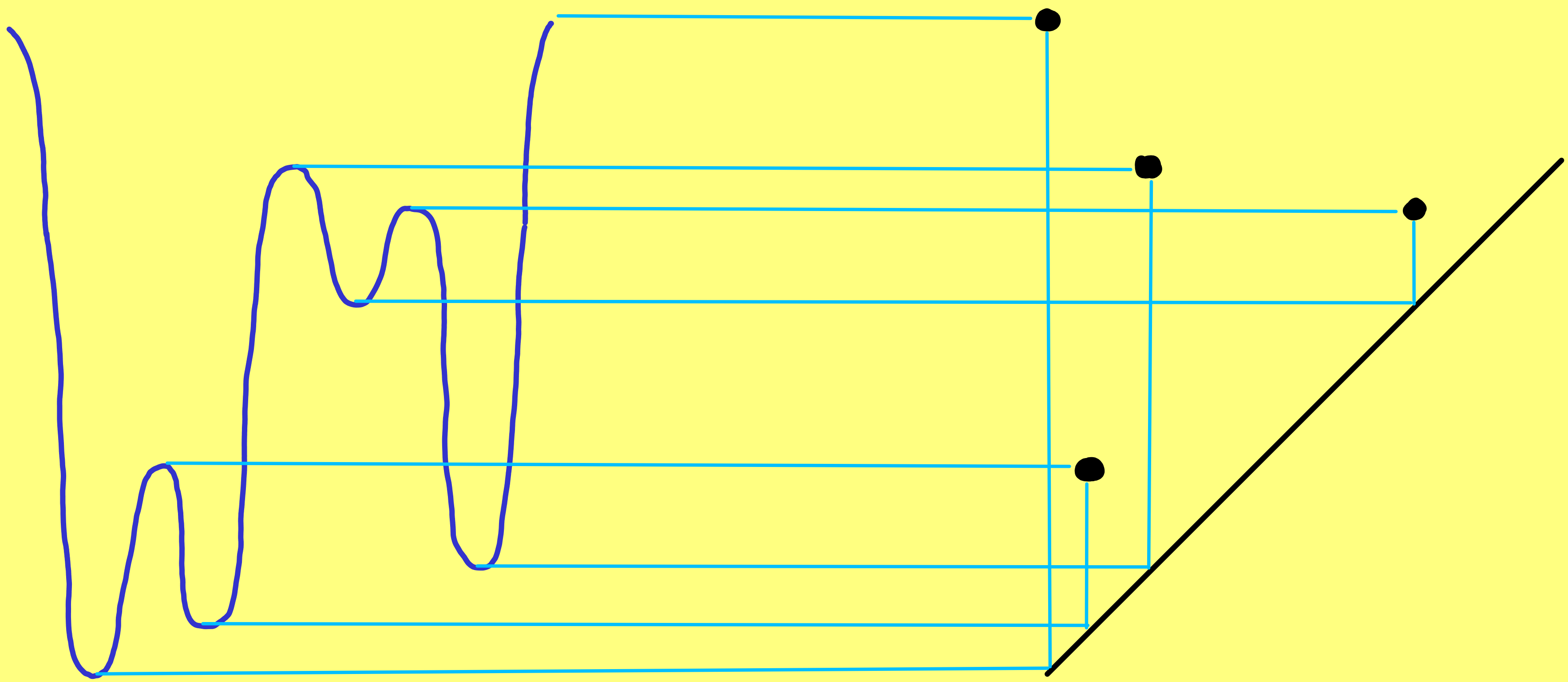
$\mathrm{Dgm}(f)$

I.2 L_∞ - STABILITY

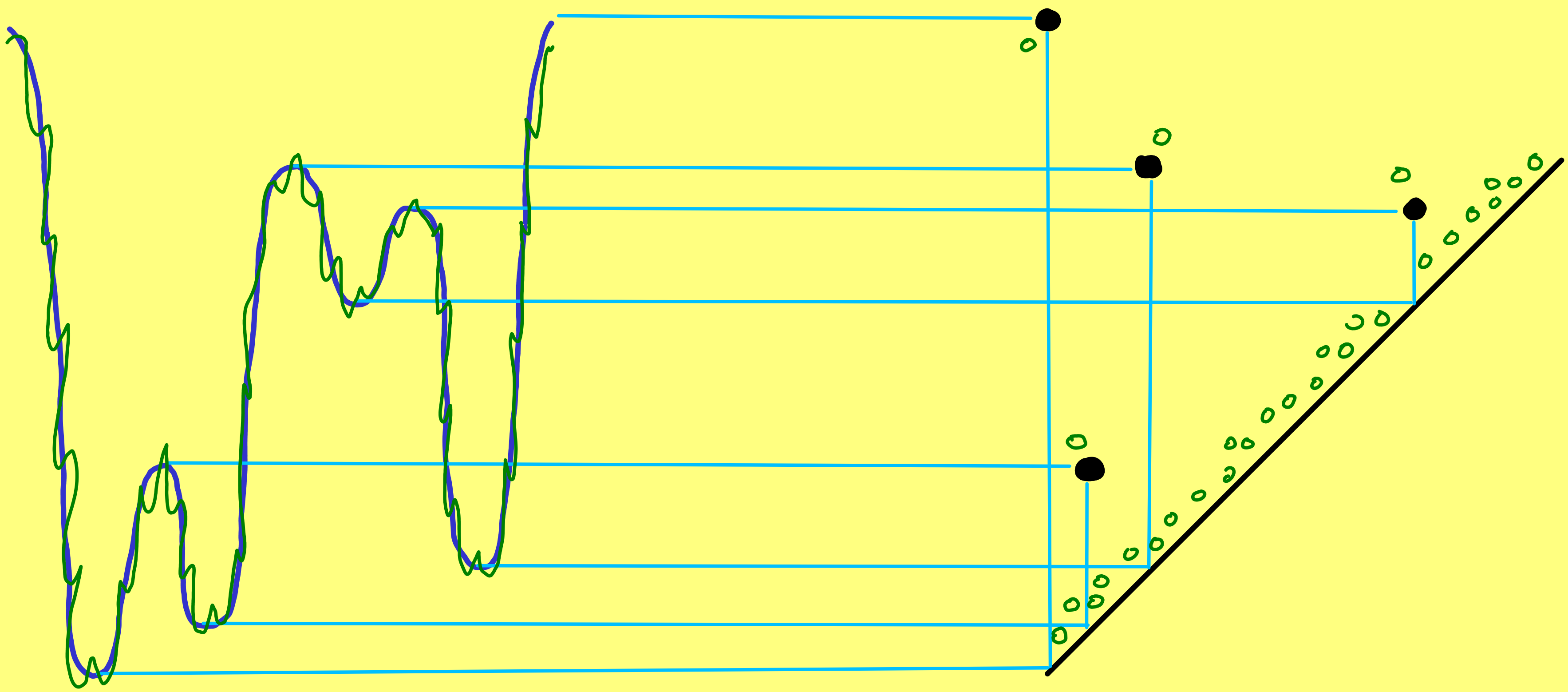
THM. For tame functions $f, g : \mathbb{X} \to \mathbb{R}$ the bottleneck distance between their diagrams is bounded by the max-difference between the functions,

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

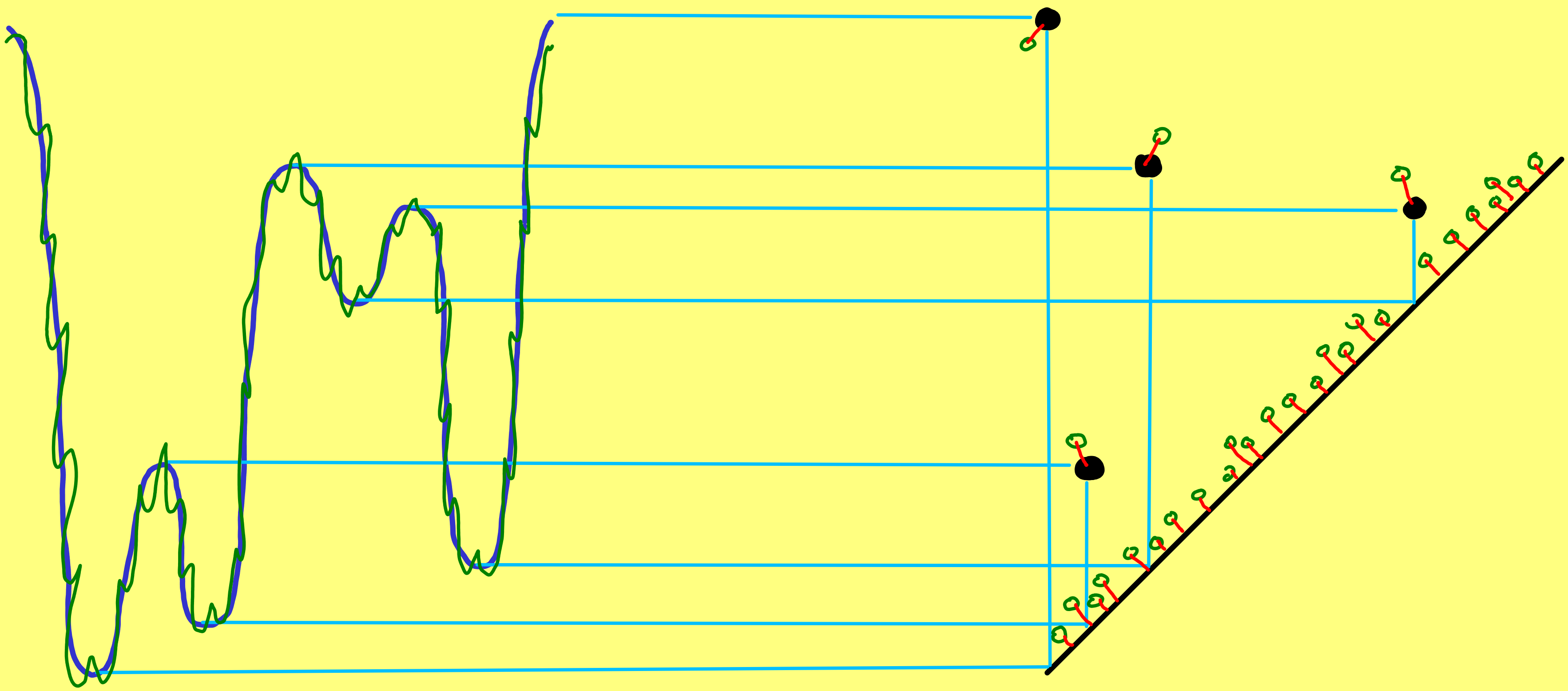
I.2 L_∞-STABILITY



I.2 L_∞-STABILITY



I.2 L_∞-STABILITY



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

I.3 L_p - STABILITY

DEF. $f : \mathbb{X} \to \mathbb{R}$ has Lipschitz constant C if $f(x) - f(y) \leq C \|x-y\|$.

I.3 L_p-STABILITY

DEF. $f : \mathbb{X} \to \mathbb{R}$ has Lipschitz constant C if $f(x) - f(y) \leq C \|x-y\|$.

DEF. The p-th Wasserstein distance between the diagrams of f and g is

$$W_p(f,g) = \inf_{\gamma : Dgm(f) \to Dgm(g)} \sum_{x \in Dgm(f)} \|x - \gamma(x)\|_1^p .$$

bijection

I.3 L_p - STABILITY

THM. Let $\mathbb{X}$ be a compact metric space and $f, g : \mathbb{X} \to \mathbb{R}$ two Lipschitz functions. Then there exist constants k and C that depend on $\mathbb{X}$ and the Lipschitz constants of f and g s.t.

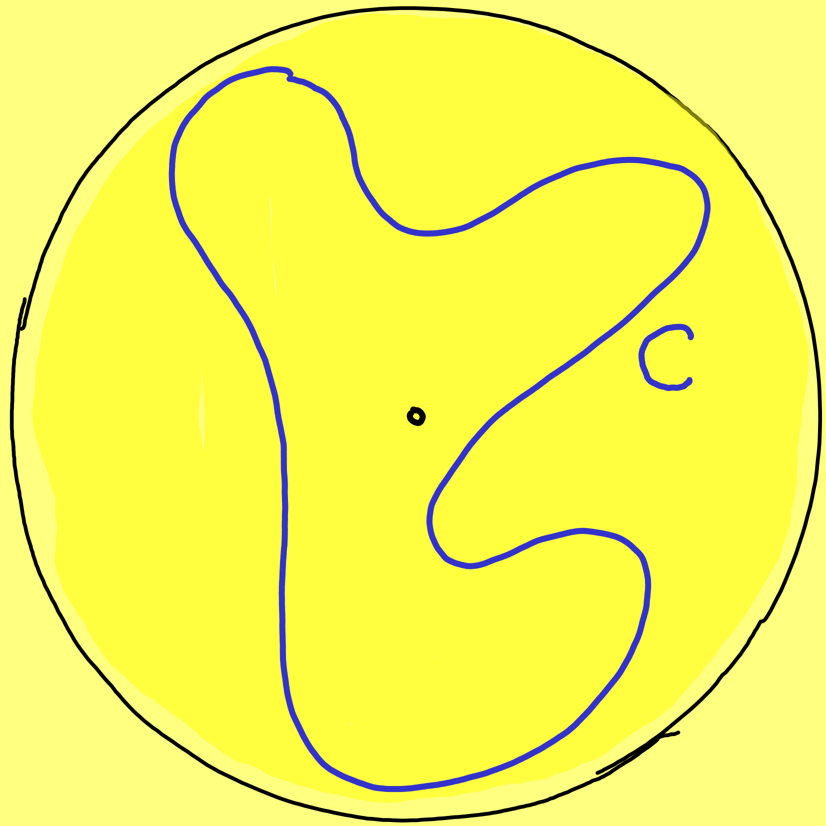
$$W_p(f, g) \leq C \|f - g\|_\infty^{p-k}$$

for every $p \geq k$.

II. CURVES

II.1 Fary's Theorem

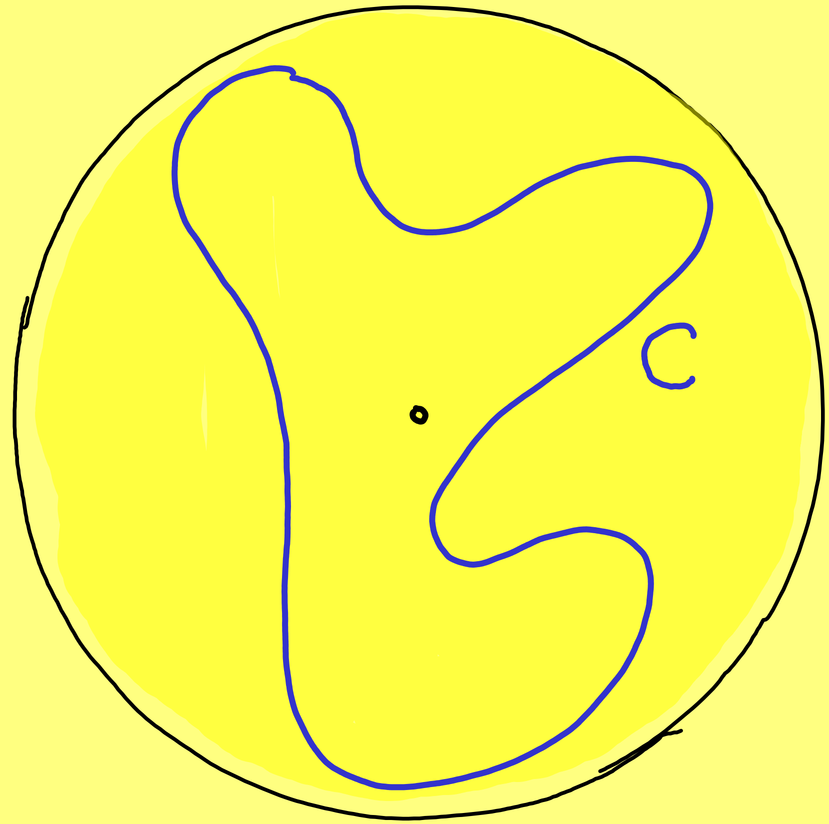
C = smooth curve inside unit disk



II.1 Fary's Theorem

C = smooth curve inside unit disk

$L(C)$ = length

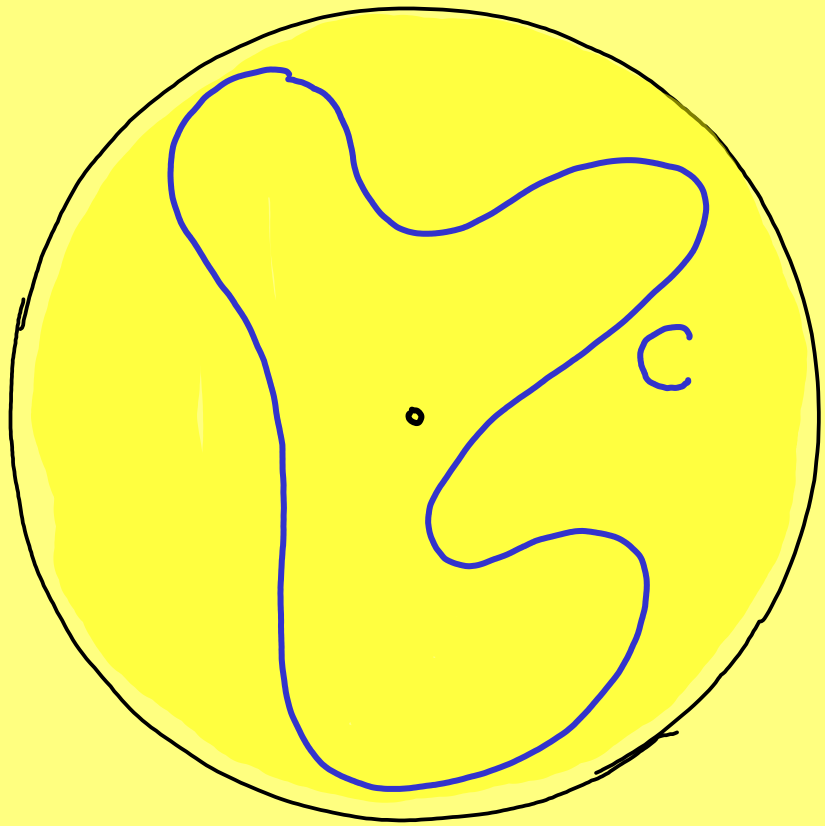


II.1 Fary's Theorem

C = smooth curve inside unit disk

$L(C)$ = length

$K(C)$ = total curvature

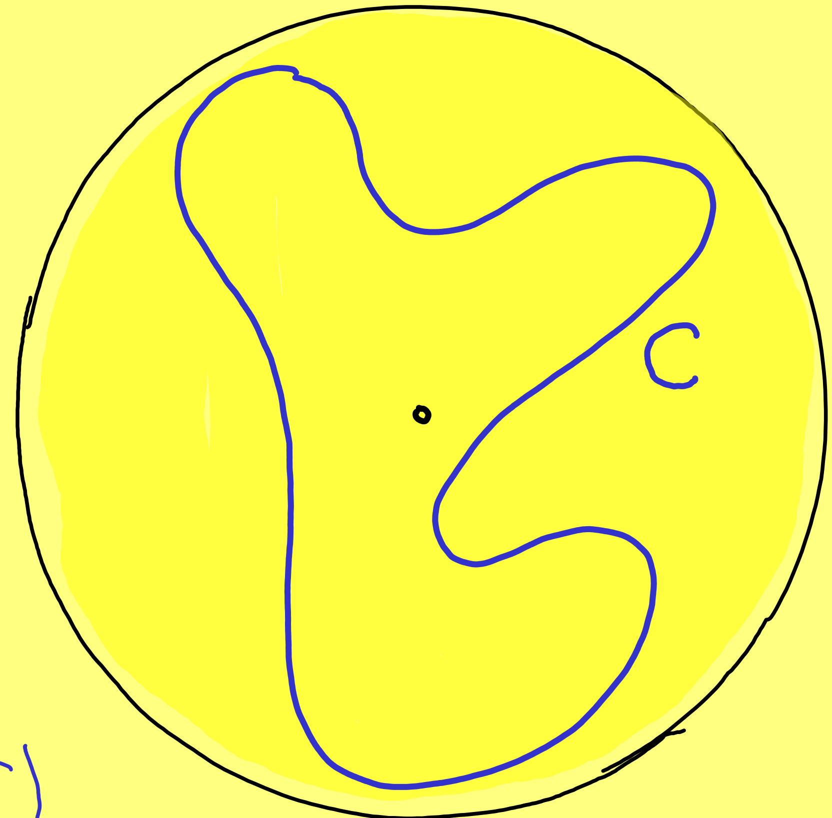


II.1 Fary's Theorem

C = smooth curve inside unit disk

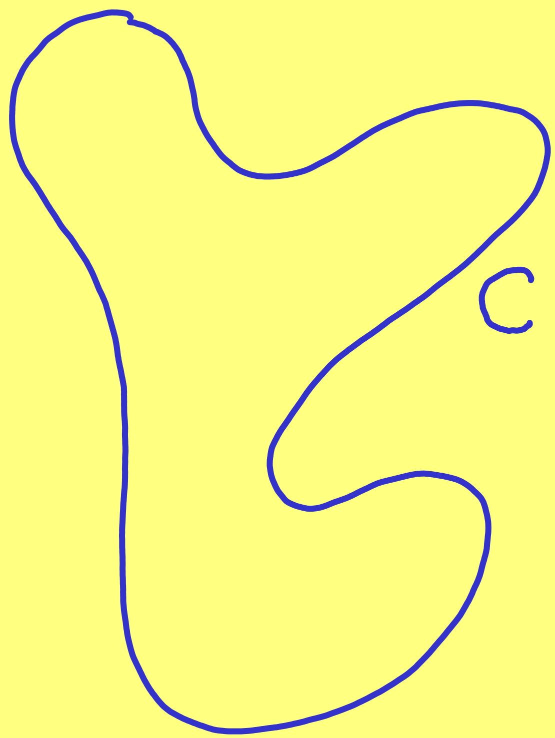
$L(C)$ = length

$K(C)$ = total curvature



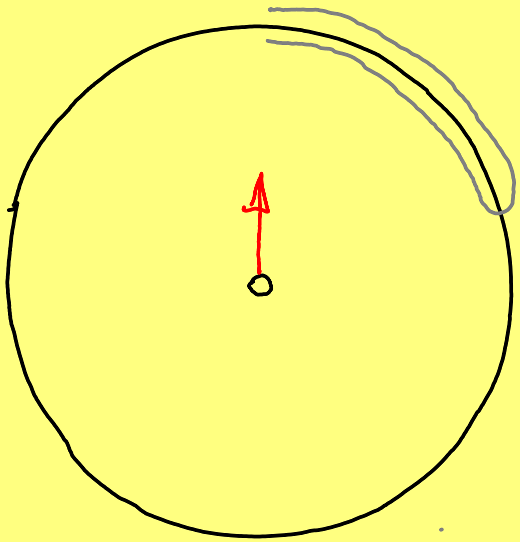
Fary: $L(C) \leq K(C)$.

II.2 Total Curvature

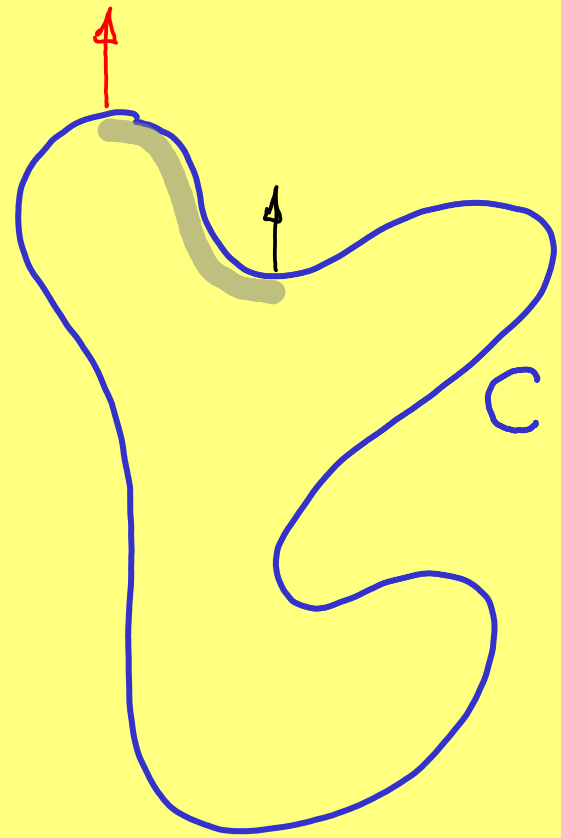


$$K(C) = \int_{x \in C} |\kappa(x)|\, dx$$

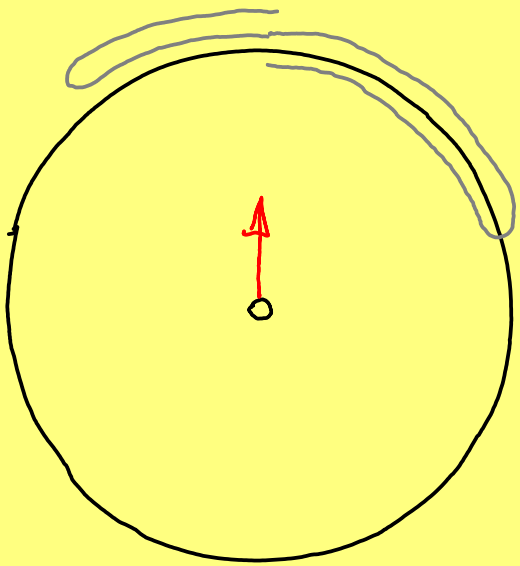
II.2 Total Curvature



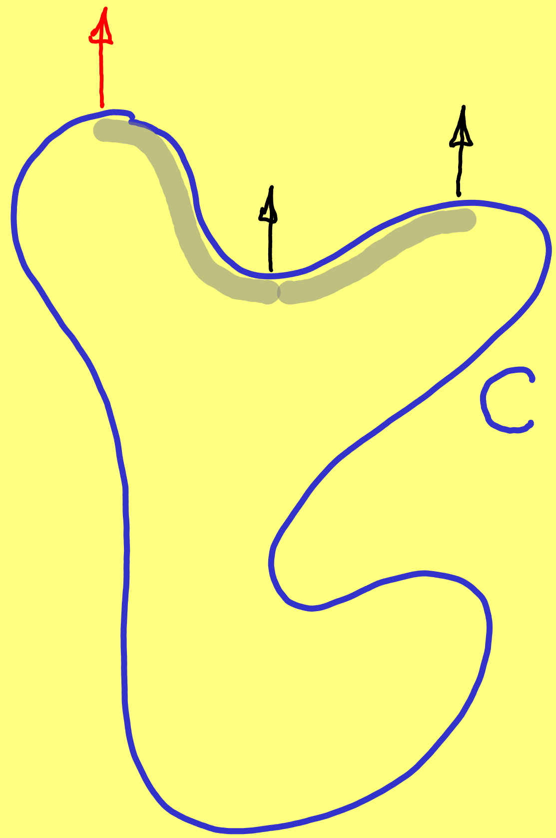
$$K(C) = \int_{x \in C} |\kappa(x)|\, dx$$



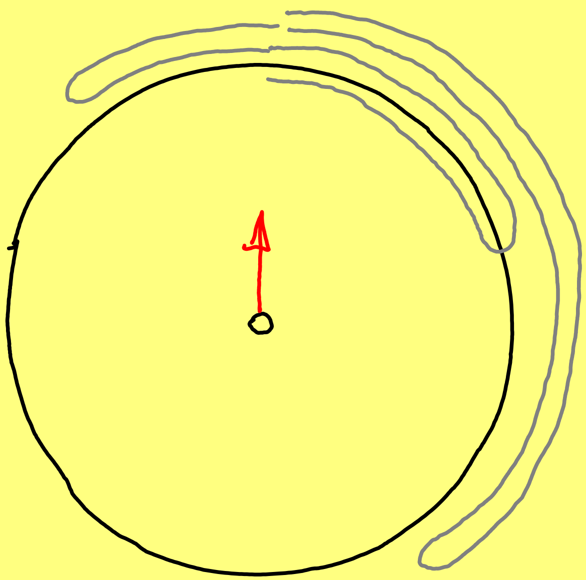
II.2 Total Curvature



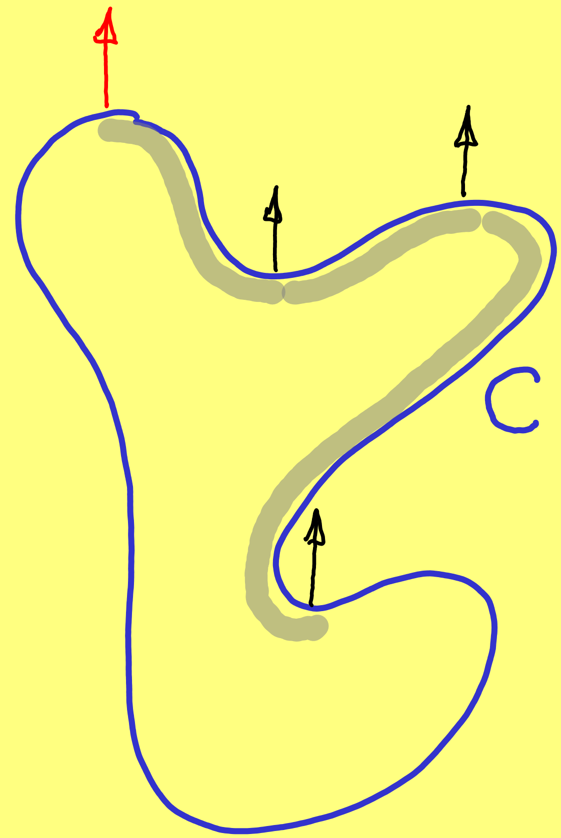
$$K(C) = \int_{x \in C} |\kappa(x)| \, dx$$



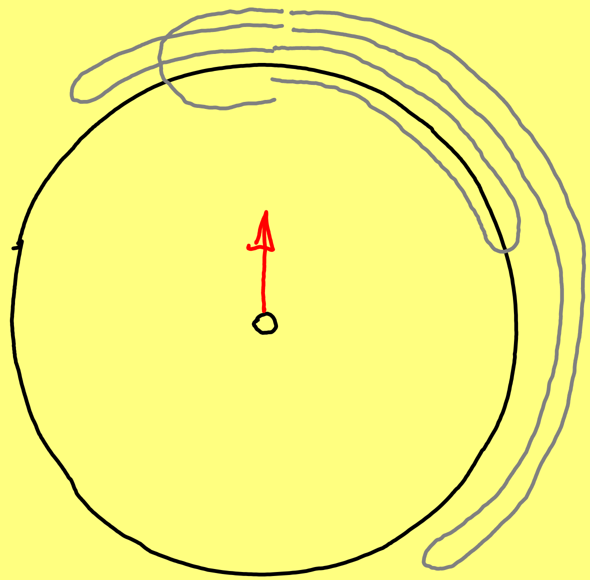
II.2 Total Curvature



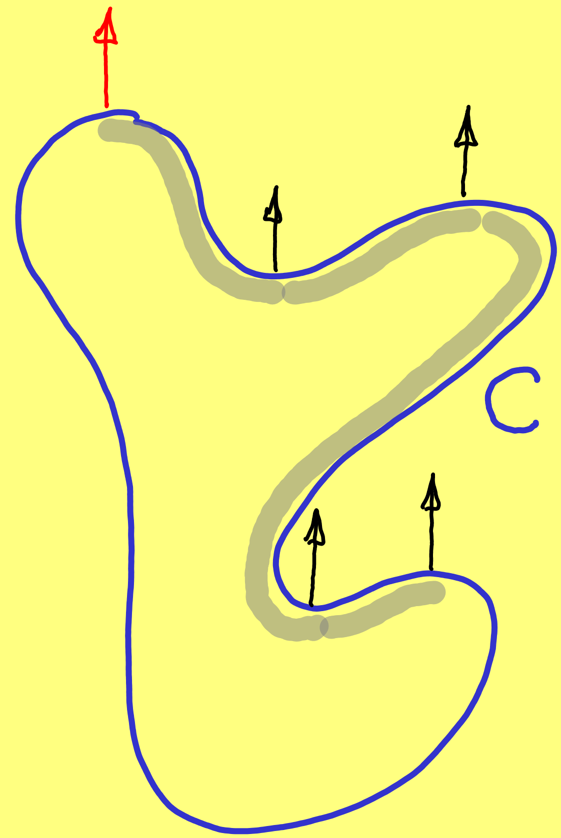
$$K(C) = \int_{x \in C} |\kappa(x)| \, dx$$



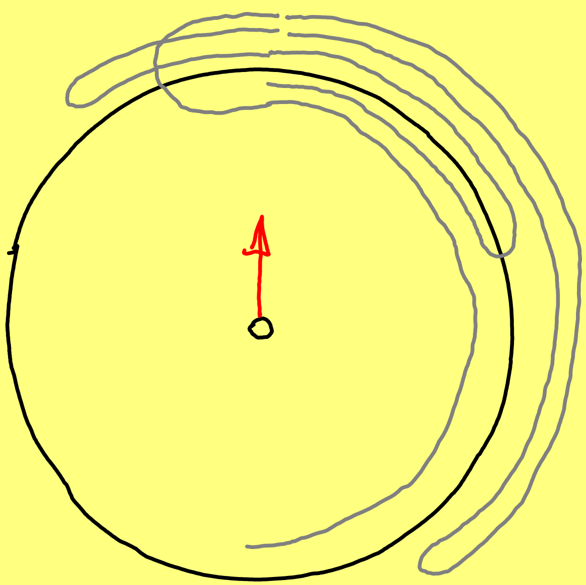
II.2 Total Curvature



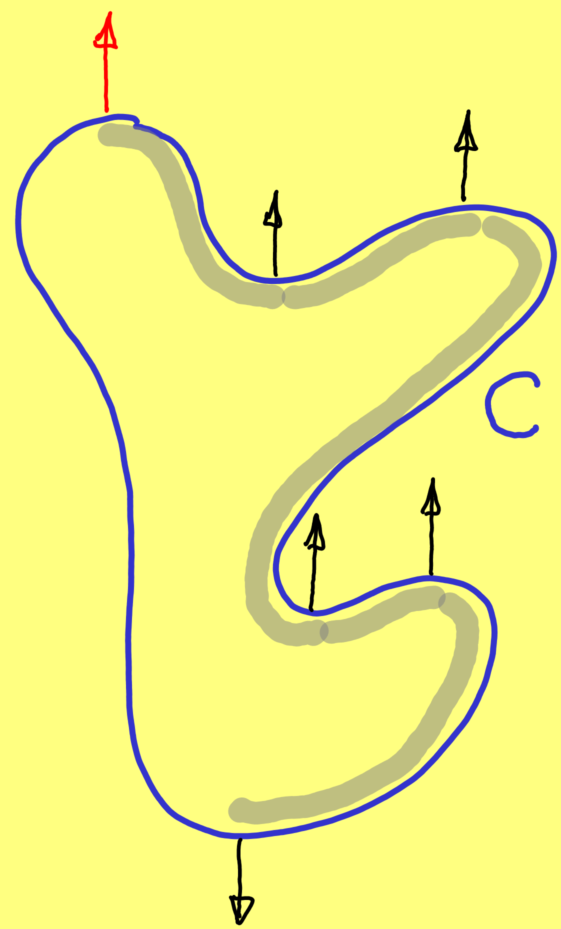
$$K(C) = \int_{x \in C} |\kappa(x)|\, dx$$



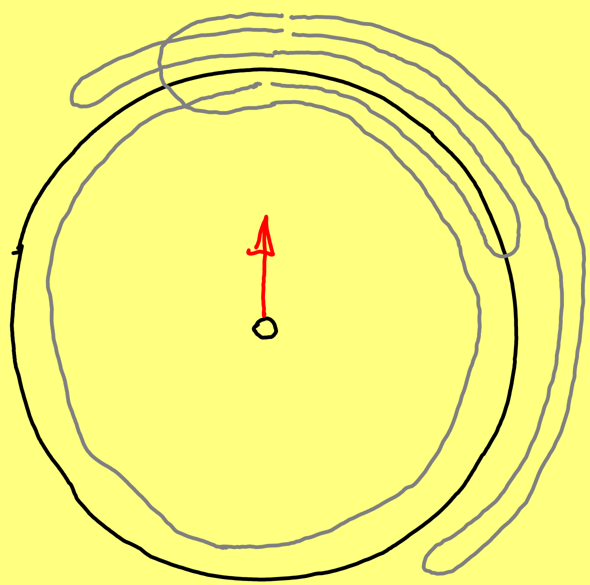
II.2 Total Curvature

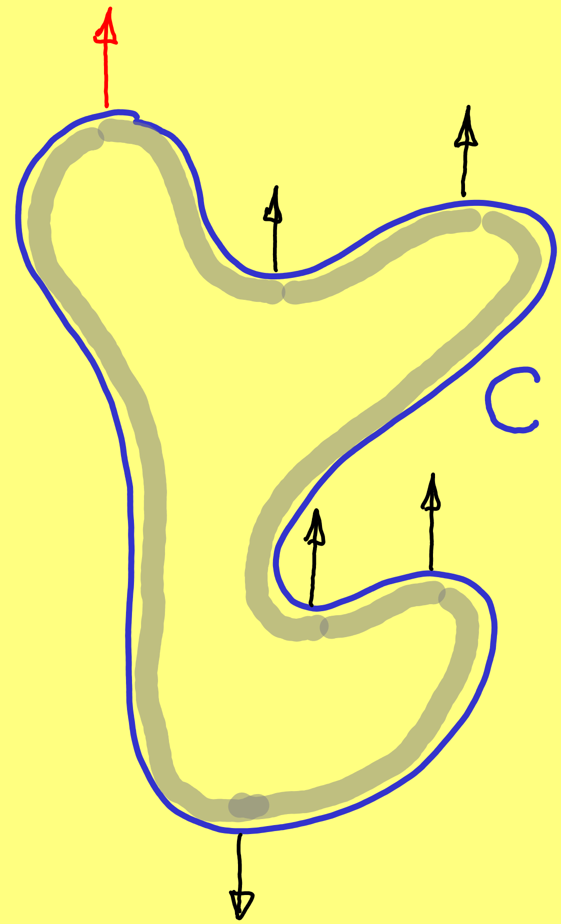


$$K(C) = \int_{x \in C} |\kappa(x)| \, dx$$



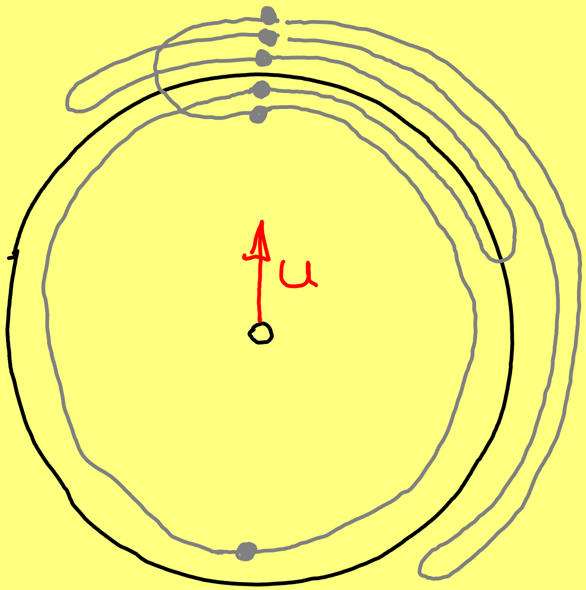
II.2 Total Curvature

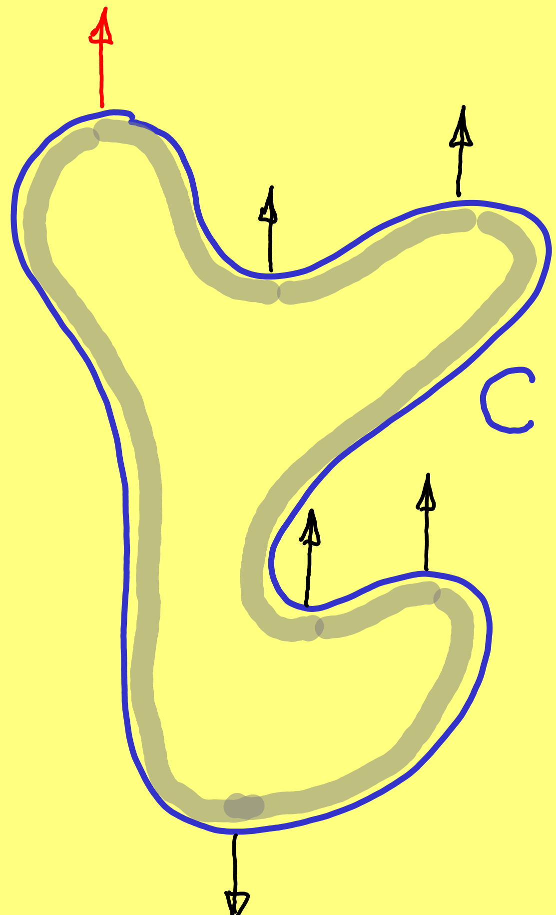




$$K(C) = \int_{x \in C} |\kappa(x)| \, dx$$

II.2 Total Curvature



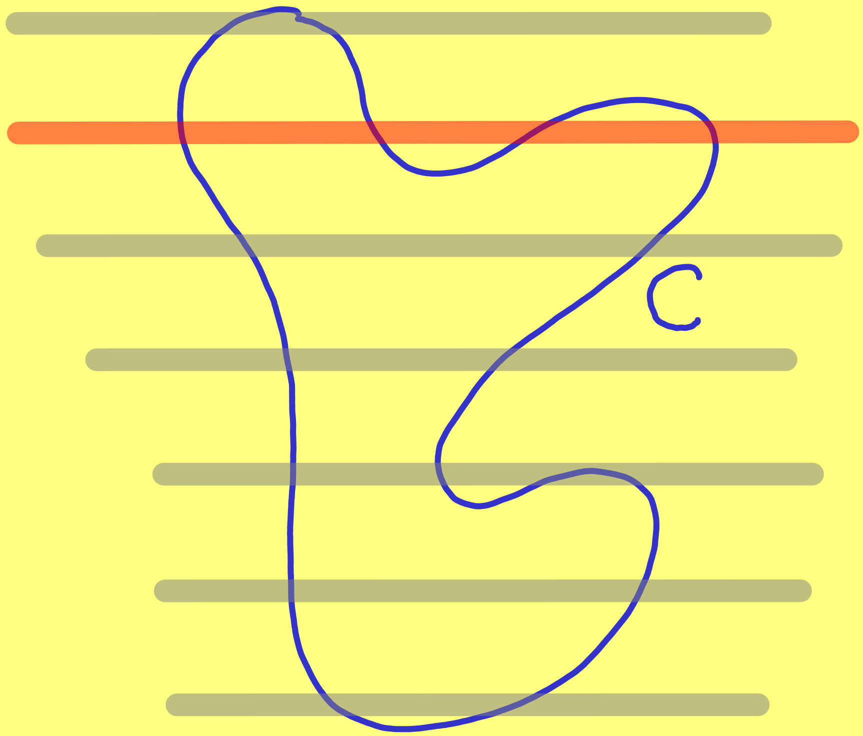


$$K(C) = \int_{x \in C} |\kappa(x)|\, dx$$

$$= \frac{1}{2} \int_{u \in S^1} (\#\,\text{crit. pts. of } f_u)\, du$$

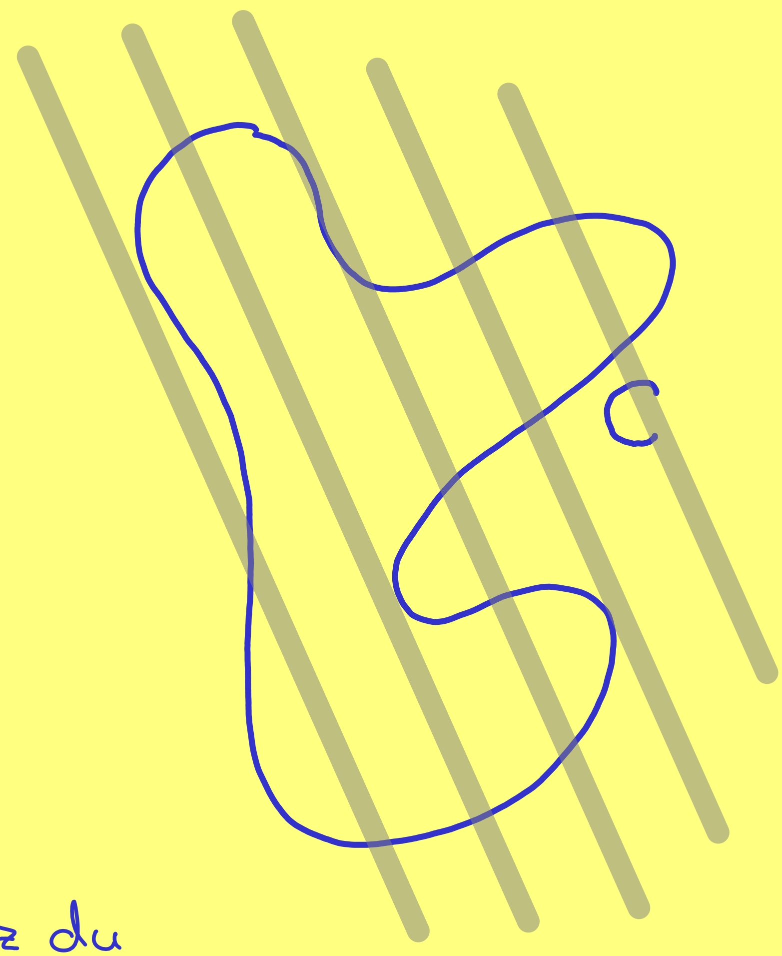
$$= \int_{u \in S^1} |\mathrm{Dgm}(f_u)|\, du$$

II.3 LENGTH



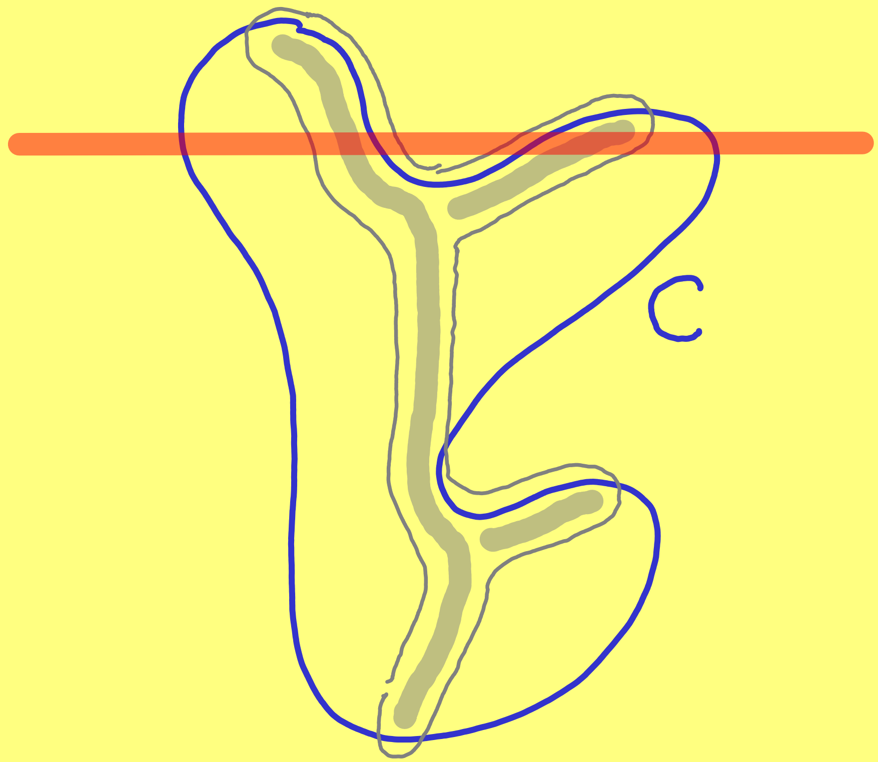
$$L(C) = \frac{1}{2} \int_u \int_z \#\mathrm{Xing}(z)\, dz\, du$$

II.3 LENGTH



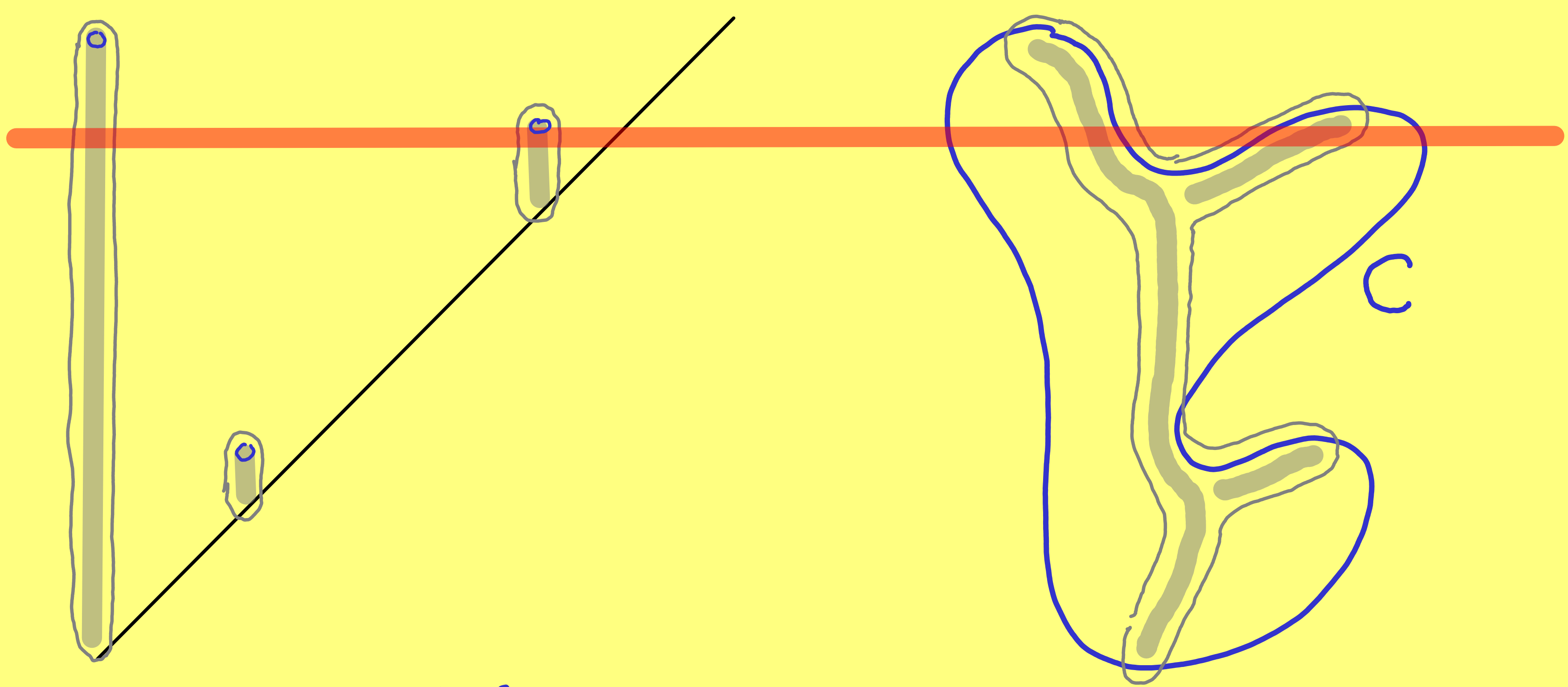
$$L(C) = \frac{1}{2} \int_u \int_z \#\mathrm{Xing}(z)\, dz\, du$$

II.3 LENGTH



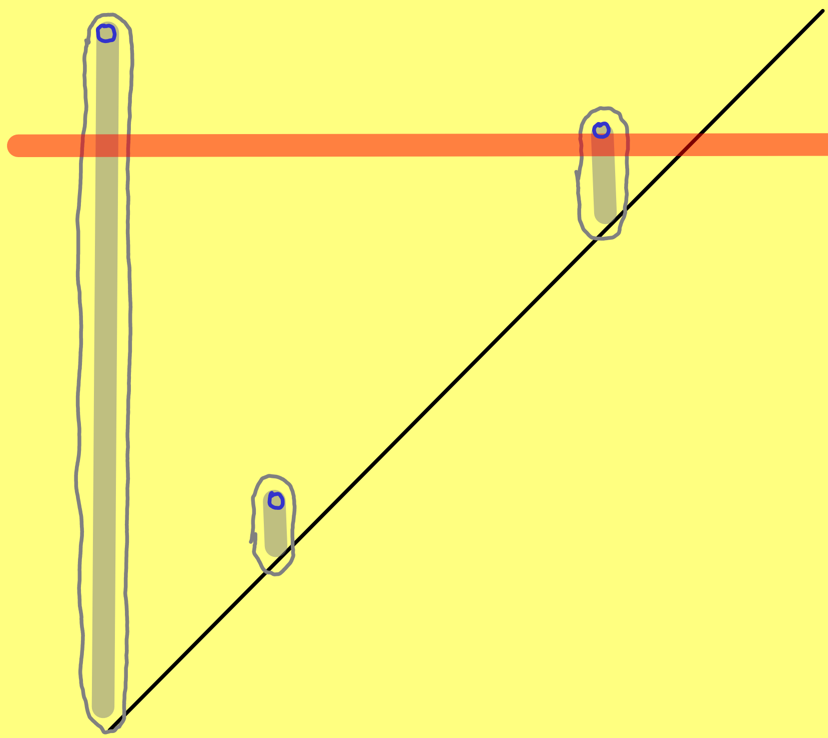
$$L(C) = \frac{1}{2} \int_u \int_z \#\mathrm{Xing}(z)\, dz\, du$$

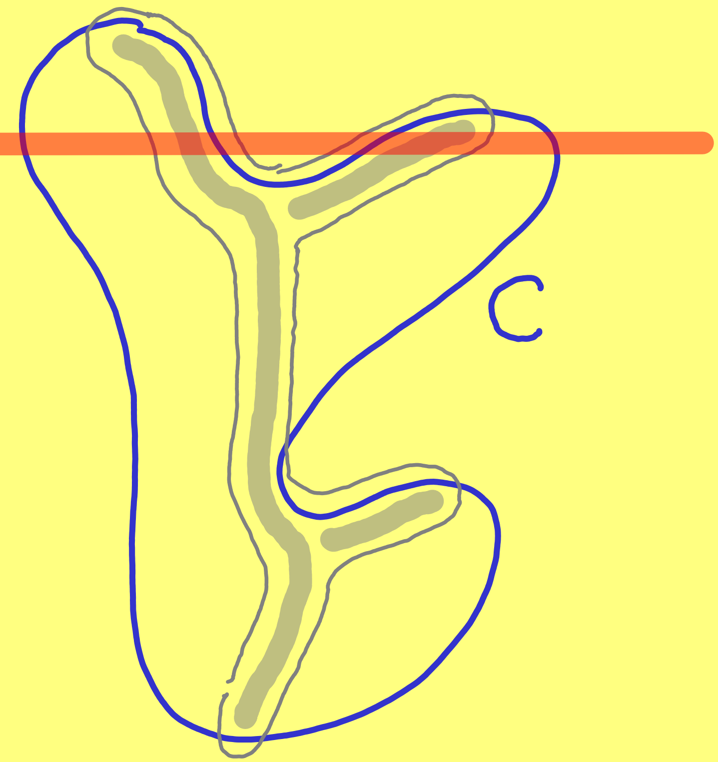
II.3 LENGTH



$$L(C) = \frac{1}{2} \int_u \int_z \#\mathrm{Xing}(z)\, dz\, du$$

II.3 Length

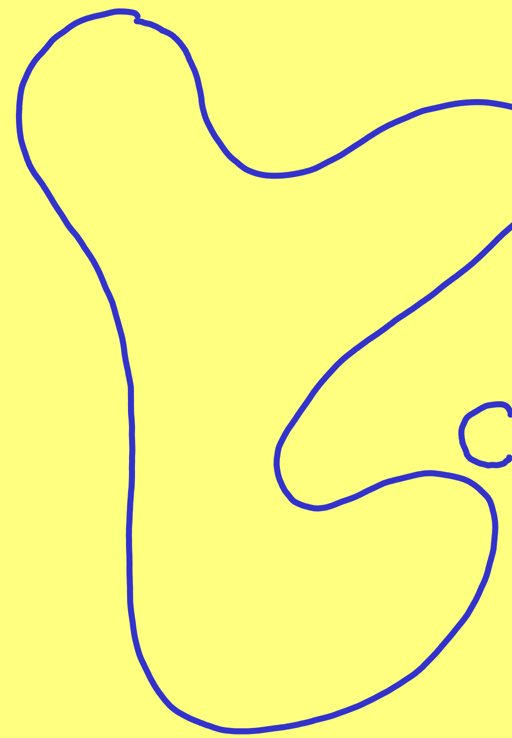




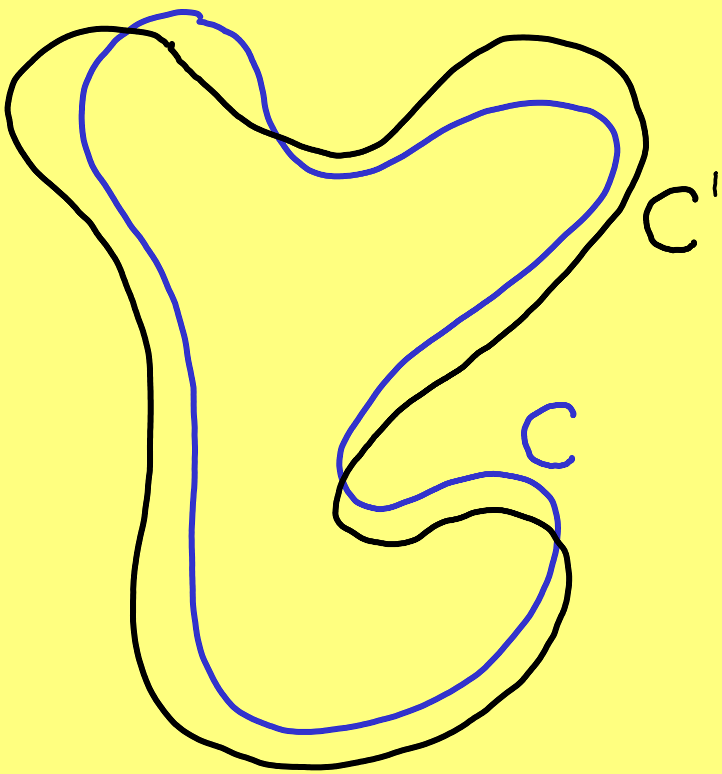
$$L(C) = \frac{1}{2} \int_u \int_z \# \mathrm{Xing}(z)\, dz\, du$$

$$= \int_u \sum_{x \in \mathrm{Dgm}(f_u)} \mathrm{pers}(x)\, du$$

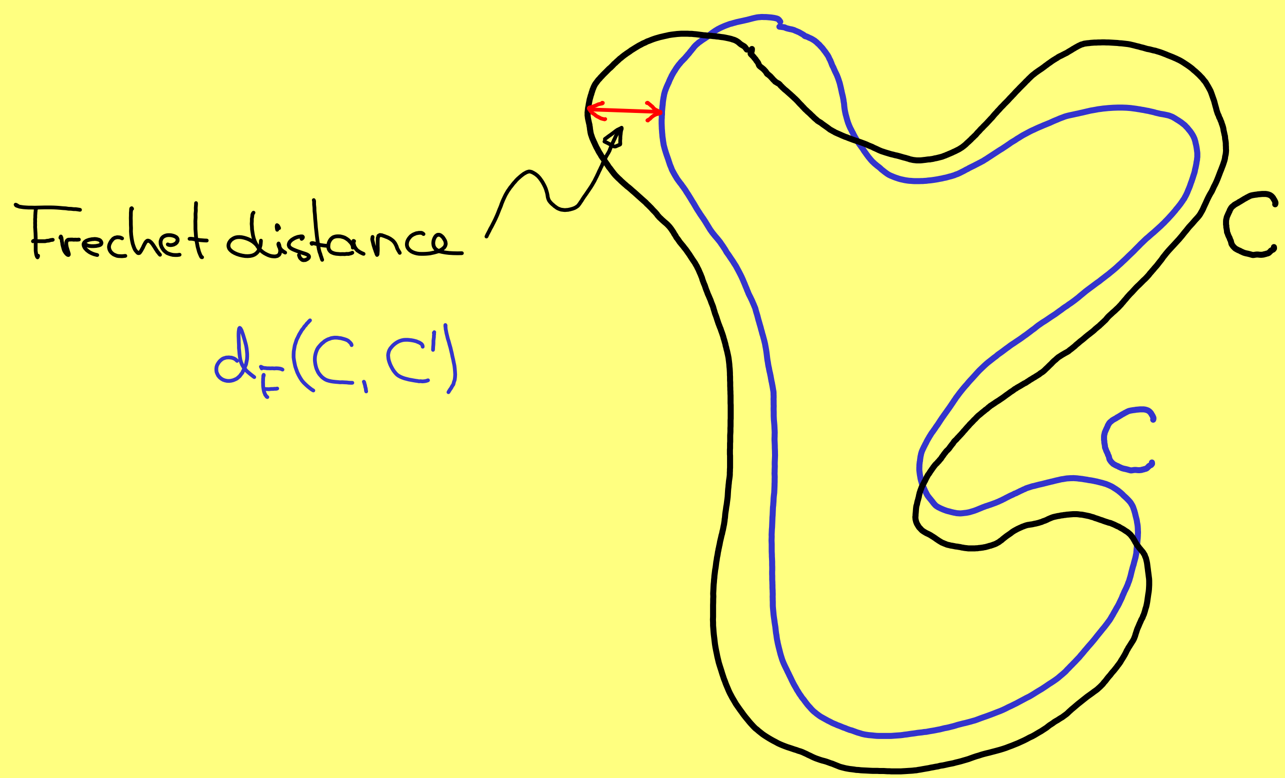
II.4 Generalized Fary Theorem



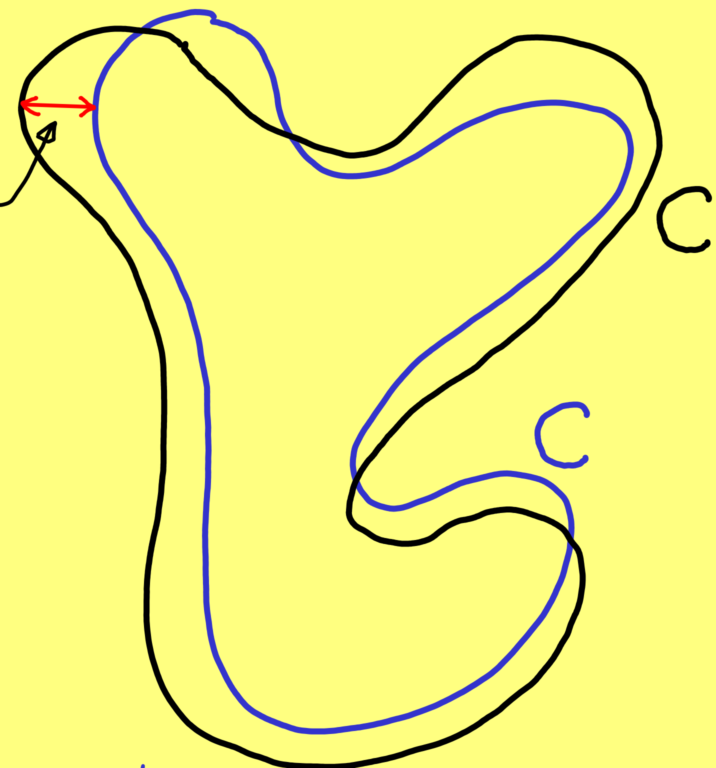
II.4 Generalized Fary Theorem



II.4 Generalized Fary Theorem



II.4 Generalized Fary Theorem

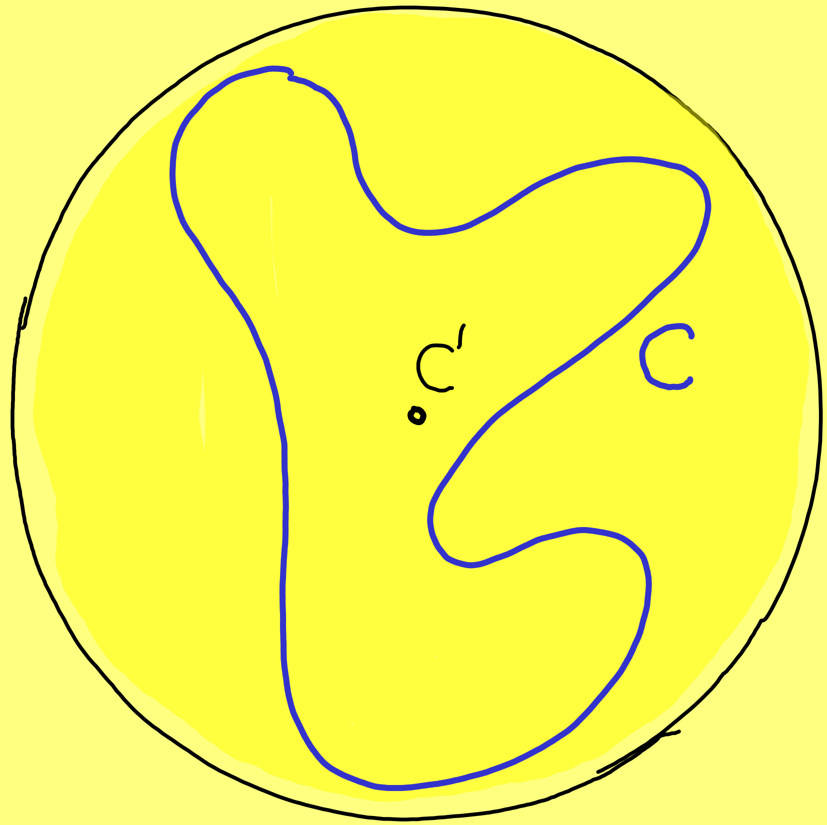


Frechet distance

$d_F(C, C')$

Thm. $|L(C) - L(C')| \leq (K(C) + K(C') - 2\pi)\, d_F(C, C').$

II.4 Generalized Fary Theorem



Thm. $|L(C) - L(C')| \leq (K(C) + K(C') - 2\pi)\, d_F(C, C')$.

III. Somites

work by Olivier Pourquié
Mary-Lee Dequéant

III.1 SEGMENTATION of vertebrate body plan

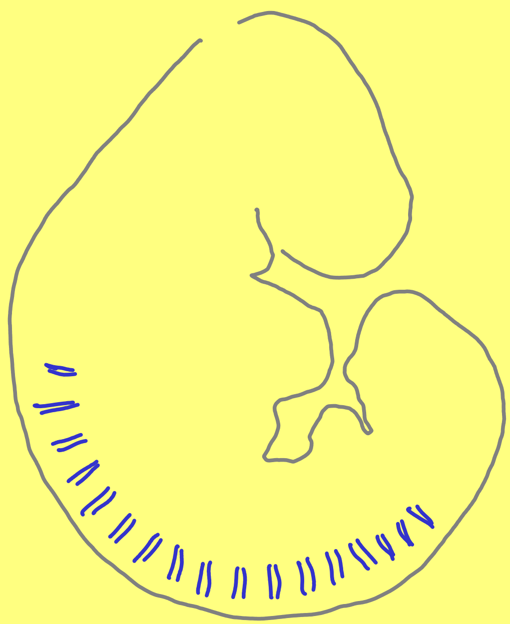
III.1 SEGMENTATION of vertebrate body plan

adult mouse

III.1 SEGMENTATION of vertebrate body plan

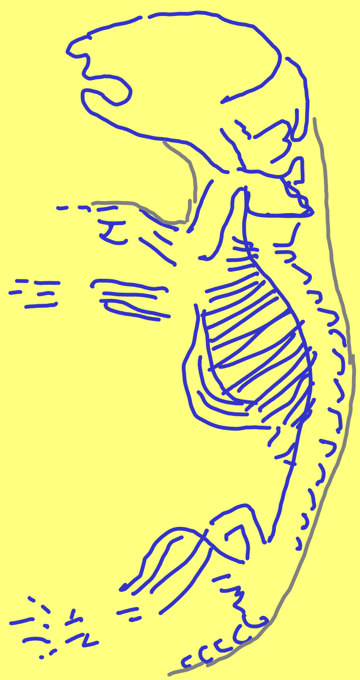
adult mouse

mouse embryo

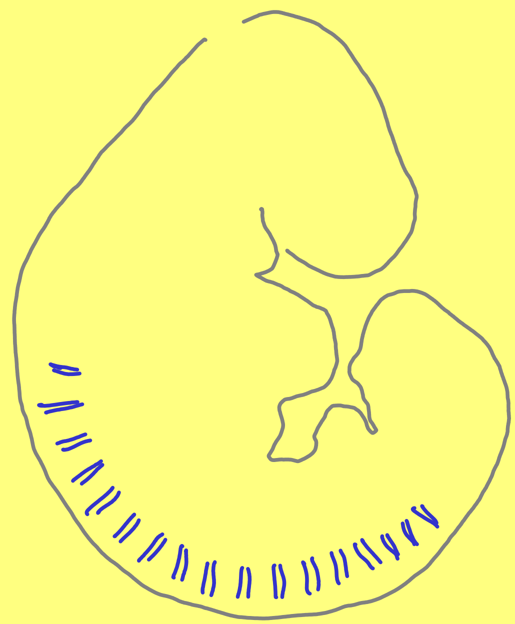


III.1 SEGMENTATION of vertebrate body plan

adult mouse

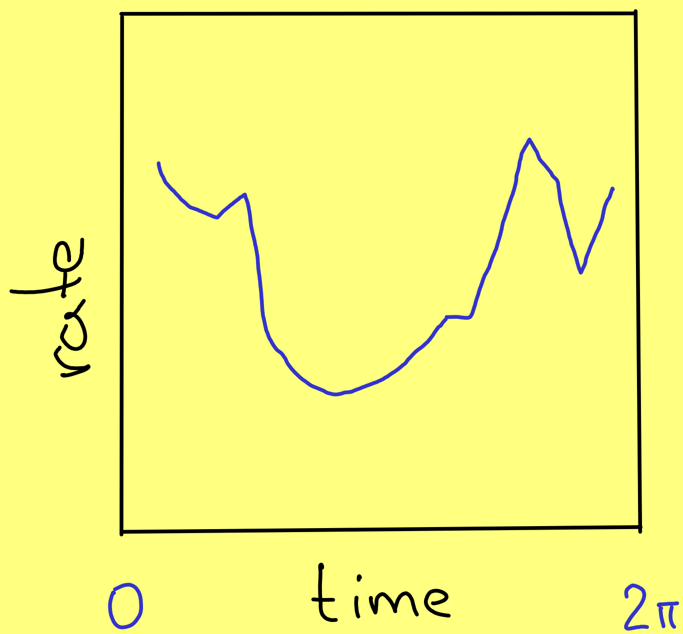


mouse embryo

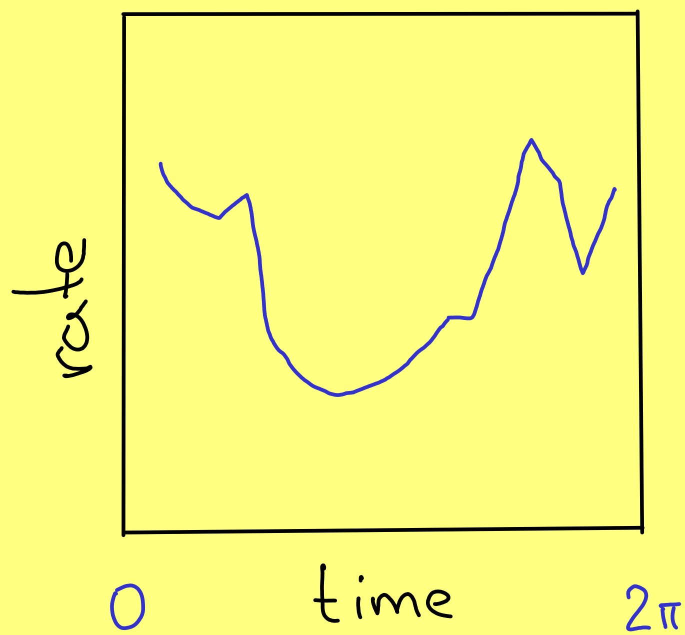


somite development is a rythmic process

III.2 Gene Expression Data



III.2 Gene Expression Data



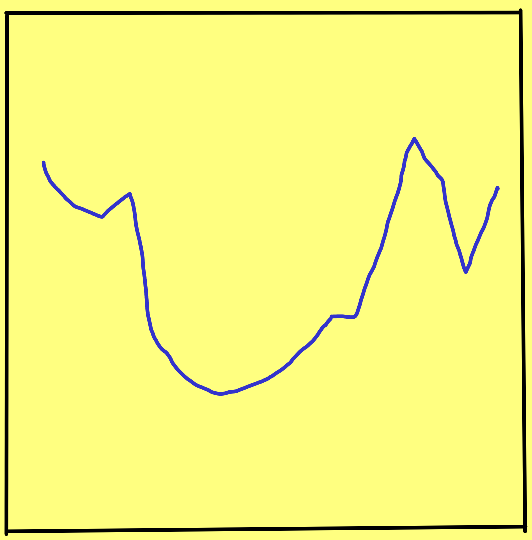
normalized to $\mathrm{amp}(f) = \max_x f(x) - \min_x f(x) = 1$

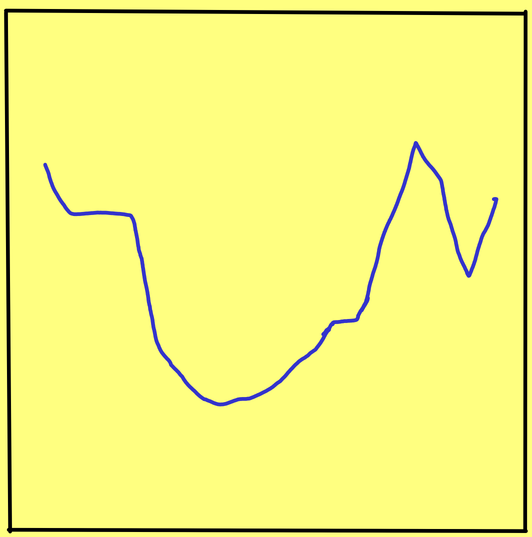
III.3 SIMPLIFICATION

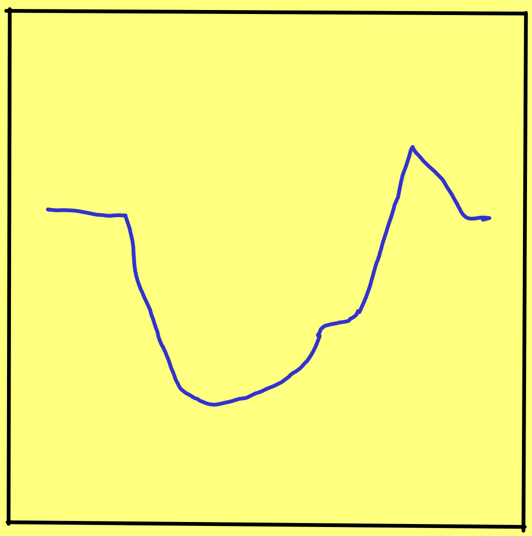
DEF. An ε-simplification of f is a function $f_\varepsilon : \mathbb{X} \to \mathbb{R}$ with $\|f - f_\varepsilon\|_\infty \le \varepsilon$ whose diagram $\mathrm{Dgm}(f_\varepsilon)$ is same as $\mathrm{Dgm}(f)$ without points of persistence at most ε.

III.3 Simplification

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

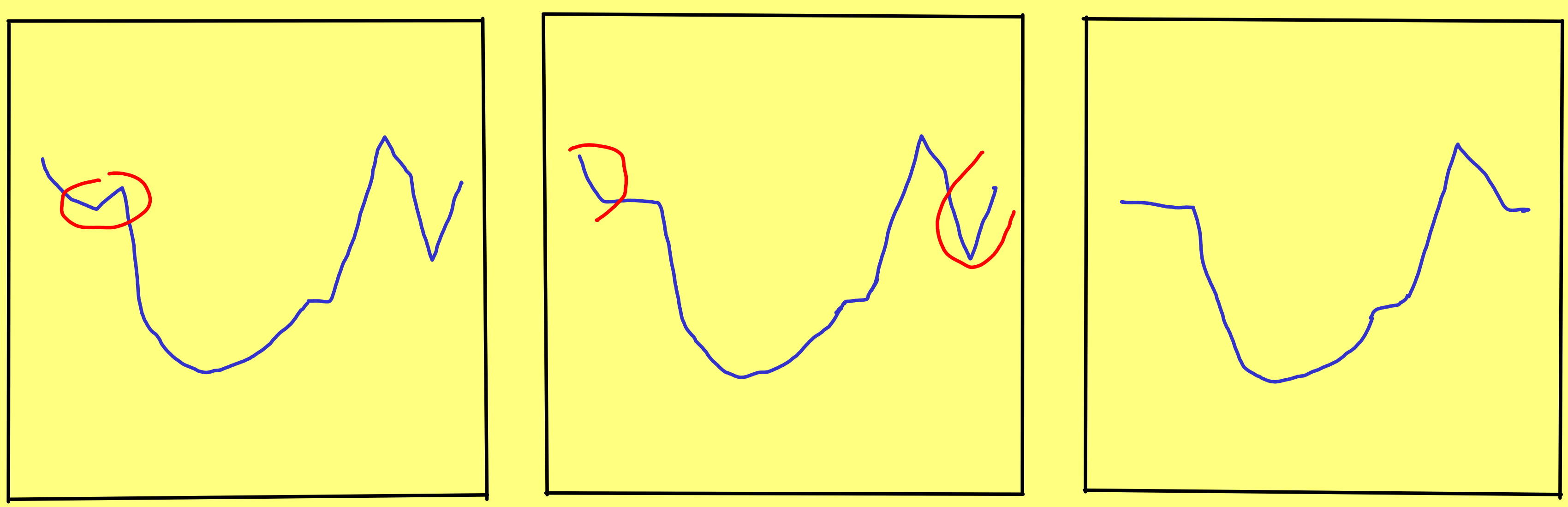




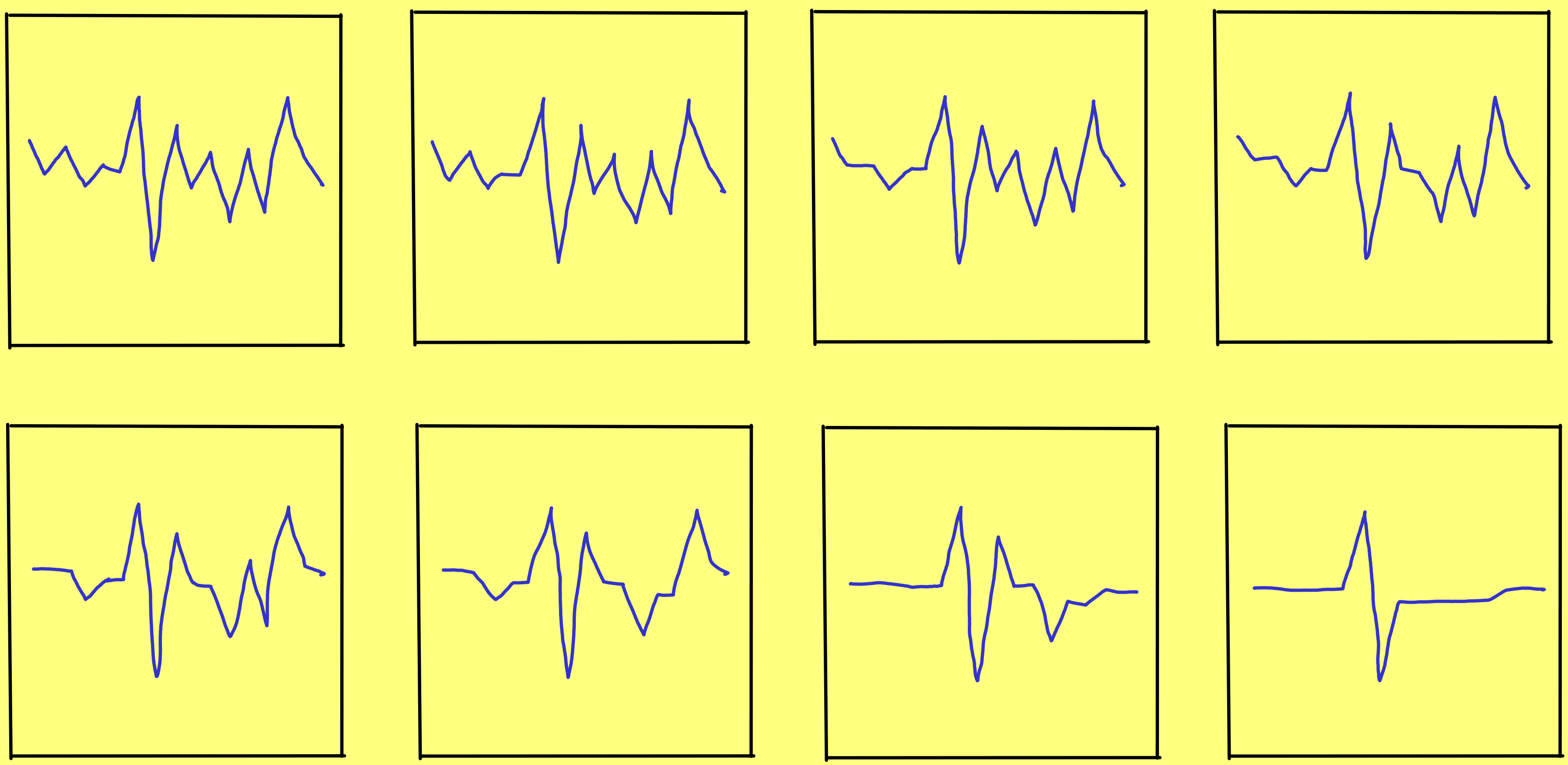


III.3 SIMPLIFICATION

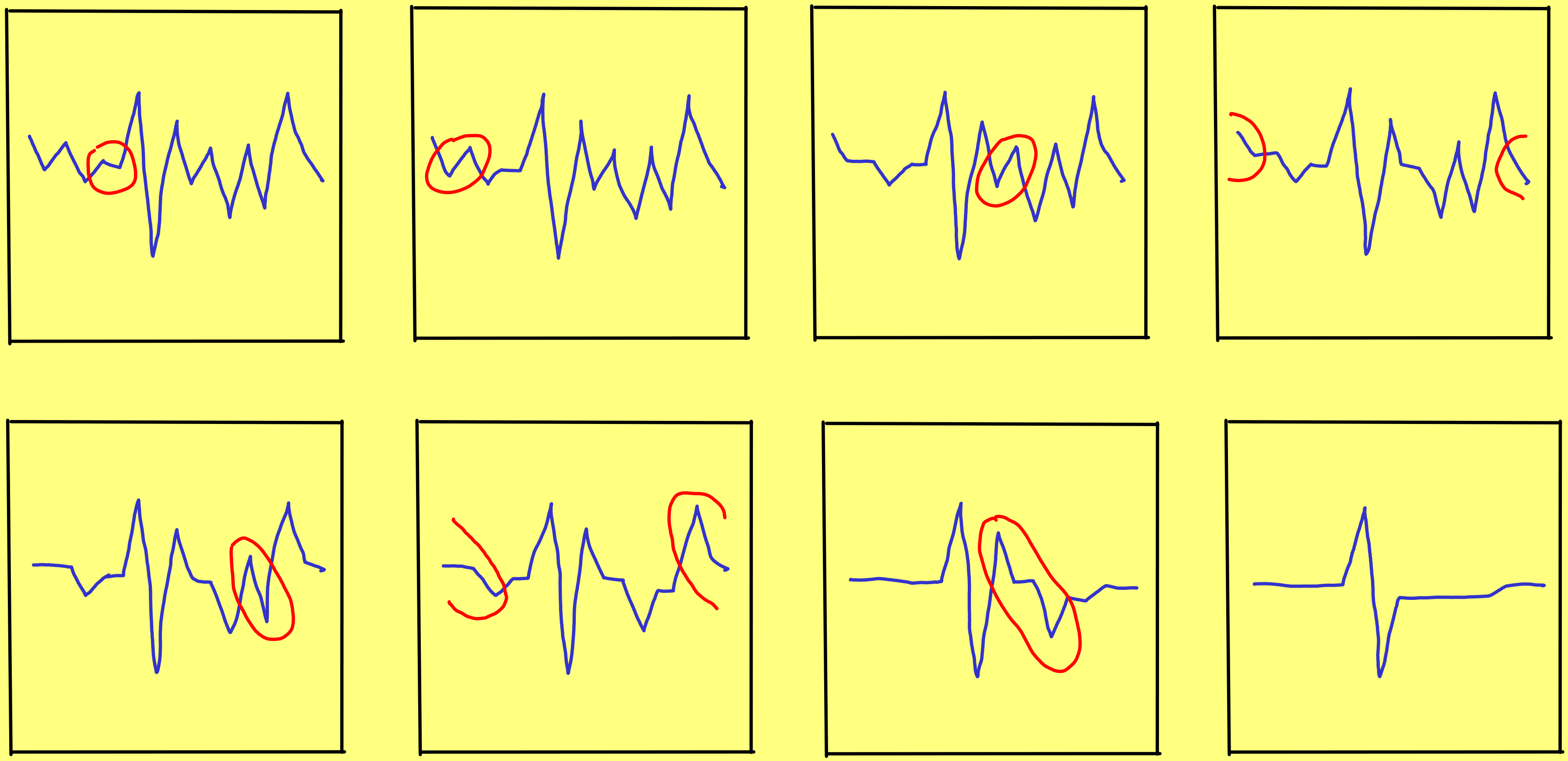
DEF. An ε-simplification of f is a function $f_\varepsilon : \mathbb{X} \to \mathbb{R}$ with $\|f - f_\varepsilon\|_\infty \leq \varepsilon$ whose diagram $\mathrm{Dgm}(f_\varepsilon)$ is same as $\mathrm{Dgm}(f)$ without points of persistence at most ε.



III.3 SIMPLIFICATION



III.3 SIMPLIFICATION



III.4 Measures

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

$M_0(f) = \frac{1}{2}$ # critical points

III.4 Measures

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

$$M_0(f) = \tfrac{1}{2} \#\text{critical points}$$

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

III.4 MEASURES

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

$$M_0(f) = \frac{1}{2} \,\#\text{critical points}$$

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

$$M_2(f) = \int_{\varepsilon=0}^{1} M_1(f_\varepsilon)\, d\varepsilon = \sum_x \operatorname{pers}(x)^2$$

etc.

I.3 L_p - STABILITY

THM. Let $\mathbb{X}$ be a compact metric space and $f, g : \mathbb{X} \to \mathbb{R}$ two Lipschitz functions. Then there exist constants k and C that depend on $\mathbb{X}$ and the Lipschitz constants of f and g s.t.

$$W_p(f, g) \leq C \|f - g\|_\infty^{p-k}$$

for every $p \geq k$.

$$\boxed{\mathbb{X} = S^1 \Longrightarrow k = 1}$$

I.3 L_p-STABILITY

THM. Let $\mathbb{X}$ be a compact metric space and $f, g : \mathbb{X} \to \mathbb{R}$ two Lipschitz functions. Then there exist constants k and C that depend on $\mathbb{X}$ and the Lipschitz constants of f and g s.t.

$$W_p(f, g) \leq C \|f - g\|_\infty^{p-k}$$

for every $p \geq k$.

III.4 Measures

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

not stable

$$M_0(f) = \frac{1}{2} \# \text{critical points}$$

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

stable

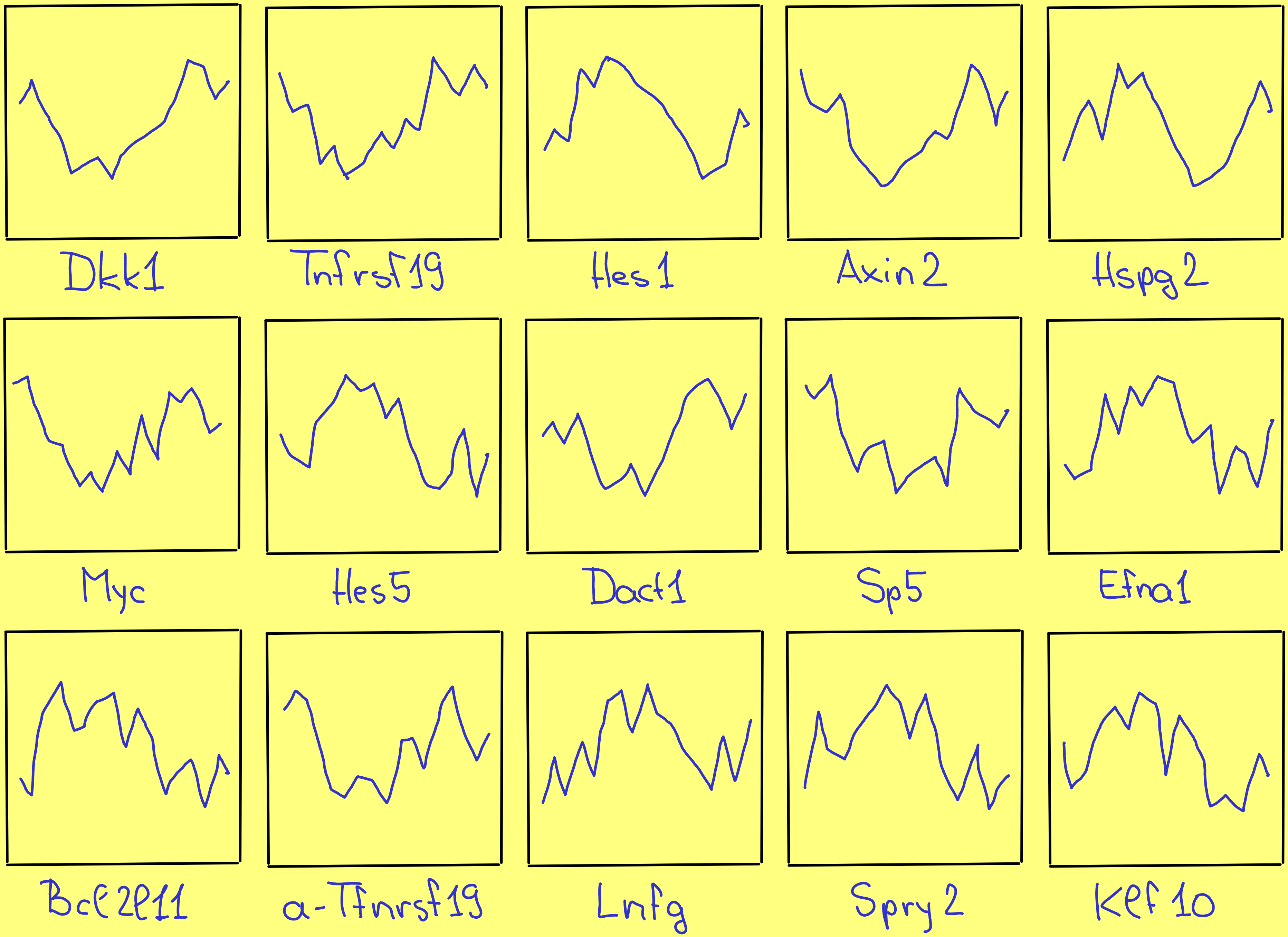
$$M_2(f) = \int_{\varepsilon=0}^{1} M_1(f_\varepsilon)\, d\varepsilon = \sum_x \text{pers}(x)^2$$

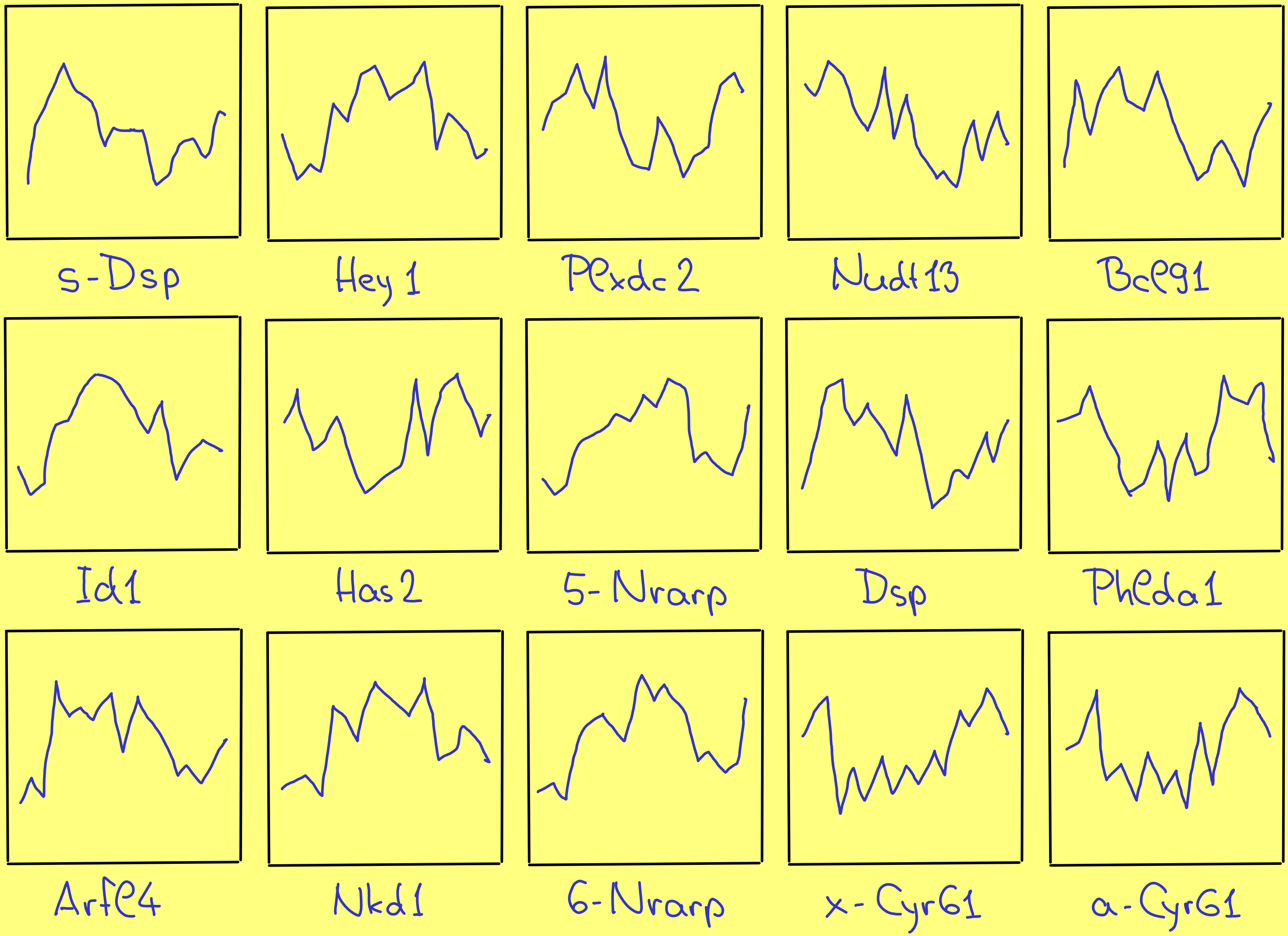
etc.

III.5 Results

using 30 genes biologically confirmed
to participate in somitogenesis

Dkk1	Tnfrsf19	Hes1	Axin2	Hspg2
Myc	Hes5	Dact1	Sp5	Efna1
Bcl2l11	a-Tnfrsf19	Lfng	Spry2	Klf10
s-Dsp	Hey1	Plxdc2	Nudt13	Bcl9l
Id1	Has2	5-Nrarp	Dsp	Phlda1
Arfl4	Nkd1	6-Nrarp	x-Cyr61	a-Cyr61


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



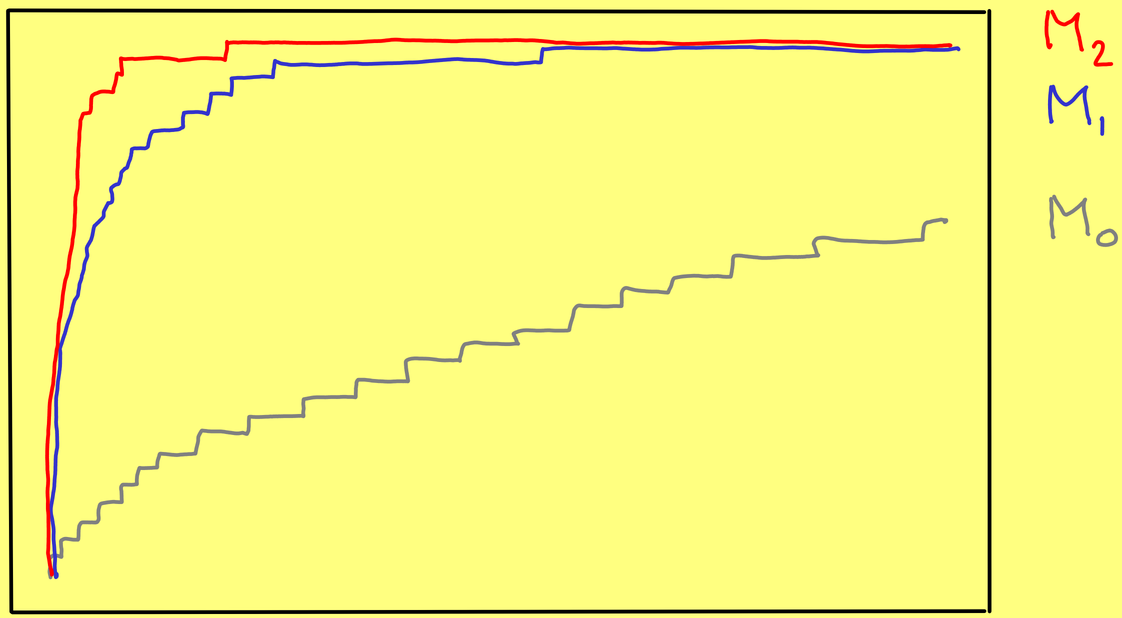
s-Dsp
Hey 1
Plxdc 2
Nudt13
Bcl9l
Id1
Has 2
5-Nrarp
Dsp
Phlda1
Arfl4
Nkd1
6-Nrarp
x-Cyr61
a-Cyr61

III.5 Results

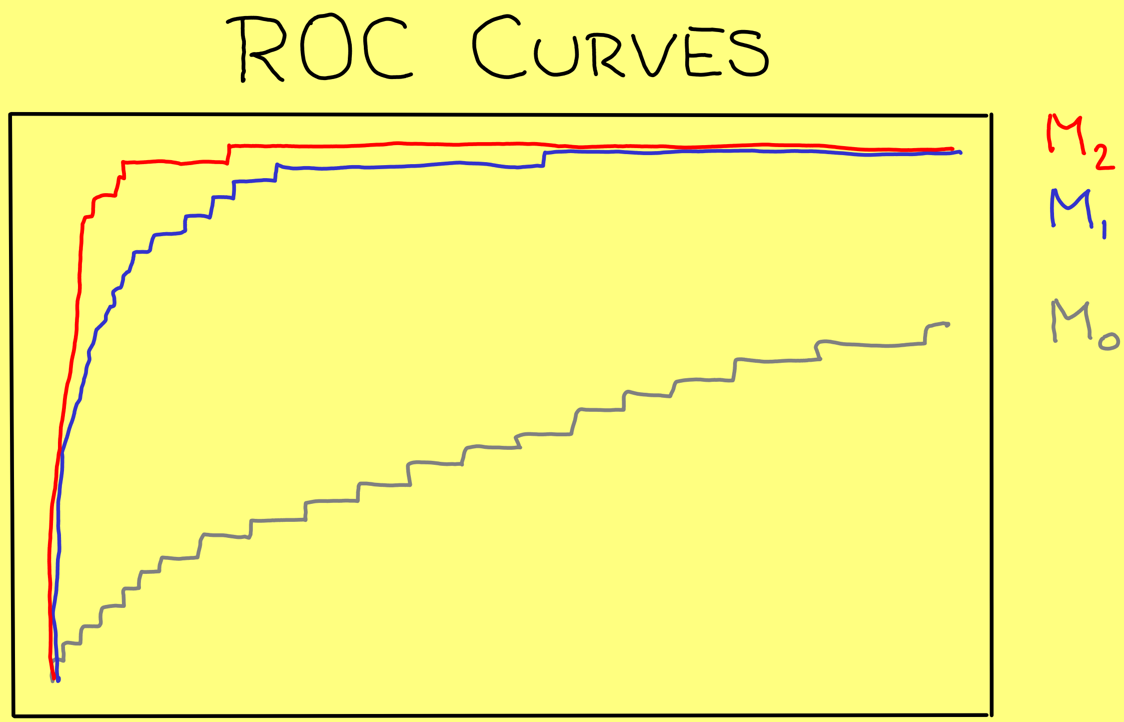
	M_1	M_2	M_3	M_4	avg
Dkk1	1	1	1	1	1
Tnfrsf19	12	2	2	2	5
Hey1	23	12	9	9	13
Nudt13	882	170	104	83	310
Arfl4	302	86	74	68	133
...	...	...	...	...	...
avg	250	109	104	117	145

III.5 Results

ROC Curves



III.5 Results



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

III.5 Results

Promising Candidates

	M_1	M_2	M_3	M_4	L
Tnfrsf9	5	4	4	6	8
Ptpn11	8	6	6	7	22
Mtm1	9	7	5	3	15
Stom	7	8	12	16	299
Star	6	9	11	13	39
...	...	...	...	...	...

IV. MOTION

IV.1 Homotopy

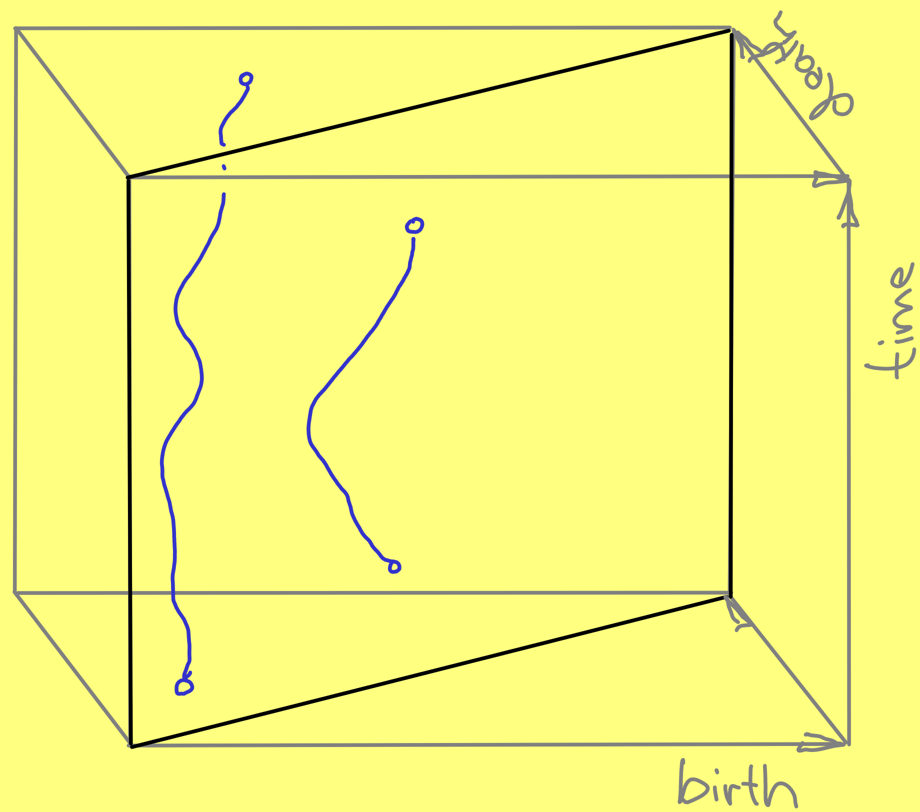
$F : \mathbb{M} \times [0,1] \to \mathbb{R}$,

$f_t(x) = F(x,t)$ leads to $\mathrm{Dgm}(f_t)$

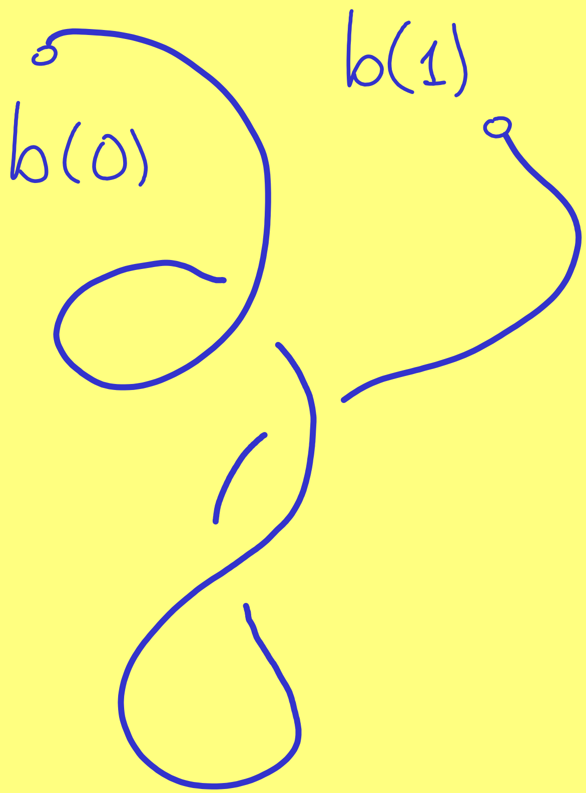
IV.1 HOMOTOPY

$F : \mathbb{M} \times [0,1] \to \mathbb{R}$,

$f_t(x) = F(x,t)$ leads to $\mathrm{Dgm}(f_t)$

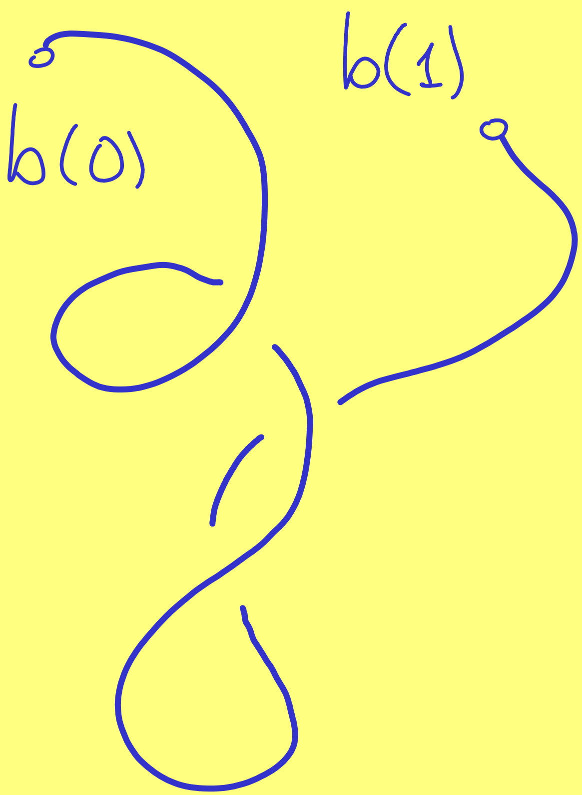


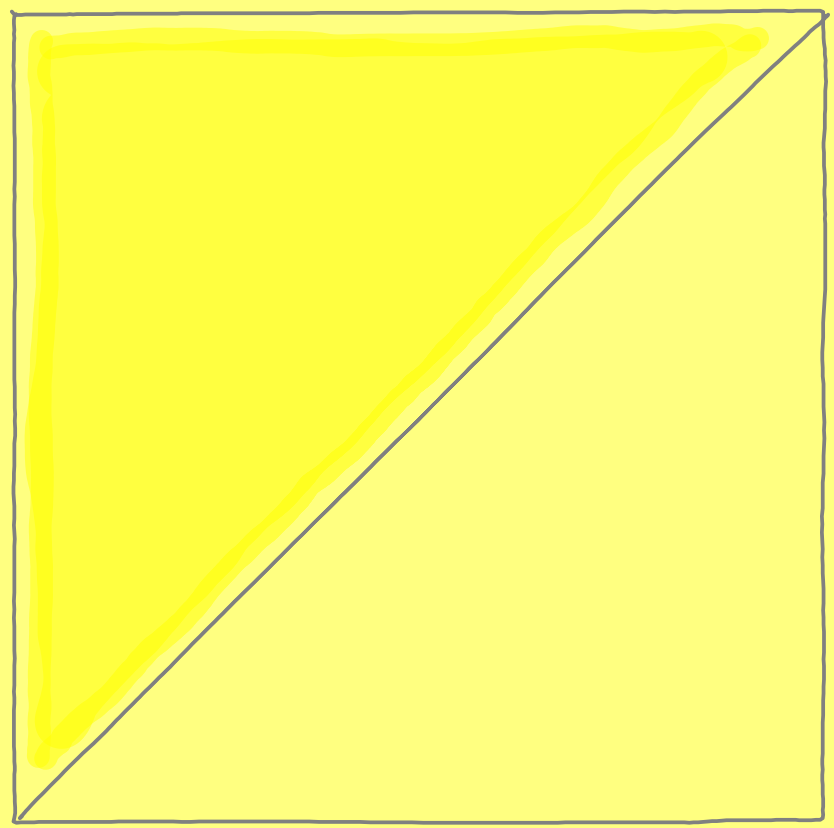
IV.2 Pairwise Distance Function



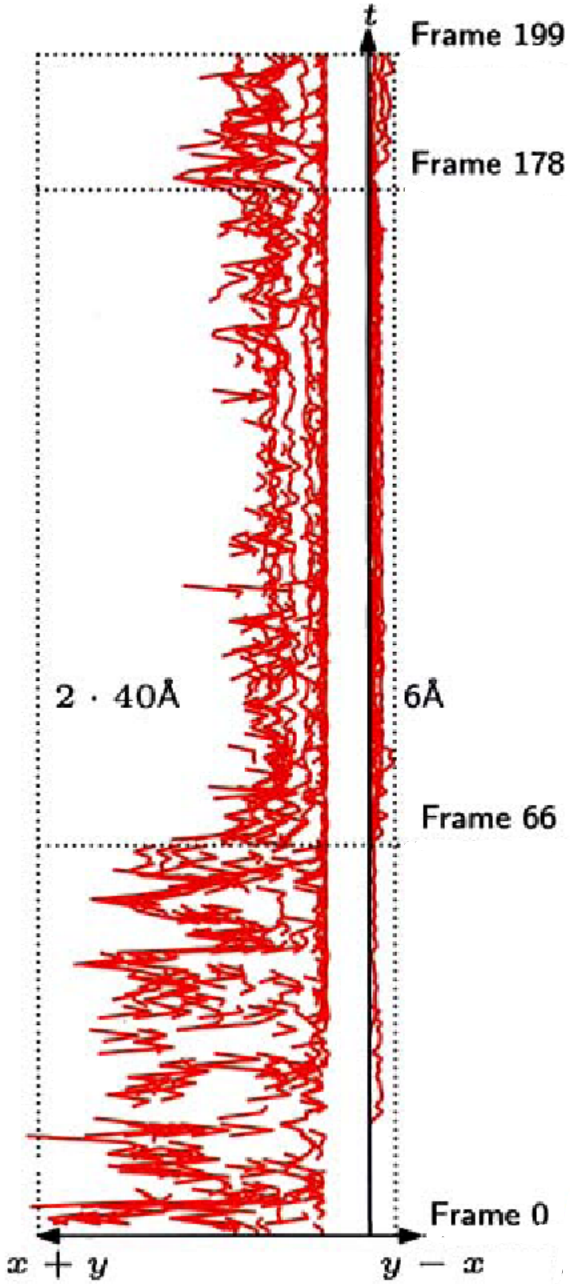
IV.2 Pairwise Distance Function

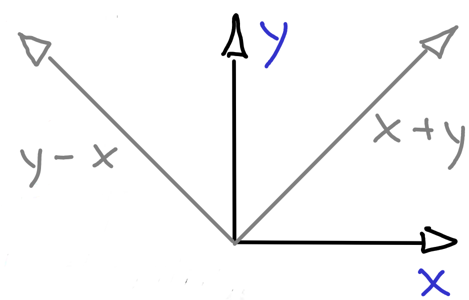
$$f(x,y) = \| b(x) - b(y) \|$$



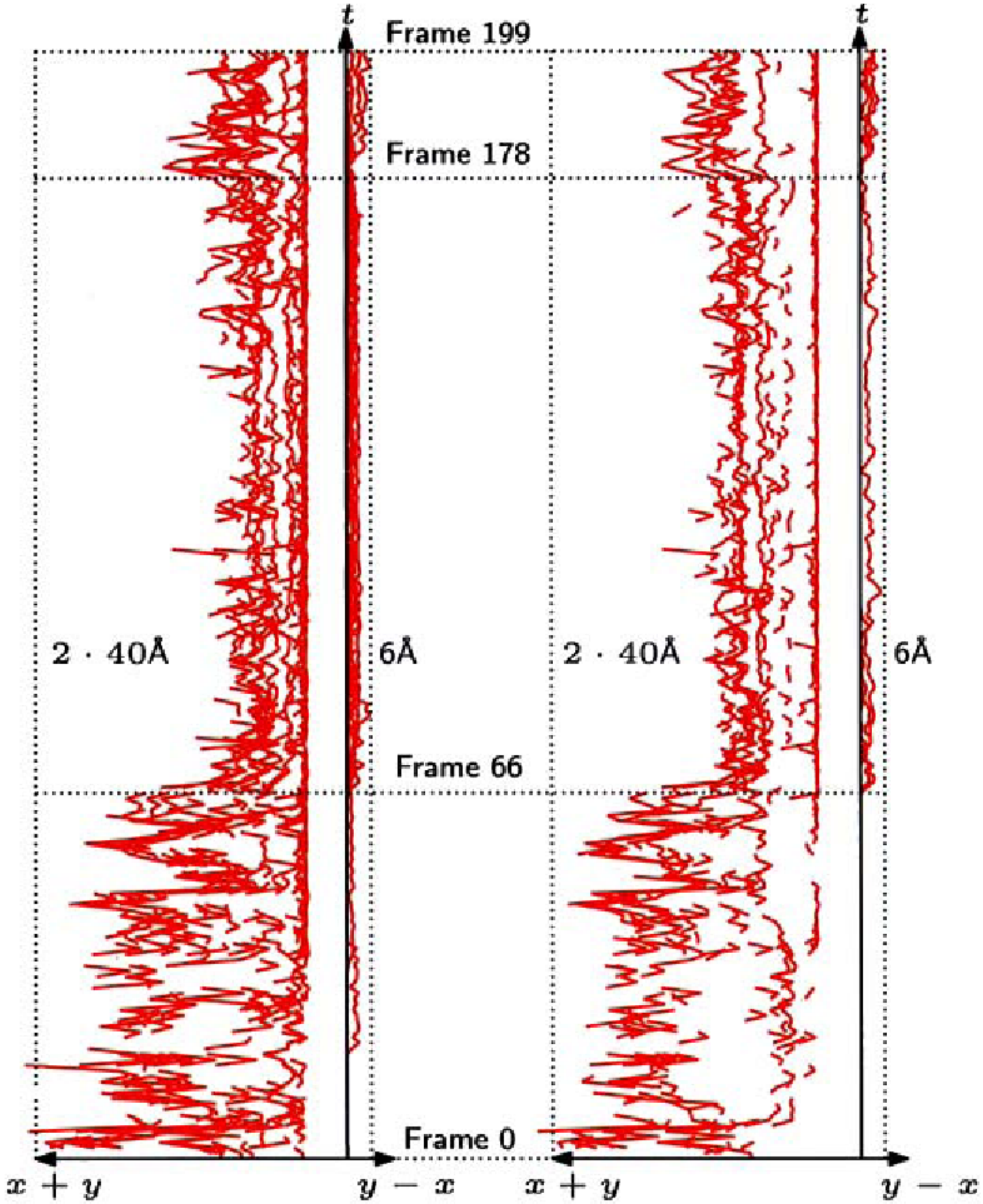


IV.3 VINEYARD





IV.3 VINEYARD



THANK YOU