ESQC 2024

Mathematical Methods Lecture 4

By Simen Kvaal

Where to find the material

SCAN ME

- Alternative 1:
  - www.esqc.org, go to“lectures”
  - Find links there
- Alternative 2:
  - Scan QR code
  - simenkva.github.io/esqc_material

Vector calculus

Functions, integration and differentiation, in one and several variables

Functions of several variables

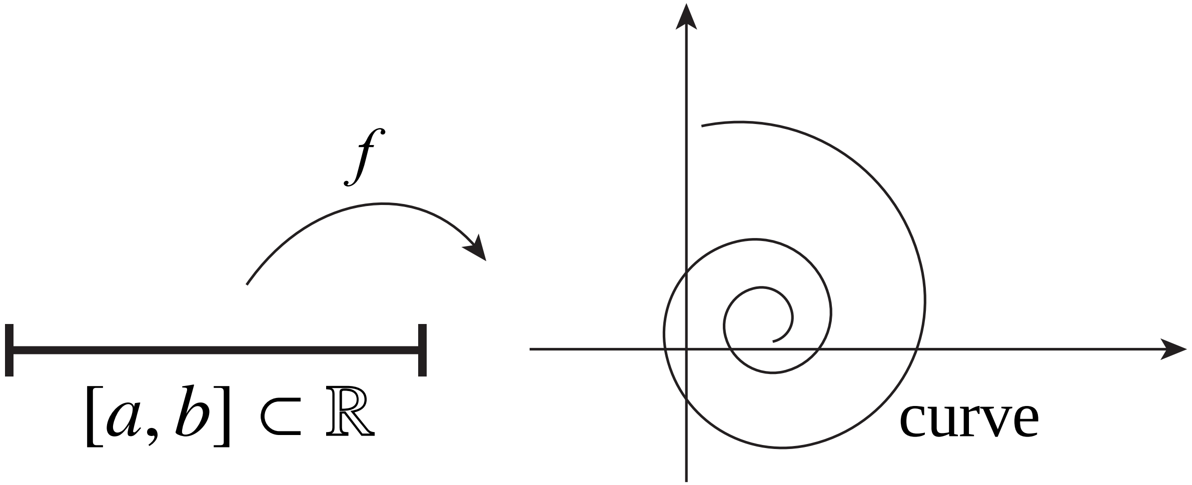
I am a vector valued function!

- We turn to the study of *vector valued* functions

$$f : \mathbb{R}^n \to \mathbb{R}^m \qquad f : \Omega(\subset \mathbb{R}^n) \to \mathbb{R}^m$$

- Such as paths

$$f : \Omega(\subset \mathbb{R}^1) \to \mathbb{R}^m$$



... scalar-valued functions

$$f : \Omega(\subset \mathbb{R}^n) \to \mathbb{R}^1$$

Compared to yesterday …

- We studied Banach spaces of functions:

$$V = \{f : \Omega \to \mathbb{F} \mid ||f|| < +\infty\}$$

- Metric measured *distance between functions*
- Now, we study the *function itself*:

$$f : \Omega(\subset \mathbb{R}^n) \to \mathbb{R}^m$$

- Now the metric measures *distance in Euclidean space*

In quantum chemistry

Hartree-Fock, for example

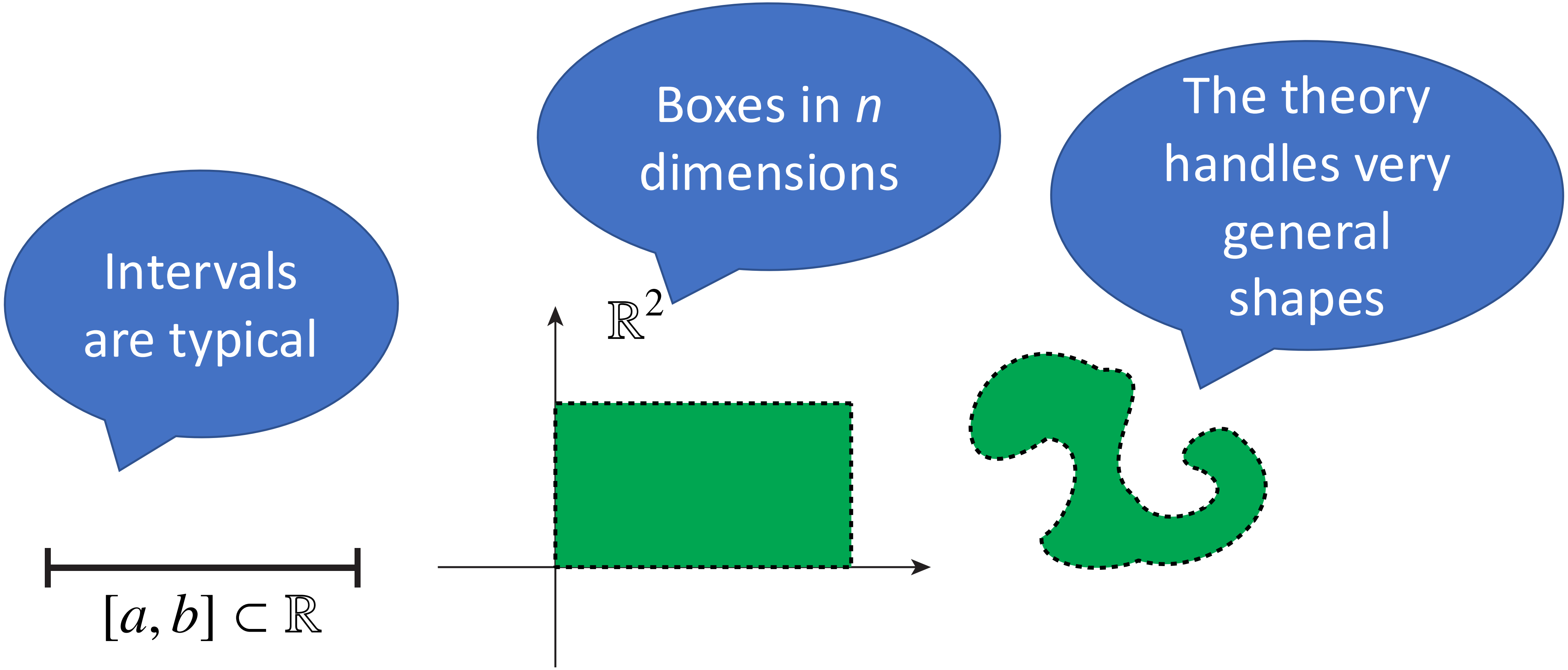
- *Most* methods can be formulated as:

$$E : \Omega(\subset \mathbb{F}^n) \to \mathbb{R}, \quad \mathbf{x} \mapsto \text{energy function}$$

Find $\mathbf{x} \in \Omega$ such that

$$E(\mathbf{x}) = \min!, \quad \text{i.e.,} \quad \nabla E(\mathbf{x}) = 0.$$

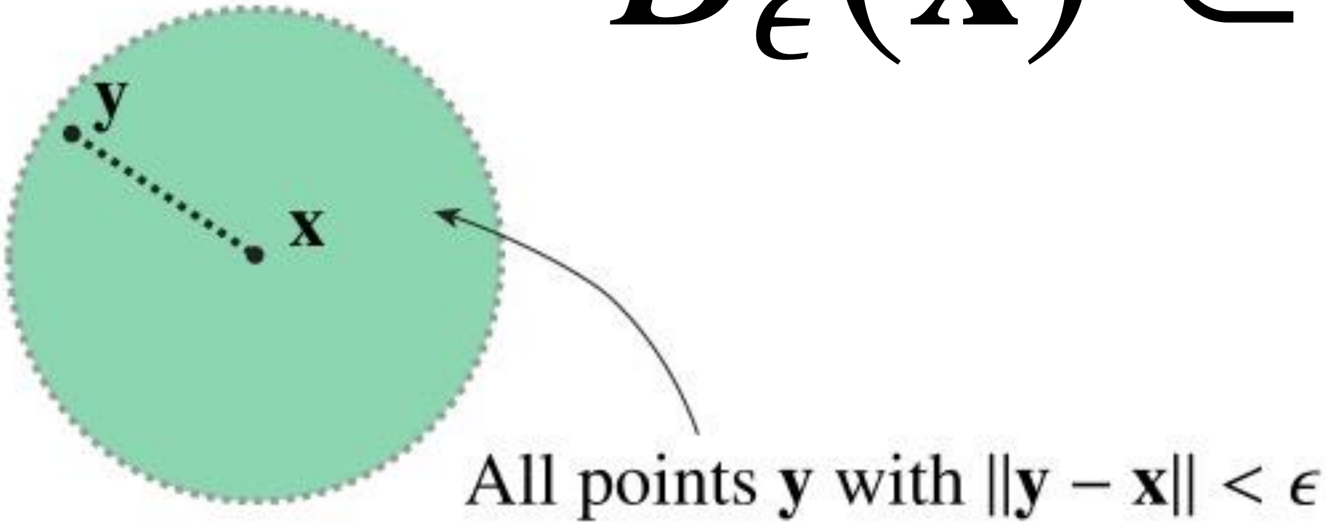
A typical domain Ω



Topology of Euclidean space

- Definition of an *epsilon-ball*

$$B_\epsilon(\mathbf{x}) \subset \mathbb{R}^n$$



Definition : Topologically important sets

1. A subset $S \subset \mathbb{R}^n$ is called *open* if, for every $\mathbf{x} \in S$, there is an $\varepsilon > 0$ such that $B_\varepsilon(\mathbf{x}) \subset S$.

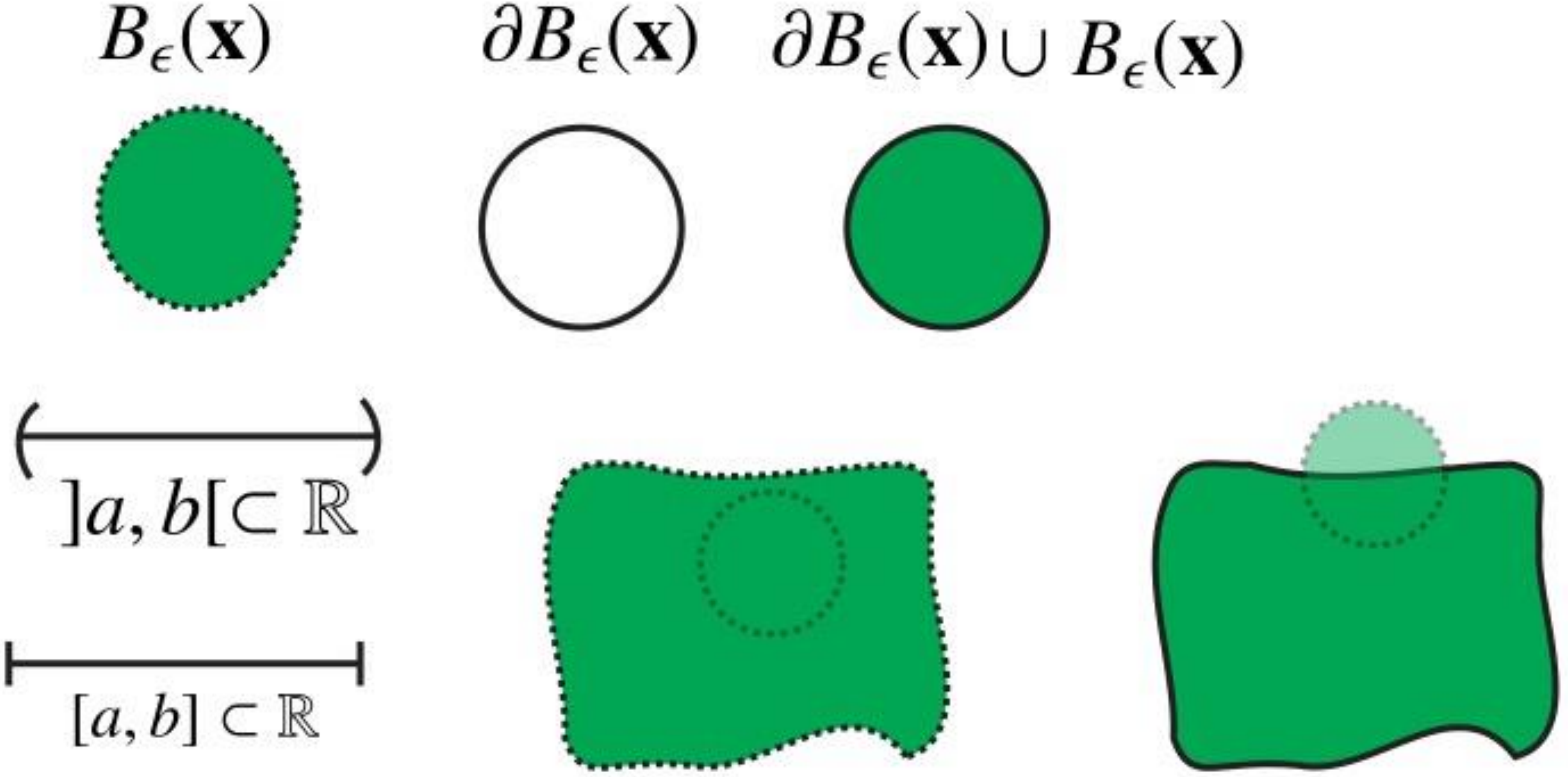
2. A subset S is called *closed* if $S^{\complement} = \mathbb{R}^n \setminus S$ is open.

3. The *closure* $\mathrm{cl}(S)$ is the smallest closed set that contains S.

4. The *interior* $\mathrm{int}(S)$ is the set of all those $\mathbf{x} \in S$ around which there exists an ε-ball in S

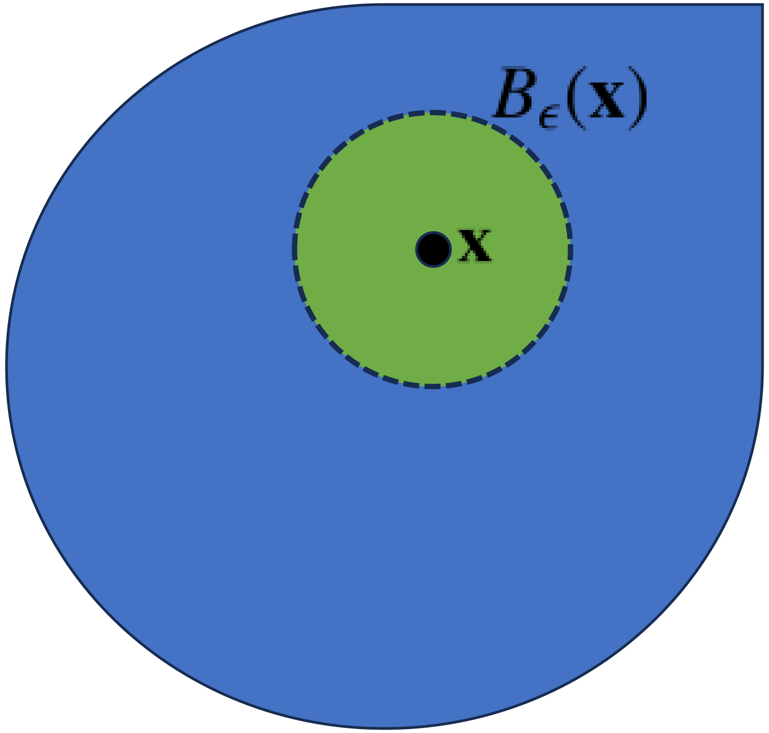
5. The *boundary* ∂S is the intersection $\mathrm{cl}(S^{\complement}) \cap \mathrm{cl}(S) = S \setminus \mathrm{int}(S)$

Examples

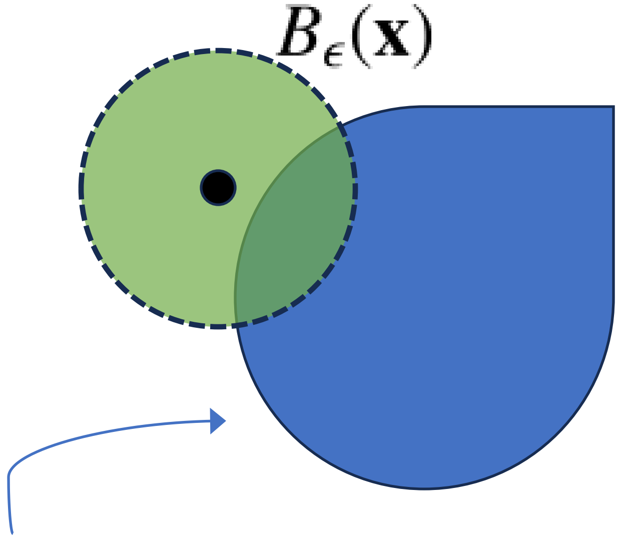

$B_\epsilon(\mathbf{x})$
$\partial B_\epsilon(\mathbf{x})$
$\partial B_\epsilon(\mathbf{x}) \cup B_\epsilon(\mathbf{x})$
$]a, b[\subset \mathbb{R}$
$[a, b] \subset \mathbb{R}$

Neighborhood of x

- Any set containing **x** and an open ball around **x**



neighborhood containing **x**



NOT a neighborhood containing **x**

Definition : Limit

Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, where Ω is open. Let $\mathbf{x}_0 \in \Omega \cup \partial\Omega$, and let N be a neighborhood of $\mathbf{b} \in \mathbb{R}^m$.

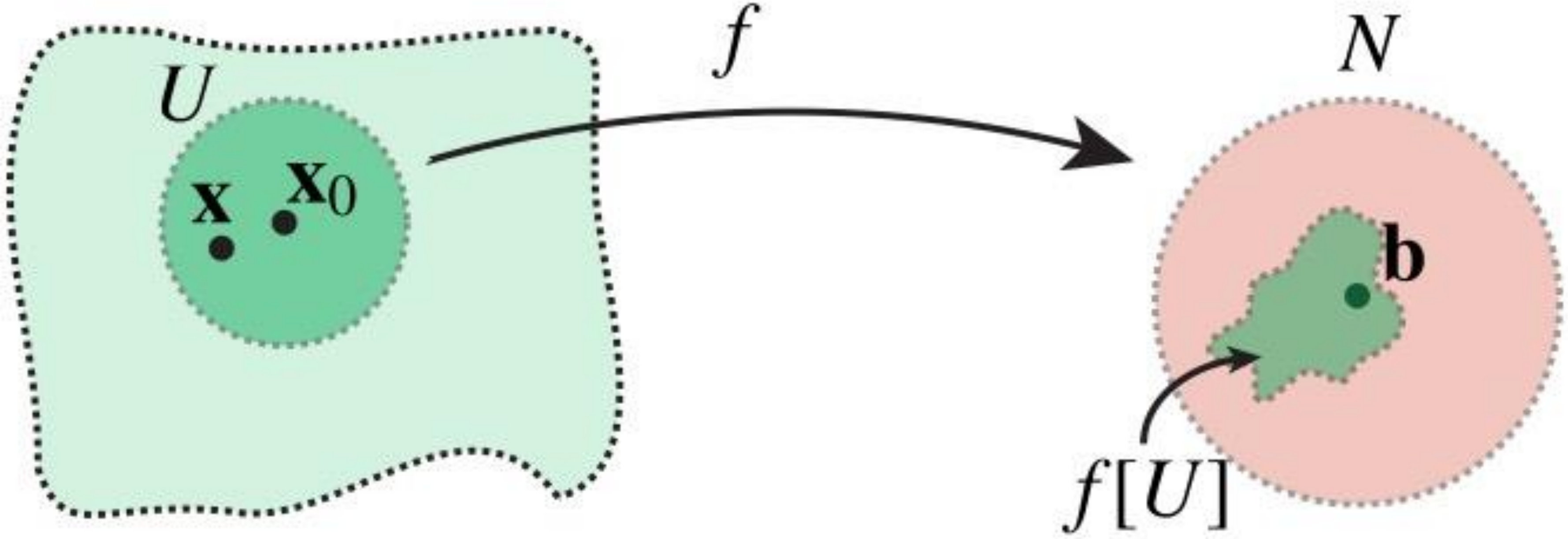
We say that f is *eventually in* N *as* $\mathbf{x}$ *approaches* $\mathbf{x}_0$, if there exists a neighborhood U of $\mathbf{x}_0$, such that $\mathbf{x} \in U$ but $\mathbf{x} \neq \mathbf{x}_0$ and $\mathbf{x} \in \Omega$ imply $f(x) \in N$.

We say that $f(\mathbf{x})$ *approaches* $\mathbf{b}$ *as* $\mathbf{x}$ *approaches* $\mathbf{x}_0$,

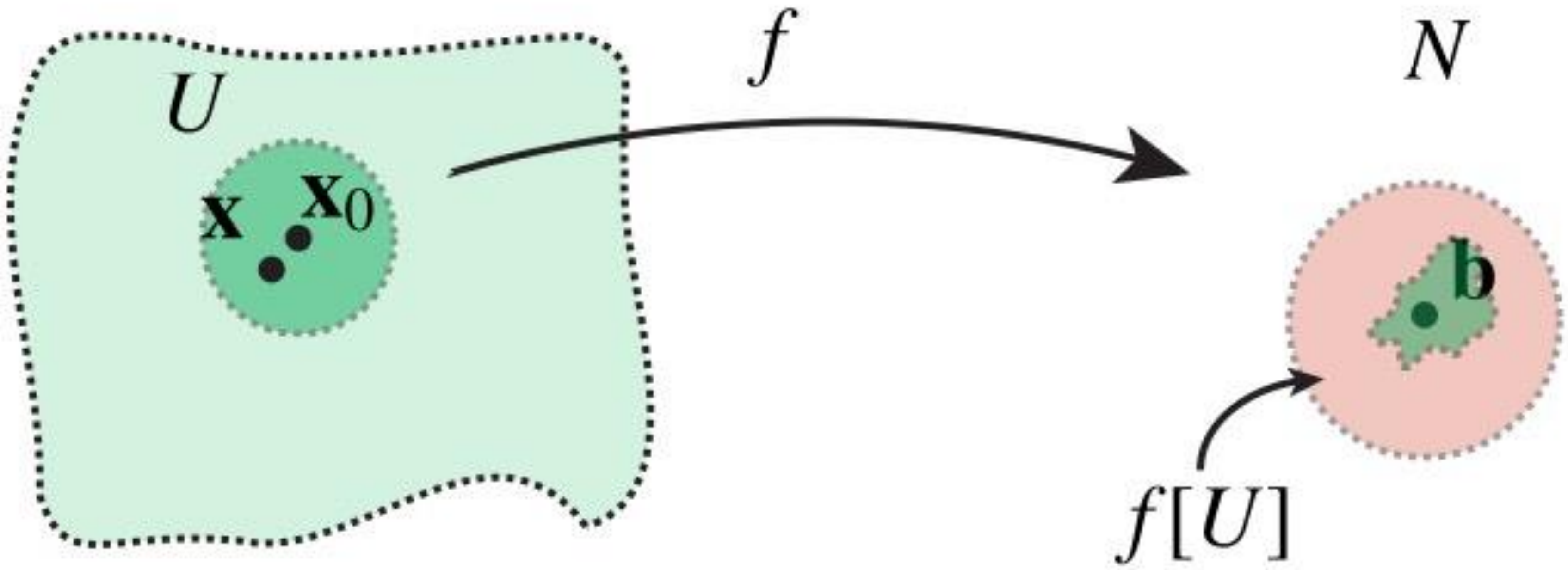
$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b} \quad \text{or} \quad f(\mathbf{x}) \to \mathbf{b} \text{ as } \mathbf{x} \to \mathbf{x}_0, \tag{1}$$

when, given any neighborhood N of $\mathbf{b}$, f is eventually in N as $\mathbf{x}$ approaches $\mathbf{x}_0$.

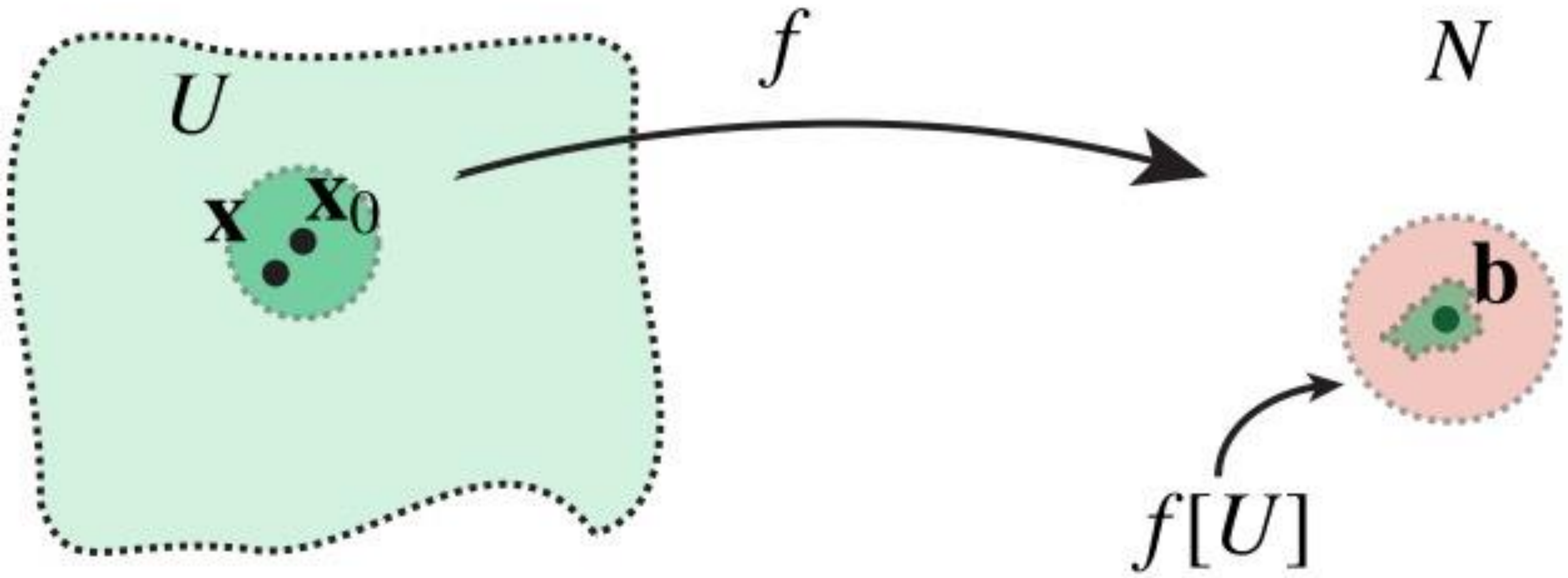
Intiuition



Intiuition



Intiuition



Definition : Continuity

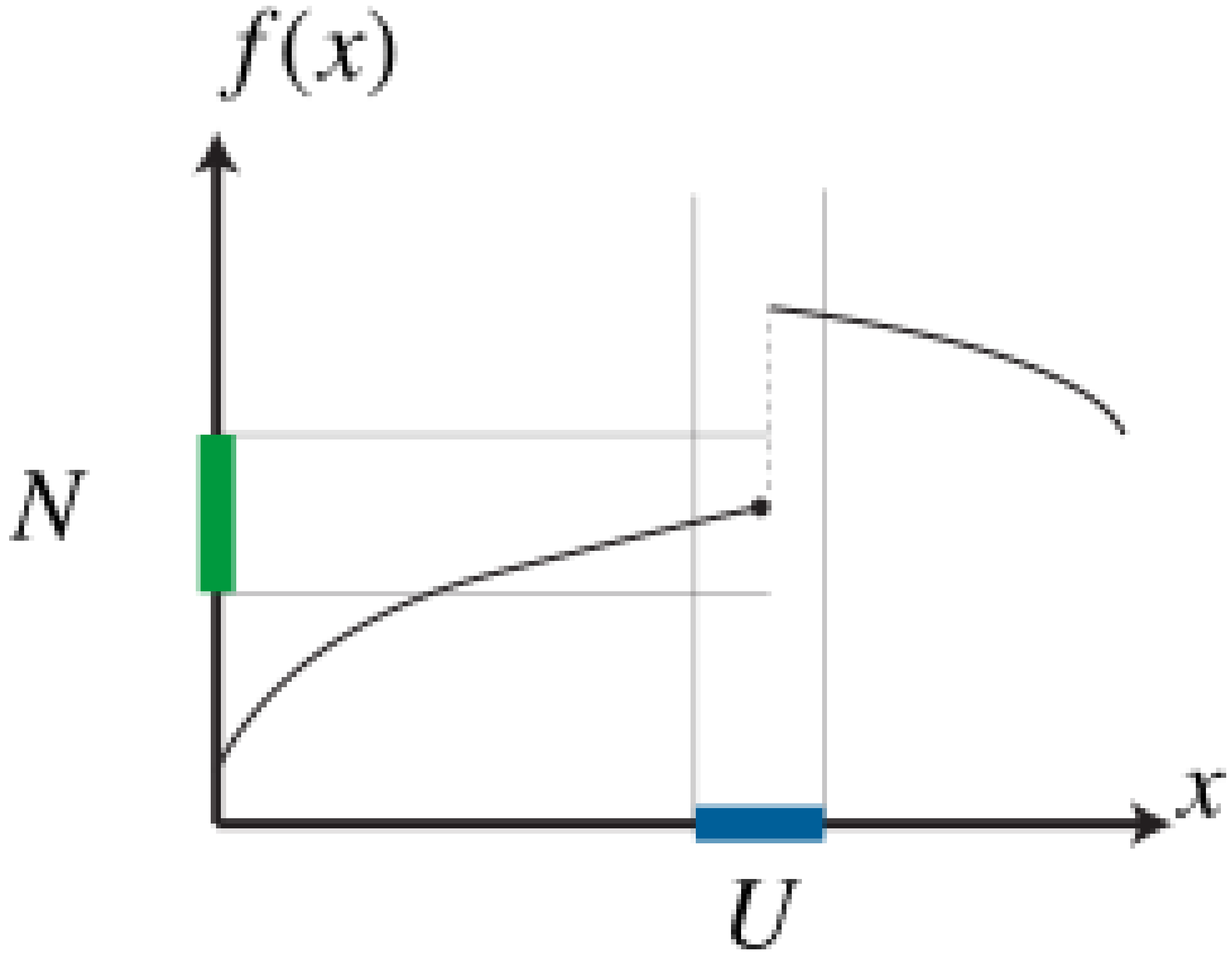
Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$. Let $\mathbf{x}_0 \in \Omega$. We say that f is *continuous at* $\mathbf{x}_0$ if

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

Multidimensional version of "unbroken graph"

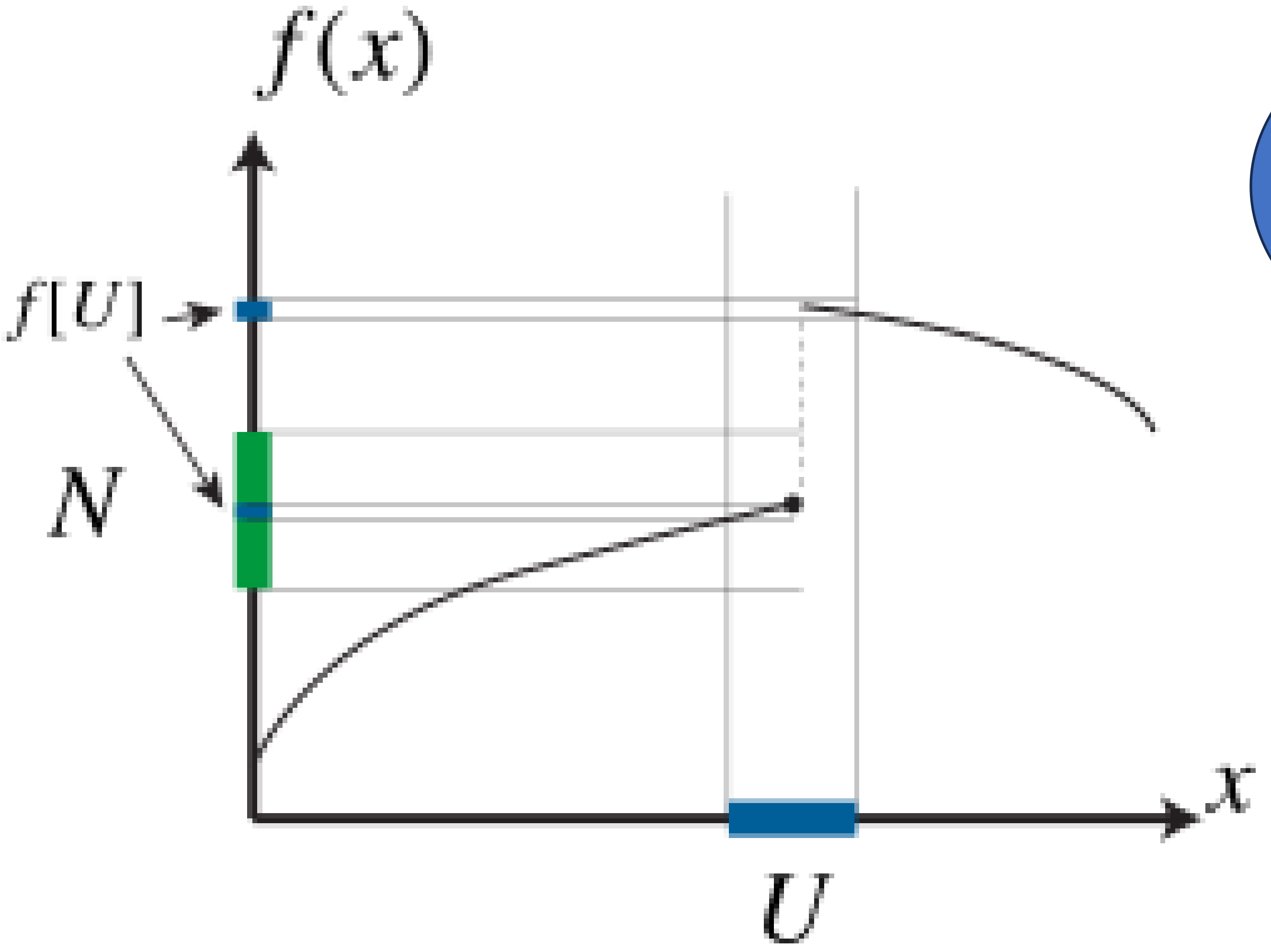
Discontinuous in 1d

$f : \mathbb{R} \to \mathbb{R}$ makes a jump



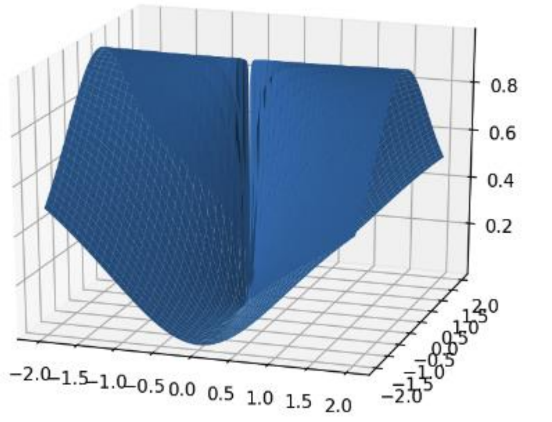
Discontinuous in 1d

$f : \mathbb{R} \to \mathbb{R}$ makes a jump



However we choose U, it will betorn apart by *f*

Example



- Is the following function continuous at (0,0)?

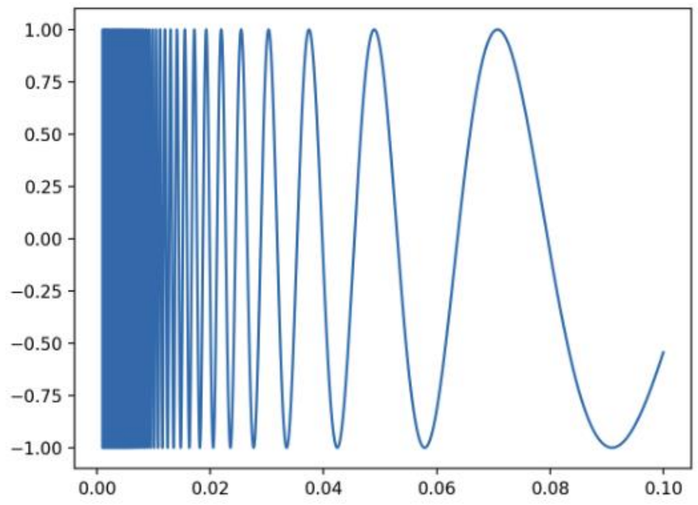
$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \frac{x^2}{x^2 + y^2}$$

- No, because the limit does not exist.
- Different limit candidates if we approach from different directions
- *The definition of limit is designed to detect such situations*

More subtle, in 1D

- Is the following function continuous?

$$f : (0, 1) \to \mathbb{R}, \quad x \mapsto \sin(1/x)$$



- Yes, since we do not include 0 in the domain
- But f has no limit at $x = 0$

Theorem : Properties of continuous functions

Let $f, g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ be functions with a common domain Ω, continuous at $\mathbf{x}_0$: Then:

1. $f + g$ and αf for any $\alpha \in \mathbb{R}$ are continuous at $\mathbf{x}_0$.
2. In the scalar-valued case $m = 1$, the product fg is continuous at $\mathbf{x}_0$
3. If $f \neq 0$ in all of Ω, then $1/f$ is continuous at $\mathbf{x}_0$
4. The component functions $f_i : \Omega \to \mathbb{R}$ are all continuous at $\mathbf{x}_0$. The converse is also true.

Theorem : Compositions of functions

Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ be continuous at $\mathbf{x}_0 \in \Omega$, and $g : \Omega' \subset \mathbb{R}^m \to \mathbb{R}^o$. Suppose $f[\Omega] \subset \Omega'$, and let g be continuous at $\mathbf{y}_0 = f(\mathbf{x}_0)$. Then $h : \Omega \subset \mathbb{R}^n \to \mathbb{R}^o$,

$$h(\mathbf{x}) = g(f(\mathbf{x}_0))$$

is continuous at $\mathbf{x}_0$.

These two theorems can be used to decide continuity of very complicated functions, once simpler functions are proven to be continuous

Examples

- polynomials in any variable
- exponential function
- sine, cosine …
- any composition of such
- careful with division!

$$f : \mathbb{R}^3 \to \mathbb{R}, \qquad f(\mathbf{x}) = \exp[-||\mathbf{x}||^4 + \cos(x_1)]x_1 x_2 x_3^4 (1 + x_1^2)^{-2}$$

Definition : Partial derivative

Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be a scalar-valued function, Ω open. The *partial derivatives* with respect to the variable x_i are defined by

$$\frac{\partial}{\partial x_i} f(\mathbf{x}) = \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{e}_i) - f(\vec{x})}{h}$$

if the limit exists.

In the case $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, the the partial derivatives are defined componentwise, i.e.,

$$\frac{\partial}{\partial x_i} f_j(\mathbf{x}).$$

Example

$$f(x, y) = xy$$

$$\frac{\partial}{\partial x} f(x, y) = \lim_{h \to 0} \frac{(x + h)y - xy}{h}$$

$$= \lim_{h \to 0} \frac{hy}{y} = \lim_{h \to 0} y = y$$

Single-variable functions

- For ”ordinary” functions $f : [a,b] \subset \mathbb{R} \rightarrow \mathbb{R}$, consider the derivative:

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

$$\frac{\Delta f}{\Delta x}$$

  if the limit exists.
- Indeed, for vector-valued functions, the partial derivative is calculated as if f was a1-variable function!
- All the other variables are "held constant"

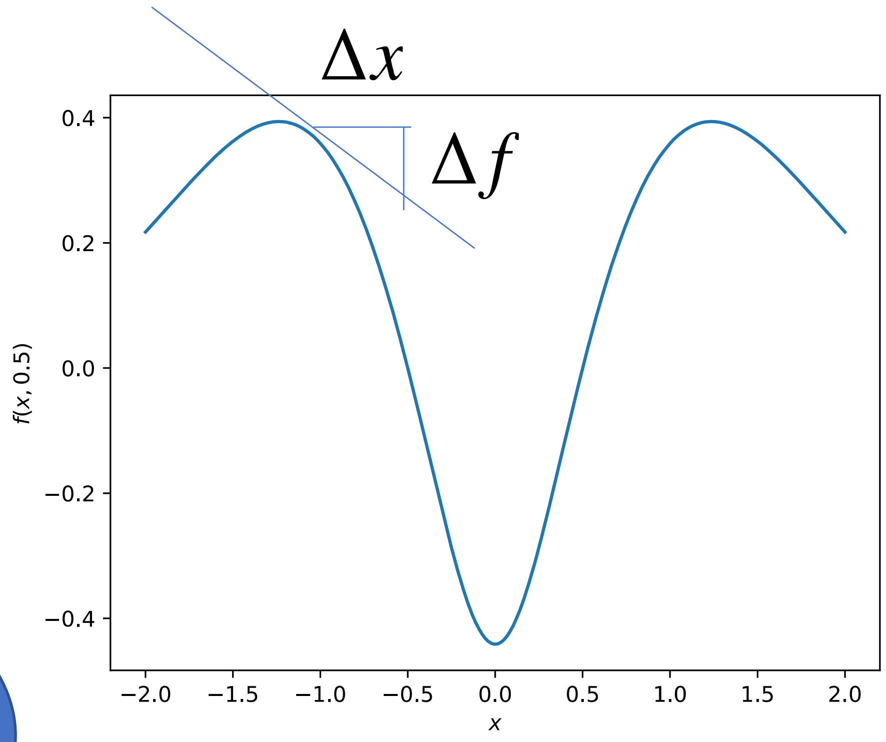
Derivative as slope

- Derivative is the *slope of tangent at x*
- When $f(x)$ has a derivative at x, the function *can be approximated*

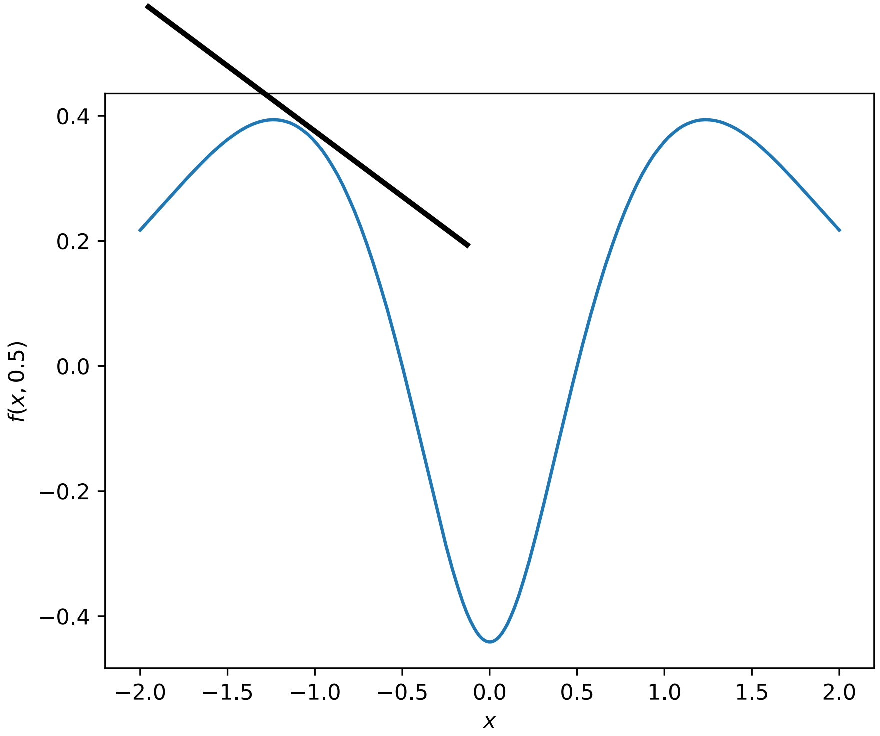
$$f(y) = f(x) + f'(x)(y - x) + \text{small error}$$

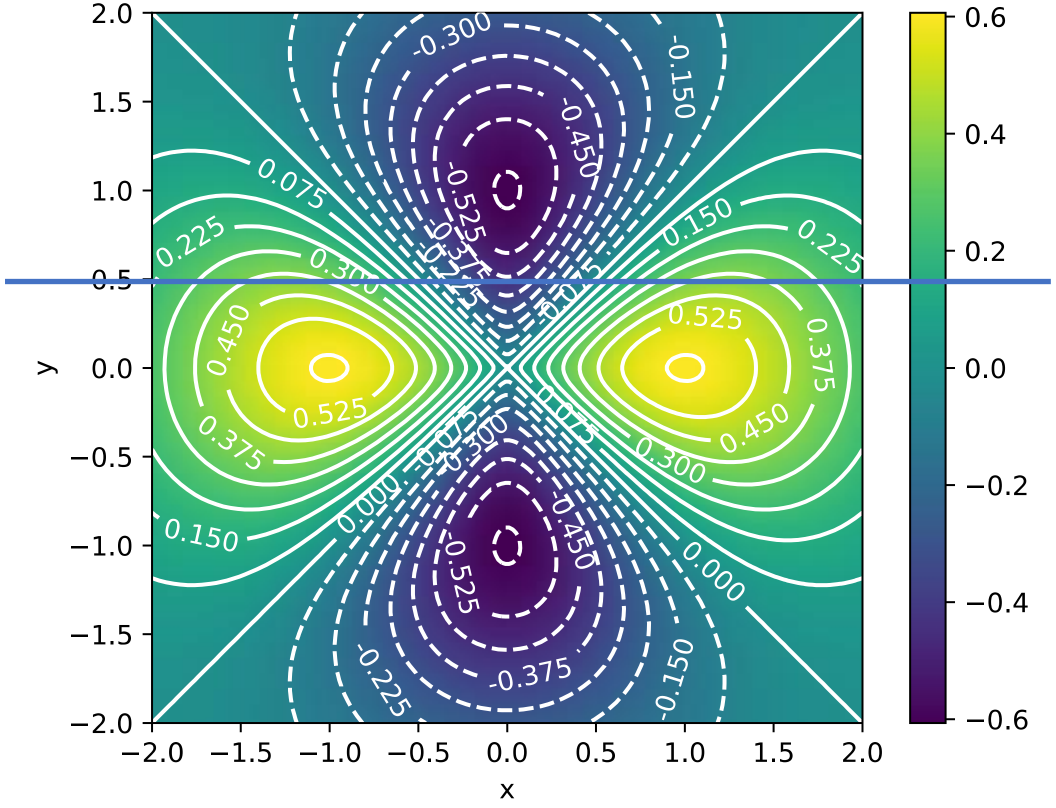
- Here y *is close to x*

Want something like this for vector-valued funcs



Derivative as slope/tangent





- Partial derivative is the rate of change as one moves in one direction

Existence of partial derivatives seems good ...

Example

let $f : \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto x^{1/3}y^{1/3}$. Since $f(x, 0) = 0$ and $f(0, y) = 0$,

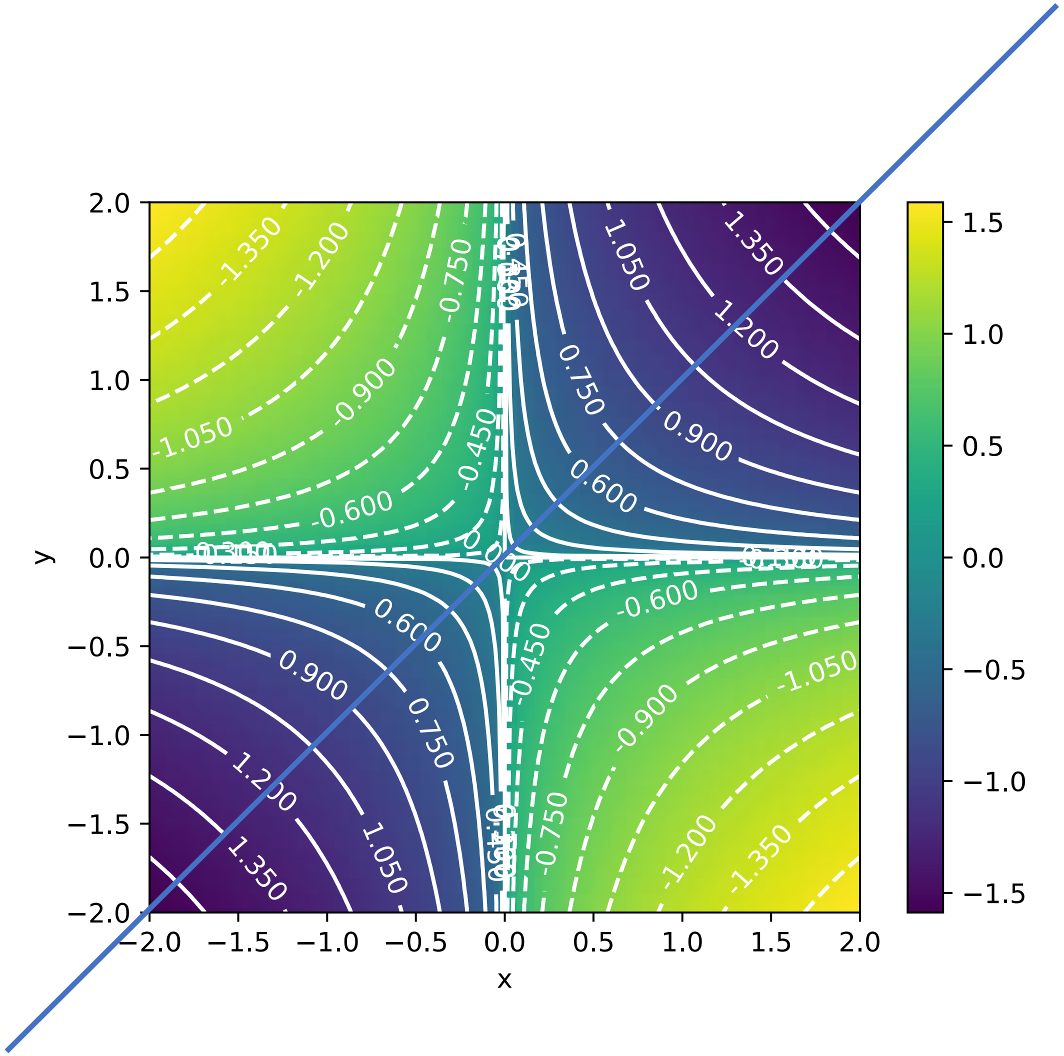
$$\frac{\partial}{\partial x} f(0,0) = \frac{\partial}{\partial y} f(0,0) = 0. \tag{1}$$

But along the “diagonal”

$$g(x) = f(x, x) = x^{2/3}. \tag{2}$$

The derivative of $g(x)$ is

$$g'(x) = \frac{2}{3}x^{-1/3} \to +\infty \quad \text{as } x \to 0 \tag{3}$$



Existence of partial derivatives seems good …

Example

let $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto$ … $^{1/3}$, …

$f(x, 0) = 0$ and $f(0, y) = 0$,

$$\frac{\partial}{\partial x} f(0,0) = \frac{\partial}{\partial y} \ldots$$

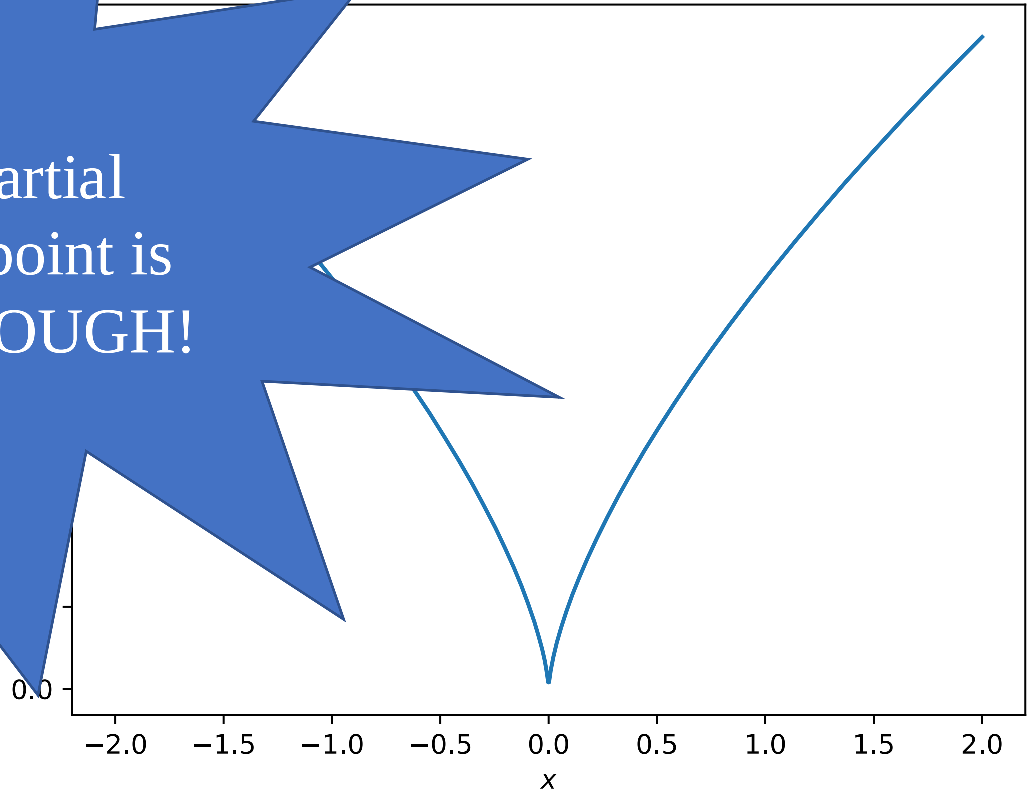
But along the “diagonal”

$$g(x) = f(x, x) = x^{2/3}.$$

The derivative of $g(x)$ is

$$g'(x) = \frac{2}{3} x^{-1/3} \to +\infty \quad \text{as } x \to 0 \qquad (3)$$

Existence of partial derivatives at a point is NOT GOOD ENOUGH!



Definition : Differentiable

Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, with Ω open. We say that f is *differentiable at* $\mathbf{x}_0 \in \Omega$ if the partial derivatives all exist at $\mathbf{x}_0$, and if

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - M(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0,$$

What does this mean?

where $M = Df(\mathbf{x}_0)$, the *derivative*, is the matrix of partial derivatives,

$$M_{ij} = \frac{\partial f_i(\mathbf{x}_0)}{\partial x_j}.$$

and where $M(\mathbf{x}-\mathbf{x}_0)$ is the matrix-vector product applied to $\mathbf{x}-\mathbf{x}_0$.

Interpretation of diffability condition

- Condition for a *first-order Taylor polynomial at* $\mathbf{x}_0$

Small
error term

$$f(\mathbf{x}) = f(\mathbf{x}_0) + Df(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|^2)$$

- Generalization of the slope of the tangent line to higher dimensions

Theorem 1

Intuitive, and good to know

If f is differetiable at $\mathbf{x}_0$, it is continuous at $\mathbf{x}_0$.

Resolves the ugly example

Theorem 2: Condition for differentiability

Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, with Ω open. Suppose the partial derivatives all exist at $\mathbf{x}_0$, and furthermore that they are all continuous *in a neighborhood* of $\mathbf{x}_0$. Then f is differentiable at $\mathbf{x}_0$.

Continuously differentiable functions

- These functions can always be approximated by first-order Polynomials

Definition 1: C^1 functions

A function whose partial derivatives exist and are continuous throughout its open domain Ω is said to be of class C^1.

Theorem : Properties of the derivative

1. Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\mathbf{x}_0 \in \Omega$, and let $c \in \mathbb{R}$. Then $h(\mathbf{x}) = cf(\mathbf{x})$ is differentiable at $\mathbf{x}_0$, and
$$Dh(\mathbf{x}_0) = cDf(\mathbf{x}_0). \quad (1)$$
Linearity

2. Let $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ be another function differentiable at $\mathbf{x}_0$. Then $h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ is differentiable at $\mathbf{x}_0$, and
$$Dh(\mathbf{x}_0) = Df(\mathbf{x}_0) + Dg(\mathbf{x}_0). \quad (2)$$

3. Let $f, g : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be *scalar-valued* functions, differentiable at $\mathbf{x}_0 \in \Omega$. Then $h(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ is differentiable at $\mathbf{x})_0$, and
$$Dh(\mathbf{x}_0) = g(\mathbf{x}_0)Df(\mathbf{x}_0) + f(\mathbf{x}_0)Dg(\mathbf{x}_0).$$
Product rule

4. As in 3, and additionally that $g > 0$ thrughout Ω. Then $h(\mathbf{x}_0) = f(\mathbf{x}_0)/g(\mathbf{x}_0)$ is differentiable at $\mathbf{x}_0$, and
$$Dh(\mathbf{x}_0) = \frac{g(\mathbf{x}_0)Df(\mathbf{x}_0) - f(\mathbf{x}_0)}{[g(\mathbf{x}_0)]^2}$$
Quotient rule

Theorem : Chain rule

Let $\Omega \subset \mathbb{R}^n$ and $\Omega' \subset \mathbb{R}^m$ be open sets, and let $g : \Omega \to \mathbb{R}^m$ with $g[\Omega] \subset \Omega'$. Let $f : \Omega' \to \mathbb{R}^o$. Thus, $h = f \circ g : \Omega \to \mathbb{R}^o$ is defined. Suppose g is differentiable at $\mathbf{x}_0 \in \Omega$, and f is differentiable at $\mathbf{y}_0 = f(\mathbf{x}_0) \in \Omega'$. Then $f \circ h$ is differentiable at $\mathbf{x}_0$ with derivative

$$D(f \circ g)(\mathbf{x}_0) = Df(\mathbf{y}_0)Df(\mathbf{x}_0),$$

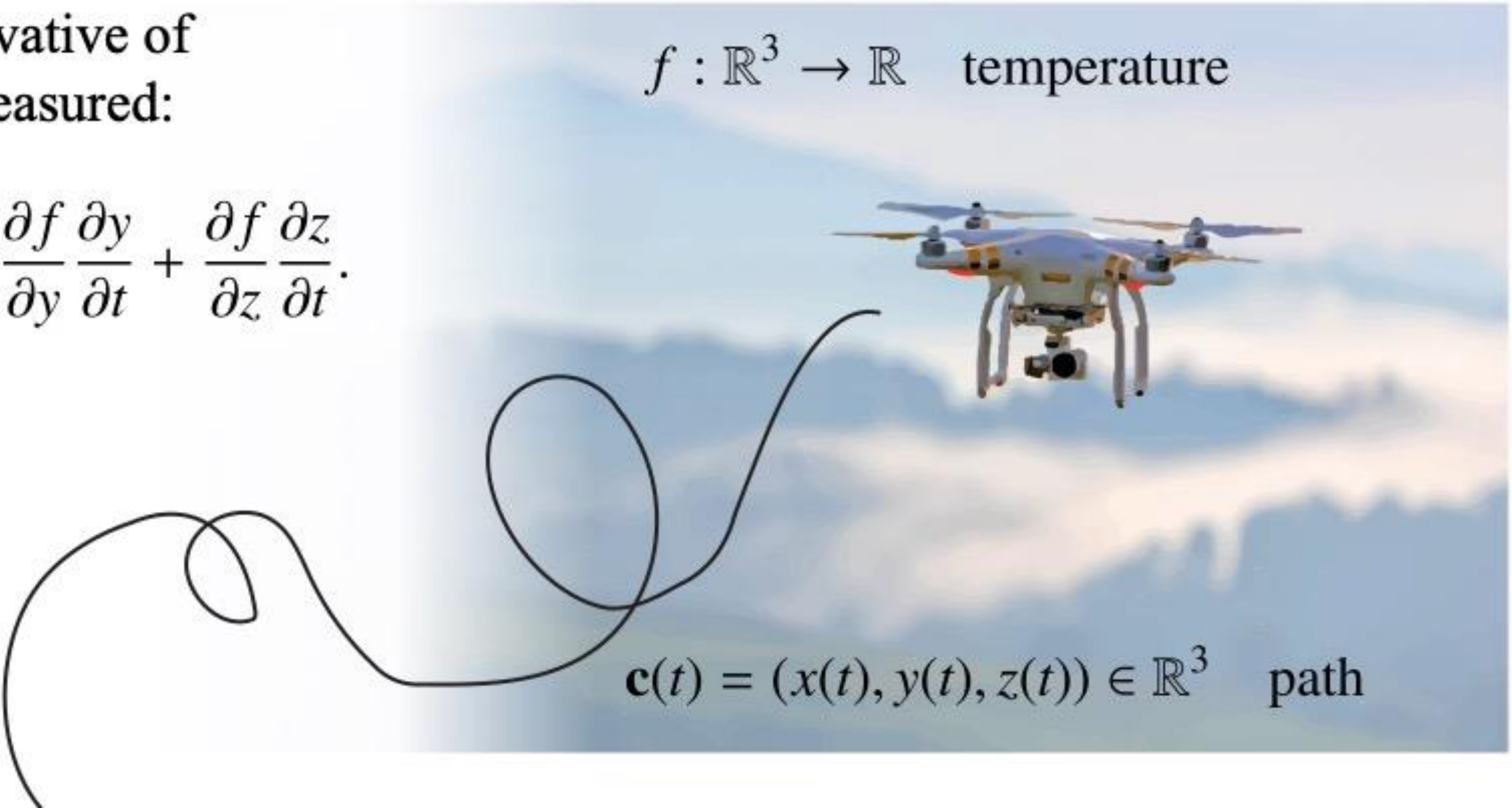
i.e., the matrix product of the Jacobian matrices.

Ex: Drone measuring temperature

$g(t) = f(\mathbf{c}(t)) \in \mathbb{R}$ temperature along path

Total time derivative of temperature measured:

$$\frac{\mathrm{d}g}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t}.$$



$f : \mathbb{R}^3 \to \mathbb{R}$ temperature

$\mathbf{c}(t) = (x(t), y(t), z(t)) \in \mathbb{R}^3$ path

Higher derivatives

- f is of class C^2 if the partial derivatives (matrix elements of Df) are of class C^1
- Matrix elements of $D(Df) = D^2f$: Iterated partial derivatives

$$[D^2 f(\mathbf{x})]_{ijk} = \frac{\partial^2}{\partial x_j \partial x_k} f_i(\mathbf{x}) \quad = \frac{\partial^2}{\partial x_k \partial x_j} f_i(\mathbf{x})$$

- Fact: If C^2, then partial derivatives *are symmetric*

Theorem : Second-order Taylor formula

Important for optimization!

et $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be of class C^2. Then we may write

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + Df(\mathbf{x}_0)\mathbf{h} + \frac{1}{2}\mathbf{h}^T D^2 f(\mathbf{x}_0)\mathbf{h} + R_2(\mathbf{h}, \mathbf{x}_0),$$

where the *remainder* satisfies $R_2(\mathbf{h}, \mathbf{x}_0)/\|\mathbf{h}\|^2 \to 0$ as $\mathbf{h} \to 0$, written

$$R_2(\mathbf{h}, \mathbf{x}_0) = o(\|\mathbf{h}\|^2).$$

Polynomial!

The symbol $D^2 f(\mathbf{x}_0)$ is the *Hessian* of f, the matrix of second-order mixed partial derivatives, a symmetric matrix.

Example

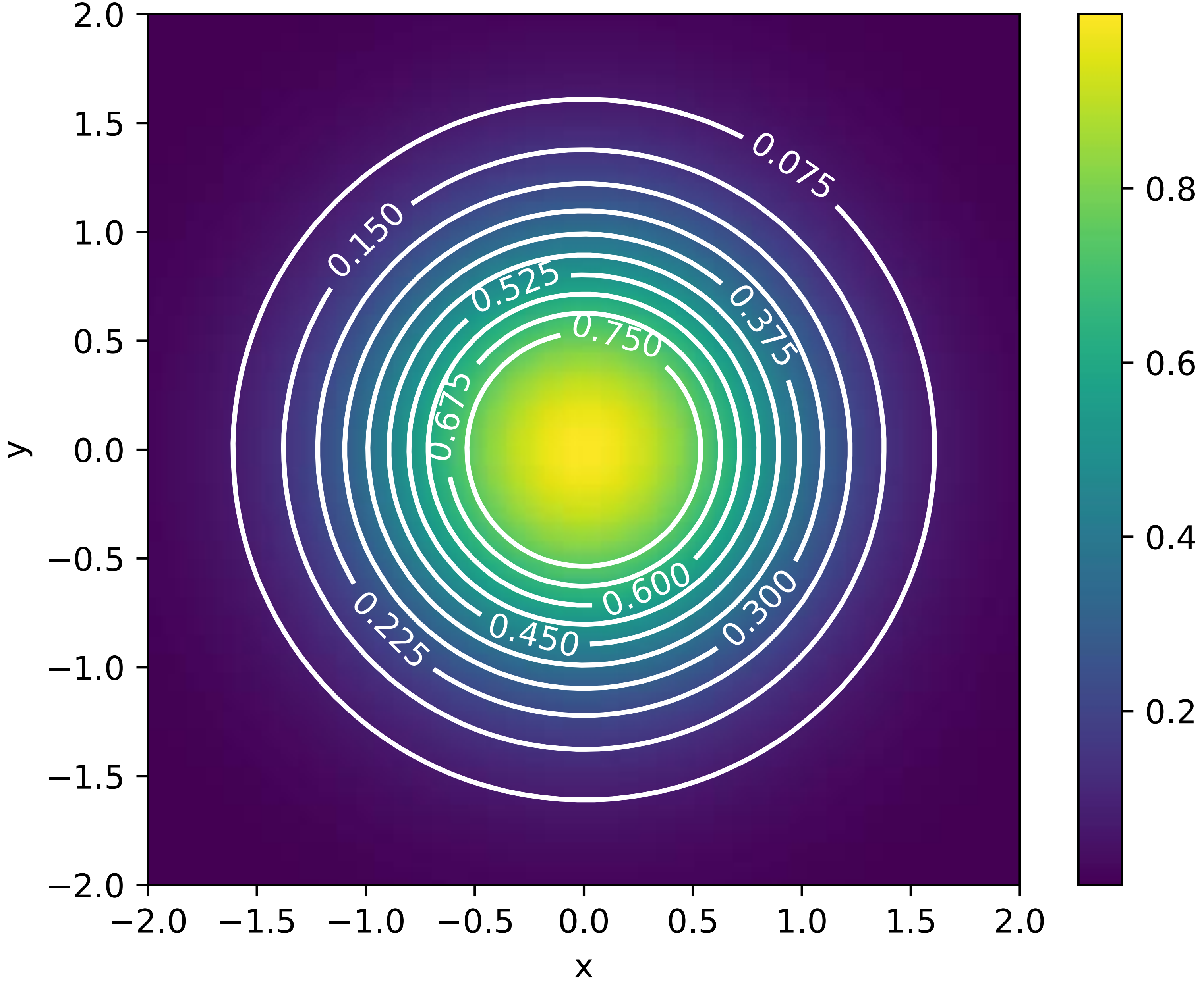
Compute the second-order Taylor polynomial of $f(x, y) = \exp(-x^2 - y^2)$ at $(0, 0)$.

$$Df(x, y) = [-2xf(x, y), -2yf(x, y)], \tag{1}$$

$$D^2 f(x, y) = \begin{bmatrix} (4x^2 - 2)f(x, y) & 4xyf(x, y) \\ 4xyf(x, y) & (4y^2 - 2)f(x, y) \end{bmatrix} \tag{2}$$

$$f(0, 0) = 1, \quad Df(0, 0) = [0, 0], \quad D^2 f(0, 0) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} \tag{3}$$

$$f(x, y) = 1 - (x^2 + y^2) + o(x^2 + y^2). \tag{4}$$



0.075
0.150
0.225
0.300
0.375
0.450
0.525
0.600
0.675
0.750
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
y
x
0.8
0.6
0.4
0.2

Local extrema

- Let $f : \Omega \to \mathbb{R}$ be twice differentiable, and $\mathbf{x}_0 \in \Omega$
- Local maximum:

$$\text{Exists } \epsilon > 0 \text{ such that } f(\mathbf{x}) \leq f(\mathbf{x}_0) \text{ for all } \mathbf{x} \in B_\epsilon(\mathbf{x}_0)$$

- Local minimum:

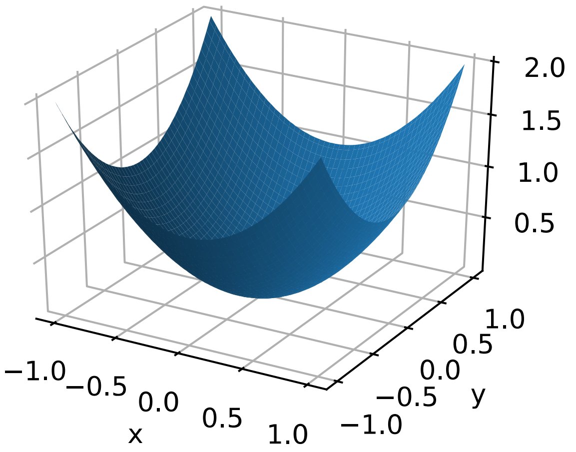
$$\text{Exists } \epsilon > 0 \text{ such that } f(\mathbf{x}) \geq f(\mathbf{x}_0) \text{ for all } \mathbf{x} \in B_\epsilon(\mathbf{x}_0)$$

- Fact: Any local extremum is a *critical point:*

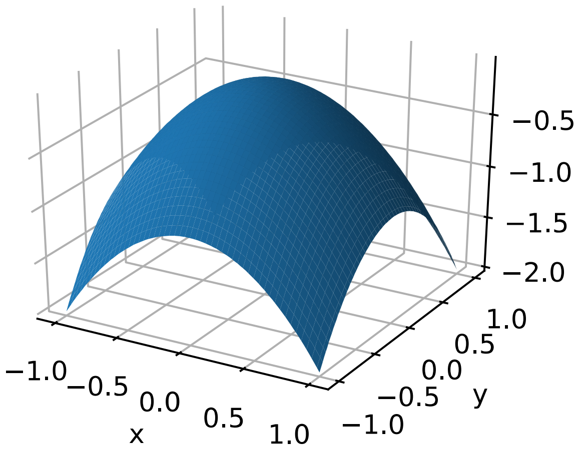
$$Df(\mathbf{x}_0) = 0$$

Critical points

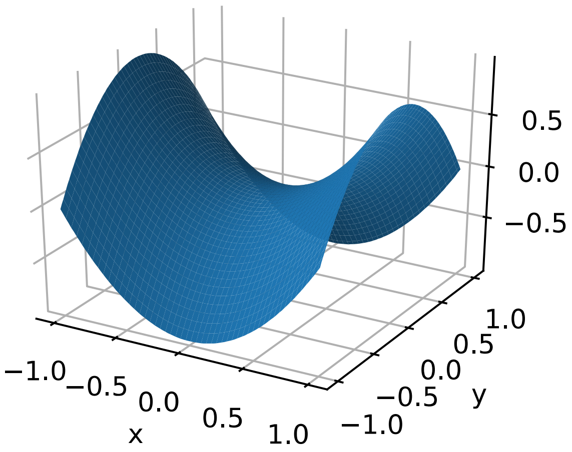
- A critical point can be a local minimum, maximum, or ***saddle point***
- Saddle points are critical points that are not a max/min



$x^2 + y^2$



$-x^2 - y^2$



$x^2 - y^2$

Example: The volcano function

Local max at the rim

Local min in the well

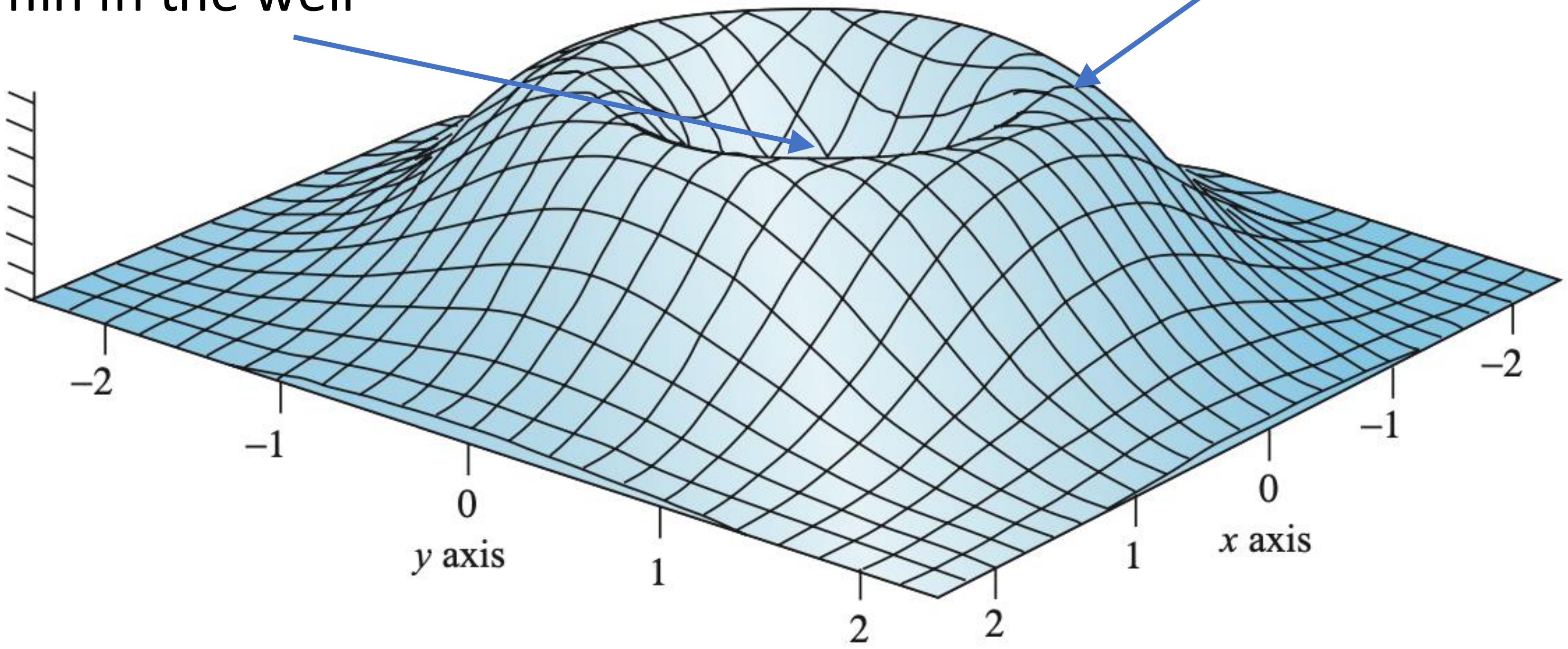


figure **3.3.4** The volcano: $z = 2(x^2 + y^2)\exp(-x^2 - y^2)$.

Picture from Marsden and Tromba, ”Vector Calculus”

Theorem : Classification of critical points

et $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, with Ω being an open domain. Let f be of class C^2. Let $H = D^2 f(\mathbf{x})$ be the second derivative (Hessian) at a critical point $\mathbf{x} \in \Omega$, i.e., $Df(\mathbf{x}) = 0$. Then we have:

1. If all the eigenvalues of H are positive, then $\mathbf{x}$ is a local minimium.
2. If all the eigenvalues of H are negative, then $\mathbf{x}$ is a local maximum.
3. If there are eigenvalues of H with both positive and negative values, but no zero eigenvalues, then $\mathbf{x}$ is a saddle point.
4. If some eigenvalues are zero, we cannot conclude based on second-order Taylor polynomials.

Further topics

- Series and convergece of series
- Integration over curves, surfaces, volumes ...
- Vector operations: curl, divergence, gradient ...
- Gauss' and Stoke's theorems for integration
- My presentation is based on →