A Theory of Sub-Barcodes

$$
\begin{tikzcd}
M \arrow[r, "F"] & N \arrow[d, "\varphi_2"] \\
M' \arrow[u, "\varphi_1"] \arrow[r, "G"] & N'
\end{tikzcd}
$$

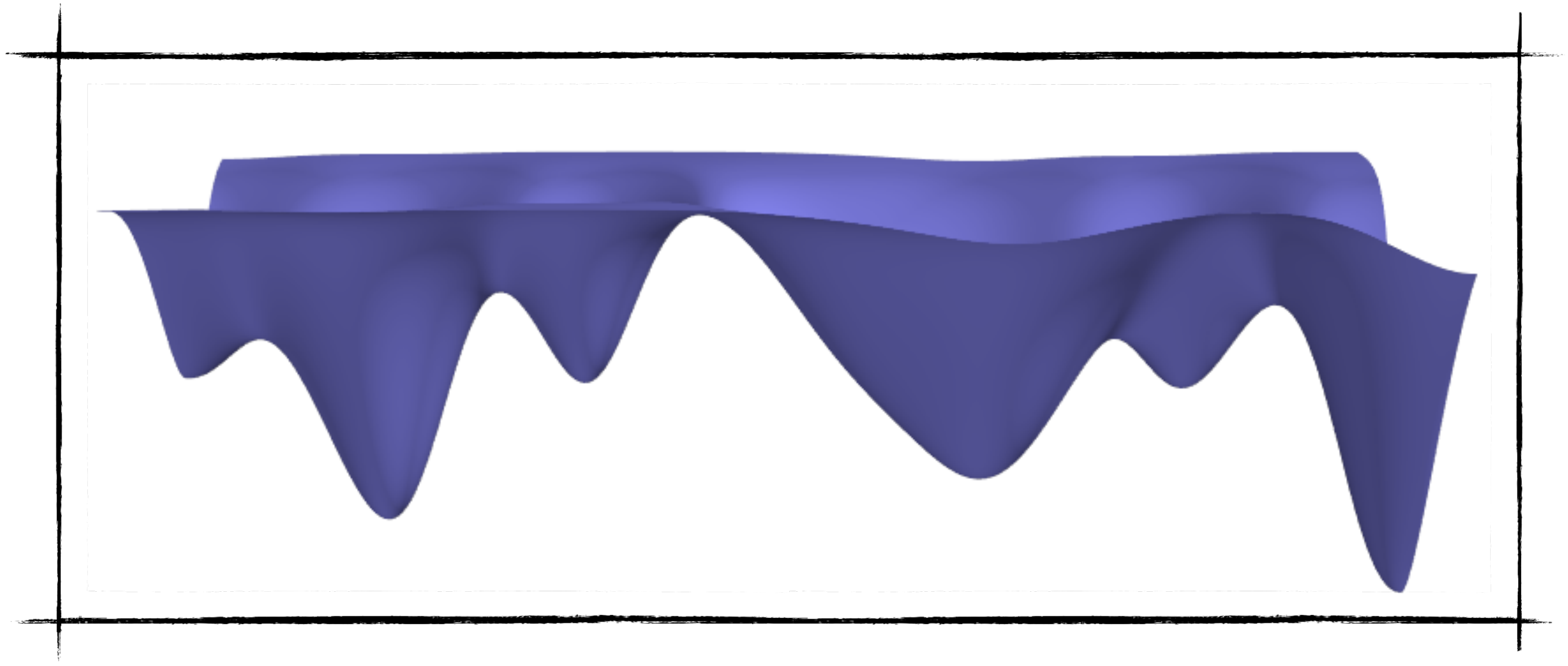
SOCG'25

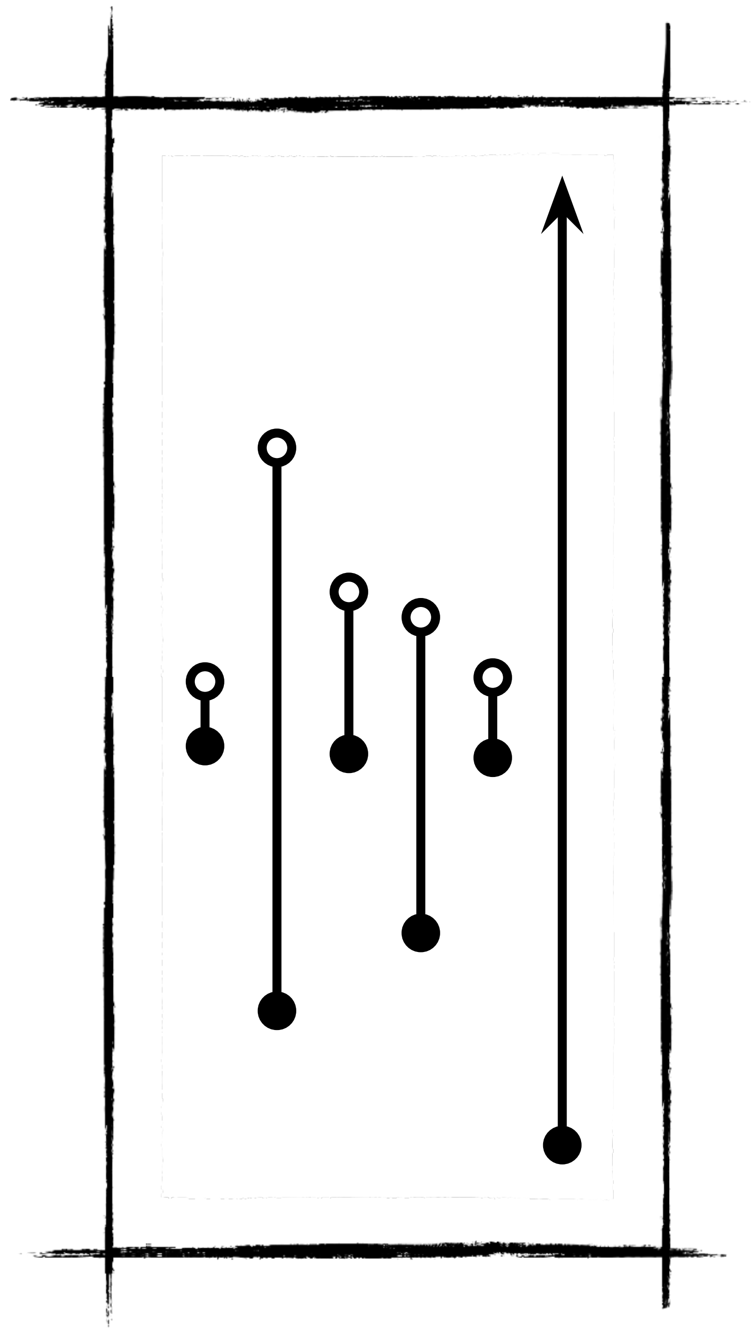
Oliver Chubet, Kirk Gardner, Don Sheehy

Persistent Homology

Input: $f : X \to \mathbb{R}$ a real valued function on a topological space

Output: $\mathrm{Bar}(f)$ a topological signature of f called the **barcode**

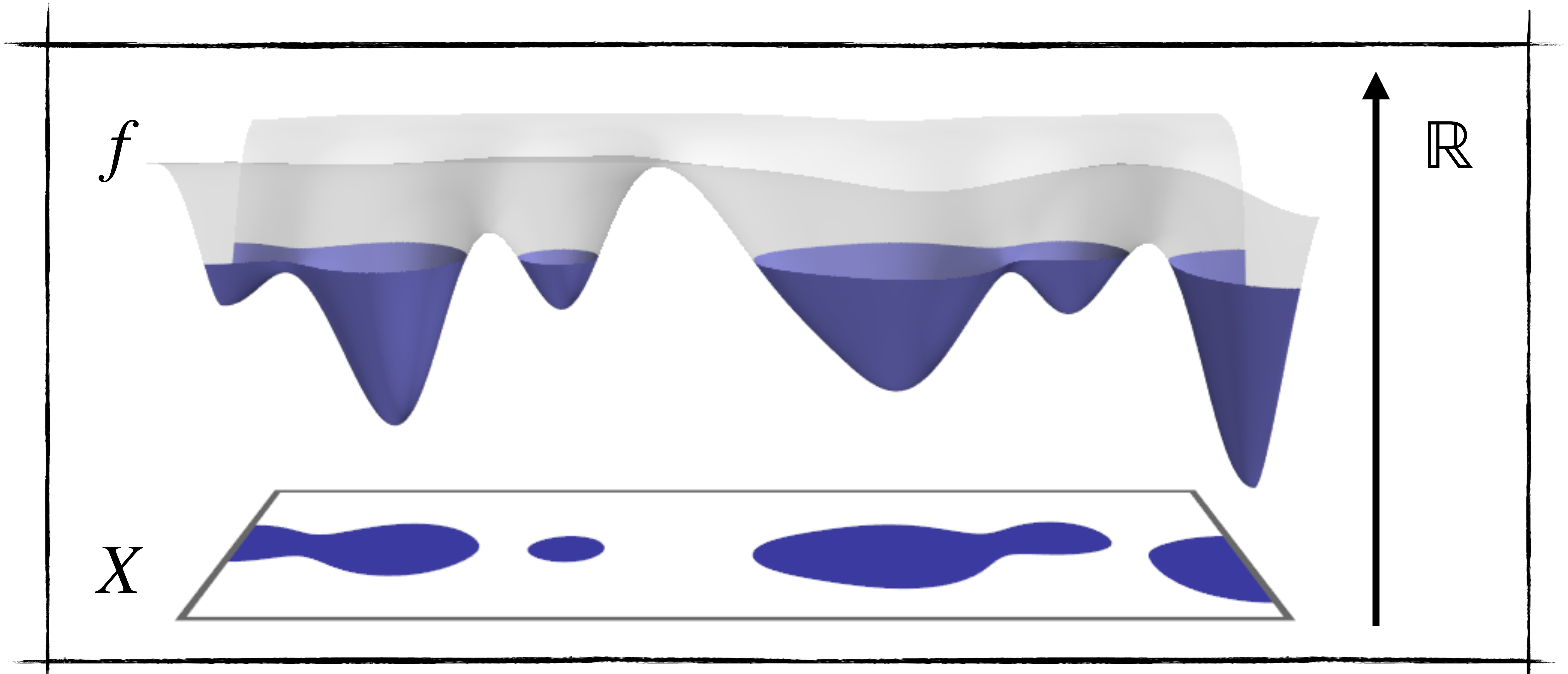




Sublevel Sets

Given a function $f : X \to \mathbb{R}$,

$$\mathsf{Sub}_f(t) := f^{-1}(-\infty, t] = \{x \in X \mid f(x) \leq t\}$$

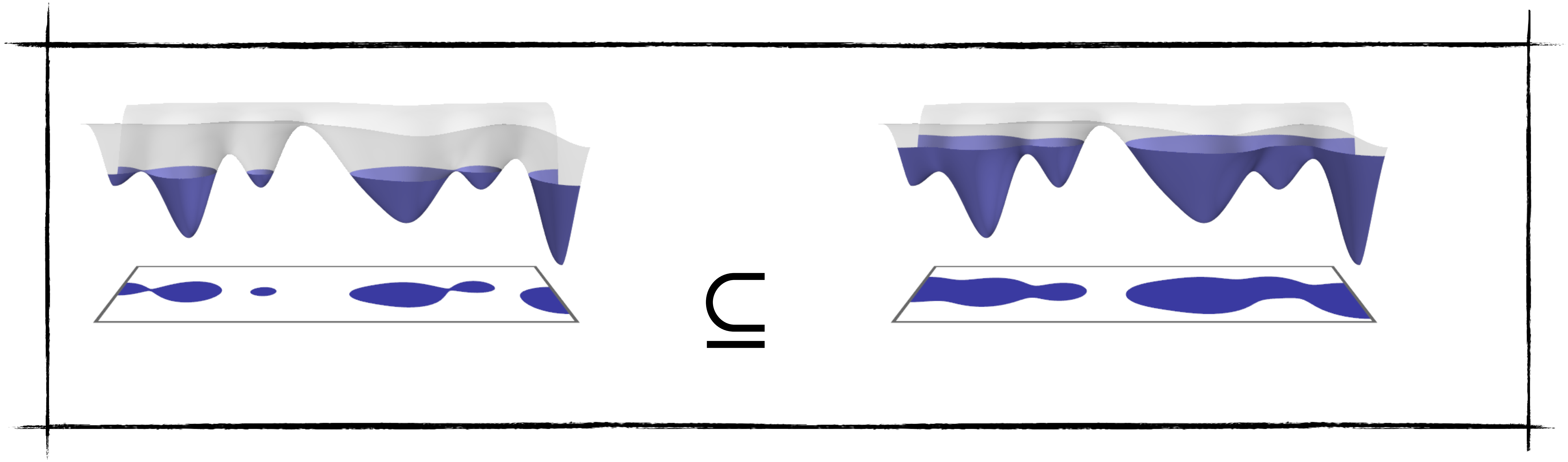


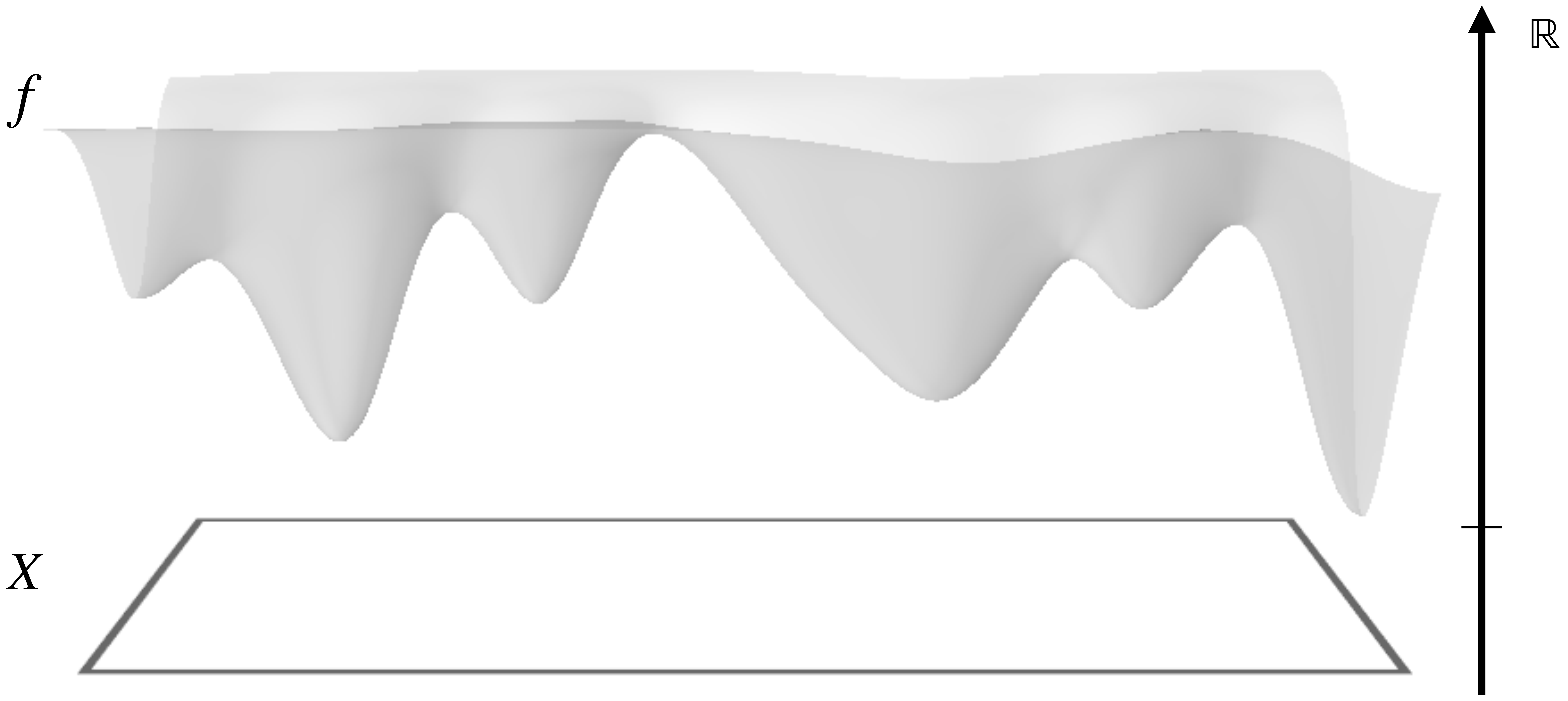
Filtrations

$\mathrm{Sub}_f : \mathbb{R} \to \mathrm{Top}$ is a filtration

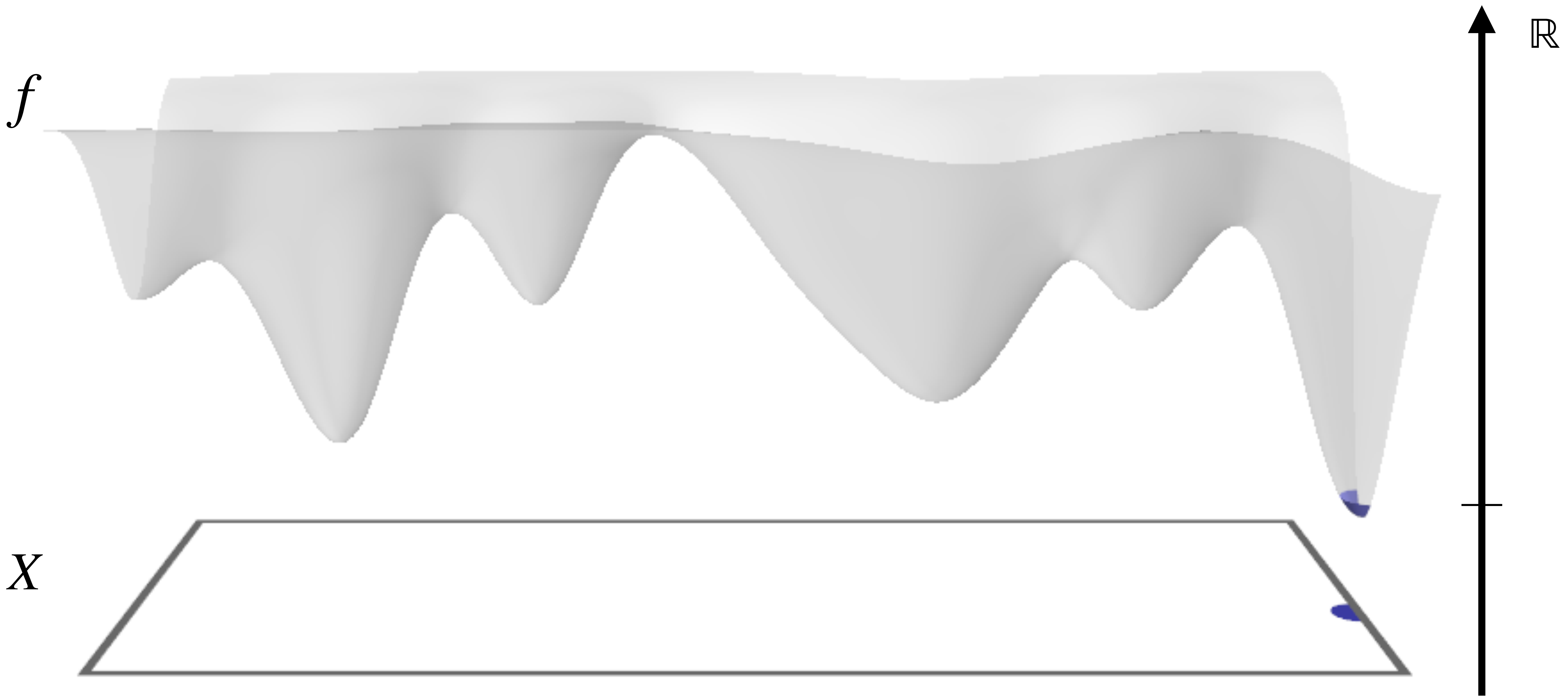
$$t \mapsto \mathrm{Sub}_f(t)$$

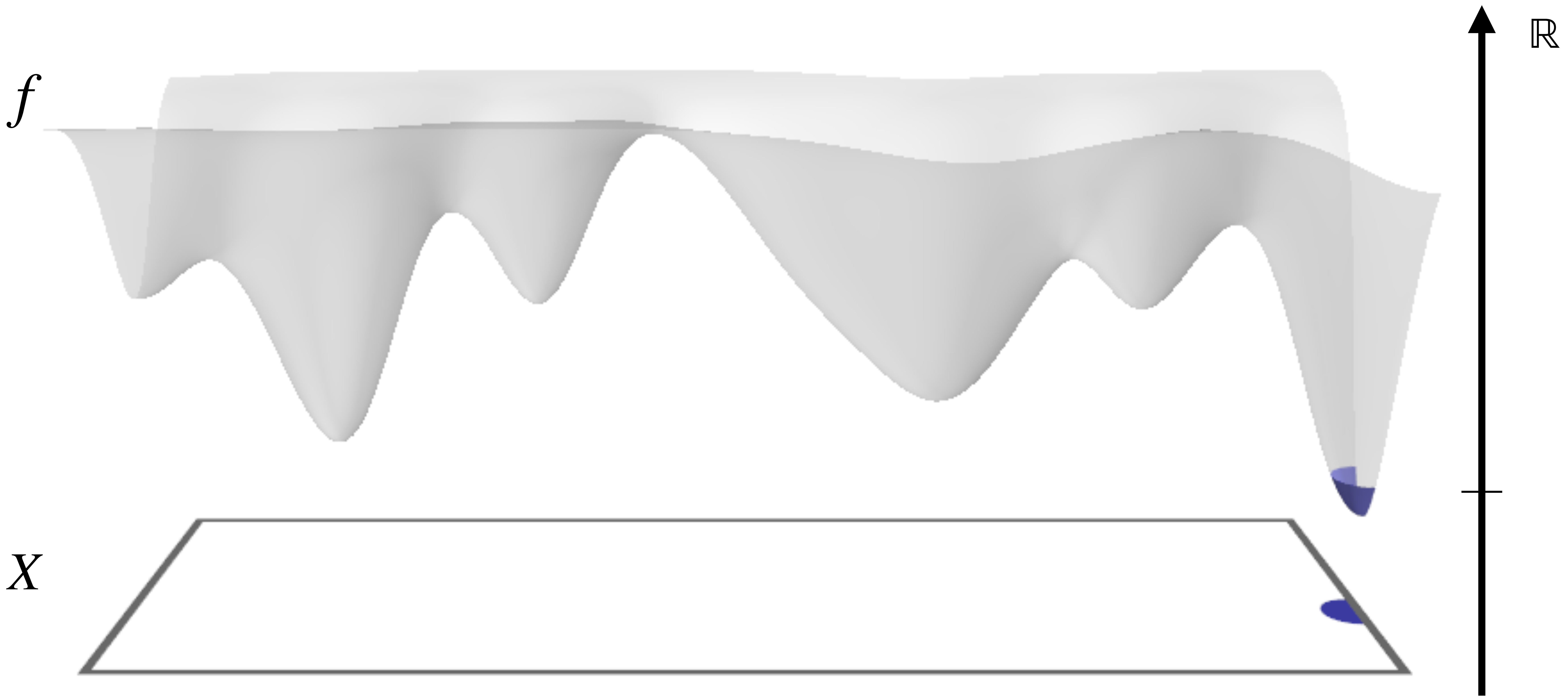
For all $s \leq t$ we have $\mathrm{Sub}_f(s) \subseteq \mathrm{Sub}_f(t)$



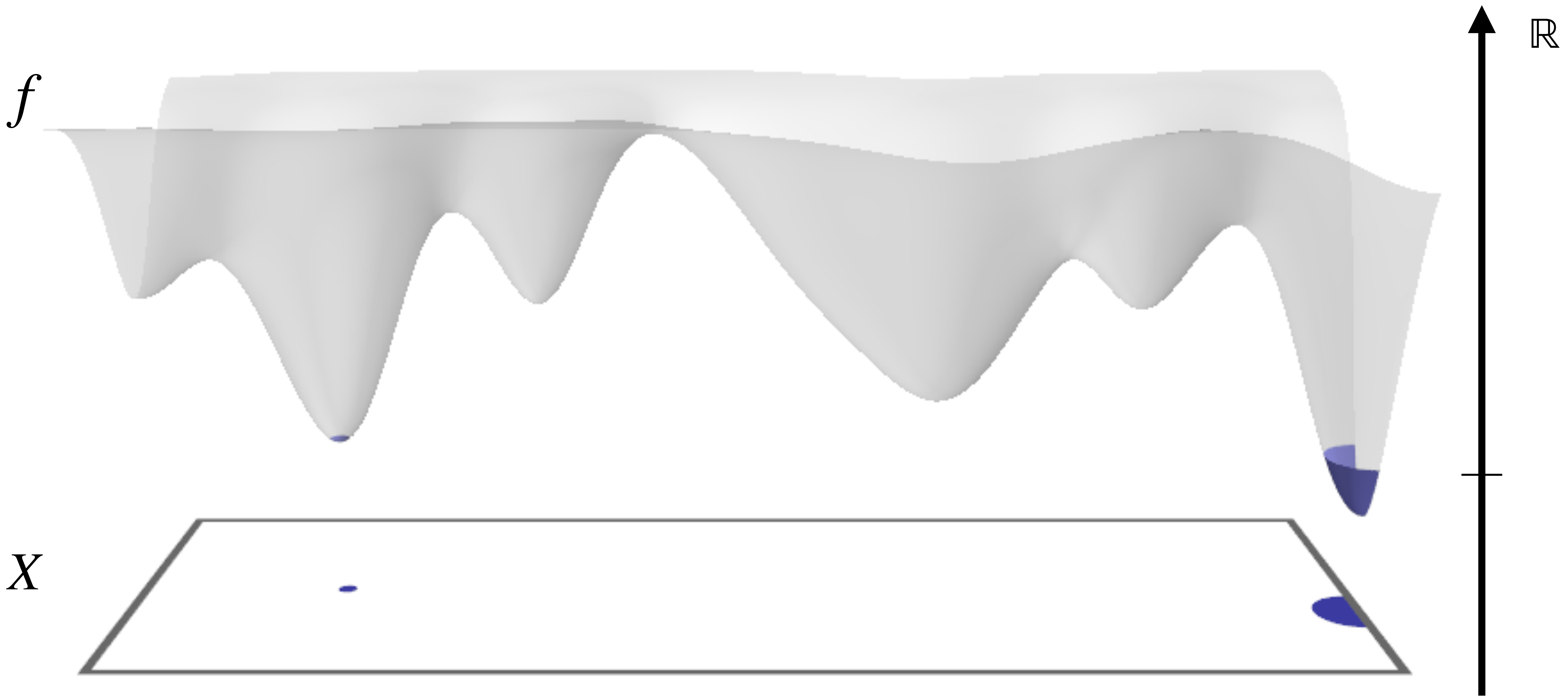


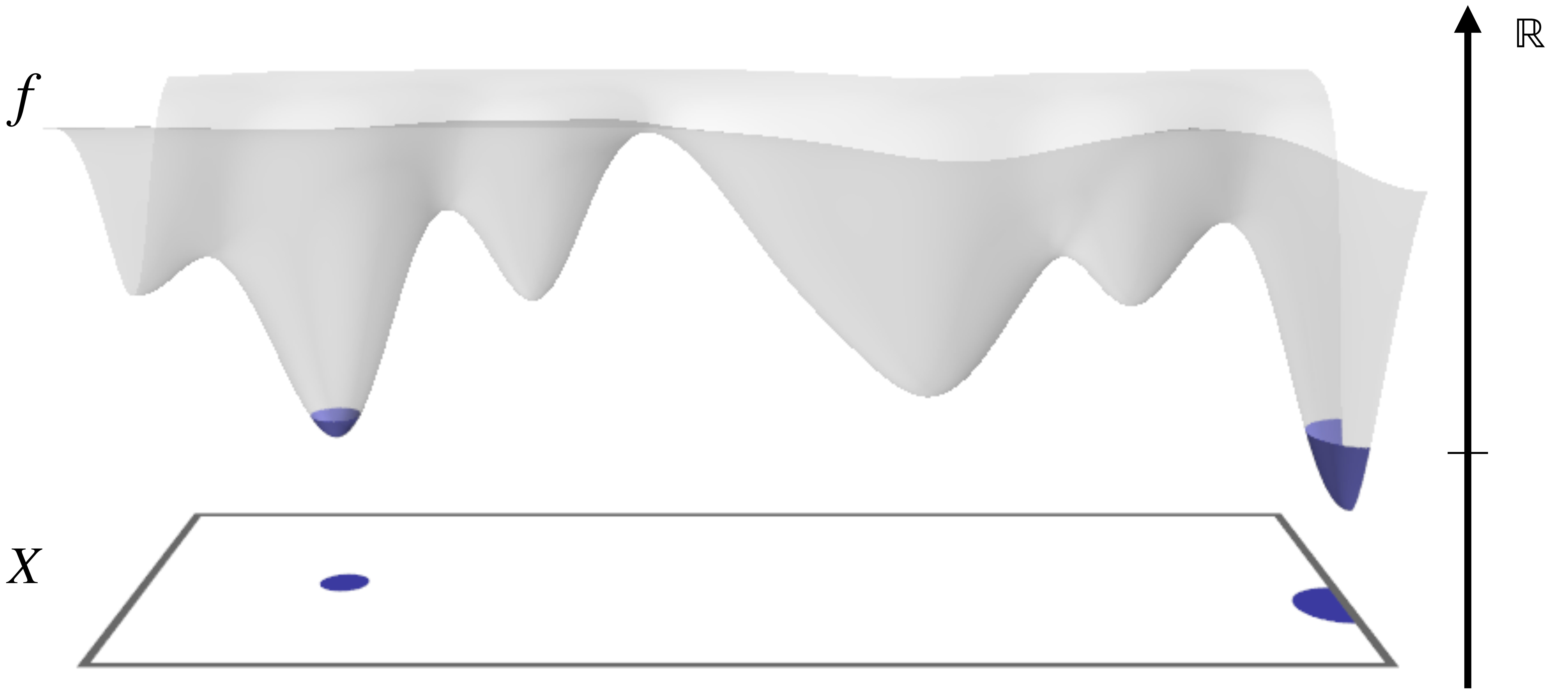
f
X
$\mathbb{R}$


f
X
$\mathbb{R}$

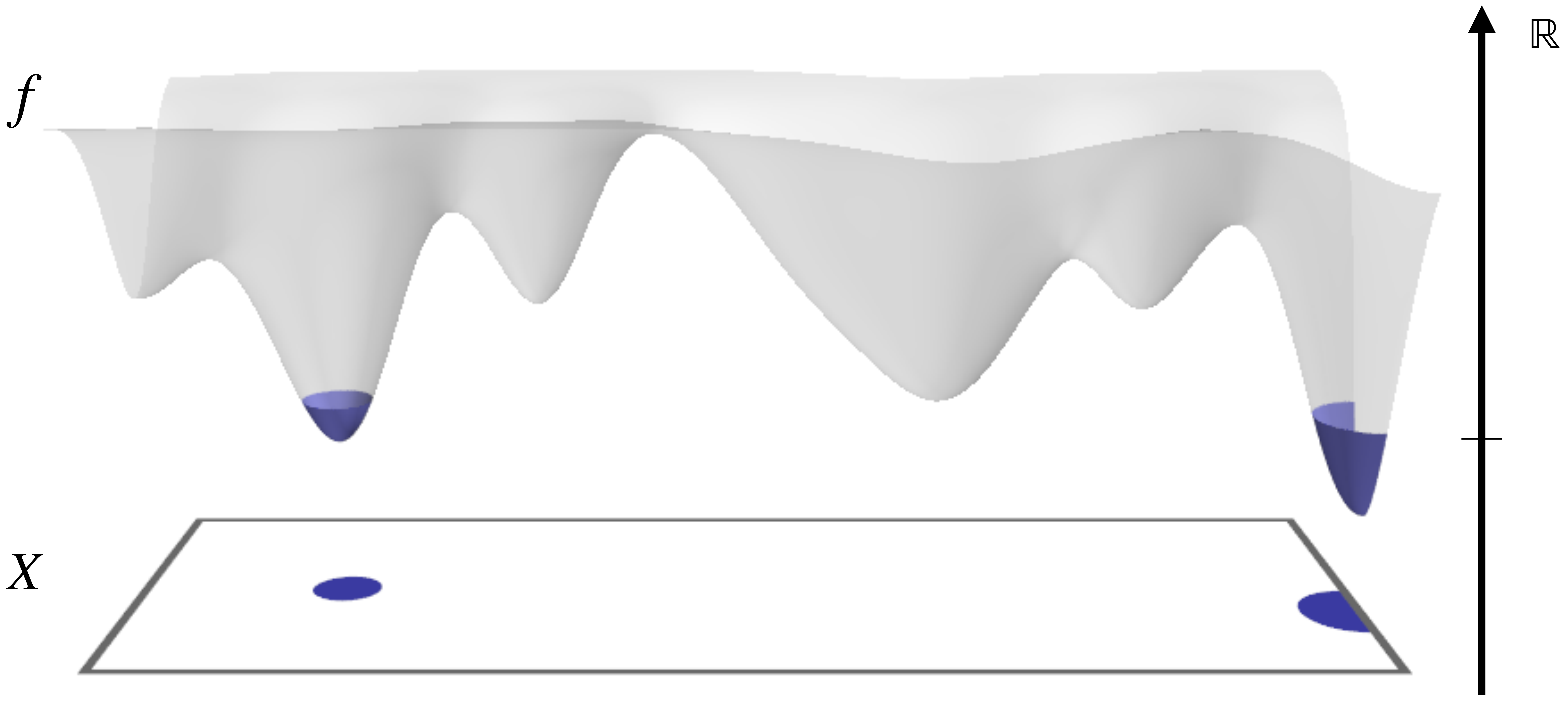


f
X
$\mathbb{R}$

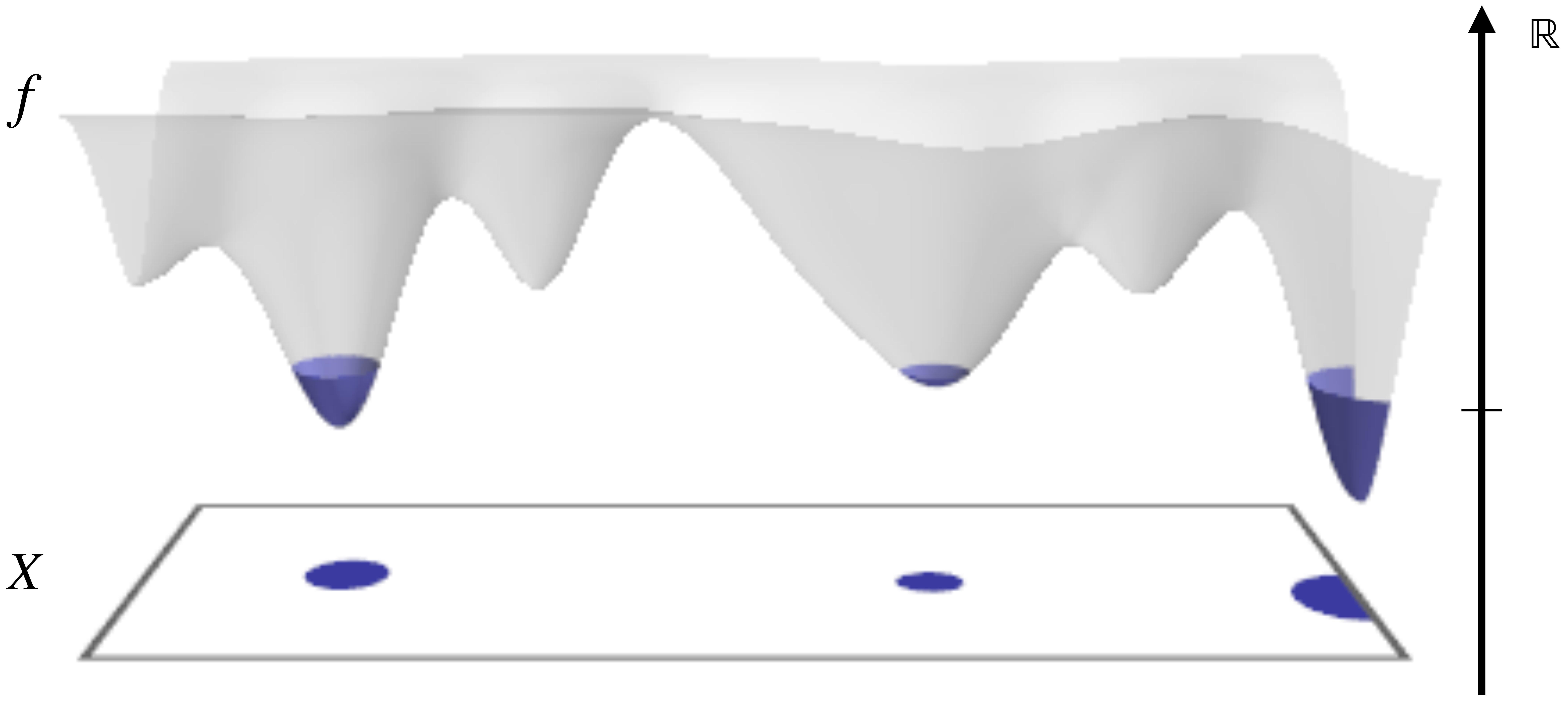

f
X
$\mathbb{R}$



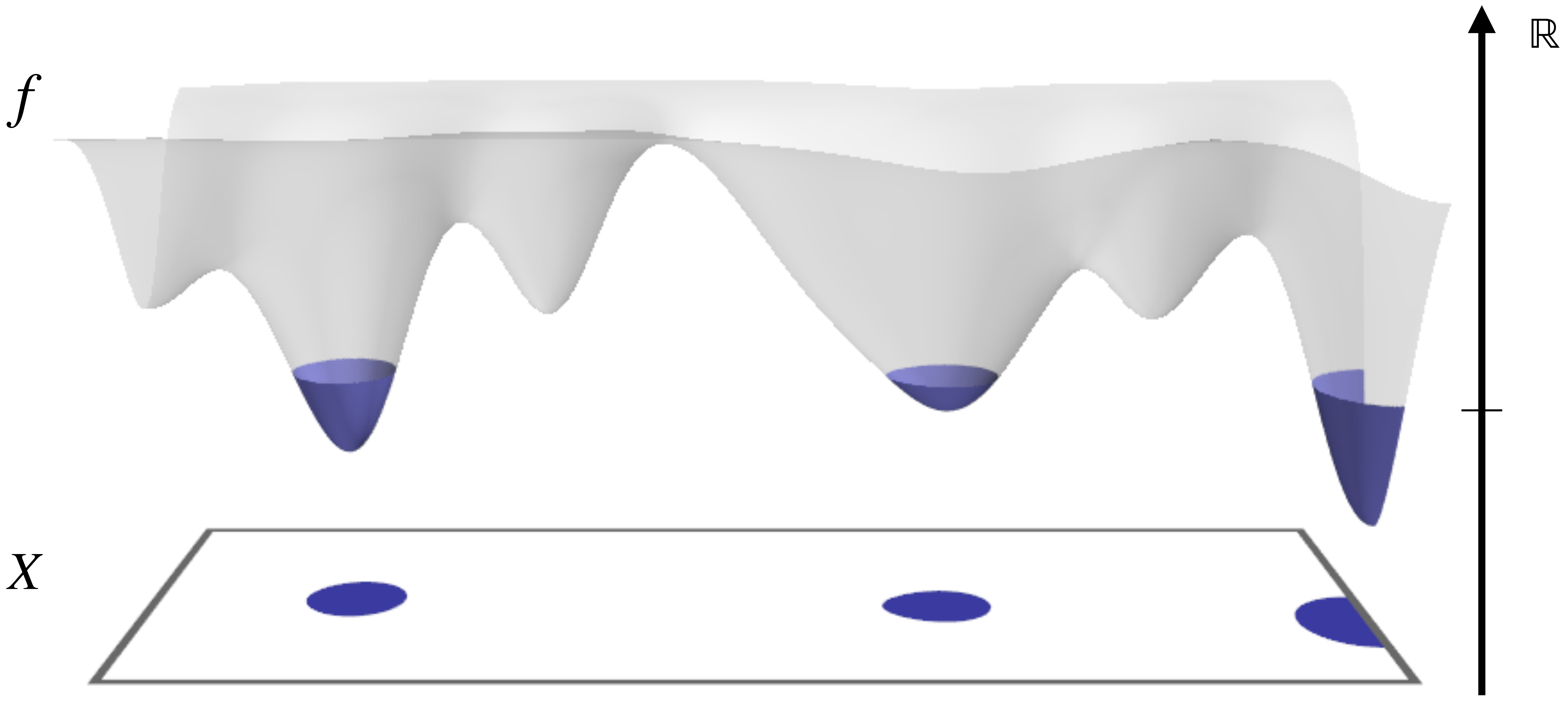
f
X
$\mathbb{R}$



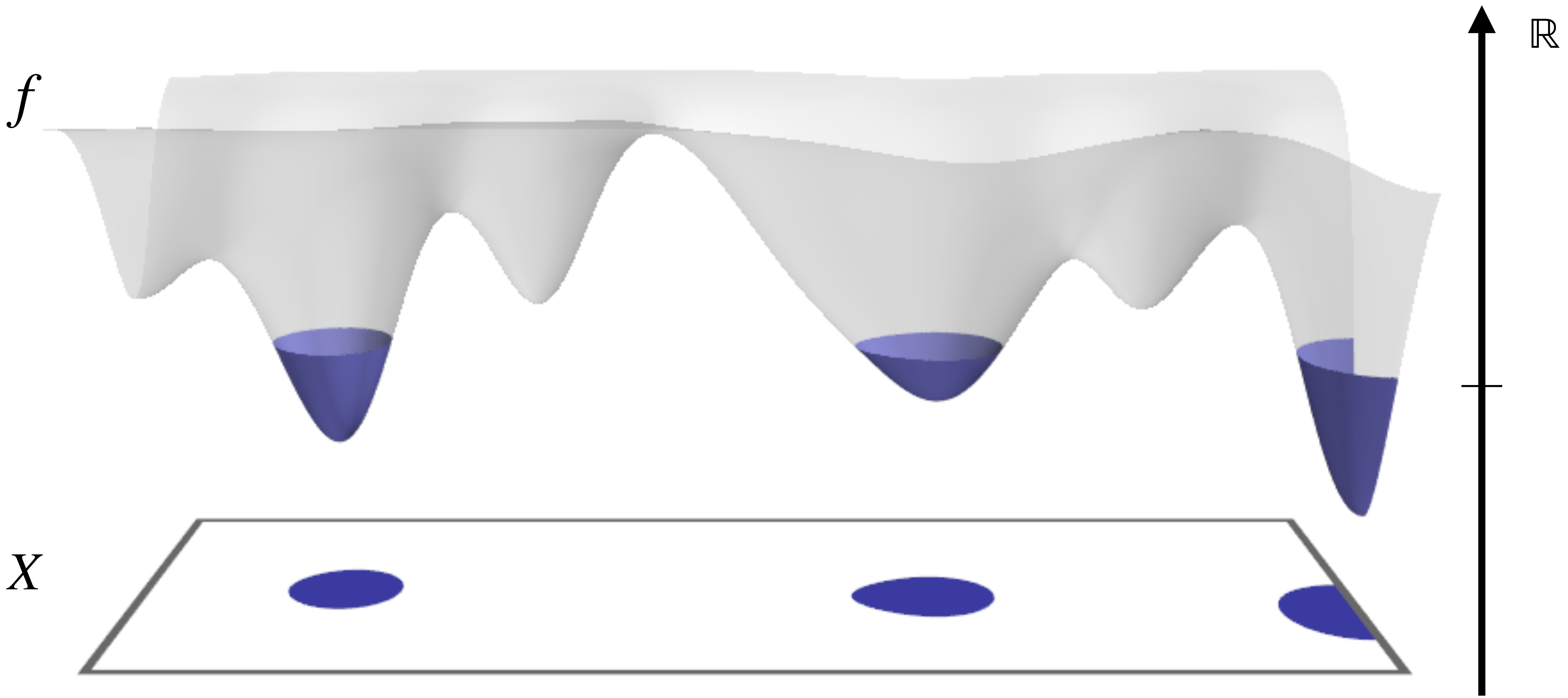
f
X
$\mathbb{R}$



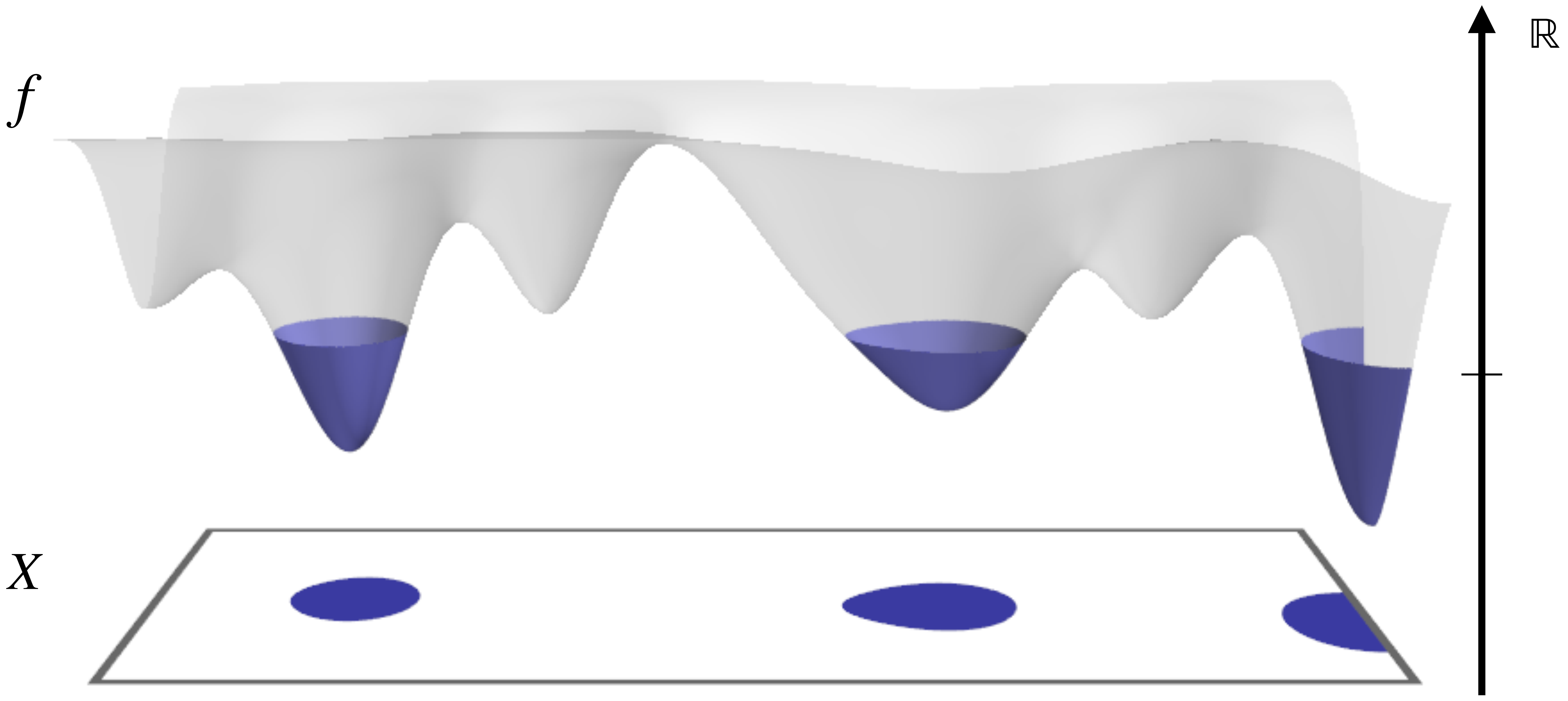
f
X
$\mathbb{R}$



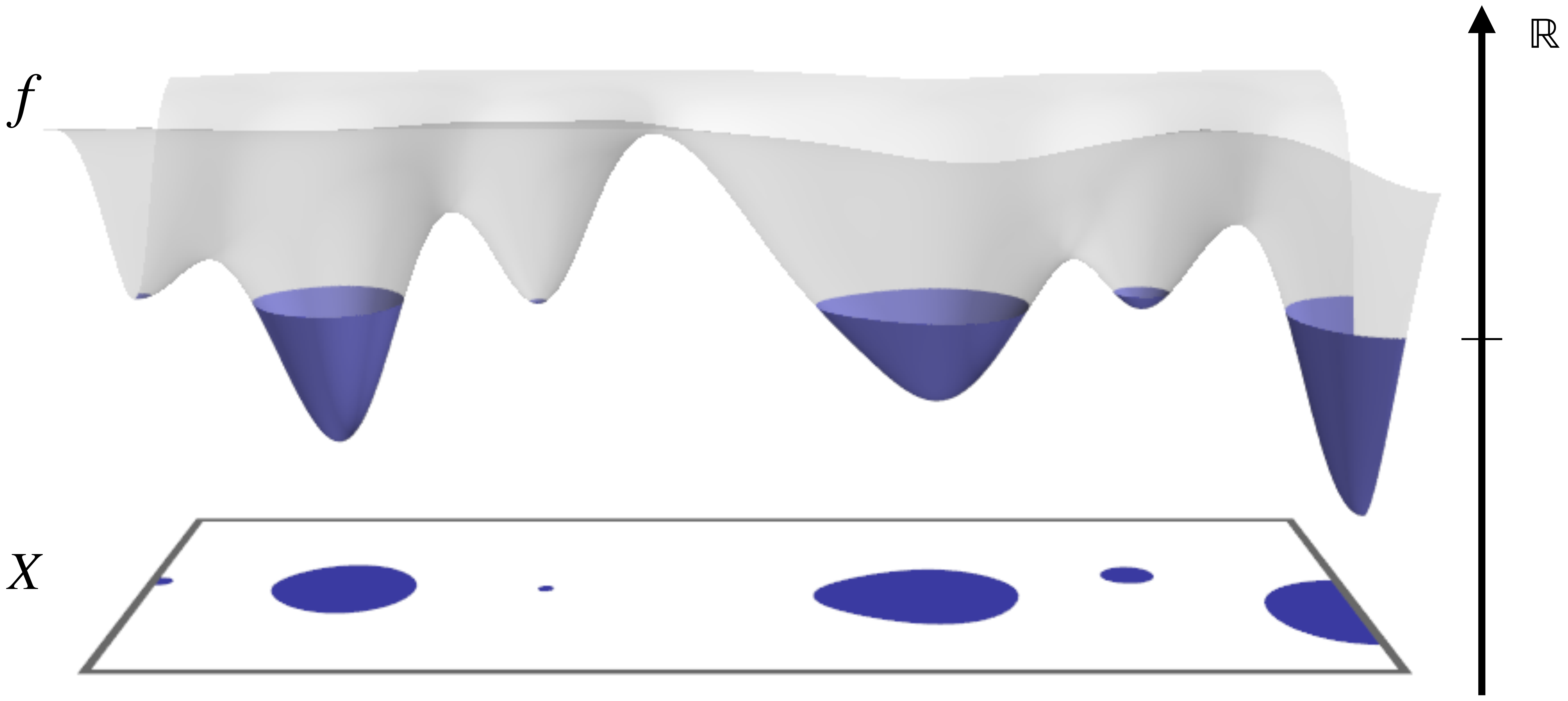
f
X
$\mathbb{R}$



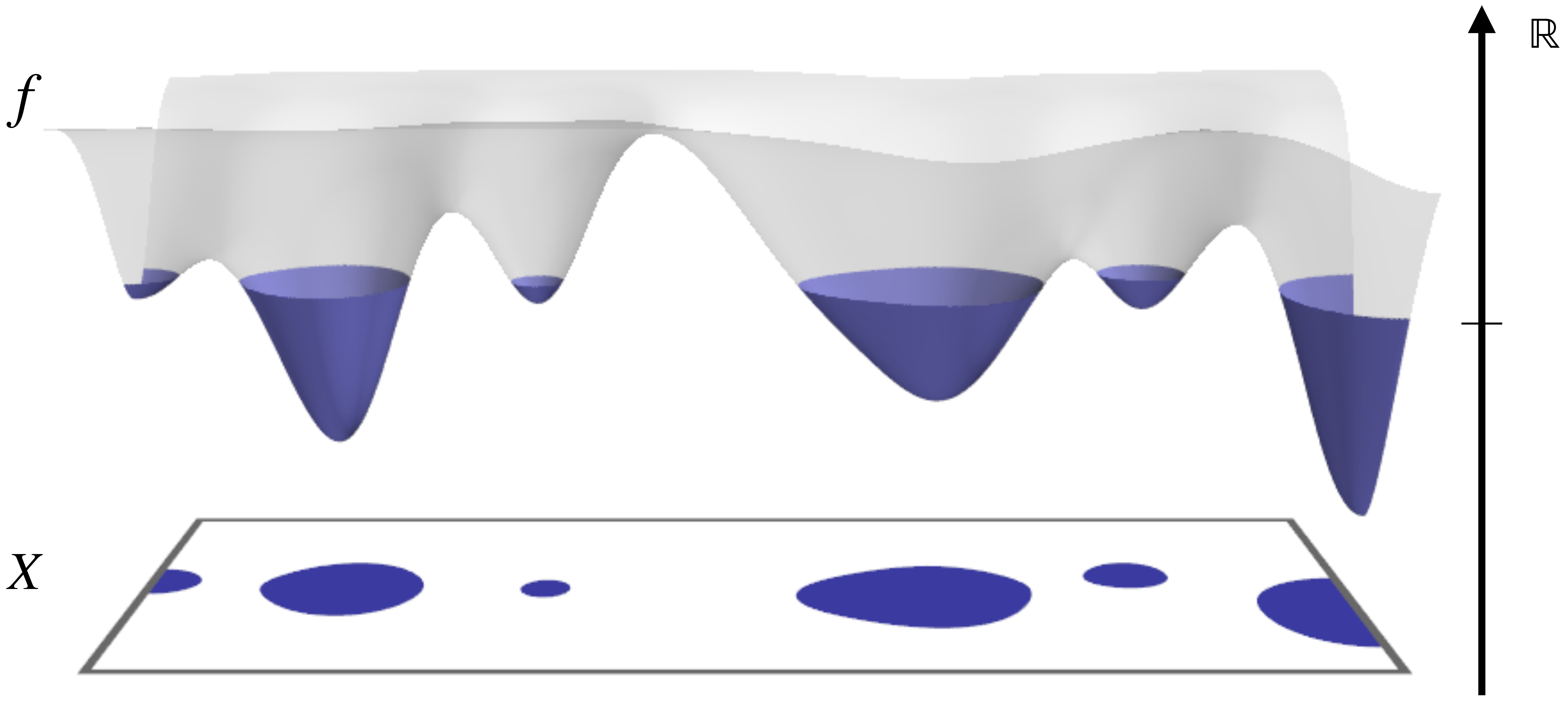
f
X
ℝ



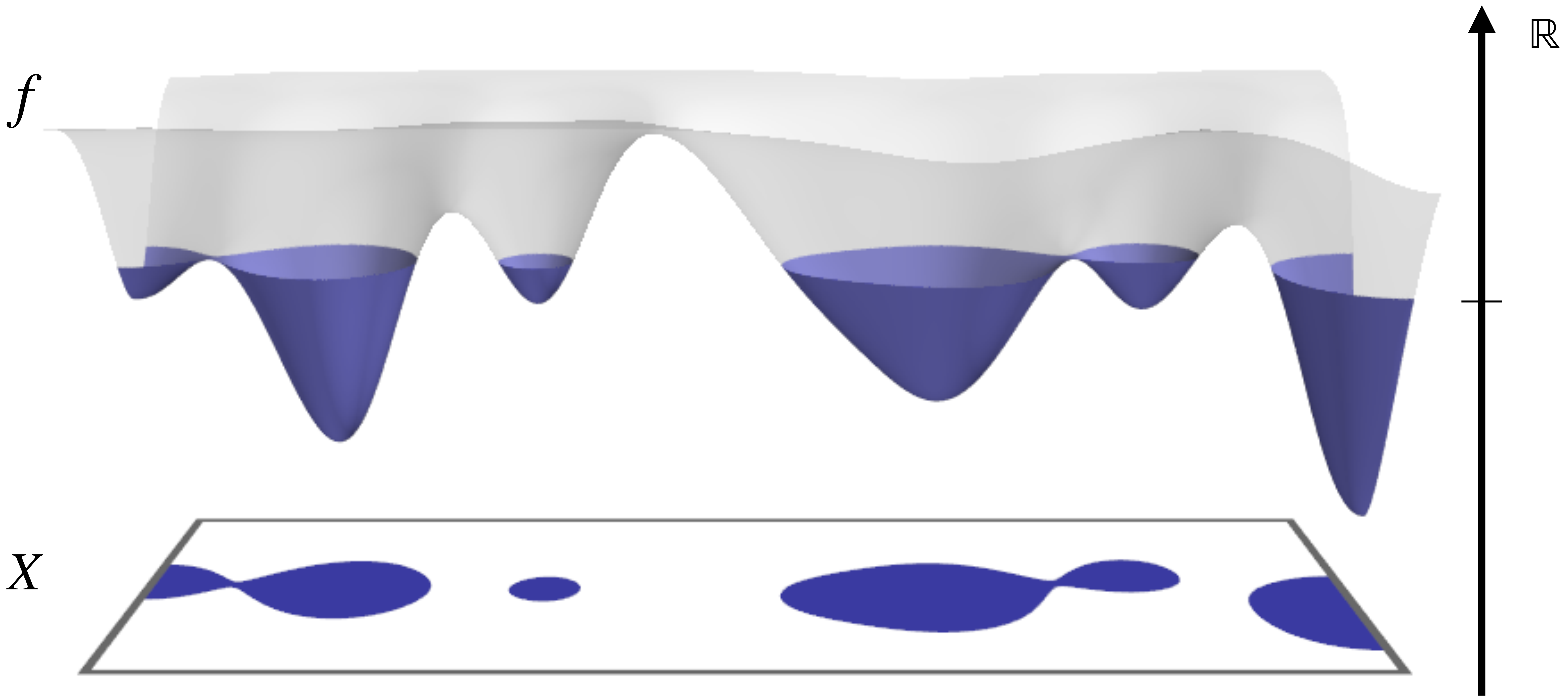
f
X
$\mathbb{R}$



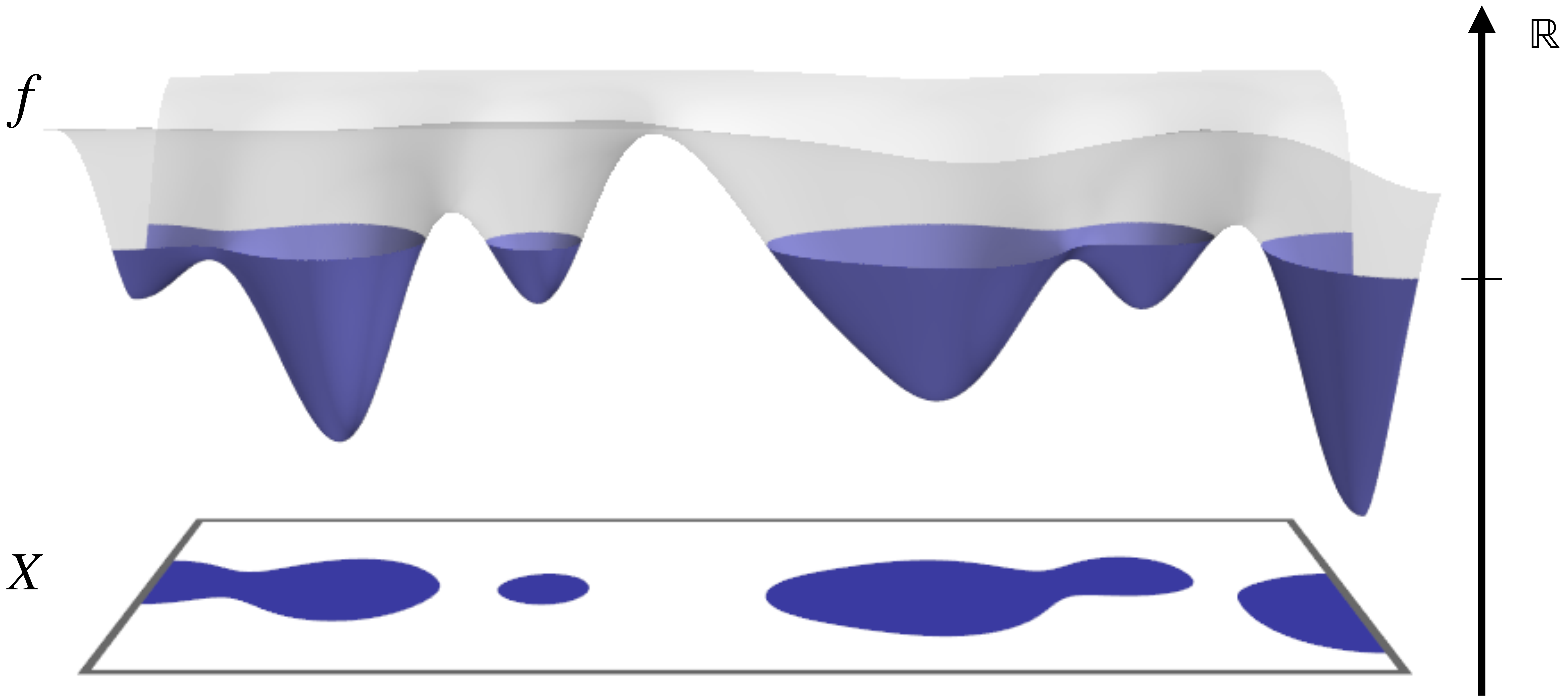
f
X
ℝ

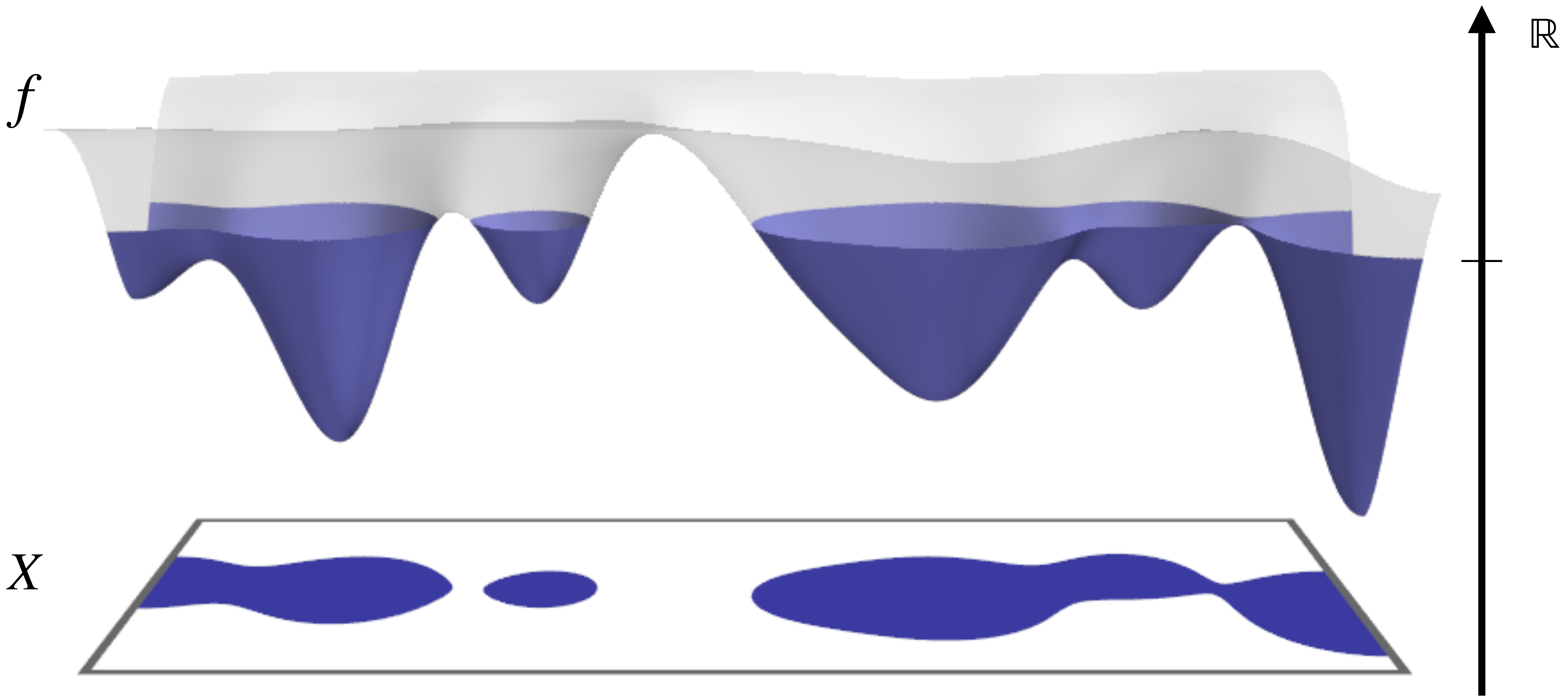


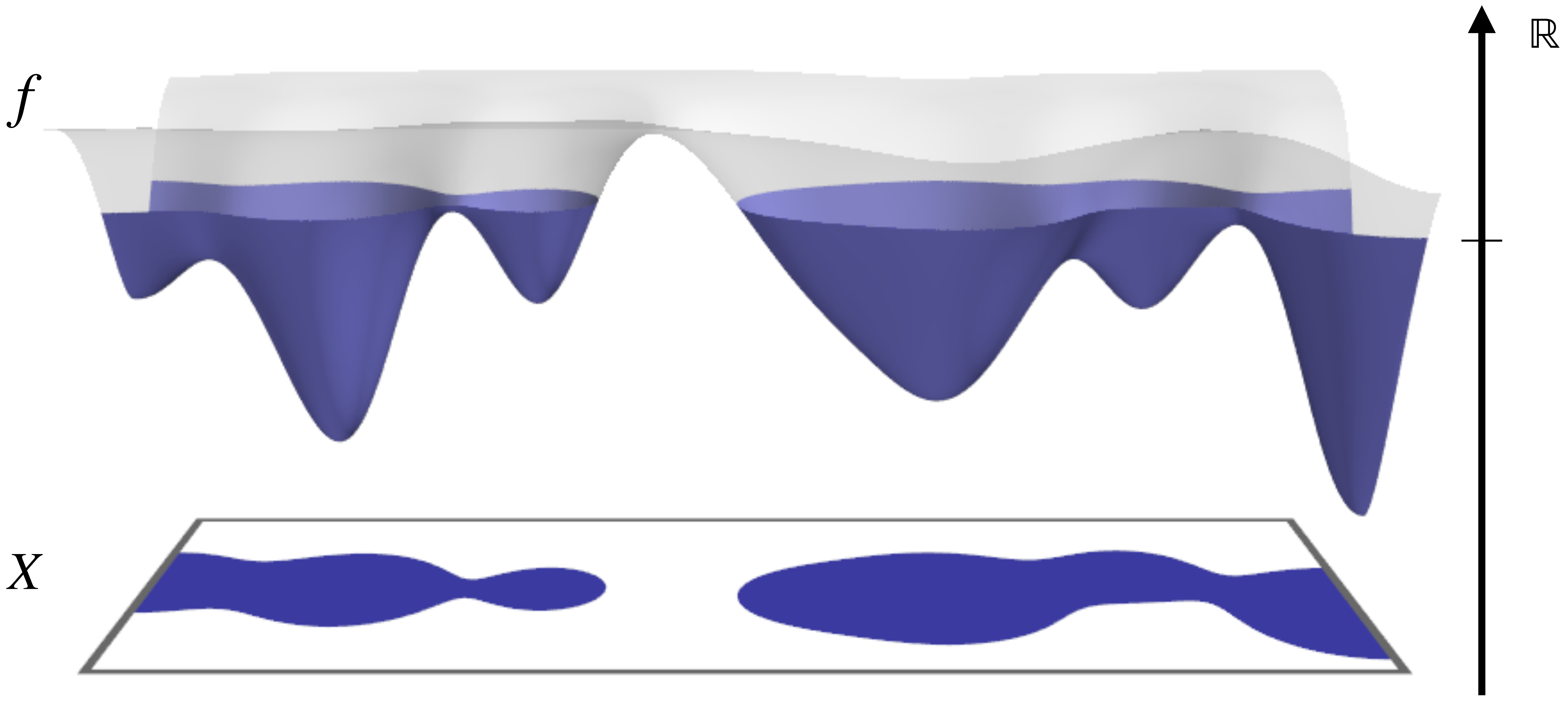
f
X
$\mathbb{R}$



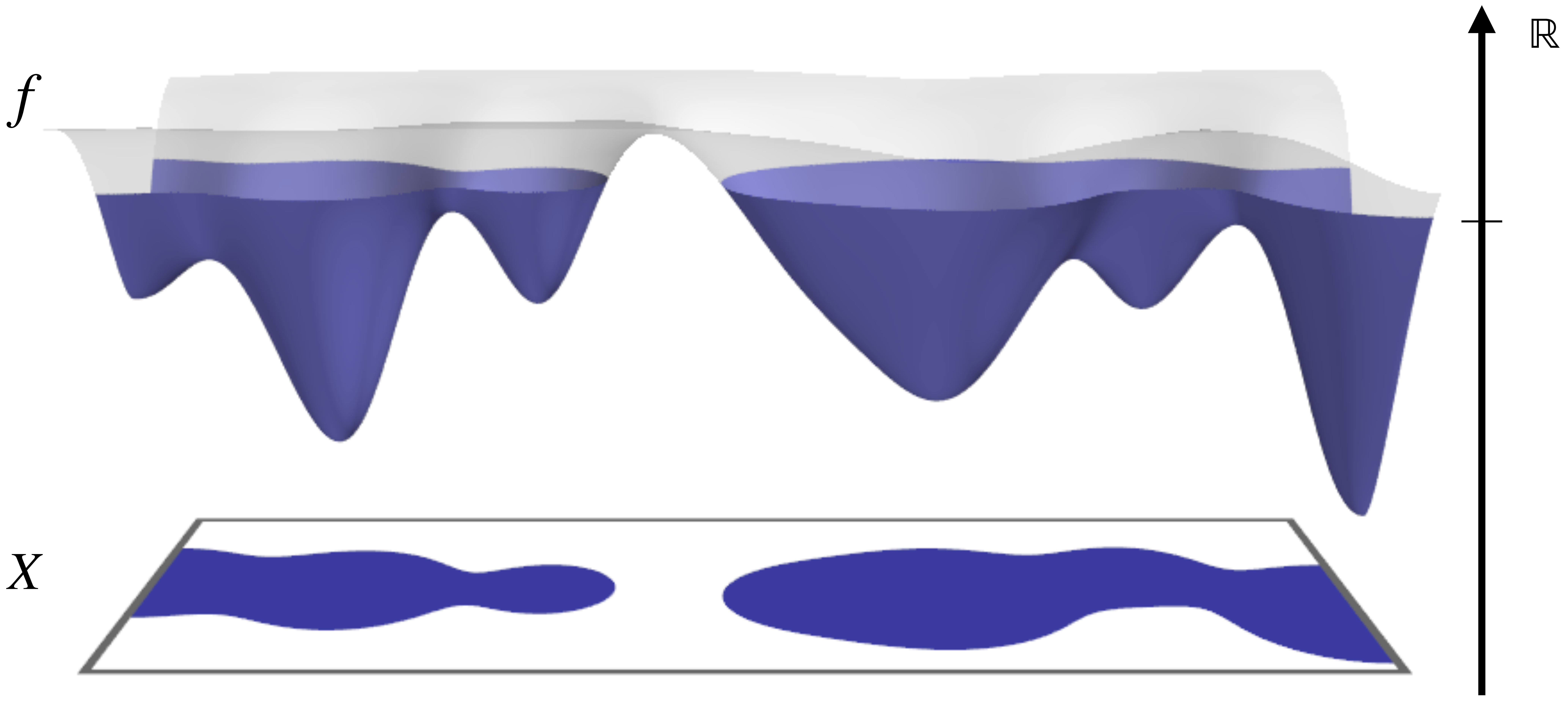
f
X
$\mathbb{R}$


f
X
ℝ

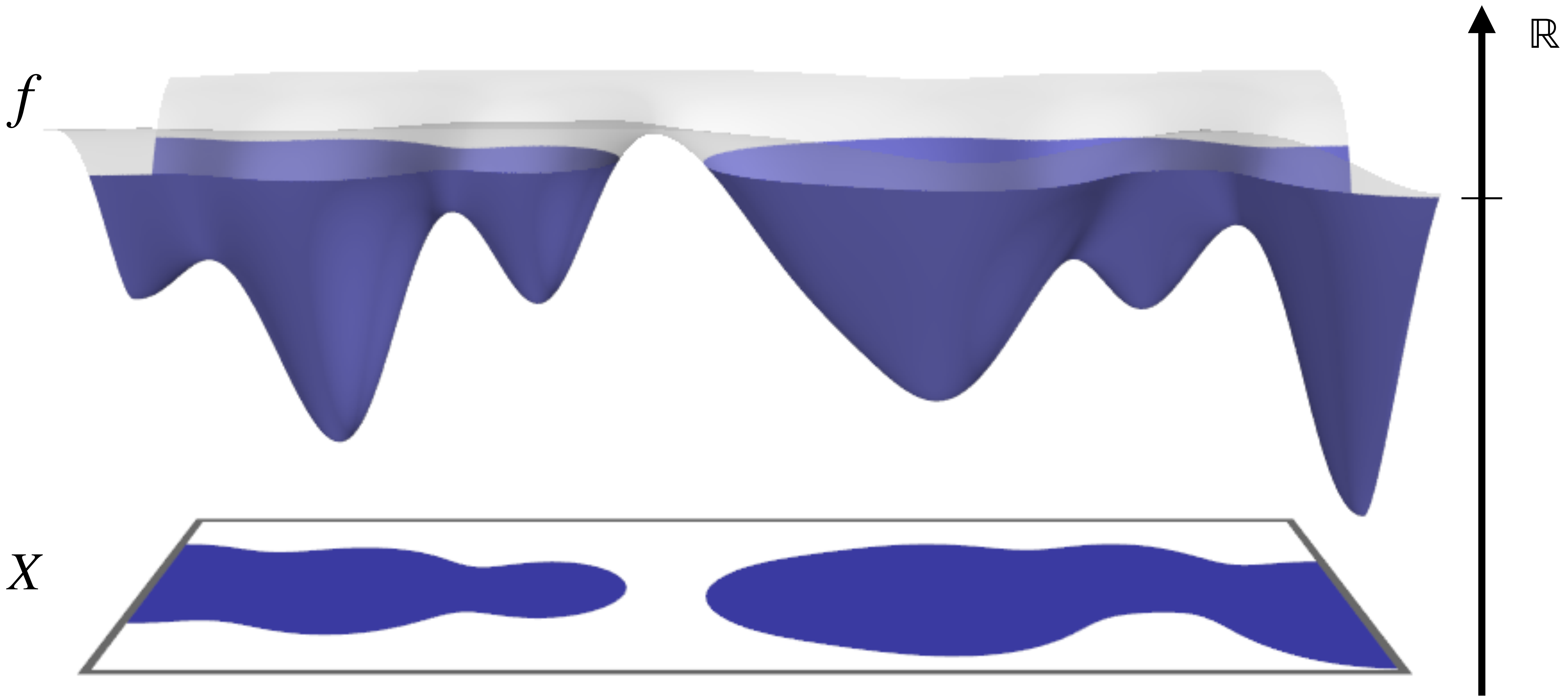

f
X
$\mathbb{R}$



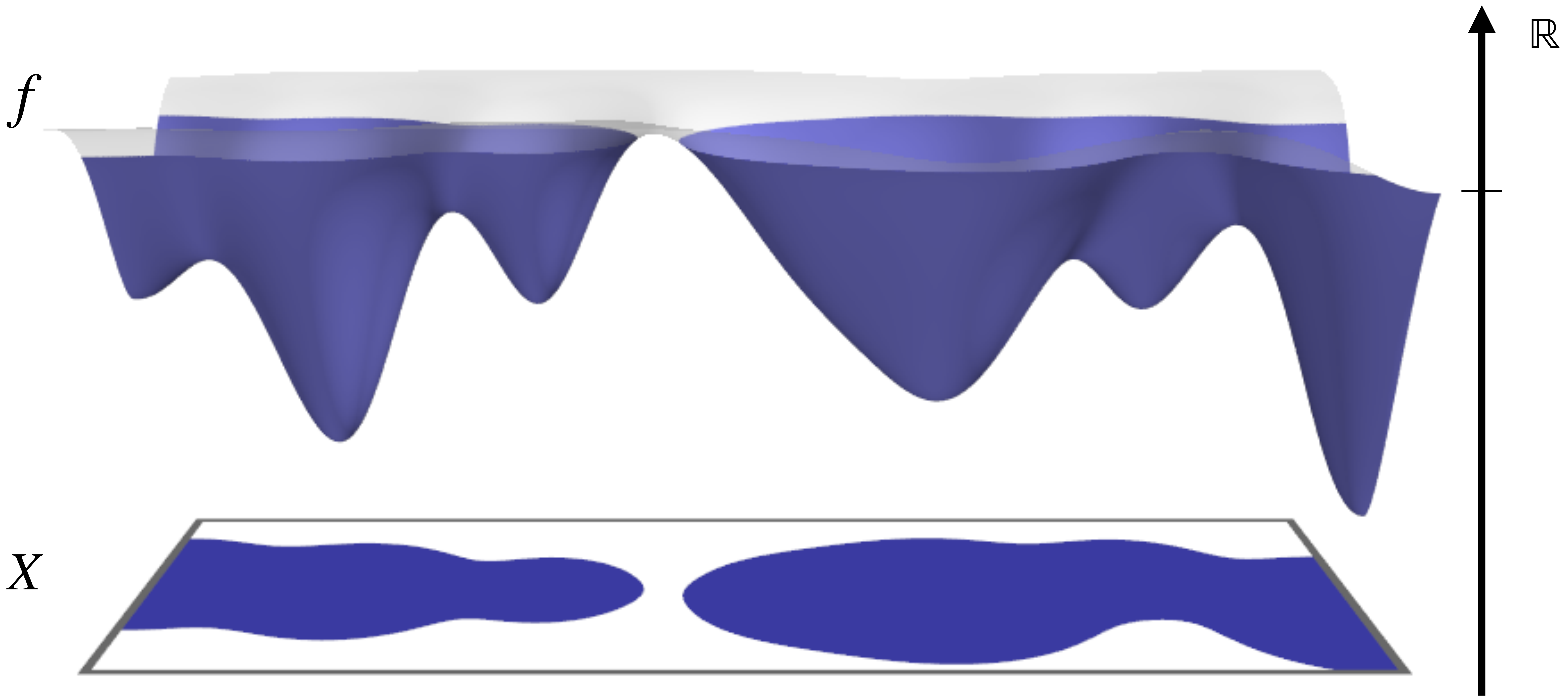
f
$\mathbb{R}$
X



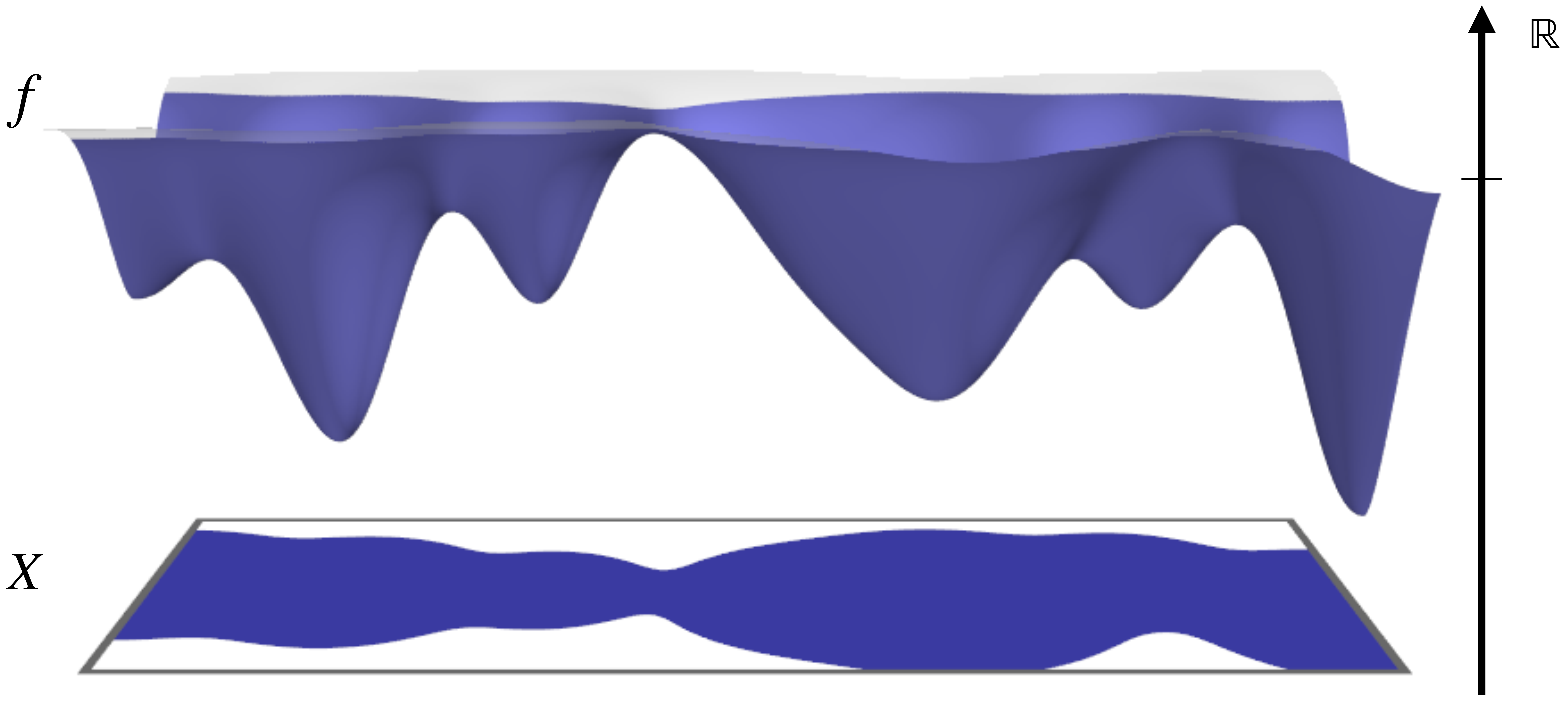
f
X
$\mathbb{R}$



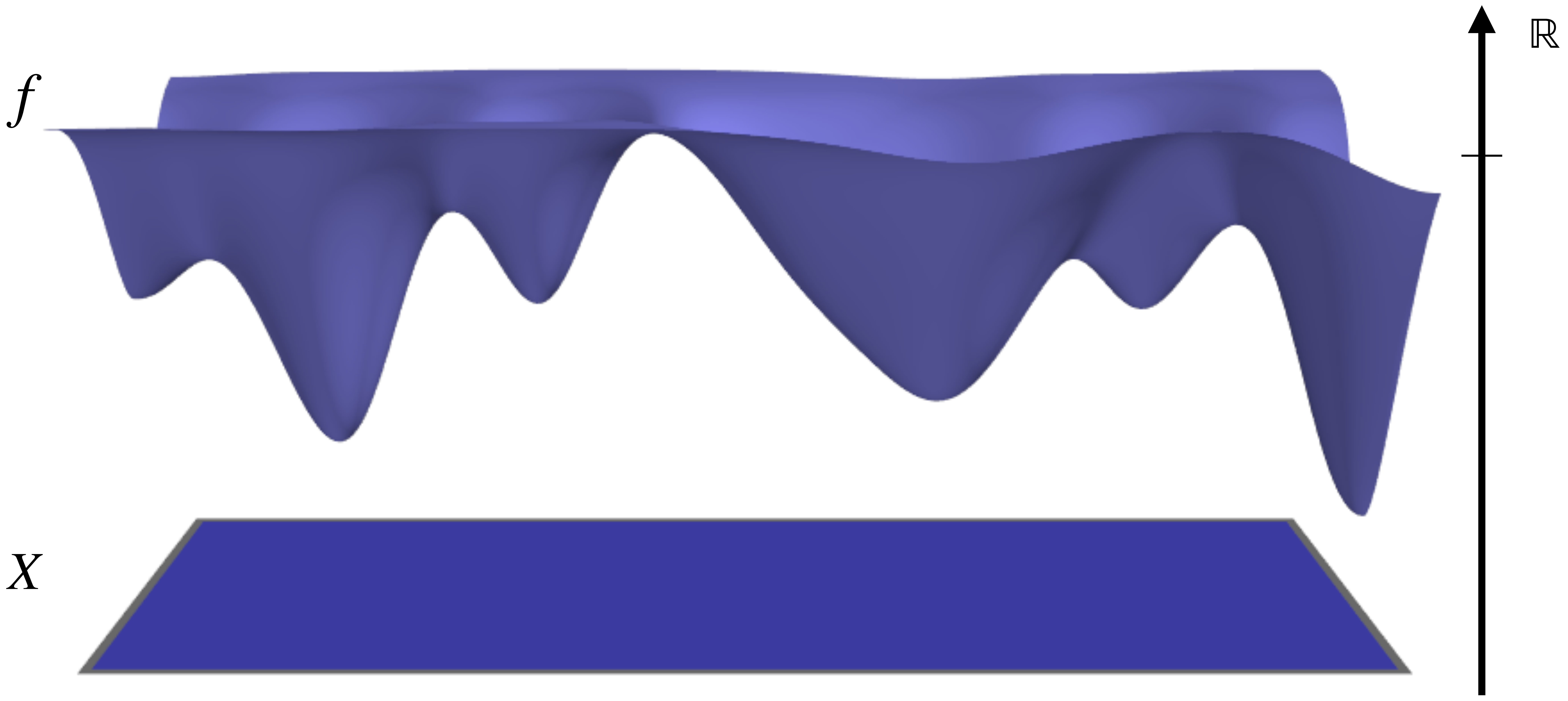
f
X
$\mathbb{R}$



f
X
$\mathbb{R}$



f
X
$\mathbb{R}$

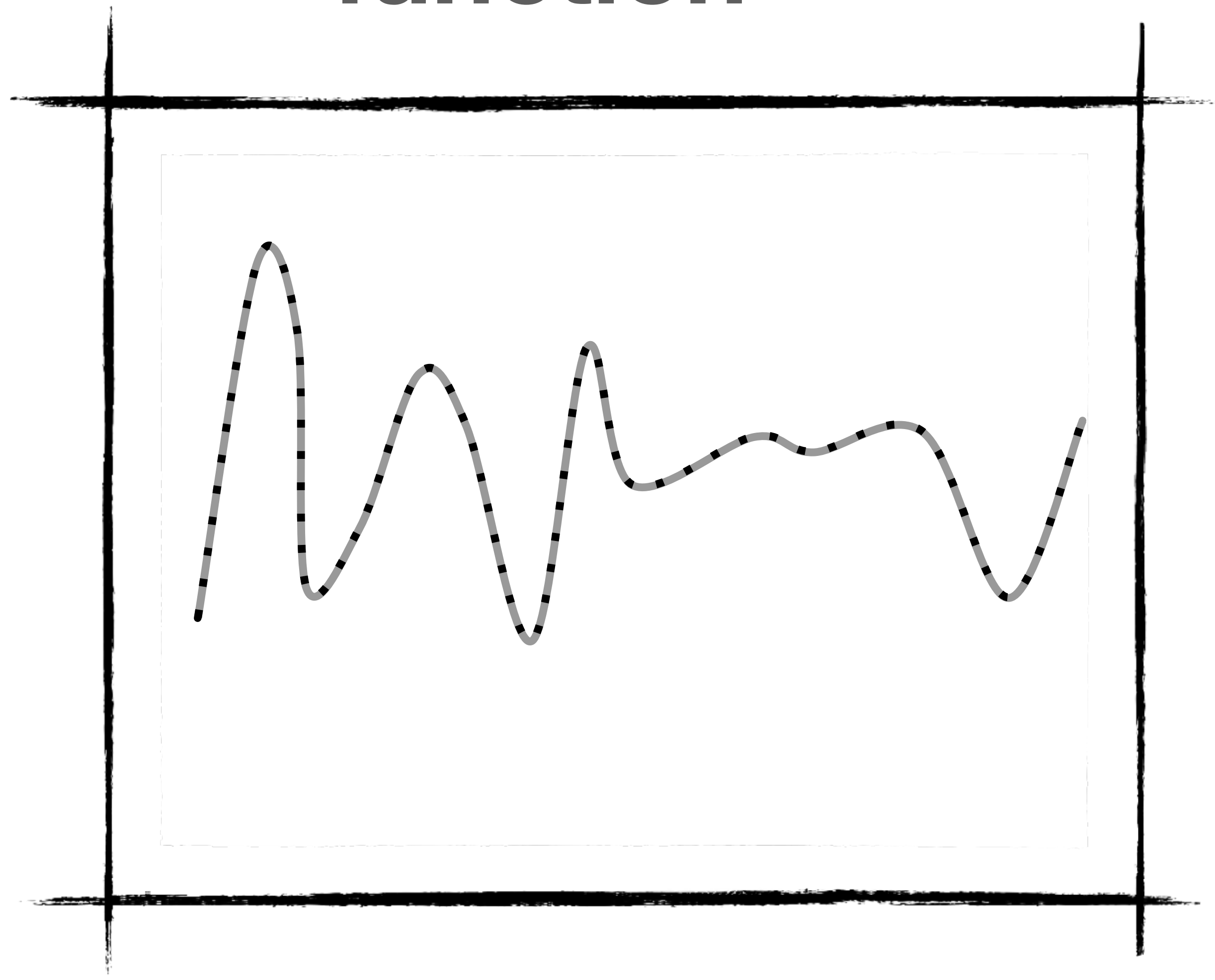


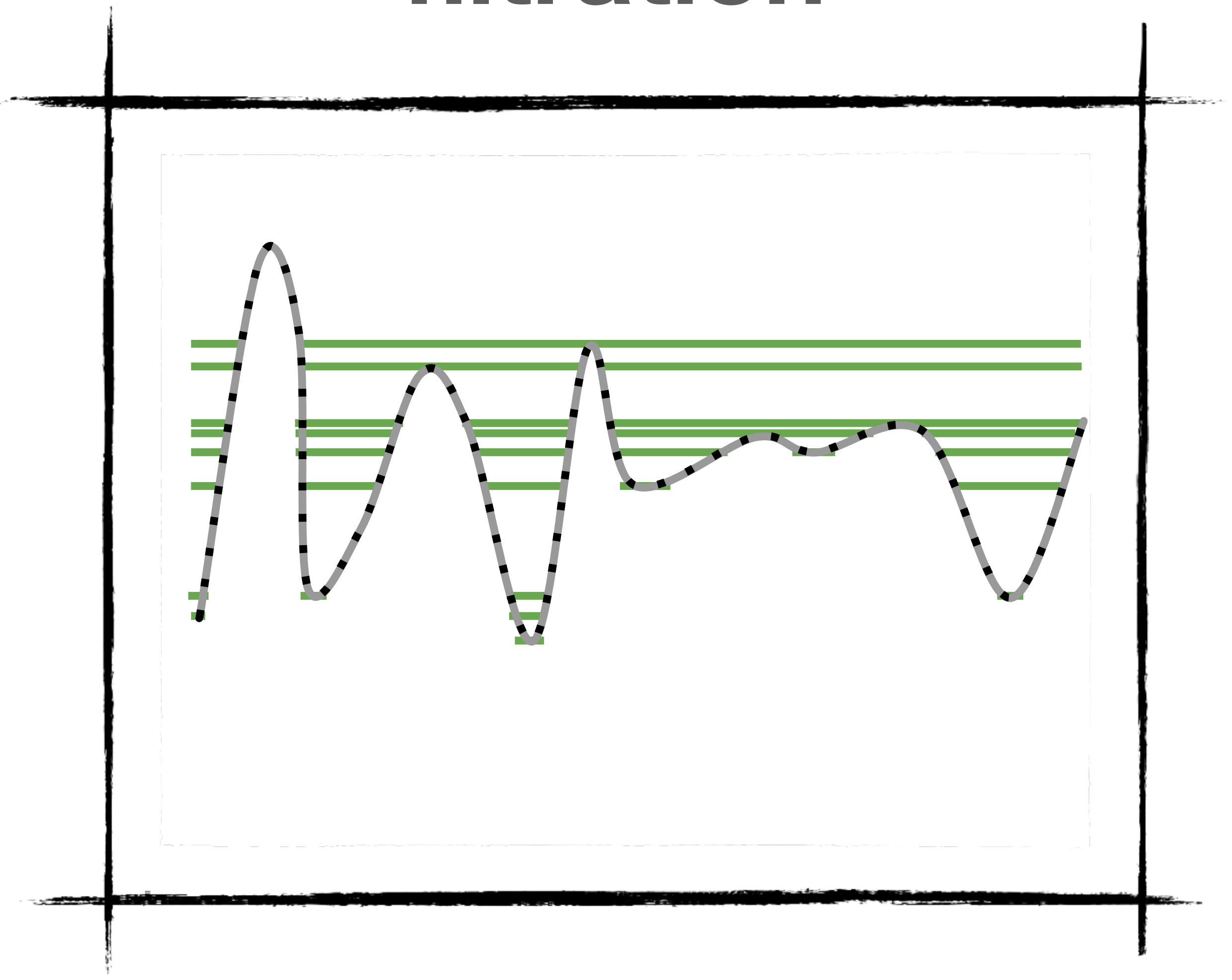
f
X
$\mathbb{R}$

“Ideal” Pipeline

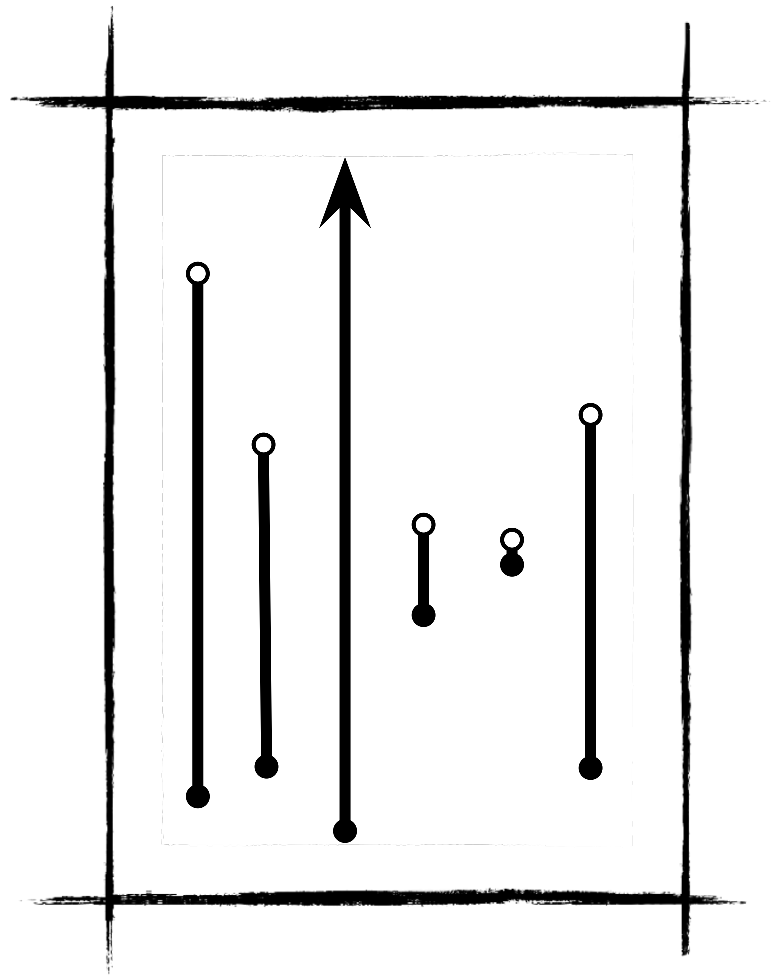
$$f : X \to \mathbb{R} \xrightarrow{\;\mathrm{Sub}\;} \mathrm{Sub}_f : \mathbb{R} \to \mathrm{Top} \xrightarrow{\;H\;} \mathrm{Pers}_f : \mathbb{R} \to \mathrm{Vec} \overset{\mathrm{Bar}}{\dashrightarrow} \mathrm{Bar}(f)$$

Input function | **Sublevel filtration** | **Persistence Module** | **Barcode**



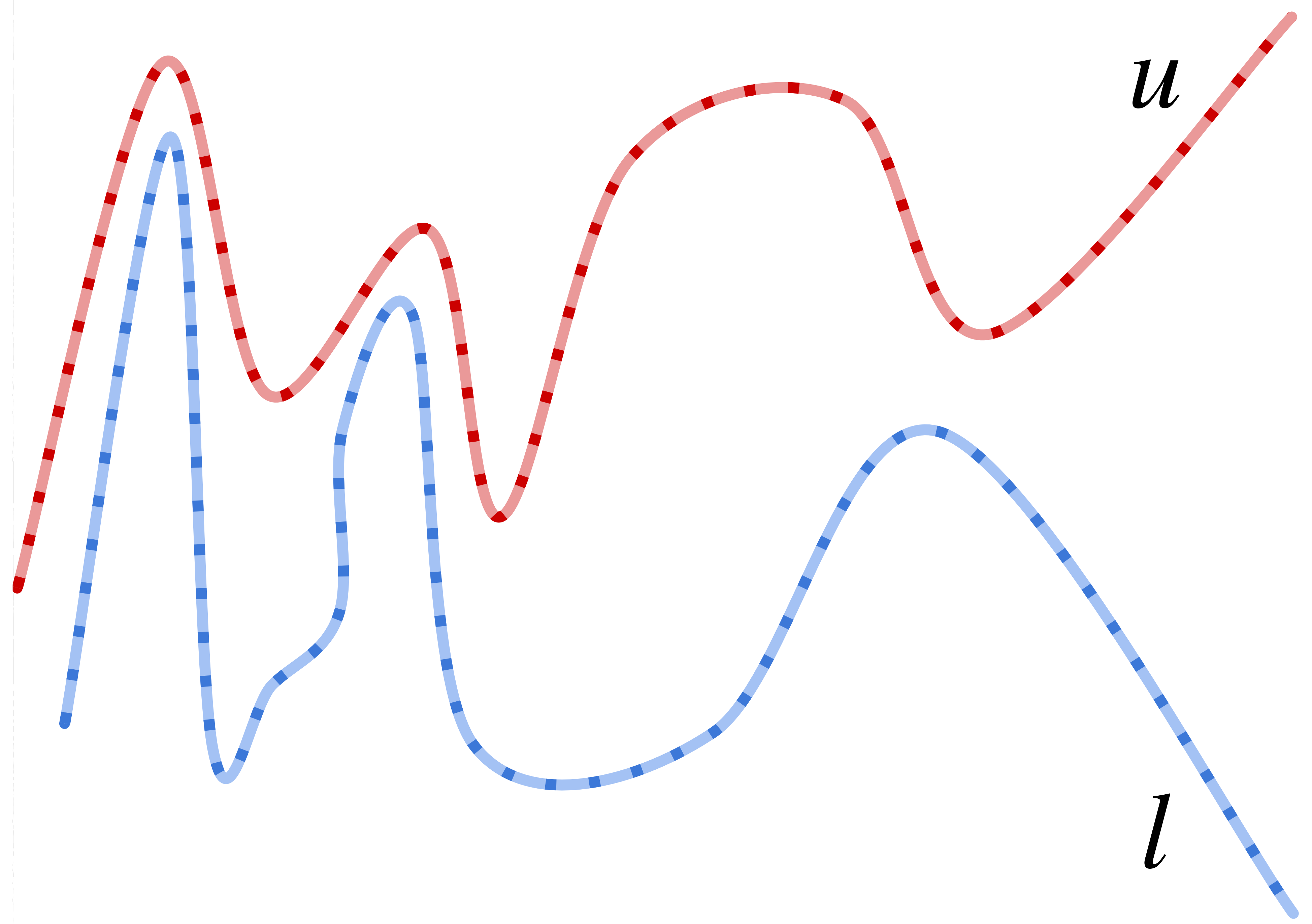


$$\begin{array}{c} H\mathrm{Sub}_f(x_n) \\ \uparrow \\ \cdots \\ \uparrow \\ H\mathrm{Sub}_f(x_1) \\ \uparrow \\ H\mathrm{Sub}_f(x_0) \end{array}$$



Barcodes of Unknown Functions

Given functions $u, l : X \to \mathbb{R}$ and $u \geq f \geq l$,
can we infer anything about $\mathrm{Pers}(f)$?



Things we know

$$u \geq f \geq l$$

$$\mathrm{Sub}_u \subseteq \mathrm{Sub}_f \subseteq \mathrm{Sub}_l$$

$$\mathrm{Pers}_u \to \mathrm{Pers}_f \to \mathrm{Pers}_l$$

Factorizations of Persistence Modules

Question

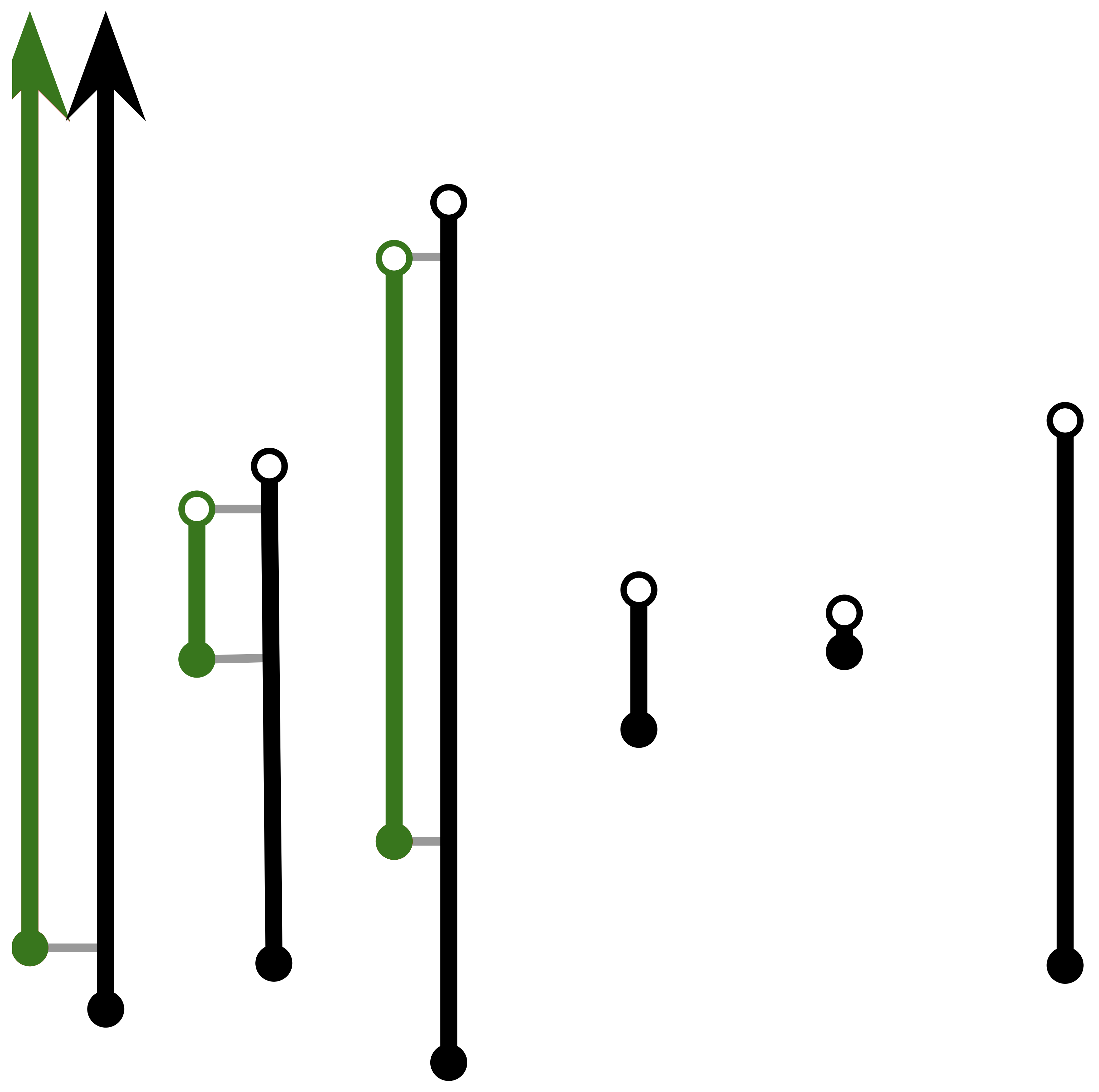
Given a factorization of persistence modules, $A \to X \to B$, what can we say about X?

Goal

Find a barcode that is **guaranteed** to be a portion of the barcode of X.

Sub-barcodes

Given a barcode B, a **sub-barcode** is a barcode that takes a (possibly empty) sub-interval of each bar in B.

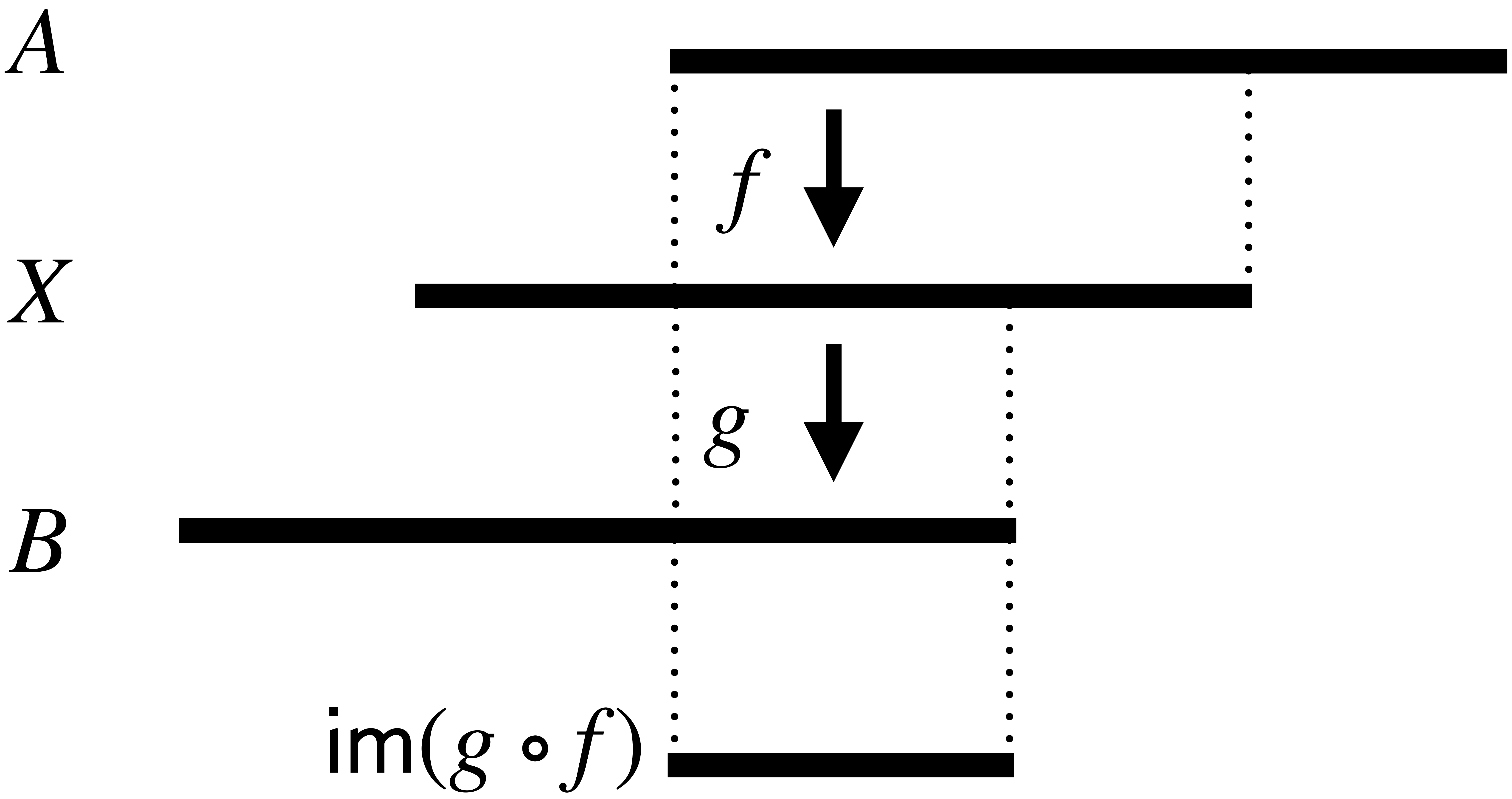


Observations

- Given a factorization of **vector spaces**, $A \xrightarrow{f} X \xrightarrow{g} B$:
  - $\mathrm{rk}(g \circ f) \leq \dim X$
  - There's an injective map, $\mathrm{im}(g \circ f) \to X$

- Given a factorization of **persistence modules**, $A \xrightarrow{f} X \xrightarrow{g} B$:
  - $\mathrm{rk}(g \circ f)_t \leq \dim X_t$ for all t
  - $\mathrm{im}(g \circ f)$ is **not** a submodule of X

Factoring Through X



There is no persistent module homomorphism between $\mathrm{im}(g \circ f)$ and X.

Induced Matchings

- **Induced matchings** (Bauer, Lesnick) show us that the barcode of the image comes from the intersection of matched bars.
- With some work we can see that the barcode of a persistence module homomorphism is always a sub-barcode

The Sub-barcode Theorem

A, X, B **are persistence modules.**

Given $A \xrightarrow{f} X \xrightarrow{g} B$**,**

$$\mathrm{Bar}(g \circ f) \sqsubseteq \mathrm{Bar}(X)$$

In More Generality

Given a factorization of persistence module homomorphisms, $G = \varphi_2 F \varphi_1$,

$$\mathsf{Bar}(G) \sqsubseteq \mathsf{Bar}(F).$$

$$\begin{tikzcd} M \arrow[r, "F"] & N \arrow[d, "\varphi_2"] \\ M' \arrow[u, "\varphi_1"] \arrow[r, "G"] & N' \end{tikzcd}$$

Sub-barcodes vs Ranks

As a partial order, sub-barcodes are more discriminating than the **rank invariant**

<u>Rank Invariant</u>

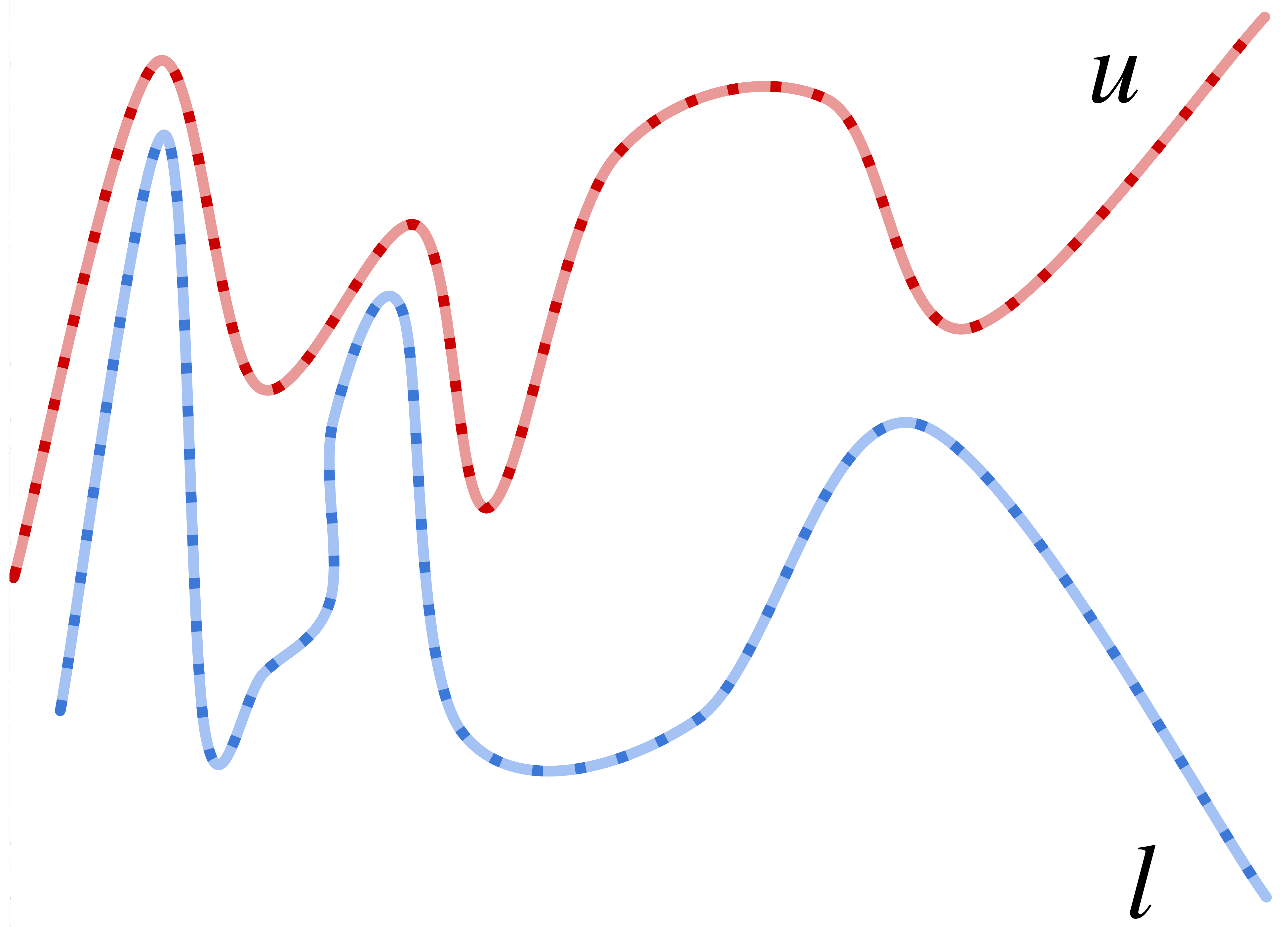
$\mathrm{rk}B \leq \mathrm{rk}A$

<u>Sub-Barcode</u>

$B \nsubseteq A$

Barcodes of Unknown Functions

Given functions $u, l : X \to \mathbb{R}$ and $u \geq f \geq l$,
can we infer anything about $\mathrm{Pers}(f)$?



<u>Things we know</u>

$$u \geq f \geq l$$

$$\mathrm{Sub}_u \subseteq \mathrm{Sub}_f \subseteq \mathrm{Sub}_l$$

$$\boxed{\mathrm{Pers}_u \to \mathrm{Pers}_f \to \mathrm{Pers}_l}$$

Can conclude that $\mathrm{Bar}(\mathrm{Pers}_u \to \mathrm{Pers}_l) \sqsubseteq \mathrm{Bar}(\mathrm{Pers}_f)$**.**

Sub-barcodes and Extensions

Suppose, $f : X \to \mathbb{R}$ **is an unknown Lipschitz function**

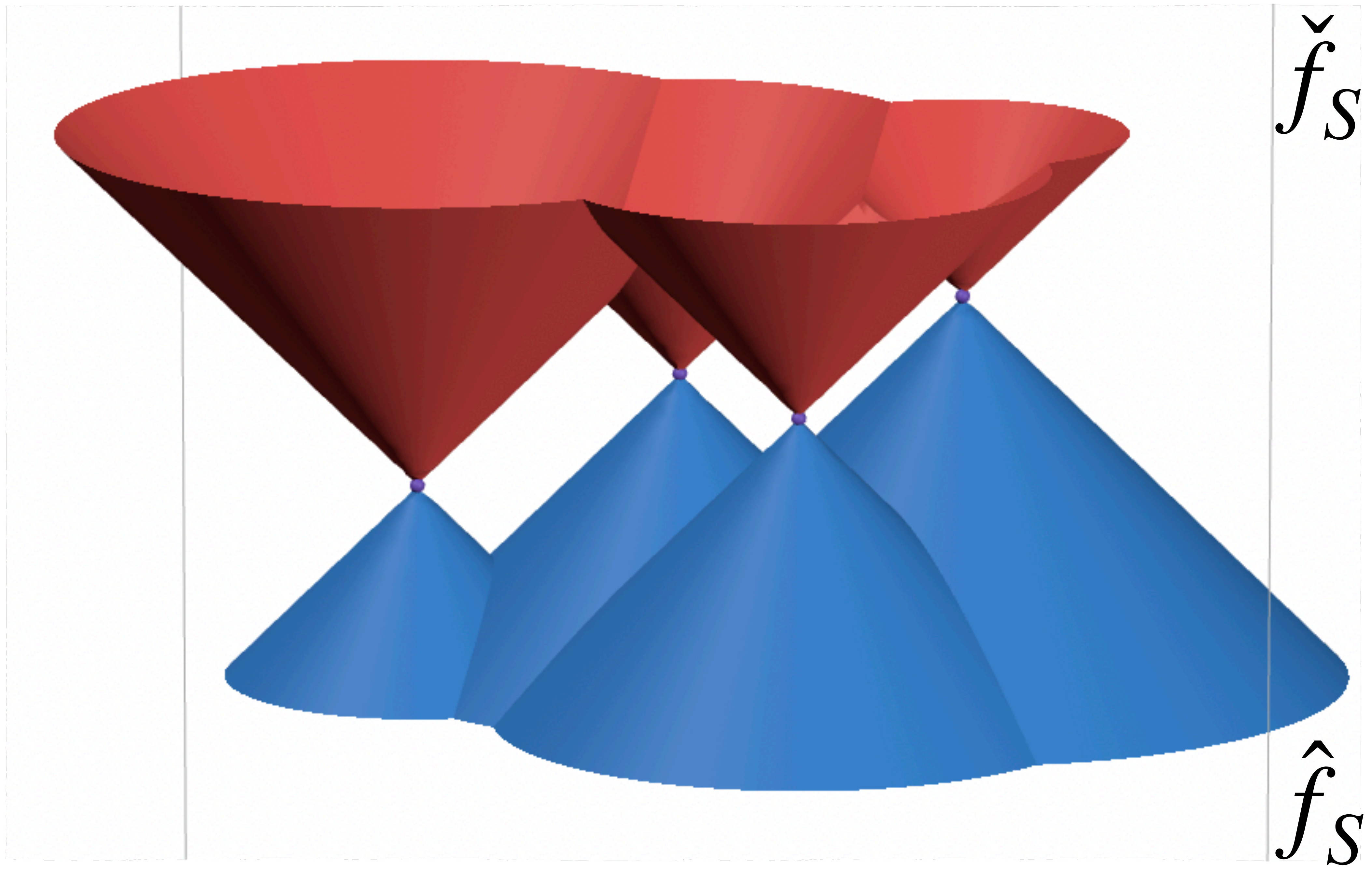
Given $f_{|S} : S \to \mathbb{R}$ **for** $S \subset X$ **a finite sample of** X**…**

A function $f : X \to \mathbb{R}$ **is** λ**-Lipschitz**

if for all $x, y \in X$**, we have**

$$f(y) - \lambda d(x, y) \leq f(x) \leq f(y) + \lambda d(x, y)$$

Lipschitz Extensions



$f : X \to \mathbb{R}$:

an unknown Lipschitz function

$S \subset X$:

a finite sample of X

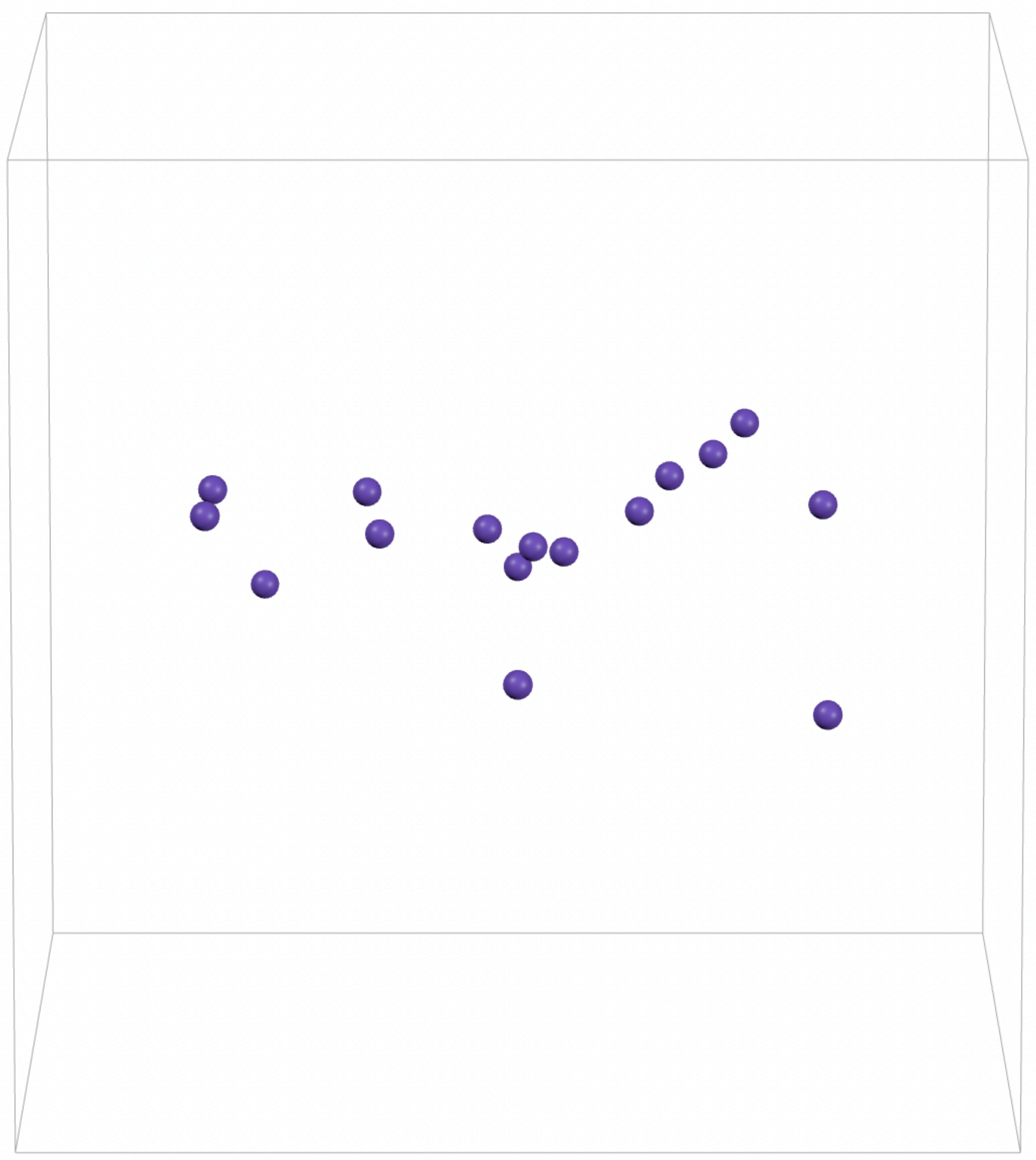
Maximum Lipschitz extension

$$\check{f}_S(x) := \min_{s \in S} f(s) + d(x, s)$$

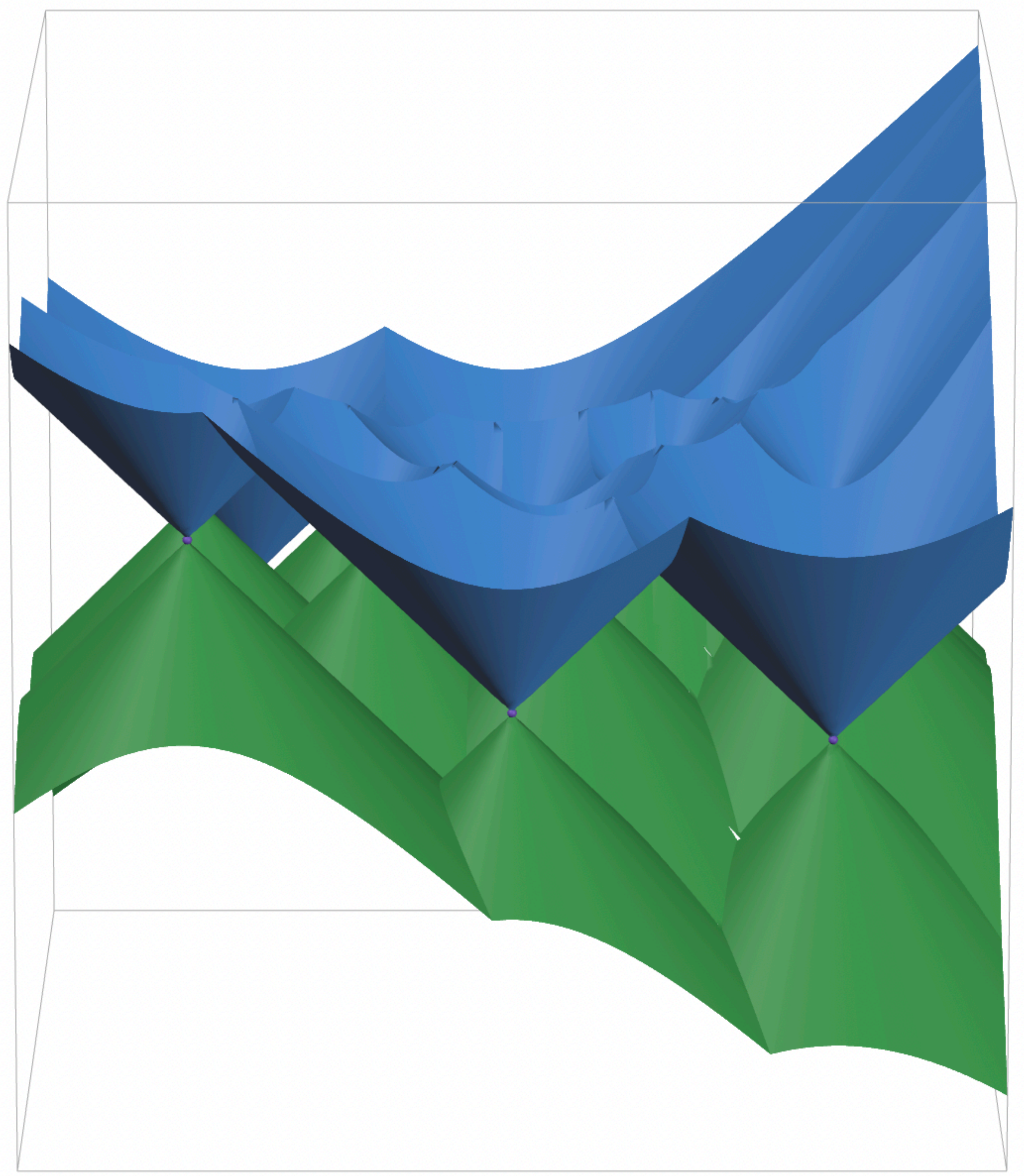
Minimum Lipschitz extension

$$\hat{f}_S(x) := \max_{s \in S} f(s) + d(x, s)$$

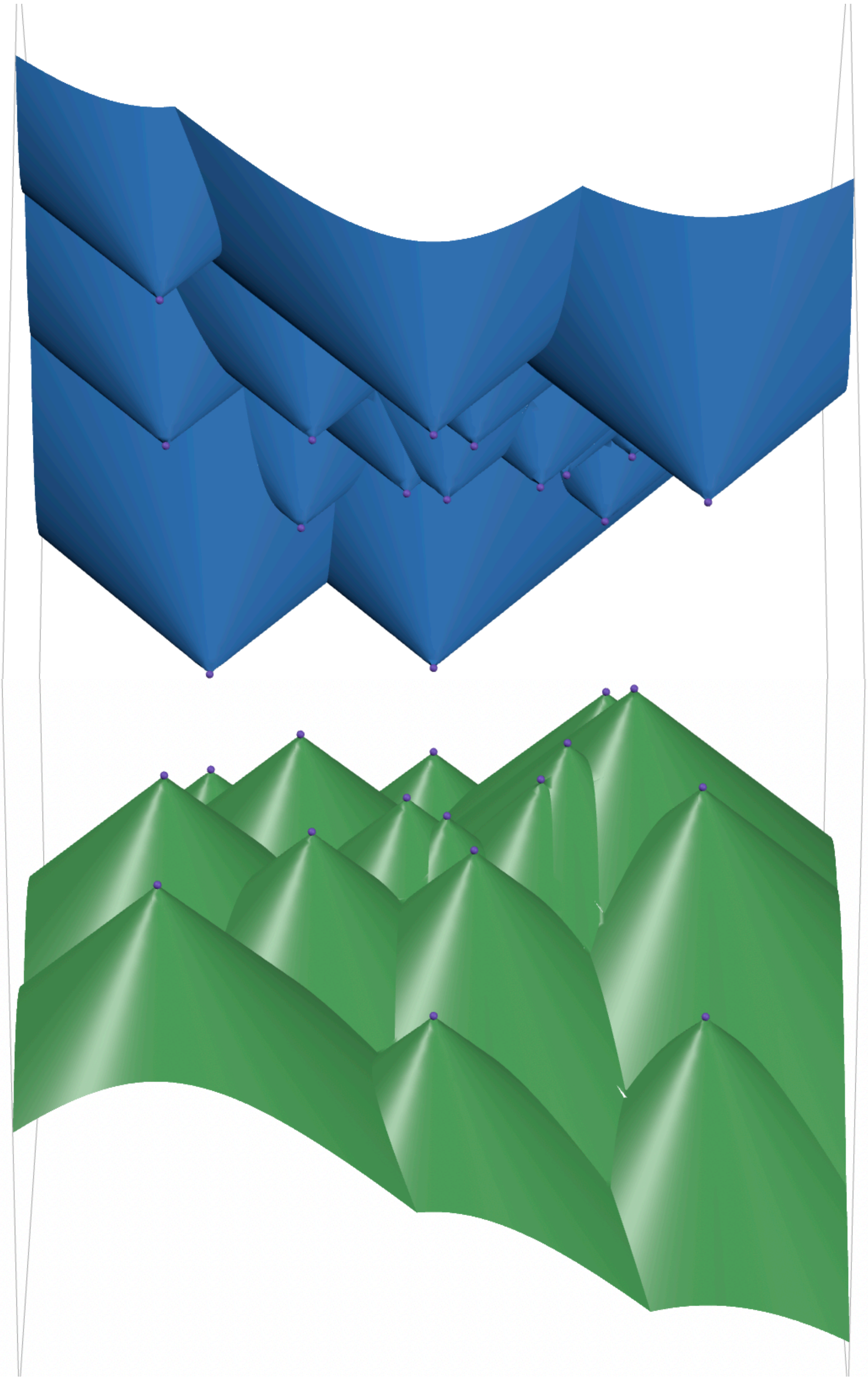
Lipschitz Extensions



Sample $S \subset X$

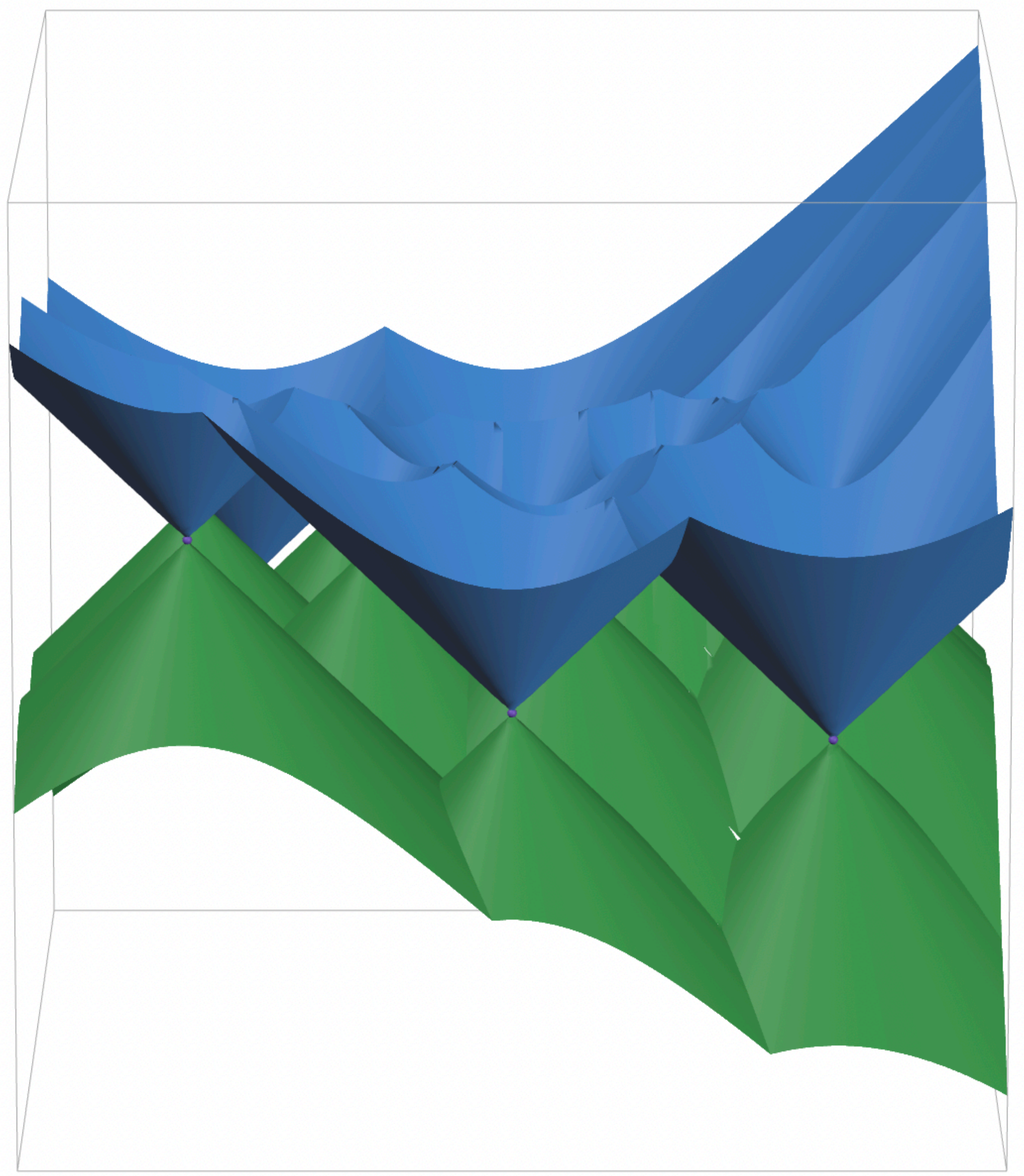


Lipschitz extension on S

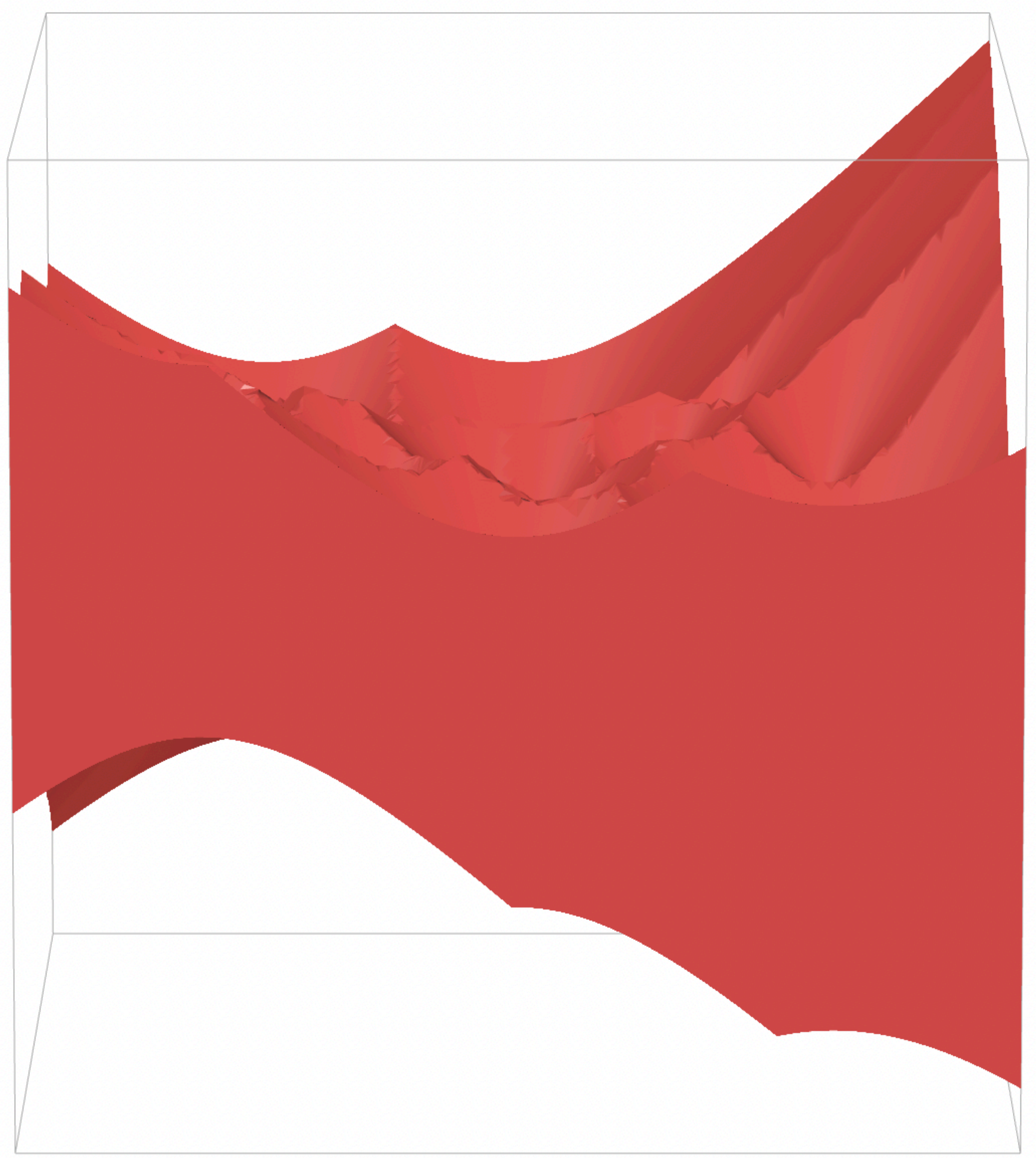


"Exploded" view

Lipschitz Extensions



Lipschitz extension on S

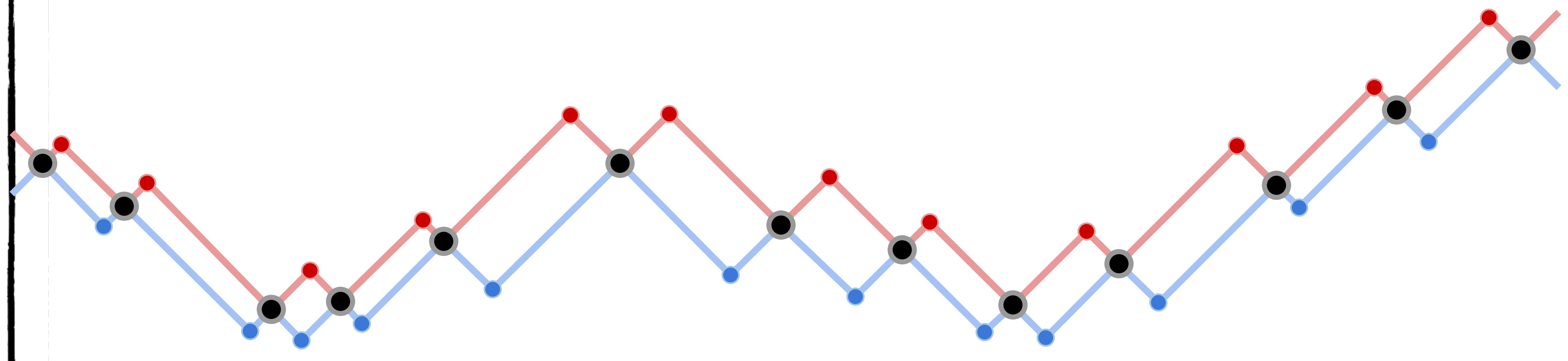


Possible function values

Lipschitz Extensions

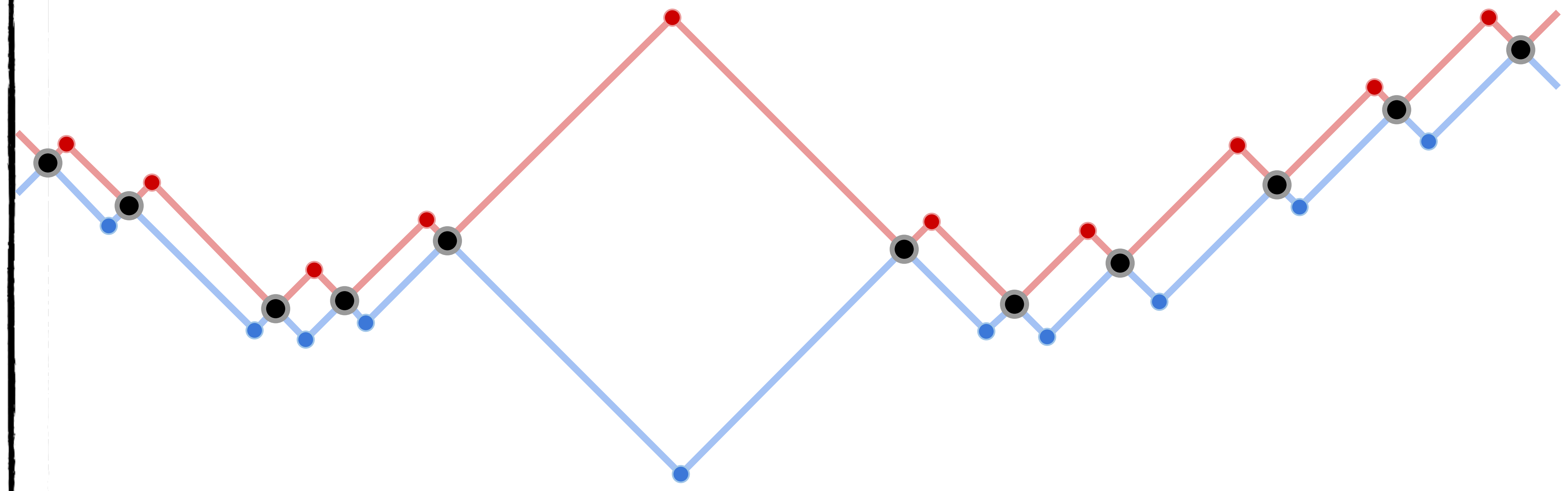
With good sampling assumptions:

Stability allows us to find an ε-close sub-barcode.

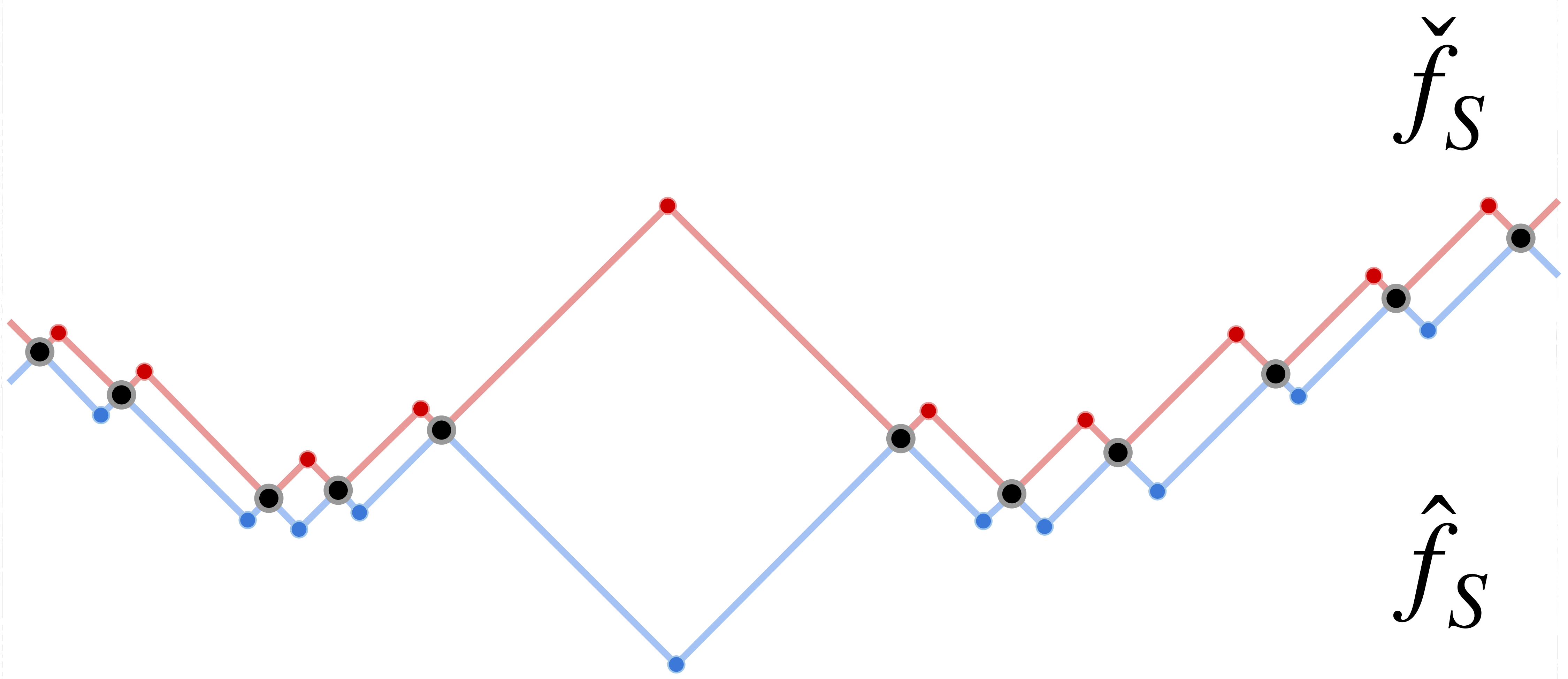


Without any sampling assumptions:

Using stability alone, all of the bars in the barcode could be considered “noise”.



Barcodes of Unknown [Lipschitz] Functions



Things we know

$$\check{f}_S \geq f \geq \hat{f}_S$$

$$\mathrm{Sub}_{\check{f}_S} \subseteq \mathrm{Sub}_f \subseteq \mathrm{Sub}_{\hat{f}_S}$$

$$\mathrm{Pers}_{\check{f}_S} \to \mathrm{Pers}_f \to \mathrm{Pers}_{\hat{f}_S}$$

$$\boxed{\mathrm{Bar}(\mathrm{Pers}_{\check{f}_S} \to \mathrm{Pers}_{\hat{f}_S}) \sqsubseteq \mathrm{Bar}(\mathrm{Pers}_f)}$$

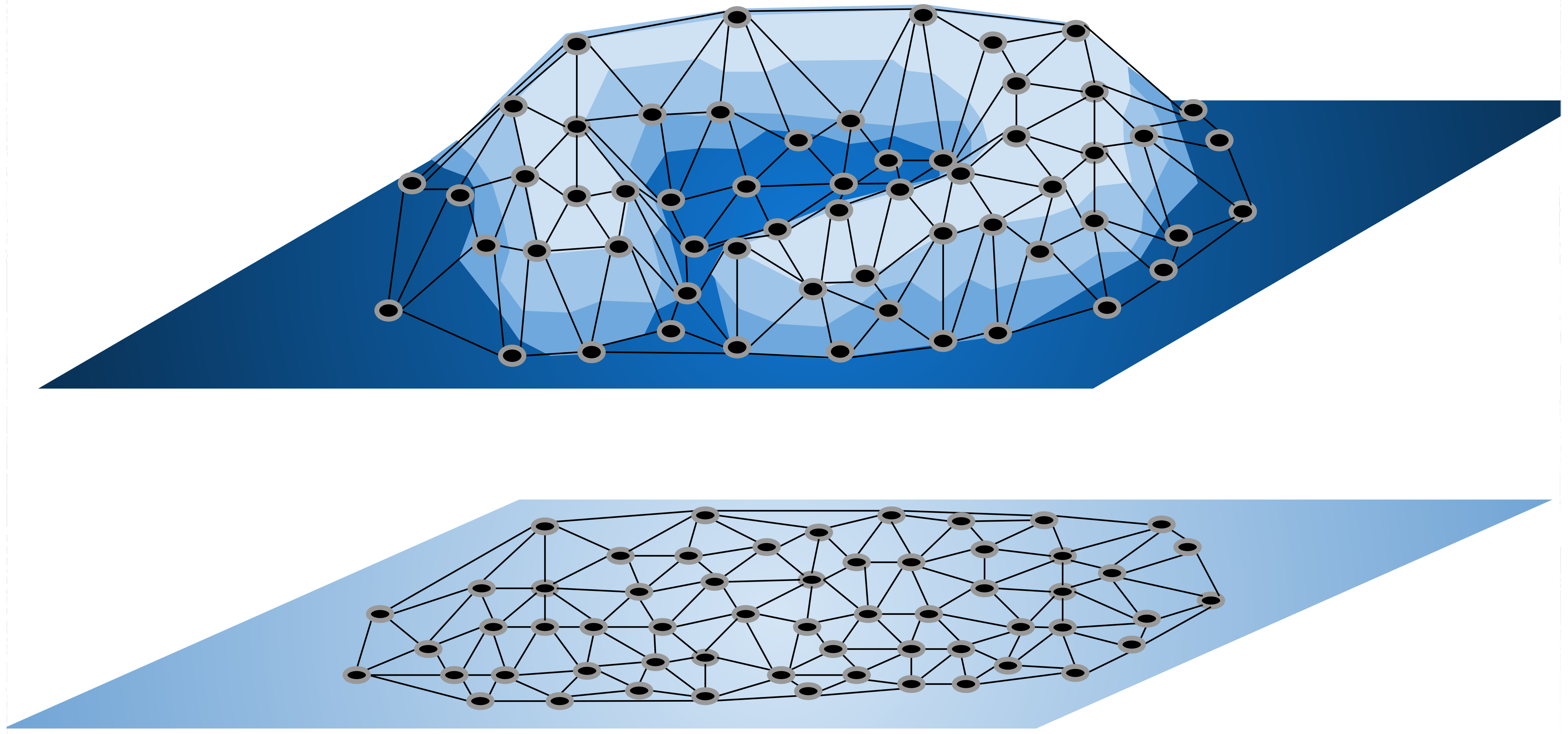
Sub-barcodes and Discretization

- Suppose, $f : X \to \mathbb{R}$ is an unknown Lipschitz function.
- The domain X is unknown, but we have an ε-sample $S \subseteq X$.
- Given $f_{|S} : S \to \mathbb{R}$ for $S \subset X$ a finite sample of X.

Discretization in $\mathbb{R}^d$

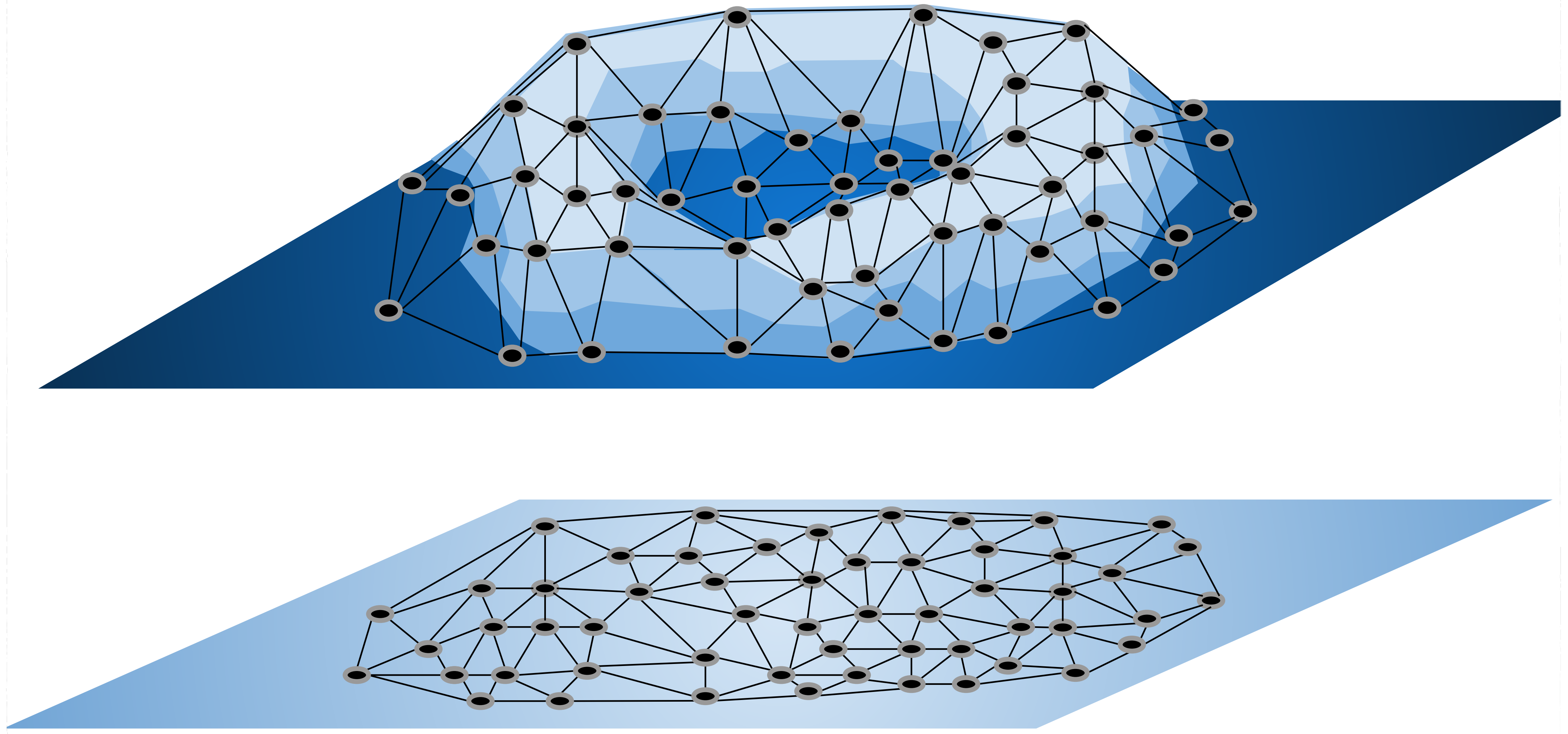
Good sampling:

We can use a **Delaunay triangulation** to estimate persistence



Slightly different sampling:

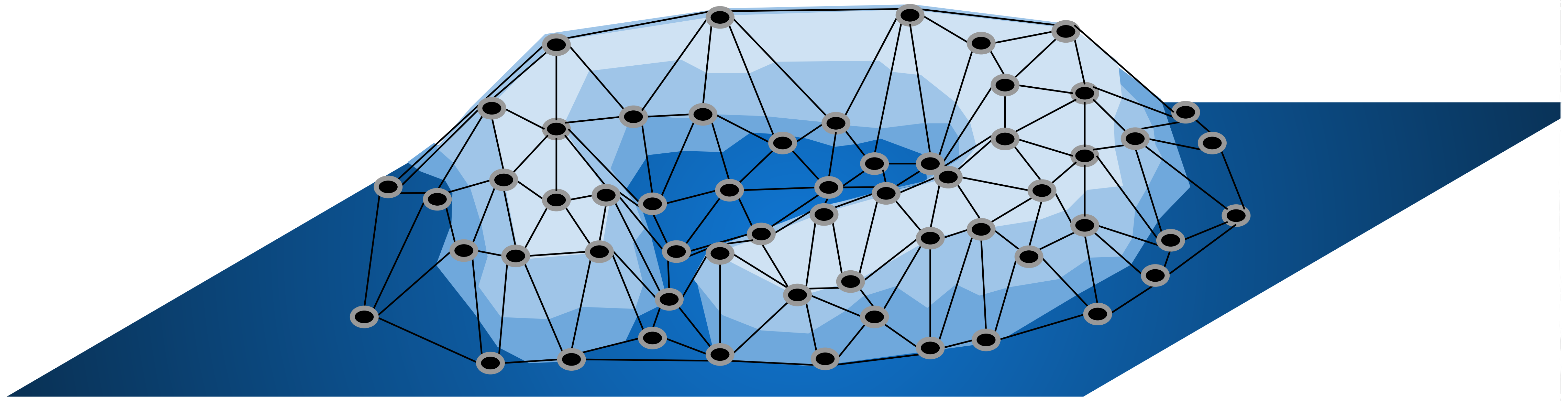
Using a Delaunay triangulation we might miss or even “hallucinate” features



Discretization in $\mathbb{R}^d$

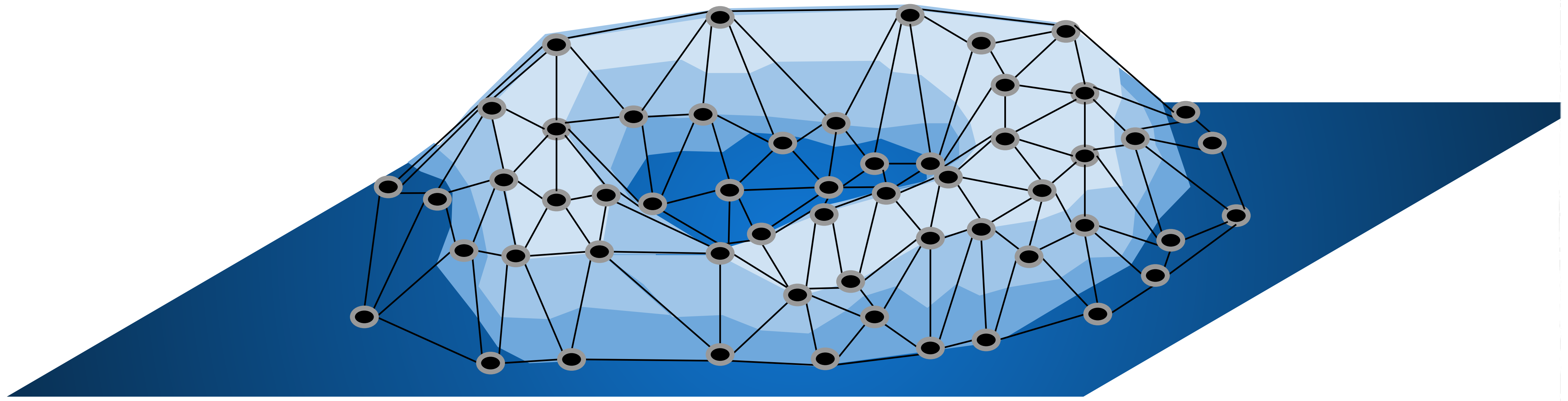
Good sampling:

We can use a **Delaunay triangulation** to estimate persistence



Slightly different sampling:

Using a linear extension we might miss or even "hallucinate" features



Using the Delaunay filtration to define Lipschitz extensions u and l,

$$\mathrm{Bar}(\mathrm{Pers}_u \to \mathrm{Pers}_l) \sqsubseteq \mathrm{Bar}(\mathrm{Pers}_f)$$

“Semi-Supervised” TDA

- **Suppose, $f : X \to \mathbb{R}$ is an unknown Lipschitz function.**
- **The domain X is unknown; we have an ε-sample $S \subseteq X$.**
- **Given $f_{|P} : P \to \mathbb{R}$ for $P \subseteq S \subset X$ a finite sample of X.**
- **u and l are defined using max- and min-Lipschitz extensions of P on S**

We can even remove sampling assumptions on X by using the barycentric sub-division of the Delaunay filtration of P

$\mathrm{Bar}(\mathrm{Pers}_u \to \mathrm{Pers}_l) \sqsubseteq \mathrm{Bar}(\mathrm{Pers}_f)$ **and**
$d_B(\mathrm{Bar}(\mathrm{Pers}_u \to \mathrm{Pers}_l), \mathrm{Bar}(\mathrm{Pers}_{\check{\mathrm{f}}_\mathrm{S}} \to \mathrm{Pers}_{\hat{\mathrm{f}}_\mathrm{S}})) \leq \varepsilon$

Summary

- We can compare barcodes from factorizations of persistence module homomorphisms.
- Sub-barcodes are more discriminating than the rank invariant.
- The output is guaranteed to be a sub-barcode of the unknown function.
- It is possible to remove sampling assumptions for the unknown function.