# Fundamental Concepts for Analysis

Functions, sequences, derivative, integral, optimization

Paul MINCHELLA, Stéphane CHRÉTIEN[1]

October 2025

Master MIASHS

[1] Some sections are adapted from the course of Arian Novruzi, Department of Mathematics and Statistics, University of Ottawa, Canada.

# Contents

# Chapter 1

# Introduction to Algebraic Notions

Before anything else, it is essential to rigorously define the objects we manipulate in mathematics: numbers.

## 1.1 The Usual Sets of Numbers

**Definition 1.1.1** (Natural Numbers). *We denote* $\mathbb{N} = \{0, 1, 2, 3, \dots\}$*. It is the set of non-negative integers.*

**Definition 1.1.2** (Integers). *We denote* $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$*. It contains all positive and negative integers, as well as* $0$*.*

**Definition 1.1.3** (Rational Numbers). *We denote*

$$\mathbb{Q} = \left\{ \tfrac{p}{q} \ : \ p \in \mathbb{Z},\, q \in \mathbb{Z}^*,\, q \neq 0 \right\}.$$

*It is the set of quotients of integers.*

**Definition 1.1.4** (Real Numbers $\mathbb{R}$). $\mathbb{R}$ *is defined as the* ***completion*** *of* $\mathbb{Q}$*. In other words, it is the set obtained by "filling in the gaps" of* $\mathbb{Q}$*. Formally,* $\mathbb{R}$ *is the unique totally ordered, complete field containing* $\mathbb{Q}$*. This means:*

- Order*:* $\mathbb{R}$ *is equipped with an order relation* $\leq$ *compatible with addition and multiplication.*
- Density*: between any two distinct real numbers, there always exists a rational number.*
- Completeness*: every increasing and bounded sequence of real numbers converges in* $\mathbb{R}$*.*

*This property of completeness distinguishes* $\mathbb{R}$ *from* $\mathbb{Q}$*.*

**Definition 1.1.5** (Complex Numbers). *We define*

$$\mathbb{C} = \{x + iy \ : \ x, y \in \mathbb{R},\ i^2 = -1\}.$$

*It is an algebraic extension of* $\mathbb{R}$ *that allows us to solve all polynomial equations.*

*Remark* 1. The hierarchy is therefore

$$\boxed{\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}}.$$

**Theorem 1.1.1** (Fundamental Theorem of Algebra)**.** *Every polynomial $P \in \mathbb{C}[X]$ of degree $n \geq 1$ admits at least one root in $\mathbb{C}$.*
*In other words, $\mathbb{C}$ is* algebraically closed*: every polynomial equation with complex coefficients has a solution in $\mathbb{C}$.*

*Remark* 2. This result implies that any polynomial $P \in \mathbb{C}[X]$ of degree $n$ can be factorized as a product of degree-1 polynomials:

$$P(X) = a_n(X - \alpha_1)(X - \alpha_2) \cdots (X - \alpha_n),$$

where the $\alpha_i \in \mathbb{C}$ are the (possibly repeated) roots of $P$. This theorem is fundamental because it justifies the importance of $\mathbb{C}$ as the natural framework of polynomial algebra.

Here is a very important definition.

**Definition 1.1.6.** *For $a, b \in \mathbb{R}$ with $a < b$, the intervals of $\mathbb{R}$ with endpoints $a$ and $b$ are denoted by*

$$(a, b), \quad [a, b), \quad (a, b], \quad [a, b].$$

The first is called the "open interval $a$–$b$", the second "right semi-open interval $a$–$b$", the third "left semi-open interval $a$–$b$", and the fourth "closed interval $a$–$b$". If one or both of $a$ and $b$ are $\pm\infty$, we denote

$$(-\infty, a), \quad (-\infty, a], \quad (b, +\infty), \quad [b, +\infty), \quad (-\infty, +\infty)$$

the corresponding intervals.

## 1.2 Fundamental Algebraic Structures

**Definition 1.2.1** (Internal Law of Composition)**.** *Let $E$ be a set. An* internal composition law *on $E$ is a map*

$$\star : E \times E \to E, \qquad (x, y) \mapsto x \star y.$$

*Examples: $+$ and $\times$ on $\mathbb{Z}$.*

**Definition 1.2.2** (Group)**.** *A* group *is a pair $(G, \star)$ where $\star$ is an internal composition law on $G$ such that:*

- ***Associativity****: $(x \star y) \star z = x \star (y \star z)$.*
- ***Neutral element****: there exists $e \in G$ such that $x \star e = e \star x = x$.*
- ***Inverses****: every $x \in G$ admits an inverse $x^{-1}$ such that $x \star x^{-1} = e$.*

*If, moreover, $x \star y = y \star x$, the group is said to be* abelian.

**Example 1.2.1.** *$(\mathbb{Z}, +)$ is an abelian group. On the other hand, $(\mathbb{Z}, \times)$ is not a group because not all integers have a multiplicative inverse in $\mathbb{Z}$.*

**Definition 1.2.3** (Ring)**.** *A* ring $(A, +, \times)$ *is a set equipped with two internal composition laws such that:*

- $(A, +)$ *is an abelian group.*
- $\times$ *is associative and has a neutral element* $1$*.*
- $\times$ *is distributive over* $+$*.*

**Definition 1.2.4** (Field)**.** *A* field *is a ring* $(K, +, \times)$ *in which every nonzero element is invertible for* $\times$*.*

**Example 1.2.2.** $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ *are fields.* $\mathbb{Z}$ *is a ring, but not a field.*

**Property 1.2.1.** *Fundamental results:*

- $\mathbb{N}$ *is not a group for* $+$ *(no inverses).*
- $(\mathbb{Z}, +)$ *is an abelian group.*
- $(\mathbb{Z}, +, \times)$ *is a ring.*
- $(\mathbb{Q}, +, \times)$*,* $(\mathbb{R}, +, \times)$ *and* $(\mathbb{C}, +, \times)$ *are fields.*

**Exercises.** Usual Numbers and Algebraic Structures.

*Exercise* 1.2.1 (Non-standard composition law on $\mathbb{Z}$)*.* For $a, b \in \mathbb{Z}$, define

$$a \star b \;=\; a + b + 1.$$

1. Show that $\star$ is an *internal* composition law on $\mathbb{Z}$.
2. Verify the associativity and commutativity of $\star$.
3. Determine the neutral element $e$ for $\star$.
4. For $a \in \mathbb{Z}$, find the inverse of $a$ for $\star$.
5. Conclude: is $(\mathbb{Z}, \star)$ a group? abelian?

**Solution.** **1.** For $a, b \in \mathbb{Z}$, $a \star b = a + b + 1 \in \mathbb{Z}$: internal stability.

**2.** *Associativity:*

$$(a \star b) \star c = (a + b + 1) + c + 1 = a + b + c + 2 = a + (b + c + 1) + 1 = a \star (b \star c).$$

*Commutativity:* $a \star b = a + b + 1 = b + a + 1 = b \star a$.

**3.** The neutral element $e$ satisfies $a \star e = a$. Then $a \star e = a + e + 1 = a$, hence $e = -1$. We also check $e \star a = a$.

**4.** The inverse $a^\star$ satisfies $a \star a^\star = e = -1$. Thus $a + a^\star + 1 = -1$, giving $a^\star = -a - 2$ (an integer). We also check $a^\star \star a = -1$.

**5.** All axioms are satisfied: $(\mathbb{Z}, \star)$ is an **abelian group** with neutral element $-1$, and the inverse of $a$ is $-a - 2$. □

---

*Exercise* 1.2.2 (Overview of groups, rings, fields, and units)*.*

1. Explain why $(\mathbb{N}, +)$ is not a group, while $(\mathbb{Z}, +)$ is one, and moreover abelian.

2. Show that $(\mathbb{Z}, \times)$ is a commutative monoid *(associative, neutral element* 1*)*, but not a group. Determine its set of *units*.

3. Show that $(\mathbb{Q}, +, \times)$ is a field.

4. In the ring $\mathbb{Z}/12\mathbb{Z}$, determine the set of *units* $U(12)$ (invertible classes for $\times$ modulo 12). Deduce that $\mathbb{Z}/12\mathbb{Z}$ is not a field.

**Solution.** **1.** In $(\mathbb{N}, +)$, there are *no inverses* for $+$: for example, 1 has no $n \in \mathbb{N}$ such that $1+n = 0$. Thus, it is not a group. However, $(\mathbb{Z}, +)$ is an abelian group: associativity and commutativity inherited from $+$, neutral element 0, inverse of $a$ equal to $-a$.

**2.** $(\mathbb{Z}, \times)$: associativity, commutativity, and neutral element 1 hold, so it is a *commutative monoid*. It is not a group since most integers do not have a multiplicative inverse in $\mathbb{Z}$ (e.g., 2 has none). The *units* (invertible elements) are exactly $\{-1, 1\}$.

**3.** In $(\mathbb{Q}, +, \times)$, $+$ makes $\mathbb{Q}$ an abelian group (neutral 0, inverse $-q$). For $\times$, every nonzero rational $\frac{p}{q}$ has inverse $\frac{q}{p}$, and distributivity links $+$ and $\times$. Hence $(\mathbb{Q}, +, \times)$ is a **field**.

**4.** A class $\overline{a} \in \mathbb{Z}/12\mathbb{Z}$ is a *unit* iff $\gcd(a, 12) = 1$. The integers $1 \leq a \leq 11$ coprime with 12 are $1, 5, 7, 11$. Thus

$$U(12) = \{\overline{1}, \overline{5}, \overline{7}, \overline{11}\},$$

which forms an abelian group under multiplication modulo 12. Since not all nonzero classes are invertible (e.g., $\overline{2}$), $\mathbb{Z}/12\mathbb{Z}$ is **not** a field. $\square$

*Remark* 3. For every prime $p$, $\mathbb{Z}/p\mathbb{Z}$ is a field.

# Chapter 2

# Basic Concepts of Functions

Functions are the central objects of analysis. A function is a *rule of correspondence* that assigns to each element of a starting set a unique element of an arrival set.

**Definition 2.0.1** (Function)**.** *Let $X$ and $Y$ be two sets. A* function *$f$ from $X$ to $Y$ is a mapping that associates to each element $x \in X$ a unique element $y \in Y$, denoted $f(x)$.*

*We write symbolically:*

$$f : X \mapsto Y$$
$$x \mapsto y = f(x).$$

## 2.1 Domain and Image

**Definition 2.1.1** (Domain)**.** *The set of $x \in X$ for which $f(x)$ is defined is called the* domain of $f$*, and is denoted*

$$\mathrm{Dom}(f).$$

**Definition 2.1.2** (Image)**.** *The set of values $f(x)$ obtained as $x$ ranges over* $\mathrm{Dom}(f)$ *is called the* image of $f$*, and is denoted*

$$\mathrm{Im}(f) = \{f(x) \mid x \in \mathrm{Dom}(f)\}.$$

## 2.2 Notions of Upper and Lower Bounds

In analysis, it is essential to know how to characterize the "bounds" of a set of real numbers.

**Definition 2.2.1** (Upper bound, lower bound)**.** *Let $A \subset \mathbb{R}$ be a non-empty set.*

- *A number $M \in \mathbb{R}$ is called an **upper bound** of $A$ if*

$$\forall x \in A, \quad x \leq M.$$

  *In this case, we say that $A$ is **bounded above**.*

- *A number $m \in \mathbb{R}$ is called a **lower bound** of $A$ if*

$$\forall x \in A, \quad x \geq m.$$

  *In this case, we say that $A$ is **bounded below**.*

**Definition 2.2.2** (Supremum and infimum)**.** *Let $A \subset \mathbb{R}$ be non-empty and bounded above.*

- *The **supremum** (or **least upper bound**) of $A$, denoted $\sup A$, is the smallest of all upper bounds of $A$. Formally:*

$$\sup A = M \iff \begin{cases} \forall x \in A,\ x \leq M, \\ \forall \varepsilon > 0,\ \exists x \in A \text{ such that } x > M - \varepsilon. \end{cases}$$

- *Similarly, if $A$ is bounded below, its **infimum** (or **greatest lower bound**), denoted $\inf A$, is the largest of all lower bounds of $A$.*

We can extend this notion to functions (that is, to their images):

**Definition 2.2.3** (Supremum of a function)**.** *Let $f : D \to \mathbb{R}$ be a function defined on a set $D \subset \mathbb{R}$. The **supremum** of $f$ on $D$ is the supremum of the set of its values:*

$$\sup_{x \in D} f(x) \ = \ \sup\{\, f(x) \ : \ x \in D\}.$$

*Similarly, the **infimum** of $f$ on $D$ is*

$$\inf_{x \in D} f(x) \ = \ \inf\{\, f(x) \ : \ x \in D\}.$$

**Example 2.2.1.**

- *If $A = (0, 1)$, then $\sup A = 1$ and $\inf A = 0$ (even though $0, 1 \notin A$).*
- *If $A = \{2, 3, 5\}$, then $\sup A = 5$ and $\inf A = 2$.*
- *If $A = \mathbb{N}$, then $A$ is not bounded above in $\mathbb{R}$, hence $\sup A = +\infty$.*

*Remark* 4. The completeness theorem of the real numbers states that every non-empty and bounded-above subset of $\mathbb{R}$ admits a supremum (and similarly for the infimum). This is an essential property that distinguishes $\mathbb{R}$ from $\mathbb{Q}$.

## 2.3 Some Elementary Functions

Some examples of basic functions are the following:

- $f(x) = c$, with $c \in \mathbb{R}$, called the *constant function*.
- $f(x) = ax + b$, with $a, b \in \mathbb{R}$, called the *affine (linear) function*.
- $f(x) = ax^2 + bx + c$, with $a, b, c \in \mathbb{R}$, called the *quadratic function*.
- $f(x) = |x|$, called the *absolute value function*.
- $f(x) = \dfrac{1}{x}$, defined on $\mathbb{R} \setminus \{0\}$.
- $f(x) = x^{\alpha}$, with $x \neq 0$, $\alpha \in \mathbb{R}$, called the *power function*.

*Exercise* 2.3.1. Let $f(x) = 3x + 7$.

1. Determine $\mathrm{Dom}(f)$.
2. Determine $\mathrm{Im}(f)$.

**Solution.** **1. Domain.** The function $f(x) = 3x + 7$ is well-defined for every $x \in \mathbb{R}$, hence $\mathrm{Dom}(f) = \mathbb{R}$.

**2. Image.** Let $y \in \mathbb{R}$ be arbitrary. We look for $x \in \mathbb{R}$ such that $f(x) = y$, that is:

$$3x + 7 = y \quad \Longleftrightarrow \quad x = \frac{y-7}{3}.$$

Such an $x$ always exists, since $y$ is any real number. Thus $\mathrm{Im}(f) = \mathbb{R}$. □

---

*Exercise* 2.3.2. Let $f(x) = \sqrt{x-1}$.

1. Determine $\mathrm{Dom}(f)$.
2. Determine $\mathrm{Im}(f)$.

**Solution.** **1. Domain.** The square root $\sqrt{x-1}$ is defined only if $x - 1 \geq 0$. Hence

$$\mathrm{Dom}(f) = [1, +\infty).$$

**2. Image.** On one hand, for all $x \in \mathrm{Dom}(f)$, we have $\sqrt{x-1} \geq 0$. Thus

$$\mathrm{Im}(f) \subset [0, +\infty).$$

On the other hand, let $y \in [0, +\infty)$. Setting $x = y^2 + 1 \geq 1$, we get

$$f(x) = \sqrt{(y^2+1) - 1} = \sqrt{y^2} = y.$$

Hence every $y \geq 0$ is attained by $f$. Therefore:

$$\mathrm{Im}(f) = [0, +\infty).$$

□

---

## 2.4 Graph of a Function

**Definition 2.4.1** (Graph)**.** *The* graph *of a function $f$ is the set of pairs $(x, f(x))$ for $x \in \mathrm{Dom}(f)$:*

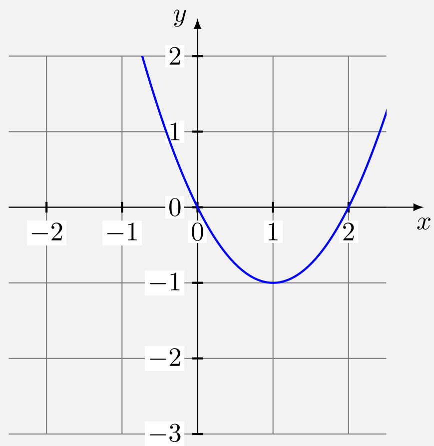
$$G(f) = \{(x, f(x)) \in \mathbb{R}^2 \mid x \in \mathrm{Dom}(f)\}.$$

*Geometrically, it is the representative curve of $f$ in the Cartesian plane.*

**Example 2.4.1.** *Let us plot the graph of the function $f(x) = x^2 - 2x$. We compute $f(x)$ for several values of $x$ as shown in the table below:*

| $x$ | $-2$ | $-1$ | $0$ | $1$ | $2$ | $3$ |
|---|---|---|---|---|---|---|
| $f(x)$ | $8$ | $3$ | $0$ | $-1$ | $0$ | $3$ |

*Then, we plot the points $(x, f(x))$ and connect them smoothly as shown in the following figure.*




*Remark* 5. We distinguish between the function itself (an abstract correspondence) and its graphical representation (a curve in the plane). This distinction allows us to work both algebraically and geometrically.

## 2.5 Increasing, Decreasing, and Monotonic Functions

**Definition 2.5.1** (Increasing and Decreasing Functions)**.** *Let $f : I \to \mathbb{R}$ be a function defined on an interval $I \subset \mathbb{R}$.*

- *$f$ is said to be* ***increasing*** *on $I$ if and only if*

$$\forall x, y \in I, \quad x < y \implies f(x) \leq f(y).$$

- *$f$ is said to be* ***strictly increasing*** *if and only if*

$$\forall x, y \in I, \quad x < y \implies f(x) < f(y).$$

- *$f$ is said to be* ***decreasing*** *(resp. strictly decreasing) if and only if*

$$\forall x, y \in I, \quad x < y \implies f(x) \geq f(y) \quad (resp.\ f(x) > f(y)).$$

**Definition 2.5.2** (Monotonic Function)**.** *A function is said to be* monotonic *on an interval $I$ if it is either increasing or decreasing on $I$.*

**Example 2.5.1** (Some examples)**.**

- *The function $f(x) = x$ is strictly increasing on $\mathbb{R}$.*
- *The function $f(x) = -x$ is strictly decreasing on $\mathbb{R}$.*
- *The function $f(x) = x^2$ is decreasing on $(-\infty, 0]$ and increasing on $[0, +\infty)$.*

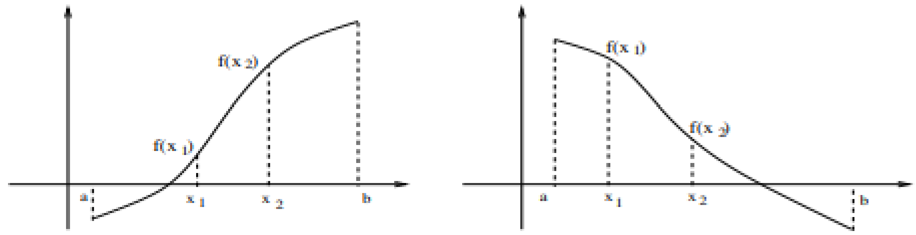



Figure 2.1: On the left: an increasing function; on the right: a decreasing function.

## 2.6 Algebraic Operations on Functions

Given two functions $f$ and $g$ defined on the same domain, and two constants $a, b \in \mathbb{R}$, one can form new functions using algebraic operations.

**Definition 2.6.1** (Linear combination, product, and quotient)**.** *We define the following functions:*

$$(af + bg)(x) = af(x) + bg(x), \quad (fg)(x) = f(x) \cdot g(x), \quad \left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)} \quad \textit{(if } g(x) \neq 0\textit{)}.$$

*Remark* 6. TWith functions, one can perform all usual algebraic operations: addition, subtraction, multiplication, division (when defined), powers, and roots.

**Example 2.6.1.** *We provide here two examples:*

- *If* $f(x) = x^2$ *and* $g(x) = x + 1$*, then*

$$(f + g)(x) = x^2 + (x + 1) = x^2 + x + 1, \qquad (fg)(x) = x^2(x + 1) = x^3 + x^2.$$

- *If* $f(x) = \sqrt{x}$ *and* $g(x) = x^2 + 1$*, then*

$$\left(\frac{f}{g}\right)(x) = \frac{\sqrt{x}}{x^2 + 1}, \quad x \geq 0.$$

## 2.7 Composition of Functions

**Definition 2.7.1** (Composition)**.** *Let* $f : X \to Y$ *and* $g : Y \to Z$ *be two functions. The* composite function $f \circ g : X \to Z$ *is defined by*

$$(f \circ g)(x) = f(g(x)), \quad \textit{for all } x \in X.$$

*Similarly,* $g \circ f$ *is defined by* $(g \circ f)(x) = g(f(x))$ *whenever this makes sense.*

**Example 2.7.1** (Practical application)**.** *Let* $t$ *denote the time elapsed since the year 2000. The population of a country is defined by* $p(t) = 50 + e^{0.01t}$ *(in millions), and the income as a function of the population is* $R(p) = 2.1 + \ln(1 + 3p)$*. Then, the composite function*

$$R \circ p(t) = R(p(t)) = 2.1 + \ln\left(1 + 3(50 + e^{0.01t})\right)$$

*gives the income as a function of time* $t$ *(expressed in years).*

**Example 2.7.2** (Computational example). *Let $f(x) = x + 1$ and $g(x) = \frac{1}{x}$, defined on $\mathbb{R} \setminus \{0, -1\}$.*

$$f \circ g(x) = f\big(g(x)\big) = f\big(\tfrac{1}{x}\big) = \tfrac{1}{x} + 1 = \tfrac{1+x}{x},$$
$$g \circ f(x) = g\big(f(x)\big) = g(x+1) = \tfrac{1}{x+1}.$$

*Note in particular (in case one had the odd idea to wonder!) that $f \circ g \neq g \circ f$: composition of functions is generally **not commutative**.*

## 2.8 Inverse, Injective, Surjective, and Bijective Functions

**Definition 2.8.1** (Inverse image). *Let $f : E \to F$ and $B \subset F$. The* inverse image *of $B$ under $f$ is the set*

$$f^{-1}(B) = \{x \in E \mid f(x) \in B\}.$$

*In particular, for $y \in F$, we write*

$$f^{-1}(\{y\}) = \{x \in E \mid f(x) = y\}.$$

**Definition 2.8.2** (Injective, Surjective, and Bijective Functions). *Let $E, F$ be two sets and $f : E \to F$ a mapping.*

- *$f$ is **injective** if*
$$\forall x, y \in E, \quad f(x) = f(y) \;\Rightarrow\; x = y,$$
*in other words, $x \neq y$ implies $f(x) \neq f(y)$.*

- *$f$ is **surjective** if*
$$\forall y \in F,\ \exists x \in E \text{ such that } f(x) = y,$$
*i.e. if the image of $f$ is exactly $F$, that is* $\mathrm{Im}(f) = F$.

- *$f$ is **bijective** if it is both injective and surjective. In this case, $f$ admits a unique* inverse function *$f^{-1} : F \to E$ such that*
$$f^{-1} \circ f = \mathrm{id}_E \qquad \text{and} \qquad f \circ f^{-1} = \mathrm{id}_F.$$

*Remark* 7. Practically, to find $f^{-1}$, set $y = f(x)$ and solve this equation for $x$ in terms of $y$. Then $x = f^{-1}(y)$.

**Example 2.8.1.** *Let* $f(x) = \sqrt{\frac{x}{1-x}}$*, defined on* $D = \{x \in \mathbb{R} \mid 0 < x < 1\}$*. To find* $f^{-1}$*, solve* $f(y) = x$*:*

$$\sqrt{\frac{y}{1-y}} = x.$$

*Squaring both sides gives:*

$$\frac{y}{1-y} = x^2 \quad \Longrightarrow \quad y = x^2(1-y).$$

*Thus:*

$$y(1+x^2) = x^2 \quad \Longrightarrow \quad y = \frac{x^2}{1+x^2}.$$

*Hence the inverse function is*

$$f^{-1}(x) = \frac{x^2}{1+x^2}, \qquad x \geq 0.$$

**Example 2.8.2.** *The function* $f : \mathbb{R} \to \mathbb{R}$ *defined by* $f(x) = 2x + 3$ *is bijective:*

- ✓ *It is injective since* $2x_1 + 3 = 2x_2 + 3 \implies x_1 = x_2$*,*
- ✓ *It is surjective since for any* $y \in \mathbb{R}$*,* $x = \frac{y-3}{2}$ *is a preimage,*
- ✓ *Its inverse function is* $f^{-1}(y) = \frac{y-3}{2}$*.*

*Remark* 8. On an **interval** $I \subset \mathbb{R}$, any **strictly monotonic** function is injective. If, moreover, $f(I) = J$, then $f : I \to J$ is a bijection and admits an inverse $f^{-1} : J \to I$.

---

*Exercise* 2.8.1 (Lambert $W$ Function on $[0, +\infty)$). **Definition (Lambert function).** The *Lambert function* is any (possibly multivalued) function $W$ satisfying

$$W(y)\, e^{W(y)} = y.$$

On the real line, certain *branches* of $W$ are defined on specific intervals (for example, $W_0$ and $W_{-1}$ on $[-e^{-1}, 0)$). In this exercise, we restrict ourselves to the interval $[0, +\infty)$ and study the equation $we^w = y$ using the following function.

Consider the function $f : [0, +\infty) \to [0, +\infty)$ defined by

$$f(x) = x\, e^x.$$

1. Show that $f$ is continuous, strictly increasing, and that it defines a bijection from $[0, +\infty)$ onto $[0, +\infty)$.
2. Deduce that, for every $y \in [0, +\infty)$, the equation

   $$w\, e^w = y$$

   admits a *unique* real solution $w \in [0, +\infty)$. We denote this solution by $W_0(y)$, and verify that $W_0 = f^{-1}$.

**Solution. 1) Continuity, strict increase, and bijectivity.**

*Continuity.* The functions $x \mapsto x$ and $x \mapsto e^x$ are continuous on $\mathbb{R}$; therefore, their product $f(x) = xe^x$ is continuous on $[0, +\infty)$.

*Strict increase (hence injectivity).* We compute

$$f'(x) = e^x(1 + x).$$

For all $x \geq 0$, we have $e^x > 0$ and $1 + x > 0$, hence $f'(x) > 0$. Therefore $f$ is strictly increasing on $[0, +\infty)$, and in particular injective.

*Image (hence surjectivity onto* $[0, +\infty)$*).* We have $f(0) = 0$ and

$$\lim_{x\to+\infty} f(x) = \lim_{x\to+\infty} xe^x = +\infty.$$

By continuity and monotonicity, the image of $[0, +\infty)$ under $f$ is

$$f\big([0, +\infty)\big) = \big[\, f(0), \lim_{x\to+\infty} f(x)\,\big] = [0, +\infty).$$

Thus $f$ is bijective from $[0, +\infty)$ onto $[0, +\infty)$.

**2) Existence and uniqueness of the real solution of $we^w = y$ on $[0, +\infty)$.**

Fix $y \in [0, +\infty)$. By bijectivity shown in question 1), there exists a unique $w \in [0, +\infty)$ such that $f(w) = y$, i.e. $we^w = y$. We define $W_0(y) := w$; then $W_0 = f^{-1}$ on $[0, +\infty)$. $\square$

---

*Remark* 9. The Lambert $W$ function is said to be *multivalued* because the equation

$$we^w = y$$

can admit several solutions $w$. On $[-e^{-1}, 0)$, there exist two distinct real branches ($W_0$ and $W_{-1}$), whereas on $[0, +\infty)$, the real solution is unique ($W_0$). For complex arguments, infinitely many branches $W_k$ appear.

This function has numerous applications in physics and engineering whenever one needs to invert a relation of the form $xe^x = y$. For instance:

- in optics and laser physics (Beer–Lambert law, beam attenuation),
- in electronics and semiconductor physics (current or potential equations),
- in statistical physics and chemical kinetics (exponential-type equations),
- in exponential relaxation models (charging/discharging of a capacitor).

# Chapter 3

# Sequences of Real Numbers

## 3.1 Definition

A *sequence of real numbers* is a succession of real numbers $u_0, u_1, u_2, u_3, \ldots$. Such a sequence is denoted $(u_n)_{n\in\mathbb{N}}$.

**Definition 3.1.1** (Numerical Sequence). *A numerical sequence is a function $u : \mathbb{N} \to \mathbb{R}$ defined by*

$$\begin{aligned} u : \mathbb{N} &\longrightarrow \mathbb{R} \\ n &\longmapsto u_n. \end{aligned}$$

## 3.2 Monotonicity and Bounds

**Definition 3.2.1** (Increasing, Decreasing, and Bounded Sequences). *Let $(u_n)_{n\in\mathbb{N}}$ be a sequence.*

- *It is **increasing** if $\forall n \in \mathbb{N},\ u_{n+1} \geq u_n$.*
- *It is **decreasing** if $\forall n \in \mathbb{N},\ u_{n+1} \leq u_n$.*
- *It is **constant** if $\forall n \in \mathbb{N},\ u_{n+1} = u_n$.*
- *A sequence that is increasing or decreasing is said to be **monotonic**.*
- *It is **bounded above** if there exists $M \in \mathbb{R}$ such that $\forall n,\ u_n \leq M$.*
- *It is **bounded below** if there exists $m \in \mathbb{R}$ such that $\forall n,\ u_n \geq m$.*
- *It is **bounded** if it is both bounded above and below, equivalently if $(|u_n|)$ is bounded above.*

## 3.3 Limit of a Sequence

The concept of a limit is fundamental: it allows us to describe mathematically how a sequence behaves as $n$ becomes very large.

**Definition 3.3.1** (Finite Limit). *Let $u^* \in \mathbb{R}$. We say that $(u_n)$ **converges to** $u^*$ if*

$$\forall \varepsilon > 0,\ \exists k \in \mathbb{N},\ \forall n \geq k, \quad |u_n - u^*| \leq \varepsilon.$$

*We then write*

$$\lim_{n\to+\infty} u_n = u^*.$$

**Definition 3.3.2** (Infinite Limit). • *We say that* $(u_n)$ ***diverges to*** $+\infty$ *if*

$$\forall M \in \mathbb{R},\ \exists k \in \mathbb{N},\ \forall n \geq k, \quad u_n \geq M.$$

*We write* $\lim_{n\to+\infty} u_n = +\infty$.

- *Similarly,* $u_n \to -\infty$ *if and only if*

$$\forall M \in \mathbb{R},\ \exists k \in \mathbb{N},\ \forall n \geq k, \quad u_n \leq M.$$

## 3.4 Operations on Limits

**Property 3.4.1** (Computation Rules). *Let* $(u_n)$ *and* $(v_n)$ *be two sequences admitting limits (finite or infinite). Whenever the expressions make sense, we have:*

$$\lim_{n\to+\infty}(u_n + v_n) = \lim_{n\to+\infty} u_n + \lim_{n\to+\infty} v_n,$$

$$\lim_{n\to+\infty}(u_n \cdot v_n) = \lim_{n\to+\infty} u_n \cdot \lim_{n\to+\infty} v_n,$$

$$\lim_{n\to+\infty} \frac{u_n}{v_n} = \frac{\lim_{n\to+\infty} u_n}{\lim_{n\to+\infty} v_n}, \quad \text{if } \lim v_n \neq 0.$$

## 3.5 Recursive Sequences

A recursive sequence is one whose terms are defined from the previous ones:

**Definition 3.5.1** (Recursive Sequence). *A sequence* $(x_n)_{n\in\mathbb{N}}$ *is called* recursive *if there exists a numerical function* $f : \mathbb{R} \to \mathbb{R}$ *such that* $(x_n)_{n\in\mathbb{N}}$ *is defined by the iterative process (or recurrence relation):*

$$\begin{cases} x_0 \text{ fixed}, \\ x_{n+1} = f(x_n), \text{ for all } n \geq 0. \end{cases}$$

A well-known example, used by our grandparents or great-grandparents, is the following:

**Example 3.5.1** (Heron's Method). *The Heron method for computing the square root of a number* $a > 0$ *consists in calculating the terms of the recursive sequence associated with the function*

$$f(x) = \frac{1}{2}\left(x + \frac{a}{x}\right).$$

**Intuitive explanation:**

- If $x_n$ is greater than $\sqrt{a}$, then $a/x_n$ is smaller than $\sqrt{a}$, and their average gets closer to the correct value.
- If $x_n$ is smaller than $\sqrt{a}$, then $a/x_n$ is greater than $\sqrt{a}$, and again the average moves toward the true square root.

Thus, the approximations become increasingly accurate at each step.

*Remark* 10. Heron's method forms the basis of modern algorithms used by computers to compute square roots efficiently.

## 3.6 Mathematical Induction

Mathematical induction is a fundamental method used to prove that a property holds for all natural numbers.

**Definition 3.6.1** (Principle of Mathematical Induction)**.** *Let $P(n)$ be a property depending on an integer $n$. To prove that $P(n)$ holds for all $n \geq 0$, it is sufficient to verify two steps:*

- ***Base case**: show that $P(0)$ is true (or $P(1)$, depending on the context).*
- ***Inductive step**: show that for all $n \geq 0$, if $P(n)$ is true, then $P(n+1)$ is also true.*

*If these two conditions are satisfied, then $P(n)$ is true for all $n \in \mathbb{N}$.*

**Example 3.6.1** (Sum of the first $n$ integers)**.** *We want to prove by induction that*

$$P(n): \quad 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}, \quad \forall n \geq 1.$$

***Base case:** for $n = 1$, we have $1 = \frac{1 \cdot 2}{2}$, which is true.*

***Inductive step:** assume the property $P(n)$ holds for some $n \in \mathbb{N}$, that is*

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

*Then, for $n + 1$:*

$$1 + 2 + \cdots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}.$$

*Hence, the formula is true for $n + 1$.*

***Conclusion:** by the principle of mathematical induction, the formula holds for all $n \geq 1$.*

## 3.7 Monotone Convergence Theorem for Sequences

Increasing (resp. decreasing) sequences are particularly simple to study, as shown by the following property:

**Theorem 3.7.1** (Monotone Convergence Theorem)**.** *Let $(u_n)_{n\geq 0}$ be a **non-decreasing** (resp. **non-increasing**) real sequence.*

- *If $(u_n)$ is **bounded above** (resp. **bounded below**), then $(u_n)$ **converges** and*

$$\lim_{n\to\infty} u_n = \sup\{u_n : n \in \mathbb{N}\} \quad \big(\textit{resp. } \lim_{n\to\infty} u_n = \inf\{u_n : n \in \mathbb{N}\}\big).$$

- *If $(u_n)$ is unbounded (increasing and unbounded, resp. decreasing and unbounded), then*

$$\lim_{n\to\infty} u_n = +\infty \quad \big(\textit{resp. } \lim_{n\to\infty} u_n = -\infty\big).$$

**Proof (bounded increasing case).** Let $S = \{u_n : n \in \mathbb{N}\}$. The sequence is bounded above, hence $S$ admits a least upper bound $L = \sup S$ (completeness of $\mathbb{R}$).

(i) $L$ **is a limit point of** $(u_n)$**.** By the definition of supremum, for every $\varepsilon > 0$ there exists $N$ such that $L - \varepsilon < u_N \leq L$.

(ii) **Monotonicity forces convergence to** $L$**.** For $n \geq N$, the monotonicity gives $u_N \leq u_n \leq L$, hence
$$0 \leq L - u_n \leq L - u_N < \varepsilon.$$
Thus $|u_n - L| < \varepsilon$ for all $n \geq N$, which proves $u_n \to L$.

The decreasing bounded case follows by replacing sup with inf and reversing inequalities. If the sequence is increasing and unbounded above, then by definition for every $M$ there exists $n$ such that $u_n > M$, which is equivalent to $u_n \xrightarrow[n\to\infty]{} +\infty$ (and similarly for the unbounded decreasing case). $\square$

*Remark* 11. In particular, if $(u_n)$ is increasing and $u_n \leq M$ for all $n$, then $\lim u_n$ **exists** and equals $\sup_n u_n$. Similarly, if $(u_n)$ is decreasing and $u_n \geq m$ for all $n$, then $\lim u_n = \inf_n u_n$.

## 3.8 Exercises on Sequences

### Geometric Sequence

*Exercise* 3.8.1 (Geometric Sequences and Partial Sum). Let the geometric sequence $x_n = c\,q^n$ with $c \neq 0$ and $q \neq 0, 1$.

1. Study the limit of $(x_n)$ depending on the values of $q$ and the sign of $c$.
2. Show that the sum of the first $n+1$ terms is
$$\sum_{k=0}^{n} x_k \;=\; \sum_{k=0}^{n} c\,q^k \;=\; c\,\frac{q^{n+1}-1}{q-1}.$$

**Solution. 1) Limits.** We distinguish the following cases:

- If $q > 1$ and $c > 0$, then $q^n \to +\infty$, hence $x_n = cq^n \to +\infty$.
- If $q > 1$ and $c < 0$, then $x_n = cq^n \to -\infty$.
- If $-1 < q < 1$, then $q^n \to 0$, hence $x_n = cq^n \to 0$.
- If $q \leq -1$, the sequence $(q^n)$ has no limit (it oscillates with unbounded amplitude if $|q| > 1$, or alternates between $\pm 1$ if $q = -1$). Therefore, $(x_n)$ has no limit.

*Remark:* for $q = 1$ (excluded here), $x_n = c$ is constant; for $q = 0$ (also excluded), $x_0 = c$ and $x_n = 0$ for $n \geq 1$.

**2) Partial sum.** Let
$$S_n = \sum_{k=0}^{n} c\,q^k = c\,(1 + q + q^2 + \cdots + q^n).$$
If $q \neq 1$, we use the classical formula
$$1 + q + \cdots + q^n = \frac{1 - q^{n+1}}{1 - q} = \frac{q^{n+1} - 1}{q - 1}.$$
Hence,
$$S_n = c\,\frac{q^{n+1}-1}{q-1}.$$
$\square$

**Study of a Certain Recursively Defined Sequence**

*Exercise* 3.8.2. Consider the sequence $(u_n)_{n\geq 0}$ defined by

$$\begin{cases} u_0 = 0, \\ u_{n+1} = \sqrt{u_n + 1}, \quad \text{for all } n \geq 0. \end{cases}$$

1. Show that $(u_n)_{n\geq 0}$ is increasing. Deduce that the sequence is well-defined for all $n$.
2. Show that $(u_n)$ is bounded above.
3. Conclude that $(u_n)$ admits a limit (by citing the monotone convergence theorem), denoted $\ell$, and show that $\ell = \dfrac{1+\sqrt{5}}{2}$ (by citing the uniqueness of the limit).

**Solution. 1) Increasing monotonicity.** Let $f(x) = \sqrt{x+1}$ defined on $[0, +\infty)$. The function $f$ is increasing on $[0, +\infty)$ (since $f'(x) = \frac{1}{2}(x+1)^{-1/2} > 0$). We have $u_0 = 0$ and $u_1 = f(u_0) = \sqrt{1} = 1$, so $u_0 \leq u_1$.

Assume, for some $n \geq 0$, that $u_n \leq u_{n+1}$. Then, since $f$ is increasing,

$$u_{n+1} = f(u_n) \leq f(u_{n+1}) = u_{n+2}.$$

By induction, the sequence is increasing: $u_n \leq u_{n+1}$ for all $n \geq 0$.

Moreover, the recurrence $u_{n+1} = \sqrt{u_n + 1}$ is well-defined as long as $u_n \geq 0$, which holds by monotonicity starting from $u_0 = 0$. Thus the sequence is well-defined for all $n$.

**2) Explicit upper bound.** We look for a constant $\alpha > 0$ such that $u_n \leq \alpha$ for all $n$. We seek a fixed point of $f$:

$$\alpha = f(\alpha) \iff \alpha = \sqrt{\alpha + 1} \iff \alpha^2 - \alpha - 1 = 0.$$

This quadratic equation has two real roots $\dfrac{1 \pm \sqrt{5}}{2}$, of which only one is positive:

$$\varphi := \frac{1+\sqrt{5}}{2} \ (> 0).$$

We show by induction that $u_n \leq \varphi$ for all $n$.

- Base case: $u_0 = 0 \leq \varphi$.
- Inductive step: if $u_n \leq \varphi$, then

  $$u_{n+1} = \sqrt{u_n + 1} \leq \sqrt{\varphi + 1} = \varphi,$$

  since $\varphi$ is a fixed point of $f$. Thus, the property holds for all $n$.

Hence, the sequence is bounded above by $\varphi$.

**3) Convergence and value of the limit.** The sequence is *increasing* (1) and *bounded above* (2). By the *monotone convergence theorem*, it converges; let $\ell = \lim_{n\to\infty} u_n$.

Taking the limit in the recurrence relation (and using the continuity of $x \mapsto \sqrt{x+1}$ on $[0, +\infty)$),

$$\ell = \sqrt{\ell + 1} \iff \ell^2 - \ell - 1 = 0.$$

The solutions of this equation are $\dfrac{1 \pm \sqrt{5}}{2}$. Since $(u_n)$ takes values in $[0, +\infty)$, any possible limit must be non-negative. By the *uniqueness of the limit* of a convergent sequence, we necessarily obtain

$$\ell = \frac{1+\sqrt{5}}{2}.$$

□

*Remark* 12. $\ell = \frac{1+\sqrt{5}}{2}$ is in fact the famous **golden ratio**!

# Chapter 4

# Continuity of a Function

The notion of limit, which involves both continuity and differentiability, is absolutely fundamental in analysis and topology. Intuitively, a limit describes a value that the function values approach as the variable gets closer and closer to a specified number.

## 4.1 Interior, Closure, and Boundary Points

**Geometric intuition.** Let $A \subset \mathbb{R}^n$ (where $n = 1, 2$, or 3, depending on the context). We distinguish several types of points according to how the set $A$ behaves in their immediate neighborhood:

- An **interior point** of $A$ is a point that has a small neighborhood *entirely contained* within $A$.
- An **adherent point** (or **closure point**) of $A$ is a point such that every neighborhood of it contains *at least one point* of $A$—in other words, $A$ "touches" this point.
- A **boundary point** of $A$ is a point whose every neighborhood intersects both $A$ and the complement $\mathbb{R}^n \setminus A$.

**Definition 4.1.1** (Formal definitions)**.** *Let $A \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$.*

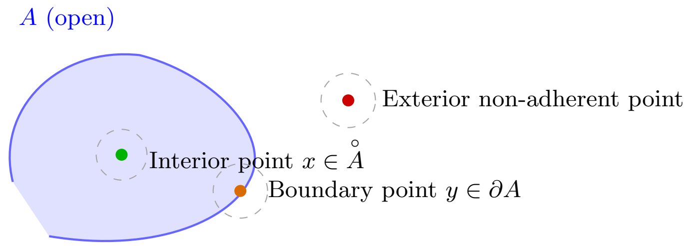
- *$x$ is an **interior point** of $A$ if there exists $r > 0$ such that the open ball $B(x, r)$ is entirely contained in $A$, i.e. $B(x, r) \subset A$.*
- *$x$ is an **adherent point** of $A$ if, for every $r > 0$, we have $B(x, r) \cap A \neq \emptyset$.*
- *$x$ is a **boundary point** of $A$ if, for every $r > 0$, we have simultaneously*

$$B(x, r) \cap A \neq \emptyset \quad \textit{and} \quad B(x, r) \cap (\mathbb{R}^n \setminus A) \neq \emptyset.$$

**Important remarks.**

- The set of all interior points of $A$ is called the **interior** of $A$, denoted by $\mathring{A}$.
- The set of all adherent (closure) points of $A$ is called the **closure** of $A$, denoted by $\overline{A}$.
- The set of boundary points is denoted by

$$\partial A = \overline{A} \setminus \mathring{A}.$$



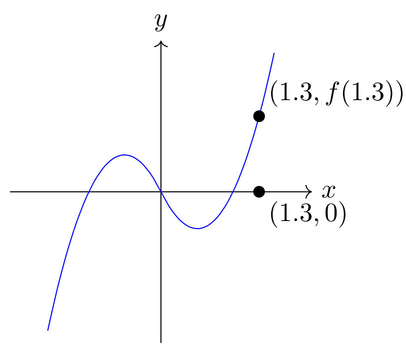

Every point of $\overline{A}$ (interior *or* boundary) is *adherent*. Here, $y \in \partial A$ is adherent, and so is $x$.

Thus:

- Every interior point belongs to $A$ and is also adherent to $A$.
- Boundary points are also adherent, but not necessarily interior.
- An exterior point, which is not adherent, has a neighborhood entirely outside $A$.

In summary, every point of $\overline{A}$—whether interior or boundary—is *adherent* to $A$.

## 4.2 Limit of a Function at a Point




In the figure above, when the values of $x$ get closer and closer to the number $x = 1.3$, either from the left or from the right, the values of $f(x)$ approach $y = f(1.3)$. In this case, we say that the limit of $f(x)$ as $x$ approaches 1.3 exists and equals $f(1.3)$.

**Definition 4.2.1** (Limit at a Point). *Let $f : A \subset \mathbb{R} \to \mathbb{R}$ be a function and let $a$ be an accumulation point of $A$. We say that $f(x)$ tends to $\ell \in \mathbb{R}$ as $x$ tends to $a$ if*

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in A, \quad 0 < |x - a| < \eta \implies |f(x) - \ell| < \varepsilon.$$

*We then write*

$$\lim_{x \to a} f(x) = \ell.$$

*Remark* 13. Two remarks:

- The condition $0 < |x - a|$ means that we do not consider the value of $f$ at $a$, but rather the behavior of $f(x)$ *in the neighborhood* of $a$.
- The number $\ell$ is called the **limit of $f$ at** $a$.

An equivalent approach, which illustrates well the idea of a function value approaching a limit step by step, is given by sequences. Let $f : A \subset \mathbb{R} \to \mathbb{R}$ and $x^*$ be an accumulation point of $A$.

**Definition 4.2.2** (Sequential Definition of the Limit)**.** *We say that $f(x)$ has limit $\ell \in \mathbb{R}$ as $x \to x^*$ if, for every sequence $(x_n)_{n\in\mathbb{N}}$ of points in $A \setminus \{x^*\}$ such that*

$$x_n \longrightarrow x^* \quad (n \to \infty),$$

*we have*

$$f(x_n) \longrightarrow \ell \quad (n \to \infty).$$

*We then write*

$$\lim_{x\to x^*} f(x) = \ell.$$

*Remark* 14. Two further observations:

- If a sequence $(x_n)$ tends to $x^*$ but $(f(x_n))$ does not admit a limit, or converges to different limits depending on the chosen sequence $(x_n)$, then the limit of $f(x)$ at $x^*$ does not exist.
- The advantage of this approach is its proximity to the intuition of sequences: to test the existence of a limit, it is enough to examine the images of all sequences converging to the point in question.

**Proposition 4.2.1** (Uniqueness of the Limit)**.** *Let $f : A \subset \mathbb{R} \to \mathbb{R}$ and $a$ be an accumulation point of $A$. If $f$ admits a limit at $a$, then this limit is unique.*

**Proof by Contradiction.** Assume that there exist $L, M \in \mathbb{R}$ such that

$$\lim_{x\to a} f(x) = L \quad \text{and} \quad \lim_{x\to a} f(x) = M,$$

and suppose for contradiction that $L \neq M$. Set $\varepsilon := \frac{|L-M|}{3} > 0$.

By the definition of the limit at $a$, there exists $\eta_1 > 0$ such that

$$0 < |x-a| < \eta_1 \;\Rightarrow\; |f(x) - L| < \varepsilon,$$

and there exists $\eta_2 > 0$ such that

$$0 < |x-a| < \eta_2 \;\Rightarrow\; |f(x) - M| < \varepsilon.$$

Let $\eta := \min(\eta_1, \eta_2) > 0$ and choose $x \in A$ such that $0 < |x-a| < \eta$ (possible since $a$ is an accumulation point of $A$). Then, simultaneously,

$$|f(x) - L| < \varepsilon \quad \text{and} \quad |f(x) - M| < \varepsilon.$$

By the triangle inequality,

$$|L - M| \leq |L - f(x)| + |f(x) - M| < \varepsilon + \varepsilon = 2\varepsilon = \tfrac{2}{3}|L - M|.$$

Thus $|L-M| < \frac{2}{3}|L-M|$, which is impossible since $|L-M| > 0$ by assumption. Contradiction. Hence our assumption is false, and we must have $L = M$. □

---

**Example 4.2.1.** *What does the expression*

$$\lim_{x\to 2}(x^2 - 4x + 6) = 2$$

*mean? It means that as $x$ approaches $2$, the expression $x^2 - 4x + 6$ approaches $\ell = 2$.*

## 4.3 Properties of Limits

Here are the fundamental properties to know:

**Property 4.3.1** (Compatibility of Limits with Usual Operations)**.** *Let $f$ and $g$ be two functions defined in a neighborhood of $x^*$. If*

$$\lim_{x\to x^*} f(x) = \kappa, \qquad \lim_{x\to x^*} g(x) = \ell,$$

*then:*

- $\lim\limits_{x\to x^*} \big(af(x) + bg(x)\big) = a\kappa + b\ell$*, for* $a, b \in \mathbb{R}$.
- $\lim\limits_{x\to x^*} (f(x))^\alpha = \kappa^\alpha$*, for* $\alpha \in \mathbb{R}$ *(provided the power makes sense).*
- $\lim\limits_{x\to x^*} (f(x)g(x)) = \kappa\ell$.
- $\lim\limits_{x\to x^*} \dfrac{f(x)}{g(x)} = \dfrac{\kappa}{\ell}$*, if* $\ell \neq 0$.
- $\lim\limits_{x\to x^*} (f \circ g)(x) = \lim\limits_{y\to\ell} f(y)$.

**Example 4.3.1.** *Suppose*

$$\lim_{x\to 2} f(x) = 4, \qquad \lim_{x\to -2} f(x) = 6, \qquad \lim_{x\to 2} g(x) = -2.$$

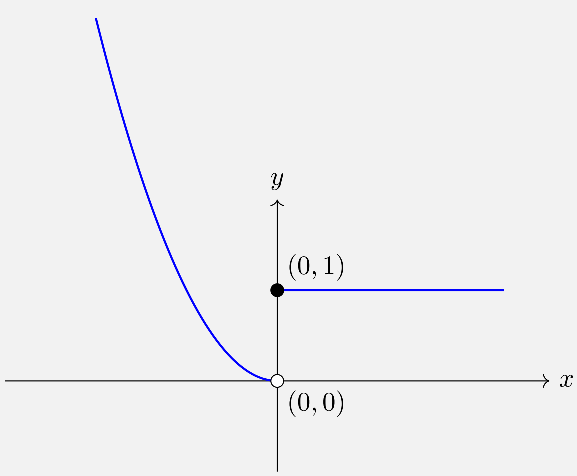
*Then:*

$$\lim_{x\to 2} \frac{f(x)}{g(x)} = \frac{4}{-2} = -2. \quad \lim_{x\to 2}(f(x))^3 = (4)^3 = 64. \quad \lim_{x\to 2}(f\circ g)(x) = \lim_{x\to -2} f(x) = 6.$$

## 4.4 Left-Hand and Right-Hand Limits

Some functions are more complex to study. As mentioned in the previous example, the value of a limit (as with many things in life) may depend on the direction from which we approach it.

**Example 4.4.1.** *Consider the function $f$ represented below:*




- *When $x \to 0^-$ (approaching from negative values), we observe that $f(x) \to 0$. Hence,*

$$\ell^- := \lim_{x\to 0^-} f(x) = 0.$$

- *When $x \to 0^+$ (approaching from positive values), we observe that $f(x) \to 1$. Hence,*

$$\ell^+ := \lim_{x\to 0^+} f(x) = 1.$$

This naturally leads to the notions of **left-hand limit** and **right-hand limit**.

**Definition 4.4.1** (Left-Hand and Right-Hand Limits)**.** *Let $f : A \subset \mathbb{R} \to \mathbb{R}$ and $x^* \in \mathbb{R}$.*

- *We say that $f$ admits a **left-hand limit** at $x^*$ equal to $\ell^-$ if*

$$\lim_{x \to x^{*-}} f(x) = \ell^-.$$

*This means that as $x$ tends to $x^*$ while remaining strictly less than $x^*$, $f(x)$ tends to $\ell^-$.*

- *We say that $f$ admits a **right-hand limit** at $x^*$ equal to $\ell^+$ if*

$$\lim_{x \to x^{*+}} f(x) = \ell^+.$$

*This means that as $x$ tends to $x^*$ while remaining strictly greater than $x^*$, $f(x)$ tends to $\ell^+$.*

## 4.5 Continuity at a Point and on a Domain

Continuity is a fundamental concept in analysis, omnipresent in many results and applications. Here is the formal definition of continuity at a point:

**Definition 4.5.1** (Continuity at a Point)**.** *Let $f : I \to \mathbb{R}$ be a function defined on an open interval $I$, and let $x^* \in I$. We say that $f$ is continuous at $x^*$ if the following three conditions are satisfied:*

1. *$f(x^*)$ exists,*
2. *$\lim_{x \to x^*} f(x)$ exists,*
3. *$\lim_{x \to x^*} f(x) = f(x^*)$.*

*Otherwise, $f$ is said to be **discontinuous** at $x^*$.*

In particular, every polynomial function is continuous. More generally, we have the following result:

**Definition 4.5.2** (Stability of Continuity under Usual Operations)**.** *Any combination of continuous functions using the usual operations (linear combination, multiplication, quotient, root, power, etc.) remains continuous on every open interval included in its domain of definition.*

Finally, this notion extends naturally to an entire interval:

**Definition 4.5.3** (Continuity on an Interval)**.**

- *If $f$ is continuous at every point $x^*$ of an open interval $I$ included in its domain of definition, then we say that $f$ is continuous on $I$ and we write $f \in C^0(I)$.*
- *$C^0(I)$ thus denotes the set of all continuous functions defined on the interval $I$.*

**Example 4.5.1.** *Consider the function*

$$f(x) = \begin{cases} \sqrt{3+x}, & x > 1, \\ x^2 + 1, & x \leq 1. \end{cases}$$

*Let us check whether this function is continuous on* $\mathbb{R}$*. Let* $x^* \in \mathbb{R}$*. We must verify whether* $f$ *is continuous at each* $x^*$ *or not. Since* $f$ *changes its expression at* $x^* = 1$*, we must consider three cases.*

- *Case* $x^* > 1$*. For* $x > 1$*, we have* $f(x) = \sqrt{3+x}$*, which is well-defined for all* $x > 1$ *and continuous at every* $x^* > 1$*.*
- *Case* $x^* < 1$*. For* $x < 1$*, we have* $f(x) = x^2 + 1$*, which is a polynomial function. Hence* $f$ *is continuous at every* $x^* < 1$*.*
- *Case* $x^* = 1$*. In this case we consider the limits:*

$$\begin{aligned} \lim_{x\to 1^+} f(x) &= \lim_{x\to 1^+} \sqrt{3+x} = \sqrt{3+1} = 2, \\ \lim_{x\to 1^-} f(x) &= \lim_{x\to 1^-} \left(x^2+1\right) = 1^2 + 1 = 2, \quad \text{and} \\ f(1) &= 1^2 + 1 = 2. \end{aligned}$$

  *Since*

$$\lim_{x\to 1^+} f(x) = \lim_{x\to 1^-} f(x) = f(1),$$

  *we conclude that* $f$ *is continuous at* $x^* = 1$*.*

*In conclusion,* $f$ *is continuous on* $\mathbb{R}$*.*

Here is another example for practice.

**Example 4.5.2.** *For which values of the parameters* $a$ *and* $b$ *is the function*

$$f(x) = \begin{cases} ax^3 - 2bx^2 - x, & x > 1, \\ ax - bx^2, & x \leq 1, \end{cases}$$

*continuous at* $x = 1$*?*
*We know that* $f$ *is continuous at* $x = 1$ *if and only if* $\lim_{x\to 1} f(x)$ *exists and equals* $f(1)$*. Since* $f$ *is given by two different expressions near* $x = 1$*, we must consider the lateral limits:*

$$\begin{aligned} \lim_{x\to 1^+} f(x) &= \lim_{x\to 1^+} \left(ax^3 - 2bx^2 - x\right) = a - 2b - 1, \\ \lim_{x\to 1^-} f(x) &= \lim_{x\to 1^-} \left(ax - bx^2\right) = a - b. \end{aligned}$$

*For* $f$ *to be continuous at* $x = 1$*, it is necessary and sufficient that*

$$\lim_{x\to 1^+} f(x) = \lim_{x\to 1^-} f(x) = f(1).$$

*Since* $f(1) = a - b$*, we obtain*

$$\begin{gathered} a - 2b - 1 = a - b, \quad \text{hence} \\ b = a^2 - a - 1. \end{gathered}$$

*Therefore, for every pair* $(a, b)$ *satisfying* $b = a^2 - a - 1$*, the function* $f$ *is continuous at* $x = 1$*.*

## 4.6 Monotone Convergence Theorem for Functions

A fundamental result applied to functions, now that we have the notion of limit:

**Theorem 4.6.1** (Monotone Convergence Theorem for Functions)**.** *Let $f : [a, +\infty[ \to \mathbb{R}$ be a* ***non-decreasing*** *(resp.* ***non-increasing****) function.*

- *If $f$ is* ***bounded above*** *(resp.* ***bounded below****), then $f$ admits a finite limit at $+\infty$, and*

$$\lim_{x\to+\infty} f(x) = \sup\{f(x) : x \geq a\} \quad \big(\text{resp. } \lim_{x\to+\infty} f(x) = \inf\{f(x) : x \geq a\}\big).$$

- *If $f$ is not bounded above (resp. below), then*

$$f(x) \xrightarrow[x\to+\infty]{} +\infty \quad \big(\text{resp. } f(x) \xrightarrow[x\to+\infty]{} -\infty\big).$$

**Idea of the Proof (Increasing and Bounded Case).** Let $L = \sup\{f(x) : x \geq a\}$ (which exists by completeness of $\mathbb{R}$).

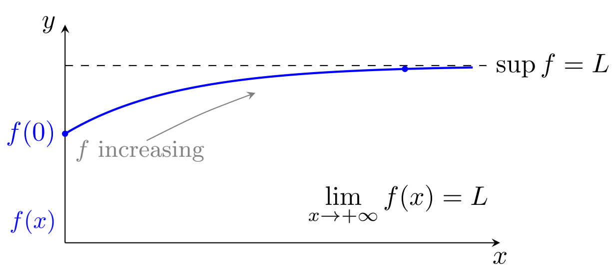
(i) For every $\varepsilon > 0$, by the definition of the supremum, there exists $X \geq a$ such that $L - \varepsilon < f(X) \leq L$.

(ii) By monotonicity, for all $x \geq X$ we have $f(X) \leq f(x) \leq L$. Hence

$$0 \leq L - f(x) \leq L - f(X) < \varepsilon.$$

Therefore $|f(x) - L| < \varepsilon$ for all $x \geq X$, which proves that $f(x) \to L$.

The decreasing bounded case is analogous, using the infimum. If $f$ is unbounded, the same reasoning shows that $f(x)$ tends to $+\infty$ (or $-\infty$). $\square$




## 4.7 Weierstrass Theorem

Since we now have the notion of continuity on a closed interval, we can state one of the most important theorems in analysis:

**Theorem 4.7.1** (Weierstrass)**.** *Every continuous function on a closed interval $[a, b]$ is* ***bounded*** *and* ***attains its bounds****: there exist $x_{\min}, x_{\max} \in [a, b]$ such that*

$$f(x_{\min}) = \min f([a, b]), \qquad f(x_{\max}) = \max f([a, b]).$$

**Idea of Proof.** The reasoning is based on a fundamental topological property of the real line.

1. The interval $[a, b]$ is **closed and bounded**, which means it is **compact** in $\mathbb{R}$ (by the Heine–Borel theorem). Intuitively, compactness expresses the idea that the set has no "holes" (it is closed) and cannot "escape to infinity" (it is bounded).

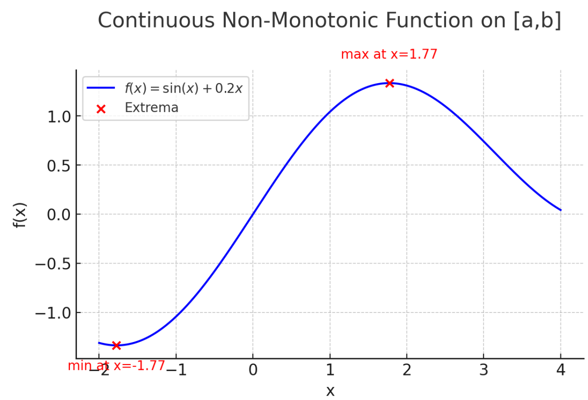
2. A fundamental result of real analysis states that *the continuous image of a compact set is compact.* Therefore, since $f$ is continuous and $[a, b]$ is compact, the image $f([a, b]) = \{f(x) \mid x \in [a, b]\}$ is also compact in $\mathbb{R}$.

3. In $\mathbb{R}$, compactness is equivalent to being both *closed* and *bounded*. Hence $f([a, b])$ is bounded (there exist $m, M \in \mathbb{R}$ with $m \leq f(x) \leq M$ for all $x \in [a, b]$) and closed (so it contains all its limit points).

4. Since $f([a, b])$ is closed and bounded, it contains its supremum and infimum. Therefore, there exist $x_{\min}, x_{\max} \in [a, b]$ such that

$$f(x_{\min}) = \min f([a, b]) \quad \text{and} \quad f(x_{\max}) = \max f([a, b]).$$

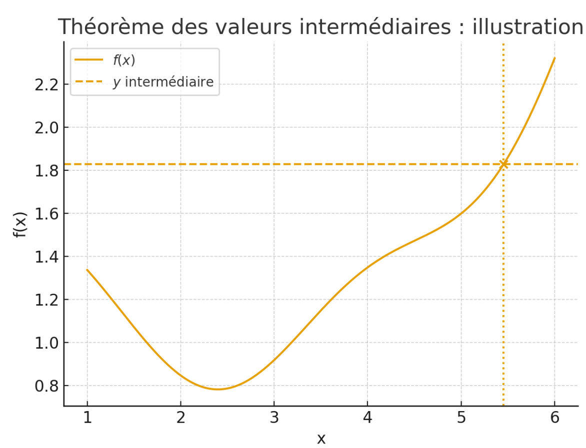
Thus, every continuous function on a closed interval attains both a maximum and a minimum value. $\square$




## 4.8 Intermediate Values and Bijection

Two intuitive yet essential theorems, with applications:

> **Theorem 4.8.1** (Intermediate Value Theorem (IVT)). *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then $f([a, b])$ is an interval. In particular, for every $y$ between $f(a)$ and $f(b)$, there exists $c \in [a, b]$ such that $f(c) = y$.*



**Theorem 4.8.2** (Bijection Theorem)**.** *Let $I$ be an interval of $\mathbb{R}$ and $f : I \to \mathbb{R}$ be* ***continuous and strictly monotone****. Then $J := f(I)$ is an interval, and $f : I \to J$ is a* ***bijection****. Moreover, the inverse function*

$$f^{-1} : J \to I$$

*is* ***continuous*** *and* ***strictly monotone*** *(in the same sense as $f$ if $f$ is increasing, and in the opposite sense if $f$ is decreasing).*

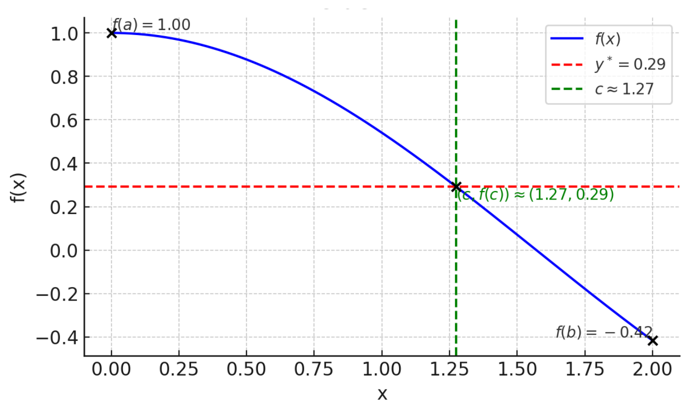



Figure 4.1: Illustration of the bijection theorem.

*Exercise* 4.8.1 (Existence and Uniqueness of a Solution via IVT and Bijection). Show that the equation

$$e^x = 3x$$

admits a unique solution in the interval $(0, 1)$.

**Solution.** Let $h(x) = e^x - 3x$ on $[0, 1]$. The function $h$ is continuous (being a combination of continuous functions). We compute:

$$h(0) = 1 > 0, \qquad h(1) = e - 3 < 0.$$

By the IVT, there exists $c \in (0, 1)$ such that $h(c) = 0$, i.e. $e^c = 3c$ (existence).

For uniqueness, note that $h'(x) = e^x - 3 < 0$ for all $x \in [0, 1]$ (since $e^x \leq e < 3$). Thus $h$ is *strictly decreasing* on $[0, 1]$ and can therefore vanish at only one point. Uniqueness follows. *We could also mention the bijection theorem here, as the link between derivative and monotonicity will be explored later in this course.*

In particular, the restriction of $h$ to $[0, 1]$ is strictly monotone and continuous; by the bijection theorem, it defines a bijection from $[0, 1]$ onto $[h(1), h(0)]$, consistent with the existence of a unique zero in $(0, 1)$. □

## 4.9 Fixed Point Theorems

### 4.9.1 Fixed Points of Continuous Functions on an Interval

**Definition 4.9.1.** *A* ***fixed point*** *of a function $f : D \to D$ (where $D \subset \mathbb{R}$) is a number $\ell \in D$ such that*

$$f(\ell) = \ell.$$

**Theorem 4.9.1** (Fixed Point Theorem for Continuous Functions)**.** *Let $f : [a, b] \to [a, b]$ be a continuous function. Then there exists at least one point $\ell \in [a, b]$ such that $f(\ell) = \ell$.*

**Idea of the proof.** Define $g(x) = f(x) - x$. Since $f$ is continuous, $g$ is also continuous on $[a, b]$. If $g(a)$ and $g(b)$ have opposite signs (for instance, $g(a) \leq 0 \leq g(b)$), then by the **Intermediate Value Theorem**, there exists $\ell \in [a, b]$ such that $g(\ell) = 0$, i.e.

$$f(\ell) = \ell.$$

Thus, every continuous self-map of a closed interval admits at least one fixed point. □

*Remark* 15. Geometrically, a fixed point corresponds to the intersection between the curve $y = f(x)$ and the diagonal line $y = x$.

### 4.9.2 Fixed Points and Monotone Sequences

Consider a sequence defined recursively by

$$u_{n+1} = f(u_n), \qquad n \geq 0,$$

where $f : [a, b] \to [a, b]$ is continuous.

**Theorem 4.9.2** (Fixed Point Theorem for Monotone Sequences)**.** *Assume that:*

- *$f : [a, b] \to [a, b]$ is continuous,*
- *the sequence $(u_n)$ defined by $u_{n+1} = f(u_n)$ is monotone (either increasing or decreasing),*
- *and $(u_n)$ is bounded.*

*Then $(u_n)$ converges to a limit $\ell \in [a, b]$, and this limit satisfies*

$$\ell = f(\ell).$$

**Idea of the proof.** Since $(u_n)$ is monotone and bounded, the **Monotone Convergence Theorem** ensures that $(u_n)$ converges to a limit $\ell \in [a, b]$. By continuity of $f$,

$$\ell = \lim_{n\to\infty} u_{n+1} = \lim_{n\to\infty} f(u_n) = f(\ell).$$

Thus, the limit of a monotone recursive sequence is a fixed point of $f$. □

*Remark* 16. This framework is often used in numerical analysis and in dynamical systems to approximate equilibrium states by successive iterations of a continuous function.

### 4.9.3 Banach Fixed Point Theorem (Contraction Principle)

The previous theorems guarantee the existence of at least one fixed point. The following result, due to Banach, also ensures **uniqueness** and provides a powerful convergence criterion.

**Theorem 4.9.3** (Banach Fixed Point Theorem)**.** *Let $f : [a, b] \to [a, b]$ be a **contraction**, i.e. there exists a constant $k \in (0, 1)$ such that*

$$|f(x) - f(y)| \leq k|x - y|, \qquad \forall x, y \in [a, b].$$

*Then:*

1. *$f$ has a unique fixed point $\ell \in [a, b]$;*
2. *For any initial value $u_0 \in [a, b]$, the sequence defined by*

$$u_{n+1} = f(u_n)$$

*converges to $\ell$.*

**Sketch of the proof.** *Uniqueness.* Suppose $f(\ell_1) = \ell_1$ and $f(\ell_2) = \ell_2$. Then

$$|\ell_1 - \ell_2| = |f(\ell_1) - f(\ell_2)| \leq k|\ell_1 - \ell_2|.$$

Since $0 < k < 1$, this implies $\ell_1 = \ell_2$.

*Convergence.* Let $\ell$ be the fixed point. For any $n \geq 0$,

$$|u_{n+1} - \ell| = |f(u_n) - f(\ell)| \leq k|u_n - \ell|.$$

By induction,

$$|u_n - \ell| \leq k^n |u_0 - \ell| \to 0 \quad \text{as } n \to \infty.$$

Hence, $u_n \to \ell$. □

*Remark* 17. The Banach theorem is fundamental in numerical analysis and functional analysis: it guarantees convergence of iterative algorithms whenever the mapping is a contraction. It also serves as a cornerstone for proving existence and uniqueness of solutions to differential and integral equations.

## Summary

| Theorem Type | Main Hypotheses | Conclusion | Nature of Result |
|---|---|---|---|
| Continuous Fixed Point | $f : [a, b] \to [a, b]$ continuous | $\exists\, \ell$ with $f(\ell) = \ell$ | Existence only |
| Monotone Sequence | $f$ continuous, $(u_n)$ monotone + bounded | $u_n \to \ell = f(\ell)$ | Existence via iteration |
| Banach Contraction | $f$ Lipschitz with $k < 1$ | $\exists!$ fixed point $\ell$ and $u_n \to \ell$ | Existence + Uniqueness + Convergence |

# Chapter 5

# Derivative of a Function at a Point and the Derivative Function

## 5.1 Motivation

Consider a vehicle traveling from Grenoble to Lyon, and let us compute its velocity. One way to calculate its speed is to consider two successive times $t$ and $t+h$ and compute the ratio between the distance traveled and the time elapsed, giving the **average velocity**:

$$\tilde{v}(t) = \frac{d(t+h) - d(t)}{h}.$$

Intuitively, the smaller $h$ is, the closer $\tilde{v}(t)$ gets to the actual speed at time $t$. This leads us to consider the limit of $\tilde{v}(t)$ as $h$ tends to 0:

$$v(t) := \lim_{h \to 0} \frac{d(t+h) - d(t)}{h}.$$

The instantaneous velocity is therefore the **instantaneous rate of change** of the distance.

More generally, for a function $f$, we define the **rate of change** between two points $x$ and $x+h$ as

$$\frac{f(x+h) - f(x)}{h},$$

and we are interested in its limit as $h \to 0$. This limit, when it exists, is called the **derivative of $f$ at $x$**.

## 5.2 Fundamental Definitions

**Definition 5.2.1** (Derivative at a Point)**.** *Let $I \subset \mathbb{R}$ be an open interval, $f : I \to \mathbb{R}$ a function, and $x^* \in I$. We say that $f$ is* ***differentiable at*** *$x^*$ if the limit*

$$\lim_{h \to 0} \frac{f(x^*+h) - f(x^*)}{h}$$

*exists and is finite.*

- *This limit is denoted $f'(x^*)$ or $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x^*)$, and it is called the* ***derivative of $f$ at*** *$x^*$.*
- *If $f$ is differentiable at every point of $I$, we say that $f$ is differentiable on $I$, and the function $f' : I \to \mathbb{R}$ is called the* ***derivative function****.*

**Definition 5.2.2** (Higher-Order Derivatives)**.** *If $f'$ is differentiable on $I$, we say that $f$ is **twice differentiable** and we define*

$$f''(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(f'(x)\right).$$

*By analogy, we define $f^{(n)}$ as the nth derivative of $f$, obtained by successive differentiations. These are called **higher-order derivatives**.*

**Examples of Derivative Computations** A few fundamental examples illustrating how to compute derivatives.

**Example 5.2.1** (Basic Derivative Examples)**.** *We compute the derivatives of several elementary functions using the definition of the derivative.*

***(1) Affine Function.*** *Let $f(x) = kx + b$. Then*

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \lim_{h\to 0}\frac{k(x+h)+b-(kx+b)}{h} = k.$$

***(2) Quadratic Function.*** *Let $f(x) = ax^2 + bx + c$. Then*

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \lim_{h\to 0}(2ax + ah + b) = 2ax + b.$$

***(3) Simple Rational Function.*** *Let $f(x) = \dfrac{1}{11-x}$. For $x \neq 11$,*

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \frac{1}{(11-x)^2}.$$

**Examples of Non-Differentiability** Here is a famous example of a non-differentiable function.

**Example 5.2.2** (The Absolute Value)**.** *Let $f(x) = |x|$. We compute the one-sided limits at 0:*

$$\lim_{h\to 0^-}\frac{f(0+h)-f(0)}{h} = \lim_{h\to 0^-}\frac{-h}{h} = -1,$$

$$\lim_{h\to 0^+}\frac{f(0+h)-f(0)}{h} = \lim_{h\to 0^+}\frac{h}{h} = 1.$$

*Since the two limits differ, $f'(0)$ does not exist. Thus, a function can be **continuous** at a point without being differentiable there.*

## 5.3 Relationship Between Continuity and Differentiability

What is the relationship between differentiability and continuity? It is quite simple: requiring a function to be differentiable at a point is a stronger condition than mere continuity, as stated below.

**Property 5.3.1.** *If a function $f$ is differentiable at $x$, then it is continuous at $x$. Indeed,*

$$\lim_{h\to 0}\big(f(x+h)-f(x)\big) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}\cdot h = f'(x)\cdot 0 = 0,$$

*hence* $\lim_{h\to 0} f(x+h) = f(x)$.

*Remark* 18. The converse is false: for example, $f(x) = |x|$ is continuous at 0 but not differentiable there.

**Example.** A Piecewise Function.

**Example 5.3.1.** *Consider*

$$f(x) = \begin{cases} 2x+1, & x < 1, \\ x^2+2, & x \geq 1. \end{cases}$$

*The function is continuous on* $\mathbb{R}$ *(by direct verification).*
*At* $x = 1$*, the one-sided derivatives are:*

$$\lim_{h \to 0^-} \frac{f(1+h) - f(1)}{h} = (2x+1)'_{|x=1} = 2, \qquad \lim_{h \to 0^+} \frac{f(1+h) - f(1)}{h} = (x^2+2)'_{|x=1} = 2.$$

*Since both limits coincide,* $f$ *is differentiable at* 1*, and* $f'(1) = 2$*.*

## 5.4 Relation with the Slope of the Tangent Line to the Graph of the Function at a Point

Let $y = f(x)$ be a function defined on an interval $I$. We wish to find the equation of the tangent line to the graph of $f$ at the point $(x_0, f(x_0))$, assuming that $f'(x_0)$ exists.

We can first consider, as an approximation, the equation of the secant line passing through the points $(x_0, y_0) = (x_0, f(x_0))$ and $(x_1, y_1) = (x_0 + h, f(x_0 + h))$ on the graph of $f$. The slope and equation of this secant line are:

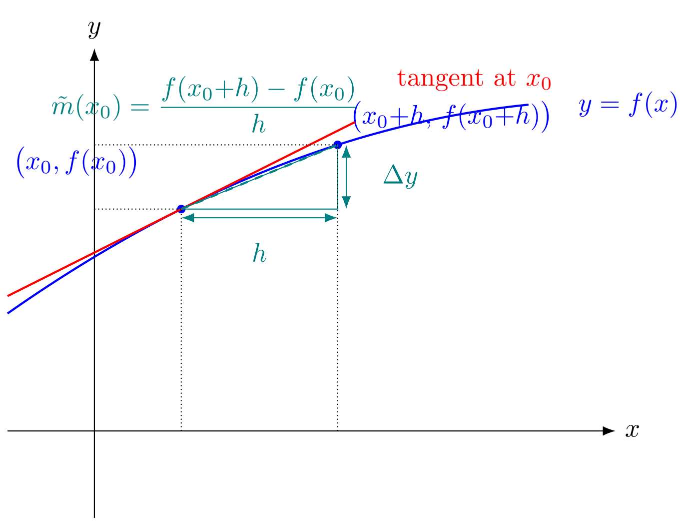
$$\tilde{m}(x_0) = \frac{y_1 - y_0}{x_1 - x_0} = \frac{f(x_0 + h) - f(x_0)}{h}, \quad y = \tilde{m}(x_0)(x - x_0) + y_0.$$

Intuitively, the slope of the tangent line to the graph of $f$ at the point $(x_0, y_0)$ should be the quantity

$$m(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0).$$

Hence, the equation of the tangent line at $(x_0, f(x_0))$ is

$$y = f'(x_0)(x - x_0) + f(x_0), \quad \text{or equivalently} \quad y = \underbrace{f'(x_0)}_{=k} x + \underbrace{(f(x_0) - x_0 f'(x_0))}_{=b}.$$



## 5.5 Differentiation Rules

Differentiation has a set of general rules that allow us to quickly compute derivatives of common functions and their combinations.

### Linear Combination

**Property 5.5.1** (Linearity)**.** *If $f$ and $g$ are two differentiable functions on $I$, and $\alpha, \beta \in \mathbb{R}$, then*

$$(\alpha f + \beta g)'(x) = \alpha f'(x) + \beta g'(x).$$

### Powers

**Property 5.5.2** (Power and Composite Power)**.** *For any $n \in \mathbb{R}$, and any differentiable function $f$,*

$$(x^n)' = nx^{n-1}, \qquad (f(x)^n)' = nf(x)^{n-1} \cdot f'(x).$$

*In particular:*

$$(1)' = 0, \quad (x)' = 1, \quad (x^2)' = 2x.$$

**Idea of the proof for $x^n$, $n \in \mathbb{N}$.** Using the definition of the derivative and the binomial expansion:

$$\lim_{h\to 0} \frac{(x+h)^n - x^n}{h} = \lim_{h\to 0} \frac{nx^{n-1}h + \cdots + h^n}{h} = nx^{n-1}.$$

$\square$

**Example 5.5.1.** *Some common derivatives illustrating the power, exponential, and logarithmic rules:*

$$(x^4)' = 4x^3, \qquad \left(x^{\frac{1}{3}}\right)' = \tfrac{1}{3}x^{-\frac{2}{3}}, \qquad (5x^{-3})' = -15x^{-4},$$

$$(e^x)' = e^x, \qquad (e^{-x})' = -e^{-x}, \qquad (e^{2x})' = 2e^{2x},$$

$$(\ln|x|)' = \tfrac{1}{x}, \qquad \left(\ln(x^2+1)\right)' = \tfrac{2x}{x^2+1}.$$

*Finally, for a more complex combination:*

$$f(x) = x^2 - 2\sqrt{x} + \tfrac{3}{\sqrt{x}} - 100 + 5\ln|x| \quad \Longrightarrow \quad f'(x) = 2x - x^{-1/2} - \tfrac{3}{2}x^{-3/2} + \tfrac{5}{x}.$$

### Product Rule

**Property 5.5.3** (Product)**.** *If $f$ and $g$ are differentiable, then*

$$(fg)' = f'g + fg'.$$

**Example 5.5.2.** *Let $f(x) = \frac{2}{3}x^3$ and $g(x) = \ln|x|$. Then*

$$(fg)' = (2x^2)(\ln|x|) + \tfrac{1}{3}x^2.$$

**Quotient Rule**

**Property 5.5.4** (Quotient). *If $f, g$ are differentiable and $g(x) \neq 0$, then*

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}.$$

**Example 5.5.3.**

$$\left(\frac{2x^3}{3\ln|x|}\right)' = \frac{2x^2(3\ln|x| - 1)}{3(\ln|x|)^2}.$$

**Composition Rule**

The result below is essential:

**Property 5.5.5** (Composite Function - Chain Rule). *If $f$ and $g$ are differentiable, then*

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x).$$

**Example 5.5.4** (Derivatives Involving Composition). *We compute the derivatives of two composite functions using the chain rule.*

***(1) Exponential Function.*** *Let $f(x) = e^{2-x+x^2}$. Then*

$$f'(x) = e^{2-x+x^2} \cdot (-1 + 2x).$$

***(2) Logarithmic Composition.*** *Let $f(x) = \big(\ln(x^2 - 1)\big)^5$. Then*

$$f'(x) = 5\big(\ln(x^2 - 1)\big)^4 \cdot \frac{2x}{x^2 - 1}.$$

## 5.6 Mean Value Theorem and Rolle's Theorem

**Theorem 5.6.1** (Rolle's Theorem). *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, differentiable on $(a, b)$, and such that $f(a) = f(b)$. Then there exists at least one $c \in (a, b)$ such that*

$$f'(c) = 0.$$

*Geometrically, this means that if a curve starts and ends at the same height, there is at least one interior point where the tangent is horizontal.*

**Idea of the proof.** Since $f$ is continuous on the closed interval $[a, b]$, it attains a maximum and a minimum (Weierstrass theorem). If $f$ is constant, then $f'(x) = 0$ everywhere. Otherwise, the maximum or minimum is reached at some interior point $c \in (a, b)$. In that case, $c$ is a local extremum and hence a critical point: $f'(c) = 0$. □

**Theorem 5.6.2** (Mean Value Theorem (MVT))**.** *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on $(a,b)$. Then there exists at least one $c \in (a,b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

*In other words, the slope of a tangent to the curve at some point $c$ equals the slope of the secant line joining the points $(a, f(a))$ and $(b, f(b))$.*

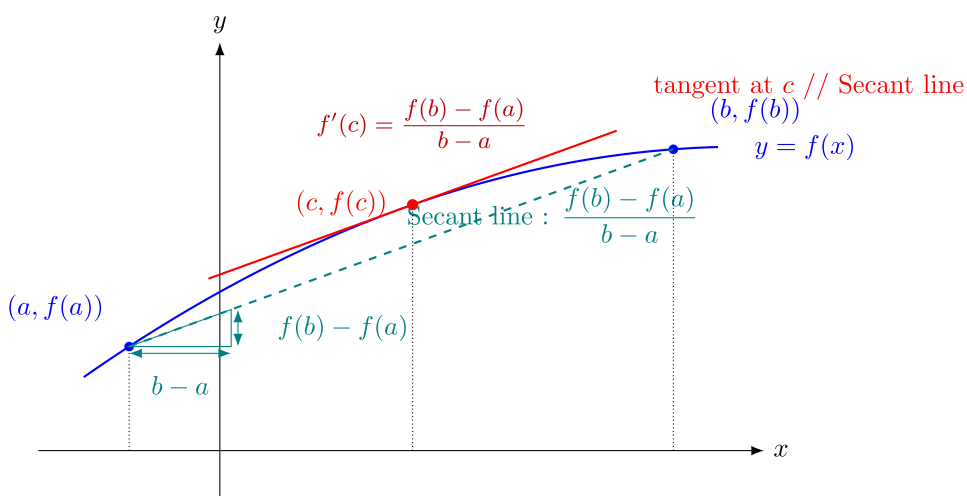
**Idea of the proof.** Consider the auxiliary function

$$g(x) = f(x) - \Big(f(a) + \frac{f(b) - f(a)}{b - a}(x - a)\Big).$$

By construction, $g(a) = g(b) = 0$. We can then apply Rolle's theorem to $g$, which gives a $c \in (a,b)$ such that $g'(c) = 0$. But

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

hence $f'(c) = \dfrac{f(b) - f(a)}{b - a}$. □




## 5.7 Monotonicity and the Sign of the Derivative

A fundamental link exists between the sign of the derivative and the monotonic behavior of a function.

**Property 5.7.1.** *Let $f$ be a differentiable function on an interval $I$. Then:*

- *If $f'(x) > 0$ for all $x \in I$, then $f$ is **strictly increasing** on $I$ (denoted $f \nearrow$).*
- *If $f'(x) < 0$ for all $x \in I$, then $f$ is **strictly decreasing** on $I$ (denoted $f \searrow$).*
- *If $f'(x) = 0$ for all $x \in I$, then $f$ is **constant** on $I$.*

*In particular, this property provides a first classification criterion for **critical points** (to be defined later).*

**Idea of the proof.** Let $a < b$ be two points of $I$. By the Mean Value Theorem (MVT), there exists $c \in (a,b)$ such that

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

- If $f'(x) > 0$ for all $x$, then in particular $f'(c) > 0$. Hence $f(b) - f(a) > 0$, which implies $f(b) > f(a)$: the function is increasing.
- If $f'(x) < 0$, we get $f(b) < f(a)$: the function is decreasing.
- If $f'(x) = 0$ everywhere, then $f(b) = f(a)$ for all $a < b$, meaning $f$ is constant.

□

---

*Exercise* 5.7.1 (Lambert $W$ Function). The Lambert function $W$ is defined as any solution (on an appropriate domain) of

$$W(z)\,e^{W(z)} = z.$$

We denote by $W_0$ the principal real branch on $[-e^{-1}, +\infty)$ and by $W_{-1}$ the secondary real branch on $[-e^{-1}, 0)$.

1. (**Inverse of** $x \mapsto xe^x$) Show that the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = xe^x$ is strictly increasing and continuous, and that its inverse on $\mathbb{R}$ is $W_0$: for all $y \in \mathbb{R}$,
$$f^{-1}(y) = W_0(y).$$
2. (**Equation solving**) Solve for $x$:
$$\text{(a)}\quad x\,e^x = a, \qquad \text{(b)}\quad e^{ax} = b\,x \quad (a > 0,\, b > 0), \qquad \text{(c)}\quad x^x = c \quad (c > 0).$$
Specify, when relevant, the conditions for the existence of real solutions and, if applicable, the number of solutions (using $W_0$ and $W_{-1}$).
3. (**Derivative of** $W$) By differentiating implicitly $W(z)e^{W(z)} = z$, show that for all $z$ such that $W(z) \neq -1$,
$$W'(z) = \frac{W(z)}{z\big(1 + W(z)\big)}.$$
4. (**Asymptotic behavior**) Show that as $x \to +\infty$,
$$W_0(x) = \ln x - \ln\ln x + o(1).$$

**Solution.** **1. Monotonicity and inverse.** We have $f'(x) = e^x(1 + x) > 0$ for all $x > -1$, and $f'(-1) = 0$, $f'(x) < 0$ for $x < -1$. Thus, $f$ is strictly decreasing on $(-\infty, -1]$ and strictly increasing on $[-1, +\infty)$, with a minimum value $f(-1) = -e^{-1}$ and $\lim_{x\to\pm\infty} f(x) = +\infty$. Hence, $f$ is bijective from $[-1, +\infty)$ onto $[-e^{-1}, +\infty)$ and its real inverse on this interval is the principal branch $W_0$. Similarly, $f$ is bijective from $(-\infty, -1]$ onto $[-e^{-1}, 0)$, with inverse branch $W_{-1}$. In particular, the multivalued inverse of $f$ over $\mathbb{R}$ is $W$, and its real inverse on $[-e^{-1}, +\infty)$ is $W_0$.

**2. Solving equations.**

(a) $xe^x = a \iff x = W(a)$. If $a \geq -e^{-1}$, there is at least one real solution: $x = W_0(a)$; if $a \in (-e^{-1}, 0)$, there are two real solutions $x = W_{-1}(a) < -1$ and $x = W_0(a) > -1$; if $a = -e^{-1}$, there is a double root $x = -1$; if $a < -e^{-1}$, there is no real solution.

(b) $e^{ax} = bx \iff e^{ax}/x = b \iff (ax)e^{ax} = ab \iff ax = W(ab)$. Therefore,
$$x = \frac{1}{a}\,W(ab).$$
Real solutions exist if and only if $ab \geq -e^{-1}$. If $0 < ab < 1$, there are two real solutions (branches $W_0$ and $W_{-1}$); if $ab = e^{-1}$ or $ab = -e^{-1}$, there is one unique solution; if $ab > 1$, there is one real solution (branch $W_0$).

(c) $x^x = c$ with $c > 0$. Let $x = e^u$ $(x > 0)$. Then $x^x = e^{u \cdot e^u} = c \iff ue^u = \ln c \iff u = W(\ln c)$. Hence,

$$x = \exp\bigl(W(\ln c)\bigr).$$

Real solutions exist iff $\ln c \geq -e^{-1}$, i.e. $c \geq e^{-e^{-1}}$. When $\ln c \in (-e^{-1}, 0)$, there are two real solutions corresponding to the branches $W_0$ and $W_{-1}$; otherwise, there is only one.

**3. Derivative of $W$.** By differentiating implicitly $W(z)e^{W(z)} = z$, we get

$$W'(z)e^{W(z)} + W(z)e^{W(z)}\, W'(z) = 1 \iff W'(z)\, e^{W(z)}\bigl(1 + W(z)\bigr) = 1.$$

Since $e^{W(z)} = \dfrac{z}{W(z)}$ (by definition), we obtain

$$W'(z)\,\frac{z}{W(z)}\bigl(1 + W(z)\bigr) = 1 \iff W'(z) = \frac{W(z)}{z(1 + W(z))},$$

valid for all $z$ such that $W(z) \neq -1$ and $z \neq 0$ (the formula extends continuously at $z = 0$ since $W(0) = 0$ gives $W'(0) = 1$).

**4. Asymptotic behavior as $x \to +\infty$.** Let $W = W_0(x)$ with $We^W = x$. Taking logarithms,

$$W + \ln W = \ln x.$$

For large $x$, $W$ is large; set a first approximation $W_1 = \ln x$, then refine by writing $W = \ln x - \ln\ln x + r(x)$. Then

$$W + \ln W = \bigl(\ln x - \ln\ln x\bigr) + \ln\bigl(\ln x - \ln\ln x + r(x)\bigr) = \ln x + o(1),$$

forcing $r(x) \to 0$. We thus obtain the classical asymptotic expansion:

$$W_0(x) \underset{x \to +\infty}{=} \ln x - \ln\ln x + o(1).$$

□

# Chapter 6

# Antiderivatives and Integrals: Computation and Applications

## 6.1 Motivations

One of the major motivations of analysis is the computation of areas of regions bounded by curves.

#### Area Under a Curve

In particular, the integral of a function measures the area under its graph between two given points. Consider the region $R$ bounded by the graph of $f$ on the interval $[a, b]$.

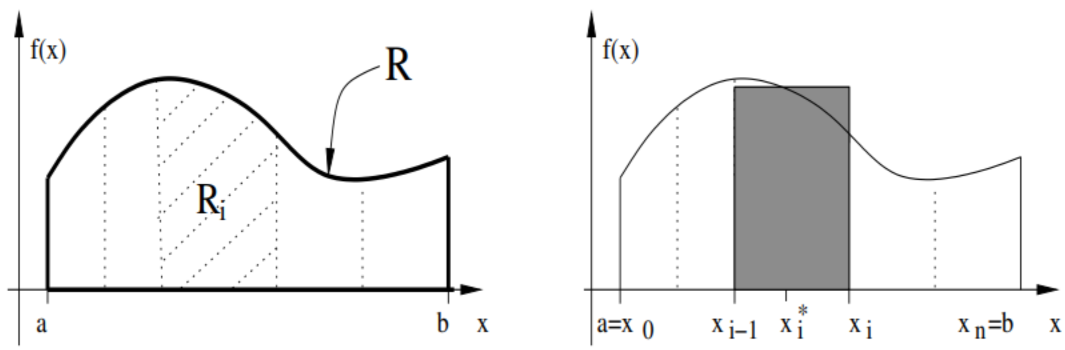



Figure 6.1: Left: the region $R$ under the graph of $y = f(x)$ on $[a, b]$ and the subregions $R_i$; Right: the same region approximated by the rectangles $R_i^*$.

A very natural way to compute the area of $R$ is to approximate it. We can divide the interval $[a, b]$ into $n$ equal subintervals, each of width $\Delta x = \frac{b-a}{n}$.

Let $x_0 = a < x_1 < \cdots < x_n = b$ be the division points, with $x_i = x_0 + i\Delta x$ for $i = 0, 1, \ldots, n$, and denote the subintervals by $[x_{i-1}, x_i]$. Let $R_i$ be the region under the graph of $f$ over $[x_{i-1}, x_i]$ (see Figure 6.1, right). Then

$$A(R) = A(R_1) + \cdots + A(R_n) = \sum_{i=1}^{n} A(R_i).$$

We approximate the area $A(R_i)$ as follows. In each interval $[x_{i-1}, x_i]$, take a point $x_i^*$ and consider the rectangle $R_i^*$ of base $[x_{i-1}, x_i]$ and height $f(x_i^*)$. Then $A(R_i)$ is approximated by

$$A(R_i^*) = (x_i - x_{i-1}) f(x_i^*) = f(x_i^*) \Delta x.$$

Thus,

$$A(R) \approx \sum_{i=1}^{n} f(x_i^*) \Delta x.$$

We then define

$$A(R) = \lim_{\Delta x \to 0} \sum_{i=1}^{n} f\left(x_i^*\right) \Delta x.$$

**Accumulation of a Quantity**

The integral is not limited to the computation of areas: it also models the **accumulation of quantities**. Consider a function $F(t)$ representing the amount of some quantity at time $t$, and let $f(t)$ denote its **instantaneous rate of change** (i.e. the derivative of $F$).

A concrete example is a bank account: $F(t)$ represents the account balance at time $t$, and $f(t)$ represents the rate of change of this balance (for instance, due to continuously compounded interest).

We wish to compute the total variation between two instants $a$ and $b$. On one hand, this variation is simply $F(b) - F(a)$. On the other hand, we can express it as a sum of elementary accumulations by dividing $[a, b]$ into $n$ equal subintervals:

$$a = t_0 < t_1 < \cdots < t_n = b,$$

and writing

$$F(b) - F(a) = \sum_{i=1}^{n} \big(F(t_i) - F(t_{i-1})\big).$$

On each subinterval $[t_{i-1}, t_i]$, the increment $F(t_i) - F(t_{i-1})$ can be approximated by $f(t_i^*)\Delta t$, where $t_i^* \in [t_{i-1}, t_i]$ and $\Delta t = t_i - t_{i-1}$. Thus:

$$F(b) - F(a) \approx \sum_{i=1}^{n} f(t_i^*)\Delta t.$$

Letting the subinterval width tend to zero ($n \to \infty$ or $\Delta t \to 0$), we obtain

$$F(b) - F(a) = \lim_{\Delta t \to 0} \sum_{i=1}^{n} f(t_i^*)\Delta t,$$

which is precisely the definition of the integral of $f$ between $a$ and $b$.

Hence,

$$F(b) - F(a) = A(R), \tag{6.1}$$

where $A(R)$ represents the area under the graph of $f(t)$ on $[a, b]$.

Similarly, replacing $b$ by a variable $t$, we obtain

$$F(t) - F(a) = A(R(t)),$$

where $A(R(t))$ denotes the area under the graph of $f$ on $[a, t]$.

Thus, **the integral links the notion of accumulated variation with that of the area under a curve**: this is the essence of the Fundamental Theorem of Calculus.

## 6.2 Antiderivative

To compute the area $A(R)$, one is naturally led to find functions $F$ such that $F'(x) = f(x)$. This observation leads us first to the study of the **indefinite integral**.

**Definition 6.2.1.** *Let $I \subset \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ a function. A function $F : I \to \mathbb{R}$ is called an **antiderivative** of $f$ on $I$ if*

$$F'(x) = f(x), \quad \text{for all } x \in I.$$

**Fundamental Properties of Antiderivatives**

**Theorem 6.2.1** (Fundamental Theorem of Calculus, Part I)**.** *Let $I \subset \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be continuous. Fix $a \in I$ and define, for every $t \in I$,*

$$F(t) \;=\; \int_a^t f(x)\,\mathrm{d}x.$$

*Then $F$ is differentiable at every point $t \in I$ where the integral is well-defined, and*

$$F'(t) \;=\; f(t).$$

**Proof.** Let $t \in I$ be such that the integral $\int_a^t f$ is well-defined (in particular, any interior point of $I$). For $h$ small enough so that $t$ and $t+h$ both belong to $I$, we have

$$\frac{F(t+h)-F(t)}{h} \;=\; \frac{1}{h}\int_t^{t+h} f(x)\,\mathrm{d}x.$$

If $h > 0$, by continuity of $f$ on the compact interval $[t, t+h]$, the integral mean value theorem yields a point $c_h \in [t, t+h]$ such that

$$\int_t^{t+h} f(x)\,\mathrm{d}x \;=\; f(c_h)\,h,$$

hence $\dfrac{F(t+h)-F(t)}{h} = f(c_h)$. Similarly, if $h < 0$, continuity on $[t+h, t]$ provides $c_h \in [t+h, t]$ with the same identity, again giving $\dfrac{F(t+h)-F(t)}{h} = f(c_h)$.

In both cases $c_h \to t$ as $h \to 0$, and by continuity $f(c_h) \to f(t)$. Therefore

$$\lim_{h\to 0} \frac{F(t+h)-F(t)}{h} \;=\; f(t),$$

which proves that $F'(t) = f(t)$. □

**Property 6.2.1** (Basic antiderivative formulas)**.** *Let $f, g$ be functions for which the following expressions make sense, and let $a, b, k \in \mathbb{R}$ with the stated restrictions. Then:*

$$\int x^n\,\mathrm{d}x \;=\; \frac{1}{n+1}\,x^{n+1} + C \qquad (n \neq -1),$$

$$\int e^{kx}\,\mathrm{d}x \;=\; \frac{1}{k}\,e^{kx} + C \qquad (k \neq 0),$$

$$\int \frac{1}{x}\,\mathrm{d}x \;=\; \ln|x| + C,$$

$$\int \big(a\,f(x) + b\,g(x)\big)\,\mathrm{d}x \;=\; a\int f(x)\,\mathrm{d}x \;+\; b\int g(x)\,\mathrm{d}x.$$

*Here $C \in \mathbb{R}$ denotes an arbitrary constant of integration.*

**Theorem 6.2.2** (Fundamental Theorem of Calculus, Part II)**.** *Let $f : [a,b] \to \mathbb{R}$ be continuous, and let $F$ be* any *antiderivative of $f$ on $[a,b]$ (that is, $F'(x) = f(x)$ for all $x \in [a,b]$). Then:*

$$\int_a^b f(x)\,\mathrm{d}x \;=\; F(b) - F(a).$$

**Sketch of proof.** By the mean value theorem for integrals, for any partition $a = x_0 < x_1 < \cdots < x_n = b$ there exists $\xi_i \in [x_{i-1}, x_i]$ such that

$$\int_{x_{i-1}}^{x_i} f(x)\,\mathrm{d}x = f(\xi_i)(x_i - x_{i-1}).$$

Summing and applying the definition of $F'(x) = f(x)$ gives

$$\sum_i f(\xi_i)(x_i - x_{i-1}) \approx \sum_i \big(F(x_i) - F(x_{i-1})\big) = F(b) - F(a).$$

As the mesh of the partition tends to 0, the Riemann sums converge to the integral, proving that

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a).$$

□

*Remark* 19 (Pedagogical intuition). Part I shows that *differentiation and integration are inverse operations* when we define the integral as an accumulation function. Part II provides the computational tool: to evaluate a definite integral, find *any* antiderivative and take the difference at the endpoints. In short:

$$\text{Derivative of the integral} = f, \qquad \text{Integral of the derivative} = F(b) - F(a).$$

Together, they establish the profound link between the local rate of change (derivative) and the global accumulation (integral).

### Integration by Parts

Recall the product rule for differentiation:

$$(fg)' = f'g + fg'.$$

Integrating both sides gives:

$$\int (fg)' = \int f'g + \int fg'.$$

By the Fundamental Theorem of Calculus,

$$\int (fg)' = fg.$$

Thus, we obtain the **integration by parts formula (IBP)**:

**Property 6.2.2** (Integration by Parts)**.** *Let $f, g$ be two functions of class $\mathcal{C}^1$ on an interval $[a, b]$. Then:*

$$\int_a^b f(x)g'(x)\,\mathrm{d}x = \big[f(x)g(x)\big]_a^b - \int_a^b f'(x)g(x)\,\mathrm{d}x,$$

*where the notation*

$$\big[f(x)g(x)\big]_a^b := f(b)g(b) - f(a)g(a)$$

*denotes the difference of the values of $f(x)g(x)$ at the endpoints.*

**Example 6.2.1.** *Let us compute* $\int xe^x \,\mathrm{d}x$.
*We set*

$$f(x) = x, \quad g'(x) = e^x \quad \Rightarrow \quad f'(x) = 1, \; g(x) = e^x.$$

*Then, applying the IBP formula:*

$$\begin{aligned}\int xe^x \,\mathrm{d}x &= f(x)g(x) - \int f'(x)g(x)\,\mathrm{d}x \\ &= xe^x - \int e^x \,\mathrm{d}x \\ &= xe^x - e^x + C.\end{aligned}$$

## 6.3 Integrals

We now formalize the notion of the integral, which, among other things, allows us to measure the area under the curve representing a function $f$ between two bounds, defined as follows:

**Definition 6.3.1** (Definite Integral)**.** *Let $I \subset \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ a* ***continuous*** *function. If $F$ is an antiderivative of $f$ on $I$ (that is, $F'(x) = f(x)$), then for $a, b \in I$ we define*

$$\int_a^b f(x)\,\mathrm{d}x := F(b) - F(a) = [F(x)]_a^b,$$

*where $a$ is the* lower bound*, $b$ the* upper bound*, and $f$ the* integrand*.*

The value of $\int_a^b f(x)\,\mathrm{d}x$ is independent of the choice of the antiderivative $F$, since if $G$ is another one, then $G(x) = F(x) + C$ for some constant $C$, and thus

$$G(b) - G(a) = (F(b) + C) - (F(a) + C) = F(b) - F(a).$$

By convention, the definition also holds for any order of bounds:

$$\int_b^a f(x)\,\mathrm{d}x = -\int_a^b f(x)\,\mathrm{d}x.$$

**Property 6.3.1** (Basic Properties of Integrals)**.** *For any continuous function $f$ and any $a, b, c \in I$:*

- $\displaystyle\int_a^a f(x)\,\mathrm{d}x = 0,$
- $\displaystyle\int_a^c f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x,$
- *For all $\alpha, \beta \in \mathbb{R}$ and functions $f, g$,*

$$\int_a^b (\alpha f(x) + \beta g(x))\,\mathrm{d}x = \alpha \int_a^b f(x)\,\mathrm{d}x + \beta \int_a^b g(x)\,\mathrm{d}x. \quad \textit{[Linearity]}$$

**Property 6.3.2** (Area Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous and nonnegative. Then the area $A$ of the region bounded by the curve $y = f(x)$ and the $x$-axis over $[a, b]$ is*

$$A = \int_a^b f(x)\,\mathrm{d}x.$$

**Theorem 6.3.1** (Fundamental Theorem of Calculus)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous. For $x \in [a, b]$, define*

$$F(x) = \int_a^x f(t)\,\mathrm{d}t.$$

*Then $F$ is continuous on $[a, b]$, differentiable on $(a, b)$, and*

$$F'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x f(t)\,\mathrm{d}t\right) = f(x), \quad \textit{for all } x \in (a, b).$$

This theorem shows that differentiation and integration are inverse operations, in the sense that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x f(t)\,\mathrm{d}t\right) = f(x), \qquad \int_a^x \frac{\mathrm{d}}{\mathrm{d}t} f(t)\,\mathrm{d}t = f(x), \ \ \text{if } f(a) = 0.$$

The definite integral can also be introduced via the notion of *Riemann sums*, which form the mathematical foundation of the concept.

Let $n \in \mathbb{N}$, $\Delta x = \dfrac{b-a}{n}$, and $x_i = a + i \cdot \Delta x$ for $i = 0, 1, \ldots, n$. For each $i$, choose a point $x_i^* \in [x_{i-1}, x_i]$. We then define the **Riemann sum**:

$$S_n = \sum_{i=1}^{n} f(x_i^*)\,(x_i - x_{i-1}) = \sum_{i=1}^{n} f(x_i^*)\,\Delta x.$$

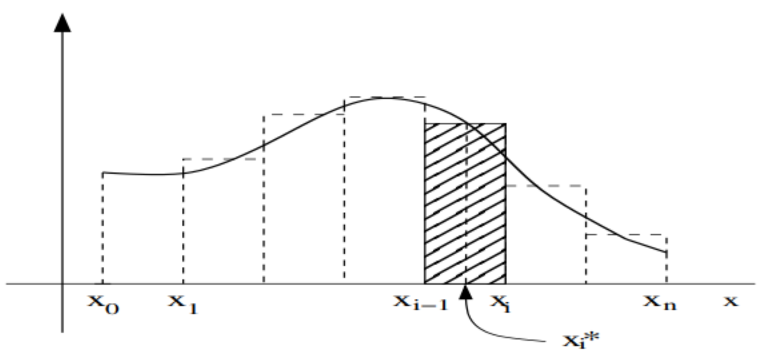



Figure 6.2: The region between the graph $y = f(x)$ and the interval $[a, b]$.

This sum represents an approximation of the area of the region between the graph of $f$ and the $x$-axis on $[a, b]$. The integral of $f$ on $[a, b]$ is then defined as the limit of these sums:

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{n\to\infty} S_n = \lim_{\Delta x \to 0} \sum_{i=1}^{n} f(x_i^*)\,\Delta x.$$

This limit exists — independently of the choice of the sample points $x_i^*$ — for broad classes of functions, in particular for continuous (or piecewise continuous) functions on $[a, b]$.

In practice, the computation of definite integrals uses the same techniques as for indefinite ones (substitution, integration by parts, etc.). The only difference is that the primitive must be evaluated at the bounds: if $F(x)$ is an antiderivative of $f$, then

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a).$$

*Exercise* 6.3.1. Determine all antiderivatives of the following functions, on an appropriate interval:

$$\begin{array}{lll} f_1(x) = 5x^3 - 3x + 7 & f_2(x) = 2\cos(x) - 3\sin(x) & f_3(x) = 10 - 3e^x + x \\ f_4(x) = \frac{5}{\sqrt{x}} + \frac{4}{x} + \frac{2}{x^2} + \frac{2}{x^3} & f_5(x) = \frac{x+5}{x^2} & f_6(x) = \frac{x^2}{5} + \frac{1}{6} \end{array}$$

**Solutions.**

1. $x \mapsto \frac{5}{4}x^4 - \frac{3}{2}x^2 + 7x + C,\ C \in \mathbb{R}$, on $\mathbb{R}$;
2. $x \mapsto 2\sin(x) + 3\cos(x) + C,\ C \in \mathbb{R}$, on $\mathbb{R}$;
3. $x \mapsto 10x - 3e^x + \frac{x^2}{2} + C,\ C \in \mathbb{R}$, on $\mathbb{R}$;
4. $x \mapsto 10\sqrt{x} + 4\ln(x) - \frac{2}{x} - \frac{1}{x^2} + C,\ C \in \mathbb{R}$, on $]0, +\infty[$;
5. $x \mapsto \ln(x) - \frac{5}{x} + C,\ C \in \mathbb{R}$ (since $f_5(x) = \frac{1}{x} + \frac{5}{x^2}$), on $]0, +\infty[$;
6. $x \mapsto \frac{x^3}{15} + \frac{x}{6} + C,\ C \in \mathbb{R}$, on $\mathbb{R}$.

□

---

*Exercise* 6.3.2. Compute the following integrals using integration by parts:

1. $I = \int_0^1 xe^x\,\mathrm{d}x$
2. $J = \int_1^e x^2 \ln x\,\mathrm{d}x$

**Solutions.**

1. Integration by parts, with:
$$\begin{aligned} u(x) &= x, & u'(x) &= 1, \\ v'(x) &= e^x, & v(x) &= e^x. \end{aligned}$$
Hence:
$$\int_0^1 xe^x\,\mathrm{d}x = 1 \cdot e^1 - 0 \cdot e^0 - \int_0^1 e^x\,\mathrm{d}x.$$
As we know $\int e^x\,\mathrm{d}x = e^x$, we find:
$$\int_0^1 xe^x\,\mathrm{d}x = e - e + 1 = 1.$$

2. Integration by parts, with:
$$\begin{aligned} u(x) &= \ln x, & u'(x) &= \tfrac{1}{x}, \\ v'(x) &= x^2, & v(x) &= \tfrac{x^3}{3}. \end{aligned}$$
Then:
$$J = \left[\frac{x^3}{3}\ln x\right]_1^e - \frac{1}{3}\int_1^e x^2\,\mathrm{d}x = \frac{e^3}{3} - \frac{1}{9}(e^3 - 1) = \frac{2e^3 + 1}{9}.$$

□

---

## 6.4 Table of Common Functions: Derivatives and Antiderivatives

| **Function** $f(x)$ | **Derivative** $f'(x)$ | **Antiderivative** $\int f(x)\,\mathrm{d}x$ |
|---|---|---|
| $k$ (constant) | $0$ | $kx + C$ |
| $x^n$ $(n \in \mathbb{R}, n \neq -1)$ | $nx^{n-1}$ | $\frac{x^{n+1}}{n+1} + C$ |
| $\sqrt{x} = x^{1/2}$ | $\frac{1}{2\sqrt{x}}$ | $\frac{2}{3}x^{3/2} + C$ |
| $\frac{1}{x} = x^{-1}$ $(x \neq 0)$ | $-\frac{1}{x^2}$ | $\ln\lvert x\rvert + C$ |
| $x^\alpha$ $(\alpha \in \mathbb{R}, \alpha \neq -1, x > 0)$ | $\alpha x^{\alpha-1}$ | $\frac{x^{\alpha+1}}{\alpha+1} + C$ |
| $e^{u(x)}$ | $e^{u(x)} \cdot u'(x)$ | $e^{u(x)}/u'(x)$ (if $u'(x) \neq 0$) |
| $\ln\lvert u(x)\rvert$ | $\frac{u'(x)}{u(x)}$ | $x\ln\lvert x\rvert - x + C$ (if $u(x) = x$) |
| $\sin x$ | $\cos x$ | $-\cos x + C$ |
| $\cos x$ | $-\sin x$ | $\sin x + C$ |
| $\tan x$ $(x \neq \frac{\pi}{2} + k\pi)$ | $1 + \tan^2 x = \frac{1}{\cos^2 x}$ | $-\ln\lvert\cos x\rvert + C$ |
| $u(x) \cdot v(x)$ | $u'(x)v(x) + u(x)v'(x)$ | — (no general antiderivative) |
| $\frac{u(x)}{v(x)}$ $(v(x) \neq 0)$ | $\frac{u'(x)v(x) - u(x)v'(x)}{v(x)^2}$ | — (no general antiderivative) |
| $[u(x)]^n$ $(n \in \mathbb{R})$ | $n[u(x)]^{n-1}u'(x)$ | — (antiderivative depends on $u$) |

**Remark:** These formulas can be combined using the rules of differential calculus (linearity, product, quotient, composition). They form the fundamental *toolbox* of differential and integral calculus.

# Chapter 7

# Application to Optimization

## 7.1 Optimization of a Function

One of the most important problems arising from the notions of differential calculus introduced earlier is to find the maximum and minimum values of a function. This is called solving an **optimization problem**.

Let $I \subset \mathbb{R}$ be an interval and $f : I \mapsto \mathbb{R}$ a function defined on $I$ with $\mathrm{Dom}(f) = I$. Assume that $f$ is continuous on $I$, i.e. $f \in C^0(I)$.

We seek the maximum and minimum values of $f$ on $I$. The maximum value, respectively minimum value, of $f$ on $I$ is a value $M = f(x_M)$, respectively $m = f(x_m)$, for some $x_M \in I$, respectively $x_m \in I$, such that

$$M = f(x_M) = \max\{f(x) \; : \; x \in I\}, \quad m = f(x_m) = \min\{f(x) \; : \; x \in I\}.$$

The maximum and minimum values of $f$ are called its **extreme values**, or **extrema**. The point $x_M$, respectively $x_m$, where the maximum (resp. minimum) is reached, is called a **maximum point**, respectively a **minimum point**.

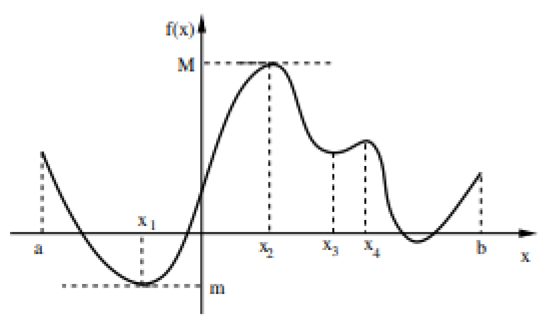



Figure 7.1: A continuous function $f \in C^0([a, b])$ and its extrema.

**Problem:** Do the extrema of $f$ on $I$ exist? And if they do, how can we find them? To answer these questions, we must introduce a few notions.

**Definition 7.1.1** (Critical Point)**.** *A point $c \in I$ is called a* ***critical point (CP)*** *of $f$ on $I$ if:*

- *$f'(c)$ exists and $f'(c) = 0$, or*
- *$f'(c)$ does not exist but $f(c)$ does.*

**Property 7.1.1.** *If $f(c)$ is an extremum, then $c$ is a critical point.*

**Proof.** Assume first that $c$ is an **interior point** of the domain of $f$ and that $f$ has a local maximum at $c$. Then there exists $r > 0$ such that for all $h$ with $|h| < r$ and $c + h$ in the interval,

$$f(c+h) \leq f(c).$$

Hence,

$$\text{if } h > 0 : \quad \frac{f(c+h) - f(c)}{h} \leq 0, \qquad \text{if } h < 0 : \quad \frac{f(c+h) - f(c)}{h} \geq 0.$$

Two cases arise:

- If $f'(c)$ **does not exist**, then by definition $c$ is a critical point and we are done.
- If $f'(c)$ **exists**, then both one-sided limits exist and equal $f'(c)$:

  $$\lim_{h \to 0^+} \frac{f(c+h) - f(c)}{h} \leq 0 \quad \text{and} \quad \lim_{h \to 0^-} \frac{f(c+h) - f(c)}{h} \geq 0.$$

  Thus,

  $$f'(c) \leq 0 \quad \text{and} \quad f'(c) \geq 0,$$

  implying $f'(c) = 0$. Therefore, $c$ is a critical point.

The case of a **local minimum** follows by reversing the inequalities: for $h > 0$, $\frac{f(c+h)-f(c)}{h} \geq 0$, and for $h < 0$, $\frac{f(c+h)-f(c)}{h} \leq 0$. Hence, either $f'(c)$ does not exist or $f'(c) = 0$, so $c$ is a critical point. □

---

*Remark* 20 (Boundary Point). If $c$ is a **boundary point** of the interval, the statement above may fail (e.g., $f(x) = x$ on $[0, 1]$ has a minimum at 0 but $f'(0) = 1$). In this case, the correct conclusion is one-sided: for a maximum at the left endpoint, $f'_+(c) \leq 0$ if the right-hand derivative exists (and analogously for other cases).

*Remark* 21. The converse is not true in general. For instance, the function $f(x) = x^3$ on $I = [-1, 1]$ has a critical point at $c = 0$, but $f(0)$ is neither a maximum nor a minimum of $f$ on $I$.

To find the extrema of $f$, we must search for its critical points in $I$. Using this notion, we can state the following result:

**Property 7.1.2** (Consequence of the Weierstrass Theorem)**.** *Let $I = [a, b]$ and $f \in C^0(I)$. Then the extrema of $f$ on $I$ exist. More precisely, if $m$ and $M$ denote respectively the smallest and largest values of $f$ on $I$, they can be determined as follows:*

1. *compute the critical points of $f$ in $I$;*
2. *evaluate $f$ at the critical points and at the endpoints $a$ and $b$, i.e. $f(a)$ and $f(b)$;*
3. *the smallest value among them is $m$, and the largest is $M$.*

**Example 7.1.1.** *Find the extrema of $f(x) = x^2 - 2x + 3$ on the interval $I = [0, 4]$.*
*Since $I$ is closed and bounded and $f$ is continuous on $I$, the Weierstrass theorem guarantees that $f$ has both a maximum and a minimum.*
*Compute the critical points: $f'(x) = 2x - 2$. Solving $f'(x) = 0$ gives $x = 1$.*
*We then evaluate $f$ at the critical and boundary points:*

$$f(0) = 3, \quad f(1) = 2, \quad f(4) = 11.$$

*Thus, $f(1) = 2$ is the minimum and $f(4) = 11$ is the maximum of $f$ on $I$.*

*Remark* 22. If $I$ is not closed and bounded, the extrema of $f$ on $I$ may not exist. For instance, for $f(x) = x$ on $I = (0, 1)$, the function attains neither a maximum nor a minimum in $I$. One might think the minimum is 0, but since $0 \notin I$, this minimum does not exist in $I$. The same reasoning applies to the maximum.

## 7.2 Local Extrema

The following definition is crucial in practice, since in many applications we are more interested in local extrema than global ones.

**Definition 7.2.1.** *Let $c \in I$, respectively $C \in I$. We say that $c$, respectively $C$, is a **local minimum point**, respectively a **local maximum point**, if $f(c)$, respectively $f(C)$, is the smallest (resp. largest) value of $f(x)$ for $x$ in a small neighborhood of $c$ (resp. $C$), i.e.*

$$f(c) \le f(x), \quad f(C) \ge f(x), \quad \text{for } x \text{ near } c, \text{ respectively } C.$$

**Property 7.2.1** (First Classification of Critical Points)**.** *Let $c \in (a, b)$ be a critical point of $f$. Then:*

- *if $f'(x) < 0$ for all $x \in [a, c]$ and $f'(x) > 0$ for all $x \in [c, b]$, then $c$ is a local minimum;*
- *if $f'(x) > 0$ for all $x \in [a, c]$ and $f'(x) < 0$ for all $x \in [c, b]$, then $c$ is a local maximum.*

**Proof.** (The signs of $f'$ are read on the open intervals $(a, c)$ and $(c, b)$, as the value at $c$ itself does not matter.)

**Case 1:** $f'(x) < 0$ on $(a, c)$ and $f'(x) > 0$ on $(c, b)$.
We show that $c$ is a *local minimum.*

Choose $\delta > 0$ such that $(c - \delta, c + \delta) \subset (a, b)$.
*Left of $c$.* For $x \in (c - \delta, c)$, by the Mean Value Theorem (MVT), there exists $\xi \in (x, c)$ such that

$$\frac{f(c) - f(x)}{c - x} = f'(\xi).$$

Since $f'(\xi) < 0$ and $c - x > 0$, we get $f(c) - f(x) < 0$, i.e. $f(c) < f(x)$.

*Right of $c$.* For $y \in (c, c + \delta)$, the MVT gives $\eta \in (c, y)$ such that

$$\frac{f(y) - f(c)}{y - c} = f'(\eta).$$

Since $f'(\eta) > 0$ and $y - c > 0$, we have $f(y) - f(c) > 0$, i.e. $f(y) > f(c)$.
Hence, for all $z \in (c - \delta, c + \delta) \setminus \{c\}$, $f(z) > f(c)$, proving that $c$ is a *local minimum.*

**Case 2:** $f'(x) > 0$ on $(a, c)$ and $f'(x) < 0$ on $(c, b)$.
The reasoning is entirely analogous.

*Left of $c$.* For $x \in (c - \delta, c)$, MVT gives $\xi \in (x, c)$ with

$$\frac{f(c) - f(x)}{c - x} = f'(\xi) > 0 \quad \Longrightarrow \quad f(c) > f(x).$$

*Right of $c$.* For $y \in (c, c + \delta)$, MVT gives $\eta \in (c, y)$ with

$$\frac{f(y) - f(c)}{y - c} = f'(\eta) < 0 \quad \Longrightarrow \quad f(y) < f(c).$$

Thus $f(z) < f(c)$ for $z$ close to $c$ and $z \neq c$, so $c$ is a *local maximum.* □

---

The table below summarizes this classification:

| $x$ | $a$ | | $c$ | | $b$ |
|---|---|---|---|---|---|
| $f'(x)$ | | $-$ | $0$ | $+$ | |
| $f$ | | $\searrow$ | local minimum | $\nearrow$ | |

| $x$ | $a$ | | $c$ | | $b$ |
|---|---|---|---|---|---|
| $f'(x)$ | | $+$ | $0$ | $-$ | |
| $f$ | | $\nearrow$ | local maximum | $\searrow$ | |

**Example 7.2.1.** *Find and classify the critical points of $f(x) = x\ln(x)$.*
*We have $\mathrm{Dom}(f) = (0, +\infty)$. Critical points satisfy $f'(x) = 0$:*

$$f'(x) = \ln(x) + 1 = 0 \quad \Rightarrow \quad \ln(x) = -1 \quad \Rightarrow \quad x = e^{-1}.$$

*Study the sign of $f'(x)$:*

- *If $x \in (0, e^{-1})$, e.g. $x = e^{-10}$, then $f'(e^{-10}) = \ln(e^{-10}) + 1 = -10 + 1 = -9 < 0$.*
- *If $x \in (e^{-1}, +\infty)$, e.g. $x = e^{10}$, then $f'(e^{10}) = \ln(e^{10}) + 1 = 10 + 1 = 11 > 0$.*

*Hence $f$ decreases on $(0, e^{-1})$ and increases on $(e^{-1}, +\infty)$. Therefore,*

$$f(e^{-1}) = e^{-1}\ln(e^{-1}) = -e^{-1}$$

*is a global minimum of $f$ on $(0, +\infty)$, since $e^{-1}$ is the only critical point.*

## 7.3 Convexity

The notion of convexity plays a central role in analysis and optimization. It is particularly important in *statistics* and *machine learning*. Indeed:

- a convex function has no "multiple valleys": every local minimum is also a global minimum;
- this makes convex optimization problems much easier and more reliable to solve (for example, linear regression, gradient descent, or cost functions in ML).

**Definition 7.3.1** (Convexity and Concavity)**.** *Let $f : I \to \mathbb{R}$ be a twice-differentiable function on an interval $I$.*

- *$f$ is said to be **convex** on $I$ if $f''(x) \geq 0$ for all $x \in I$.*
- *$f$ is said to be **concave** on $I$ if $f''(x) \leq 0$ for all $x \in I$.*

*Exercise* 7.3.1. Find the critical points of the following functions and classify them (minimum, maximum, inflection point, or non-differentiable point):

$$(1)\ f(x) = \tfrac{x}{x^2+4}, \quad (2)\ f(x) = \tfrac{\ln x}{x}, \quad (3)\ f(x) = |x-7|,$$
$$(4)\ f(x) = \tfrac{x}{x^2-1}, \quad (5)\ f(x) = x^2 - |x-2|, \quad (6)\ f(x) = x^3 - 9x^2 + 8x - 7.$$

**Solution.**

- Each function is analyzed by finding critical points via $f'(x) = 0$ (or nonexistence of $f'$).
- Then, we use the sign of $f''(x)$ (if it exists) or a monotonicity analysis to classify them: minimum, maximum, or cusp point (as for $|x-7|$).
- Example: for $f(x) = \frac{x}{x^2+4}$, we compute

$$f'(x) = \frac{(x^2+4) - 2x^2}{(x^2+4)^2} = \frac{4-x^2}{(x^2+4)^2}.$$

  Thus $f'(x) = 0 \iff x = \pm 2$. Then $f''(x)$ shows that $x = -2$ is a local minimum and $x = 2$ a local maximum.

□

*Exercise* 7.3.2. Determine the intervals of increase and decrease of the following functions:

$$(1)\ f(x) = \tfrac{x-1}{x^2+2}, \quad (2)\ f(x) = |x-1|, \quad (3)\ f(x) = \begin{cases} 4x-1 & x \leq 1, \\ x^2+2 & x > 1, \end{cases}$$
$$(4)\ f(x) = \tfrac{x}{1+x^2} - |x|, \quad (5)\ f(x) = \tfrac{x^2+1}{x}, \quad (6)\ f(x) = x^2 - 3|x-2|,$$
$$(7)\ f(x) = \tfrac{1}{x^2-4}, \quad (8)\ f(x) = x^3 - 3x^2 - 9x - 7.$$

**Solution.**

- Compute the derivative $f'(x)$ of each function.
- Study the sign of $f'(x)$ over the intervals determined by the critical points (solutions of $f'(x) = 0$ or where $f'$ does not exist).
- Example: for $f(x) = \frac{x-1}{x^2+2}$, we obtain

$$f'(x) = \frac{(x^2+2)\cdot 1 - (x-1)\cdot 2x}{(x^2+2)^2} = \frac{-x^2+2x+2}{(x^2+2)^2}.$$

  The sign of $f'(x)$ is that of the polynomial $-x^2+2x+2 = -(x^2-2x-2)$, which changes near the roots $x = 1 \pm \sqrt{3}$.
- From this, we deduce the intervals of increase and decrease.

□

---

*Exercise* 7.3.3 ((Economic Application).). The demand function for a product is given by $p(x) = e^{-2x}$, where $x$ denotes the quantity (number of units) and $p(x)$ the price per unit. What unit price maximizes the total revenue?

**Solution.**

1. The total revenue is $R(x) = x\,p(x) = xe^{-2x}$, for $x > 0$.
2. Compute:
$$R'(x) = (1-2x)e^{-2x}, \qquad R''(x) = 4(x-1)e^{-2x}.$$
3. The critical points satisfy $R'(x) = 0 \iff x = \frac{1}{2}$.
4. Since $R''(\frac{1}{2}) = 4\left(-\frac{1}{2}\right)e^{-1} < 0$, this corresponds to a local maximum. It is also a global maximum since $R(x) \to 0$ as $x \to +\infty$.
5. The corresponding unit price is $p(\frac{1}{2}) = e^{-1}$.

□

## 7.4 Second Derivative Test for Local Extrema

**Statement.** Let $f : \mathbb{R} \to \mathbb{R}$ be a twice-differentiable function on an interval $I$, and let $x^\star \in I$ be such that

$$f'(x^\star) = 0.$$

We wish to determine the nature of the critical point $x^\star$ (local minimum, local maximum, or inflection point).

**Taylor Expansion.** Since $f$ is twice differentiable, its second-order Taylor expansion around $x^\star$ can be written as:

$$f(x) \underset{x \to x^\star}{=} f(x^\star) + \tfrac{1}{2} f''(x^\star)(x - x^\star)^2 + o\big((x - x^\star)^2\big).$$

The linear term $f'(x^\star)(x - x^\star)$ vanishes since $f'(x^\star) = 0$. Hence, for $x$ close to $x^\star$, we have the local approximation:

$$f(x) - f(x^\star) \approx \tfrac{1}{2} f''(x^\star)(x - x^\star)^2.$$

**Interpretation.** The sign of $f''(x^\star)$ determines the concavity of $f$ near $x^\star$:

- If $f''(x^\star) > 0$, the coefficient of $(x - x^\star)^2$ is positive. The graph of $f$ is **locally convex**, and $x^\star$ is a **local minimum**.
- If $f''(x^\star) < 0$, the coefficient is negative. The graph of $f$ is **locally concave**, and $x^\star$ is a **local maximum**.
- If $f''(x^\star) = 0$, the quadratic term disappears, and higher-order derivatives must be examined to determine the nature of the critical point. This case may correspond to an **inflection point**.

**Geometric Intuition.** The second derivative $f''(x^\star)$ measures the **curvature** of the function near the critical point: a positive curvature ($f'' > 0$) indicates a "bowl" shape corresponding to a local minimum, while a negative curvature ($f'' < 0$) indicates a "hill" shape corresponding to a local maximum. This one-dimensional result generalizes naturally to higher dimensions through the study of the *Hessian matrix*.

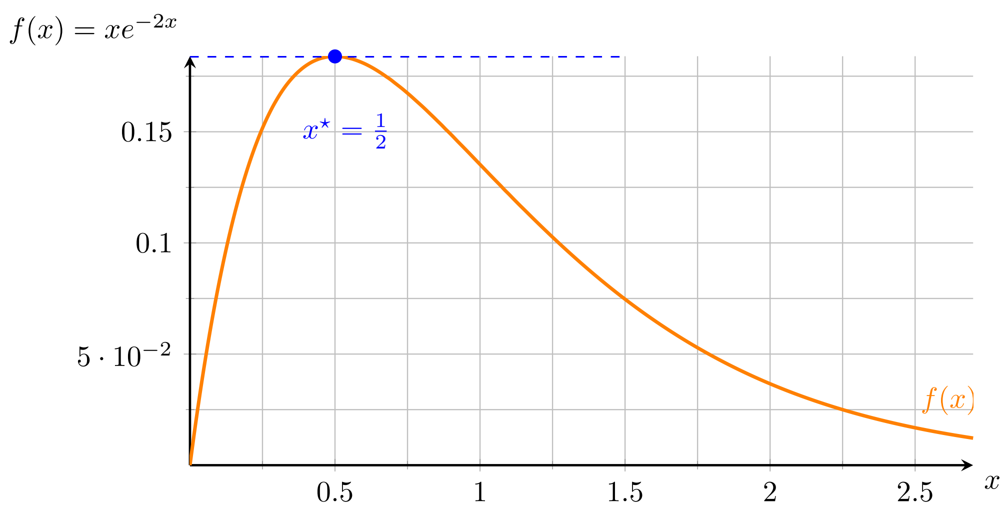



Figure 7.2: Graph of $f(x) = xe^{-2x}$ showing a local maximum at $x^\star = \frac{1}{2}$.

# Chapter 8

# Multivariable Derivatives and Optimization with Several Variables

Phenomena and real-life activities are often complex, and single-variable functions are insufficient to describe them accurately. For instance, a company may produce several products, say $n$ ($n \in \mathbb{N}^*$), in quantities $x_1, x_2, \ldots, x_n$. The total cost $C$ then depends on all these variables. If $n = 2$, one could have, for example,

$$C = 3x_1 + 10x_2 + 7x_1x_2^3.$$

This situation naturally leads to the notion of functions of several variables. At the end of this chapter, the student will:

✓ understand the concept of multivariable functions,

✓ know the definitions of continuity and partial derivatives for such functions,

✓ be able to apply tools from multivariable calculus to solve optimization problems in two dimensions.

## 8.1 Basic Definitions

We first consider $\mathbb{R} \times \mathbb{R}$, i.e. the set of pairs of real numbers:

$$\mathbb{R} \times \mathbb{R} = \{(x, y) \mid x, y \in \mathbb{R}\}.$$

We then write $(x, y) \in \mathbb{R}^2$, where $x$ and $y$ are called the coordinates of the pair. Similarly, we consider

$$\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) \mid x, y, z \in \mathbb{R}\},$$

which we denote by $\mathbb{R}^3$. This reasoning naturally extends to $\mathbb{R}^d$, the set of all $d$-tuples of real numbers.

The notion of distance between points in $\mathbb{R}^n$ is fundamental.

**Definition 8.1.1** (Euclidean Distance)**.** *The Euclidean distance between two points $p$ and $q$ is defined as follows:*

- *In one dimension: if $p = x_1, q = x_2$, then $d(p, q) = |x_1 - x_2|$.*
- *In two dimensions: if $p = (x_1, y_1), q = (x_2, y_2)$, then $d(p, q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$.*
- *In three dimensions: if $p = (x_1, y_1, z_1), q = (x_2, y_2, z_2)$, then*

$$d(p, q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$

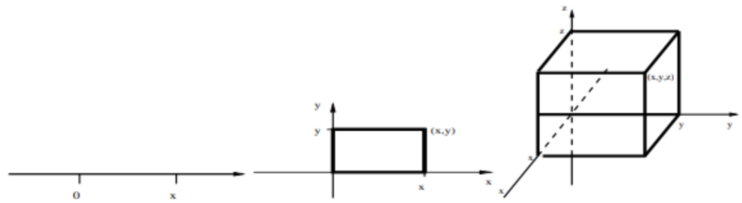



Figure 8.1: Geometric representation of elements in $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^3$.

## 8.2 Functions of Two Variables

**Definition 8.2.1.** *A function of two variables is a mapping*

$$\begin{aligned} f: \quad & \mathbb{R}^2 \to \mathbb{R}, \\ & (x,y) \mapsto z = f(x,y). \end{aligned}$$

*Here,* $\mathbb{R}^2$ *is the domain, and* $\mathbb{R}$ *is the codomain. We define:*

$$\begin{aligned} \mathrm{Dom}(f) &= \{(x,y) \in \mathbb{R}^2 \mid f(x,y) \text{ is well-defined}\}, \\ \mathrm{Im}(f) &= \{f(x,y) \mid (x,y) \in \mathrm{Dom}(f)\}. \end{aligned}$$

Alternative notations such as $(u,v)$ for the variables and $g$ for the function may be used, but in this course we will generally keep $(x,y)$ and $f$.

**Definition 8.2.2** (Graph of a Function of Two Variables)**.** *The graph of* $f$ *is the set of points*

$$G(f) = \{(x,y,f(x,y)) \mid (x,y) \in \mathrm{Dom}(f)\},$$

*that is, a surface in three-dimensional space.*

**Example 8.2.1.** *Let* $f: \mathbb{R}^2 \to \mathbb{R}$ *be defined by* $f(x,y) = x - y$*. We have* $\mathrm{Dom}(f) = \mathbb{R}^2$*, since* $f(x,y)$ *is defined for all* $(x,y) \in \mathbb{R}^2$*. Moreover,* $\mathrm{Im}(f) = \mathbb{R}$ *because for any* $z \in \mathbb{R}$*, one can choose* $(x,y) = (z,0)$ *and obtain* $f(z,0) = z$*.*

**Example 8.2.2.** *Let* $f: \mathbb{R}^2 \to \mathbb{R}$ *be defined by* $f(x,y) = \sqrt{y-x}$*. We have*

$$\mathrm{Dom}(f) = \{(x,y) \in \mathbb{R}^2 \mid y \geq x\}, \qquad \mathrm{Im}(f) = [0, +\infty).$$

*Indeed, if* $z \geq 0$*, one can take* $(x,y) = (0, z^2)$ *and obtain* $f(0,z^2) = z$*.*

We can plot the graphs of simple functions as follows:

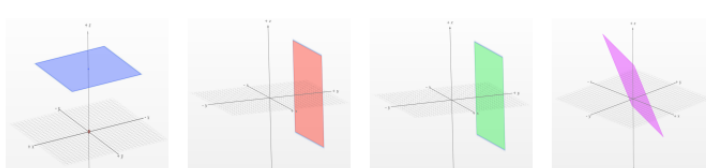


Figure 8.2: Graphs of $z = C$, $x = C$, $y = C$, and $z = ax + by + c$.

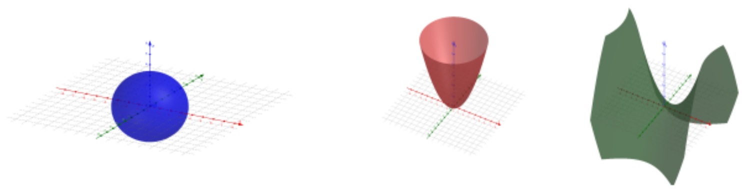


Figure 8.3: Graphs of $x^2 + y^2 + z^2 = r^2$, $z = x^2 + y^2$, and $z = x^2 - y^2$.

## 8.3 Continuity of Functions of Two Variables

Let $z = f(x, y)$ be a function of two variables defined for $(x, y) \in D \subset \mathbb{R}^2$, with $D = \text{Dom}(f)$.

*We make a short detour to recall the definition of sequences of points and their convergence.*

**Definition 8.3.1.** *A sequence of points $(p_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$ is defined by*

$$p_n = (x_{1,n}, x_{2,n}, \dots, x_{d,n}), \qquad n \in \mathbb{N},$$

*where $(x_{1,n})_{n\in\mathbb{N}}, \dots, (x_{d,n})_{n\in\mathbb{N}}$ are sequences of real numbers.*

**Definition 8.3.2.** *A sequence $(p_n)_{n\in\mathbb{N}}$ of points in $\mathbb{R}^d$ is said to converge to a point $p^*$ if*

$$d(p_n, p^*) \longrightarrow 0 \quad \text{as } n \to +\infty.$$

*In this case, we write $p_n \to p^*$.*

### Continuity

The following definition generalizes the single-variable case.

**Definition 8.3.3.** *We say that $f$ is continuous at $(x_0, y_0) \in D$ if*

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0).$$

*If $f$ is continuous at every point of $D$, we say that $f$ is continuous on $D$ and write $f \in C^0(D)$.*

**Property 8.3.1.** *Any function of two variables obtained through standard operations (sums, products, compositions with polynomial, trigonometric, exponential, or logarithmic functions) is continuous in its domain of definition.*

## 8.4 Derivatives and Partial Derivatives

The notion of partial derivative is based on the usual derivative of a single-variable function.

**Definition 8.4.1.** *Let $f \in C^0(D)$ and $(x, y) \in D$. The partial derivative of $f$ with respect to $x$ at $(x, y)$, denoted by $\frac{\partial f}{\partial x}(x, y)$, is defined as*

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

*whenever this limit exists.*

*Similarly, the partial derivative of $f$ with respect to $y$ at $(x, y)$, denoted by $\frac{\partial f}{\partial y}(x, y)$, is defined as*

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

*whenever this limit exists.*

**Example 8.4.1.** *Let us compute $f_x$ and $f_y$ for the following functions:*

1. *$f(x, y) = ax + by + c$. We have*

$$f_x(x, y) = a, \qquad f_y(x, y) = b.$$

2. *$f(x, y) = ax^2 + bxy + d\frac{x+y}{x-y}$. We obtain*

$$f_x(x, y) = 2ax + by - \frac{2dy}{(x-y)^2}, \qquad f_y(x, y) = bx + \frac{2dx}{(x-y)^2}.$$

## 8.5 Gradient of a Function

The gradient of a function $f$ at a point (say $p = (x, y)$ in two dimensions) is the vector whose components are the partial derivatives with respect to each variable.

**Definition 8.5.1.** *The gradient of $f$ at $p = (x, y)$ is defined by*

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x}(x, y) \\ \frac{\partial f}{\partial y}(x, y) \end{bmatrix}.$$

## 8.6 Tangent Plane to the Graph of $f$ at a Point $p_0$

For the tangent plane to the graph of a function $z = f(x, y)$, suppose that $f$ admits partial derivatives at $p_0 = (x_0, y_0)$.

**Definition 8.6.1.** *The equation of the tangent plane to the graph of $f$ at the point $\big(x_0, y_0, f(x_0, y_0)\big)$ is given by*

$$z = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + f(x_0, y_0).$$

The tangent plane can be interpreted as the best affine approximation of the graph of $f$ at that point, which facilitates its geometric visualization.

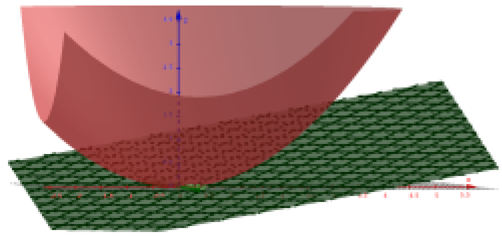


Figure 8.4: The graph of $z = f(x, y)$ and its tangent plane at $(x_0, y_0, f(x_0, y_0))$.

## 8.7 Chain Rule for Multivariable Functions

Differentiation of composite functions plays a central role in multivariable calculus. In this section, we establish the general form of the *chain rule*, first for compositions of the form

$$f : \mathbb{R}^p \to \mathbb{R}, \qquad g : \mathbb{R} \to \mathbb{R}^p,$$

and then in its most general version when both $f$ and $g$ are multivariable maps.

### 8.7.1 The Setting: Composition $f \circ g$ with $f : \mathbb{R}^p \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}^p$

Let $g : \mathbb{R} \to \mathbb{R}^p$ be a vector-valued function,

$$g(t) = \begin{bmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_p(t) \end{bmatrix},$$

and let $f : \mathbb{R}^p \to \mathbb{R}$ be a scalar function of $p$ variables,

$$f(x_1, \ldots, x_p).$$

The composition

$$h(t) = f(g(t)) = f\big(g_1(t), \ldots, g_p(t)\big)$$

is therefore a scalar function of a single variable $t$.

**Goal:** Compute the derivative $h'(t)$ in terms of $f$ and $g$.

**Definition 8.7.1** (Differentiability of the Composition)**.** *If $f$ is differentiable at $g(t)$ and each component $g_i$ is differentiable at $t$, then the composition $h = f \circ g$ is differentiable at $t$, and its derivative is given by*

$$h'(t) = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}\big(g(t)\big)\, g_i'(t).$$

**Idea of the proof.** Since $f$ is differentiable at $g(t)$, we have the first-order expansion:

$$f(g(t + \Delta t)) = f(g(t)) + \nabla f(g(t))^\top \big(g(t + \Delta t) - g(t)\big) + o(\|g(t + \Delta t) - g(t)\|).$$

Dividing by $\Delta t$ and letting $\Delta t \to 0$, we obtain:

$$h'(t) = \nabla f(g(t))^\top g'(t),$$

which corresponds exactly to the coordinate formula above. $\square$

*Remark* 23. The chain rule expresses the derivative of the composition as the product of two derivatives:

$$h'(t) = \nabla f(g(t))^\top g'(t).$$

Intuitively, it says that the rate of change of $f$ with respect to $t$ is the sum of its partial rates of change with respect to each variable $x_i$, weighted by the rate of change of $x_i$ with respect to $t$.

**Example 8.7.1.** *Let*

$$f(x_1, x_2) = x_1^2 + 3x_2, \qquad g(t) = \begin{bmatrix} e^t \\ \sin t \end{bmatrix}.$$

*Then*

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 3 \end{bmatrix}, \quad g'(t) = \begin{bmatrix} e^t \\ \cos t \end{bmatrix}.$$

*Hence*

$$h'(t) = \nabla f(g(t))^\top g'(t) = [2e^t,\ 3] \begin{bmatrix} e^t \\ \cos t \end{bmatrix} = 2e^{2t} + 3\cos t.$$

### 8.7.2 Intuitive Interpretation

The formula

$$h'(t) = \nabla f(g(t))^\top g'(t)$$

has a natural geometric interpretation.

- The vector $g'(t)$ represents the *velocity* of the point $g(t)$ moving in $\mathbb{R}^p$ as $t$ changes.
- The gradient $\nabla f(g(t))$ indicates the direction in which $f$ increases most rapidly.
- Their scalar product therefore measures how quickly $f$ changes along the trajectory defined by $g(t)$.

In other words, $h'(t)$ is the rate of change of $f$ *along the path* $g(t)$.

### 8.7.3 General Chain Rule: $f : \mathbb{R}^p \to \mathbb{R}$ and $g : \mathbb{R}^q \to \mathbb{R}^p$

**Definition 8.7.2** (Jacobian Matrix)**.** *Let $g : \mathbb{R}^q \to \mathbb{R}^p$ be a differentiable function,*

$$g(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{bmatrix}, \qquad x = (x_1, \dots, x_q)^\top.$$

*The **Jacobian matrix** of $g$ at $x$ is the $p \times q$ matrix*

$$J_g(x) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(x) & \cdots & \frac{\partial g_1}{\partial x_q}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_p}{\partial x_1}(x) & \cdots & \frac{\partial g_p}{\partial x_q}(x) \end{bmatrix}.$$

*This describes the best linear approximation of the function $g$ near the vector $x \in \mathbb{R}^q$, showing how small changes in the input variables produce changes in all output components.*

*Remark* 24. Each *row* of the Jacobian corresponds to the gradient of one component function:

$$J_g(x) = \begin{bmatrix} (\nabla g_1(x))^\top \\ \vdots \\ (\nabla g_p(x))^\top \end{bmatrix}.$$

This convention ensures that for a scalar-valued function $f : \mathbb{R}^p \to \mathbb{R}$, the chain rule can be written compactly as

$$\nabla(f \circ g)(x) = J_g(x)^\top \nabla f(g(x)).$$

We now extend the previous result to the more general case where both $f$ and $g$ are multivariable functions.

**Theorem 8.7.1** (General Chain Rule)**.** *Let*

$$f : \mathbb{R}^p \to \mathbb{R}, \qquad g : \mathbb{R}^q \to \mathbb{R}^p,$$

*and define $h = f \circ g$, i.e. $h(x) = f(g(x))$ for $x \in \mathbb{R}^q$. If $f$ is differentiable at $g(x)$ and $g$ is differentiable at $x$, then $h$ is differentiable at $x$ and*

$$\nabla h(x) = J_g(x)^\top \nabla f(g(x)),$$

*where $J_g(x)$ is the Jacobian matrix of $g$ at $x$.*

**Idea of the proof.** The total differential of $f$ satisfies

$$\mathrm{d}f = \nabla f(g(x))^\top \, \mathrm{d}g.$$

But since $\mathrm{d}g = J_g(x)\,\mathrm{d}x$, we obtain

$$\mathrm{d}h = \nabla f(g(x))^\top J_g(x)\,\mathrm{d}x.$$

Transposing both sides yields the desired formula for the gradient:

$$\nabla h(x) = J_g(x)^\top \nabla f(g(x)).$$

$\square$

*Remark* 25. The compact form

$$\nabla h = J_g^\top \nabla f$$

is the multidimensional analogue of the one-dimensional rule

$$(f \circ g)' = (f' \circ g) \cdot g'.$$

The Jacobian matrix $J_g$ generalizes the derivative $g'$ to several variables, and the transposed product plays the same compositional role.

**Example 8.7.2.** *Let*

$$f(x_1, x_2, x_3) = x_1^2 + x_2 x_3, \quad g(u, v) = \begin{bmatrix} u^2 + v \\ e^u \\ \sin v \end{bmatrix}.$$

*Then*

$$\nabla f(x_1, x_2, x_3) = \begin{bmatrix} 2x_1 \\ x_3 \\ x_2 \end{bmatrix}, \quad J_g(u, v) = \begin{bmatrix} 2u & 1 \\ e^u & 0 \\ 0 & \cos v \end{bmatrix}.$$

*Thus*

$$\nabla h(u, v) = J_g(u, v)^\top \nabla f(g(u, v)) = \begin{bmatrix} 2u & e^u & 0 \\ 1 & 0 & \cos v \end{bmatrix} \begin{bmatrix} 2(u^2 + v) \\ \sin v \\ e^u \end{bmatrix}.$$

*In summary:* The chain rule preserves its conceptual form across dimensions:

$$\text{Derivative of the composition} = \text{Product of derivatives (via the Jacobian).}$$

It is the fundamental link that unifies one-dimensional calculus and multivariable differential analysis.

## 8.8 Optimization of Functions of Two Variables

Optimization lies at the core of statistics, machine learning, artificial intelligence, and many other fields. It is therefore essential to understand how to approach a problem depending on several variables. The case of two variables remains simple to visualize, but the ideas generalize naturally to any number of variables.

Let $z = f(x, y)$ be the term of a continuous function $f$ on a bounded and closed domain $D$. We seek to determine the extremal values of $f$ on $D$, that is,

$$m = \min\{f(x, y) \mid (x, y) \in D\}, \qquad M = \max\{f(x, y) \mid (x, y) \in D\}.$$

The value $m$ is called the *global minimum of $f$ on $D$*, and $M$ the *global maximum of $f$ on $D$*. Together, they are the *global extrema* of $f$ on $D$.

**Theorem 8.8.1** (Weierstrass). *Let $D \subset \mathbb{R}^2$ be a bounded and closed domain, and let $f \in C^0(D)$. Then the global extrema of $f$ on $D$ exist and are attained.*

To determine the extrema, we introduce the notion of critical points.

**Definition 8.8.1.** *A point $(x_0, y_0) \in D$ is called a critical point of $f$ if*

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad \textit{and} \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

*Equivalently, $(x_0, y_0)$ is critical if the gradient $\nabla f(x_0, y_0)$ is zero.*

If $f$ admits an extremum at $(x_0, y_0)$, then $(x_0, y_0)$ must necessarily be a critical point (as in the one-dimensional case). However, the converse is not generally true.

**Definition 8.8.2.** *Let $(x_0, y_0) \in D$ and a neighborhood $V \subset D$ of $(x_0, y_0)$.*

- $f(x_0, y_0)$ *is a* local maximum *if*

$$\forall\, (x, y) \in V,\ f(x, y) \leq f(x_0, y_0).$$

- $f(x_0, y_0)$ *is a* local minimum *if*

$$\forall\, (x, y) \in V,\ f(x_0, y_0) \leq f(x, y).$$

- *Any local maximum or minimum is called a* local extremum.
- $(x_0, y_0)$ *is a* saddle point *if it is neither a local maximum nor a local minimum.*

The classification of critical points relies on the Hessian matrix.

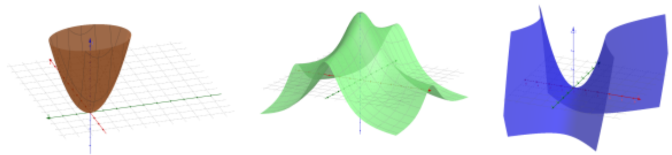


Figure 8.5: Examples: global minimum (left), local maxima (center), saddle point (right).

**Definition 8.8.3.** *The Hessian matrix of $f$ at $p = (x, y)$ is*

$$\nabla^2 f(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{bmatrix}.$$

*We say that the Hessian is positive definite if*

$$v^\top \nabla^2 f(x,y) v > 0 \quad \textit{for all } v \in \mathbb{R}^2,$$

*that is, if all its eigenvalues are strictly positive. It is negative definite if all its eigenvalues are strictly negative.*

**Property 8.8.1.** *Let $p^* = (x^*, y^*)$ be a critical point of $f$.*

- *If the Hessian at $p^*$ is positive definite, then $p^*$ is a local minimum.*
- *If the Hessian at $p^*$ is negative definite, then $p^*$ is a local maximum.*

**Idea of the proof.** We aim to understand why the sign of the Hessian matrix at a critical point determines the nature of that extremum.

1. Assume that $f$ is twice differentiable and expand $f$ near $p^* = (x^*, y^*)$ using the second-order Taylor expansion:

$$f(x,y) \underset{(x,y)\to(x^*,y^*)}{=} f(x^*,y^*) + \underbrace{\nabla f(x^*,y^*)}_{=0} \cdot \begin{bmatrix} x - x^* \\ y - y^* \end{bmatrix} + \tfrac{1}{2} \begin{bmatrix} x - x^* & y - y^* \end{bmatrix} \nabla^2 f(x^*,y^*) \begin{bmatrix} x - x^* \\ y - y^* \end{bmatrix} + o\big(\|\,(x,y) - (x^*,y^*)\,\|\big)$$

2. Since $p^*$ is a *critical point*, the linear term disappears ($\nabla f(x^*, y^*) = 0$). Thus, we essentially have:

$$f(x,y) - f(x^*,y^*) \approx \tfrac{1}{2}\, v^\top \nabla^2 f(x^*,y^*) v, \quad \text{where } v = \begin{bmatrix} x - x^* \\ y - y^* \end{bmatrix}.$$

3. If the Hessian is *positive definite*, then $v^\top \nabla^2 f(x^*, y^*) v > 0$ for all $v \neq 0$. Therefore, $f(x, y) > f(x^*, y^*)$ in a neighborhood of $p^*$: $p^*$ is a *local minimum*.

4. If the Hessian is *negative definite*, then $v^\top \nabla^2 f(x^*, y^*) v < 0$ for all $v \neq 0$. Hence, $f(x, y) < f(x^*, y^*)$ near $p^*$: $p^*$ is a *local maximum*.

5. Finally, if the Hessian is *indefinite* (that is, its eigenvalues have different signs), then depending on the direction of $v$, the quadratic term can be positive or negative. Thus, $f(x, y) - f(x^*, y^*)$ takes both positive and negative values: in some directions, $f$ increases, and in others, it decreases. In this case, $p^*$ is a ***saddle point***.

*Conclusion:* The sign of the Hessian generalizes, in two (or more) dimensions, the role of the second derivative in one dimension. □

**Example 8.8.1.** *Let $x$ (resp. $y$) denote the demand for a product $P$ (resp. $Q$), with unit prices*

$$p(x, y) = 100 - 3x - y, \qquad q(x, y) = 180 - x - 4y.$$

***Goal:*** *determine the values of $x, y$ that maximize the total revenue.*
*The total revenue is*

$$R(x, y) = xp(x, y) + yq(x, y) = 100x + 180y - 3x^2 - 2xy - 4y^2.$$

***Step 1: Critical points.*** *We solve*

$$R_x(x, y) = 100 - 6x - 2y = 0, \quad R_y(x, y) = 180 - 2x - 8y = 0.$$

*We find $(x, y) = (10, 20)$.*

***Step 2: Classification.*** *We compute*

$$R_{xx} = -6, \quad R_{xy} = -2, \quad R_{yy} = -8, \quad D = R_{xx}R_{yy} - R_{xy}^2 = (-6)(-8) - (-2)^2 = 44 > 0.$$

*Since $R_{xx} < 0$, the Hessian is negative definite: $(10, 20)$ is therefore a local maximum of the revenue.*

***Step 3: Interpretation.*** *The corresponding prices are*

$$(p, q) = (100 - 3 \cdot 10 - 20,\ 180 - 10 - 4 \cdot 20) = (50, 90).$$

*Thus, the revenue is maximized for $x = 10$ and $y = 20$.*