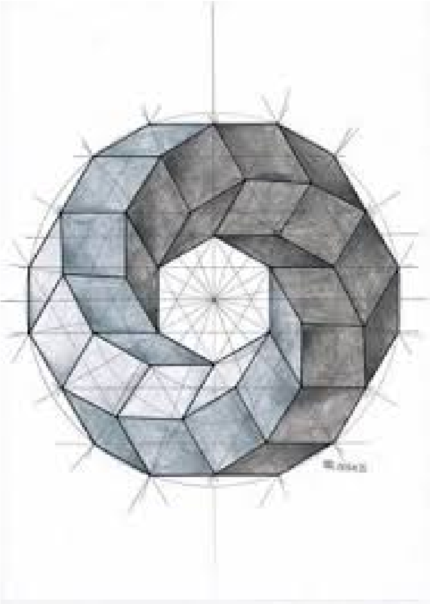
UNIVERSITY MOHAMED EL BACHIER EL IBRAHIMI
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE
DEPARTMENT OF MATHEMATICS

UNIVERSITE MOHAMED EL BACHIR EL IBRAHIMI
BORDJ BOU ARRERIDJ

Faculté des
Mathématiques et
d'Informatique

LECTURE NOTES

Analysis IV



Lecturer : FARES BENSAID

2024/2025

Contents

I Topology

II Functions of several variables

III Multiple Integrals

Preface

These lecture notes are designed for second-year undergraduate students pursuing a Bachelor's degree in Mathematics. The course provides a comprehensive introduction to the theory and applications of multivariable functions, equipping students with the foundational tools necessary for advanced studies in analysis, geometry, and mathematical physics.

Students will develop essential skills such as understanding and applying continuity, partial derivatives, and differentiability for functions of several variables. They will explore generalizations of the mean value theorems and Taylor's formula to multiple variables, learn techniques for calculating extrema — including constrained extremum problems — and master the computation of double and triple integrals with various coordinate transformations.

A solid foundation in single-variable calculus and introductory real analysis is recommended for tackling the material. Familiarity with limits, continuity, differentiability, Riemann integration, sequences and series of functions, and basic topology concepts from courses like Analysis 1 and Analysis 2 will greatly enhance comprehension.

The content is structured into three chapters, progressing from the topology of Euclidean space to the core properties of multivariable functions, and finally to multiple integrals. Theoretical explanations are complemented by illustrative examples and exercises to reinforce understanding. The aim is not only to present rigorous mathematical concepts but also to foster an intuition for working in multiple dimensions.

We hope these notes serve as a valuable resource, guiding students towards a deeper appreciation of multivariable analysis and its vast applications.

I

Topology

1. Topology of $\mathbb{R}^n$

Throughout this course, we denote by $\mathbb{R}^n$ ($n \geq 1$ being an integer) the Cartesian product

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}}_{n \text{ factors}}$$

In other words, $\mathbb{R}^n$ is the set of ordered n-tuples $(x_1, \ldots, x_n)$ of real numbers. We consider the following operations:

- **Addition:** $\forall x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n : x + y = (x_1 + y_1, \ldots, x_n + y_n)$
- **Scalar multiplication:** $\forall x = (x_1, \ldots, x_n) \in \mathbb{R}^n, \forall \lambda \in \mathbb{R} : \lambda x = (\lambda x_1, \ldots, \lambda x_n)$

Equipped with these two operations, $\mathbb{R}^n$ is a vector space over $\mathbb{R}$ of dimension n and has as its canonical basis the n vectors $e_1 = (1, 0, \ldots), e_2 = (0, 1, 0, \ldots)$ and $e_n = (0, \ldots, 0, 1)$.

1.1 Norms

Definition 1.1.1 A norm on $\mathbb{R}^n$ is any function $\mathcal{N} : \mathbb{R}^n \to [0, +\infty[$ satisfying the following properties:

1. **Separation:** For all $x \in \mathbb{R}^n, \mathcal{N}(x) = 0 \Longleftrightarrow x = 0$.
2. **Homogeneity:** For all $x \in \mathbb{R}^n$ and all $\lambda \in \mathbb{R}, \mathcal{N}(\lambda x) = |\lambda| \mathcal{N}(x)$.
3. **Triangle inequality:** For all $x, y \in \mathbb{R}^n, \mathcal{N}(x + y) \leq \mathcal{N}(x) + \mathcal{N}(y)$.

Equipped with a norm, $\mathbb{R}^n$ is called a normed vector space. The notation $\|x\|$ is commonly used for $\mathcal{N}(x)$, which recalls the analogy with the absolute value in $\mathbb{R}$.

■ **Example 1.1** In dimension $n = 1$, it can be verified that if $\mathcal{N}$ is a norm on $\mathbb{R}$, then there exists a unique $\beta > 0$ such that $\mathcal{N}(x) = \beta |x|$. ■

■ **Example 1.2** Common Norms

Let $p \in [1, +\infty[$. On $\mathbb{R}^n$, the following norms are often used, defined for $x = (x_1, x_2, \ldots, x_n)$ by

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}, \quad \|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}.$$

If $p = 2$, it is called the Euclidean norm. ■

Proposition 1.1.1 Let $\mathcal{N}$ be a norm on $\mathbb{R}^n$. Then

- $\mathcal{N}$ is a 1-**Lipschitz** function, that is, for all $x, y \in \mathbb{R}^n$,

$$|\mathcal{N}(x) - \mathcal{N}(y)| \leq \mathcal{N}(x-y).$$

- $\mathcal{N}$ is a **convex** function, that is:

$$\forall \lambda \in [0,1], \forall x, y \in \mathbb{R}^n : \mathcal{N}(\lambda x + (1-\lambda)y) \leq \lambda \mathcal{N}(x) + (1-\lambda)\mathcal{N}(y).$$

We will now state a fundamental theorem that expresses that the choice of norm does not affect the topological properties of functions on $\mathbb{R}^n$ (continuity, limit, etc.).

Definition 1.1.2 Two norms $\mathcal{N}_1$ and $\mathcal{N}_2$ on $\mathbb{R}^n$ are said to be equivalent if there exist two positive constants α and β such that

$$\alpha \mathcal{N}_1(x) \leq \mathcal{N}_2(x) \leq \beta \mathcal{N}_1(x), \quad \forall x \in \mathbb{R}^n.$$

Theorem 1.1.2 All norms on $\mathbb{R}^n$ are equivalent.

■ **Example 1.3** Common Norms
The following three norms

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} \quad \text{and} \quad \|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\},$$

are pairwise equivalent, and for all $x \in \mathbb{R}^n$, we have

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty, \quad \|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty \quad \text{and} \quad \|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2.$$

■

Definition 1.1.3 — Dot Product. A dot product (or scalar product, inner product) on $\mathbb{R}^n$ is any symmetric, positive-definite bilinear form, that is, any function

$$\begin{aligned} p : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x,y) &\mapsto p(x,y), \end{aligned}$$

which satisfies for all $x, y, z \in \mathbb{R}^n$ and all $a, \beta \in \mathbb{R}$

- $p(x,y) = p(y,x)$.
- $p(x,x) \geq 0$ and $p(x,x) = 0 \Longleftrightarrow x = 0$.
- $p(ax + \beta y, z) = ap(x,z) + \beta p(y,z)$.

Equipped with a dot product, $\mathbb{R}^n$ is called a Euclidean space.

■ **Example 1.4** The usual dot product of $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, denoted $\langle x, y \rangle$ or $x \cdot y$, is defined by

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

The norm associated with this dot product is the Euclidean norm on $\mathbb{R}^n$, that is, for all $x \in \mathbb{R}^n$:

$$\|x\|_2 = \sqrt{\langle x, x \rangle}.$$

■

1.2 Open and Closed Subsets of $\mathbb{R}^n$.

Definition 1.2.1 Balls

In $\mathbb{R}^n$ equipped with an arbitrary norm $\|x\|$:

The open ball centered at $a \in \mathbb{R}^n$ with radius $r > 0$ is the set

$$B(a,r) = \{x \in \mathbb{R}^n : \|x-a\| < r\}.$$

2. The closed ball centered at $a \in \mathbb{R}^n$ with radius $r > 0$ is the set

$$\bar{B}(a,r) = \{x \in \mathbb{R}^n : \|x-a\| \leqslant r\}.$$

3. The sphere centered at $a \in \mathbb{R}^n$ with radius $r > 0$ is the set

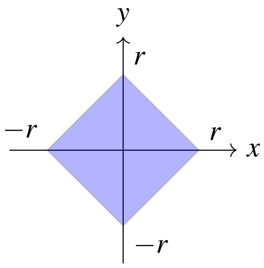
$$S(a,r) = \{x \in \mathbb{R}^n : \|x-a\| = r\}.$$

In the case where $a = 0$ and $r = 1$, these are called the unit balls or unit spheres.

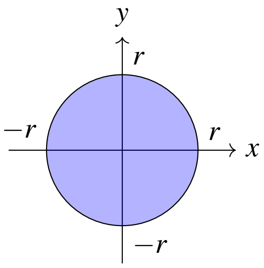
R In dimension $n = 1$, the open ball $B(a,r)$ is the open interval $]a-r, a+r[$, while the closed ball is $[a-r, a+r]$.

■ **Example 1.5** We will graphically represent in $\mathbb{R}^2$ the balls centered at $O = (0,0)$ with radius r for the three common norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$.
The closed ball centered at $O = (0,0)$ with radius $r > 0$ is the set of points

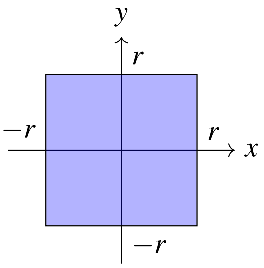
$$\bar{B}(O,r) = \{(x,y) \in \mathbb{R}^2 : \|(x,y)-(0,0)\| \leq r\} = \{(x,y) \in \mathbb{R}^2 : \|(x,y)\| \leq r\}.$$



Relative to $\|\cdot\|_1$: $|x|+|y| \leq r$.



Relative to $\|\cdot\|_2$: $\sqrt{x^2+y^2} \leq r$.



Relative to $\|\cdot\|_\infty$: $\max(|x|,|y|) \leq r$.

■

Definition 1.2.2 — Bounded Subset. A subset A of $\mathbb{R}^n$ is said to be bounded if it is contained in a closed ball of $\mathbb{R}^n$. In other words:

$$A \text{ is bounded} \iff \exists a \in \mathbb{R}^n, \exists r > 0 : A \subset \bar{B}(a,r).$$

Definition 1.2.3 — Neighborhood.
Let $\mathscr{V}$ be a subset of $\mathbb{R}^n$ and $x \in \mathscr{V}$. We say that $\mathscr{V}$ is a neighborhood of x if $\mathscr{V}$ contains an open ball centered at x. In other words:

$$\mathscr{V} \text{ is a neighborhood of } a \iff \exists r > 0 : B(x,r) \subset \mathscr{V}.$$

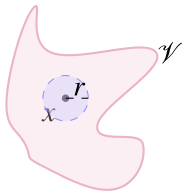


Figure 1.1: A neighborhood

Definition 1.2.4 Open Subset A subset $A \subset \mathbb{R}^n$ is said to be open if for every point $x \in A$, A contains an open ball centered at x. We also say that A is an open set of $\mathbb{R}^n$. In other words:

$$A \text{ is open} \iff \forall x \in A, \exists r > 0 : B(x,r) \subset A.$$

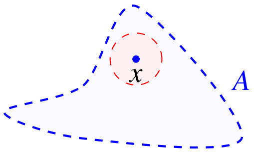


Figure 1.2: An open set

Definition 1.2.5 — Continuous Function. A function $f : X \to Y$ is said to be **continuous** if the inverse image of every open subset of Y is open in X. In other words, if $V \in \mathscr{F}_Y$, then its inverse image $f^{-1}(V) \in \mathscr{F}_X$.

Proposition 1.2.1 A function $f : X \to Y$ is continuous if and only if for each $x \in X$ and each neighborhood N of $f(x)$ in Y, the set $f^{-1}(N)$ is a neighborhood of x in X.

Theorem 1.2.2

- Let φ be a continuous function on $\mathbb{R}^n$. Let a be a real number. The sets:

  $$\{M \in \mathbb{R}^n \mid \varphi(M) < a\} \quad \text{and} \quad \{M \in \mathbb{R}^n \mid \varphi(M) > a\}$$

  are sets referred to as **open sets** in $\mathbb{R}^n$.
- $\mathbb{R}^n$ and $\varnothing$ are referred to as **open sets** in $\mathbb{R}^n$.

Example 1.6

1. Let A be a point in $\mathbb{R}^n$ and $r > 0$. Justify that the set $\mathscr{B}(A,r) = \{M \in \mathbb{R}^n \mid AM < r\}$ (ball) is open.
2. Among the different types of intervals in $\mathbb{R}$, which ones are open?
3. In $\mathbb{R}^2$, show that the half-plane $x + y < 1$ is open.

■

Theorem 1.2.3 — Compatibility with set operations.

- A finite or infinite union of open sets in $\mathbb{R}^n$ is an open set in $\mathbb{R}^n$.
- A finite intersection of open sets in $\mathbb{R}^n$ is an open set in $\mathbb{R}^n$.
- If $(I_1, \dots, I_n)$ are open intervals in $\mathbb{R}$, then the Cartesian product $I_1 \times \dots \times I_n$ is an open set in $\mathbb{R}^n$.

Example 1.7

1. Justify that $]-1,2[\times]-3,1[$ is open in $\mathbb{R}^2$.
2. Show that $\mathscr{O}_1 = \{(x,y,z) \in \mathbb{R}^3 \mid x > -2,\ y < 1,\ -1 < z < 2\}$ is open in $\mathbb{R}^3$.
3. Show that $\mathscr{O}_2 = \left\{(x,y) \in \mathbb{R}^2 \mid x+2 > 0,\ 2x+y-1 < 0,\ x-y < 0\right\}$ is open in $\mathbb{R}^2$.
4. Let A be a point in $\mathbb{R}^n$. Show that $\displaystyle\bigcap_{k\in\mathbb{N}} \left\{M \in \mathbb{R}^n \mid AM < \frac{1}{k+1}\right\}$ is not open in $\mathbb{R}^n$.

■

Definition 1.2.6 — Closed Subset. A subset $A \subset \mathbb{R}^n$ is said to be closed if and only if its complement $A^c = \mathbb{R}^n \setminus A$ is open.

Definition 1.2.7 — Pseudo-definition (=admitted theorem).

- Let φ be a continuous function on $\mathbb{R}^n$. Let a be a real number. The sets:

  $$\{M \in \mathbb{R}^n \mid \varphi(M) \leqslant a\} \quad \text{and} \quad \{M \in \mathbb{R}^n \mid \varphi(M) \geqslant a\}$$

  are sets referred to as **closed sets** in $\mathbb{R}^n$.
- $\mathbb{R}^n$ and $\varnothing$ are referred to as **closed sets** in $\mathbb{R}^n$.

Example 1.8

1. Let A be a point in $\mathbb{R}^n$ and $r > 0$. Justify that the set $\mathscr{B}_f(A,r) = \{M \in \mathbb{R}^n \mid AM \leqslant r\}$ (closed ball) is closed.
2. Among the different types of intervals in $\mathbb{R}$, which ones are closed?
3. In $\mathbb{R}^2$, show that the half-plane $x + y \geqslant 1$ is closed.

■

Theorem 1.2.4 — Compatibility with set operations.

- A finite or infinite intersection of closed sets in $\mathbb{R}^n$ is a closed set in $\mathbb{R}^n$.
- A finite union of closed sets in $\mathbb{R}^n$ is a closed set in $\mathbb{R}^n$.
- If $(I_1, \dots, I_n)$ are closed intervals in $\mathbb{R}$, then the Cartesian product $I_1 \times \cdots \times I_n$ is a closed set in $\mathbb{R}^n$.

■ **Example 1.9**

1. Justify that $[-1,2] \times [-3,1]$ is closed in $\mathbb{R}^2$.
2. Show that $\mathscr{F}_1 = \{(x,y,z) \in \mathbb{R}^3 \mid x \geqslant -2,\ y \leqslant 1,\ -1 \leqslant z \leqslant 2\}$ is closed in $\mathbb{R}^3$.
3. Show that $\mathscr{F}_2 = \left\{(x,y) \in \mathbb{R}^2 \mid x+2 \geqslant 0,\ 2x+y-1 \geqslant 0,\ x-y \leqslant 0\right\}$ is closed in $\mathbb{R}^2$.
4. Let A be a point in $\mathbb{R}^n$. Show that $\bigcap_{k\in\mathbb{N}} \left\{M \in \mathbb{R}^n \mid AM \leqslant \dfrac{1}{k+1}\right\}$ is closed, but $\bigcup_{k\in\mathbb{N}} \left\{M \in \mathbb{R}^n \mid AM \leqslant 1 - \dfrac{1}{k+1}\right\}$ is open in $\mathbb{R}^n$.

■

1.3 Interior - Closure - Boundary - Isolated Point - Accumulation Point

Definition 1.3.1 — Interior. Let A be any subset of $\mathbb{R}^n$. The **interior** of A, denoted $\mathring{A}$, is the set defined by

$$\mathring{A} = \{x \in A : \exists r > 0, B(x,r) \subset A\}.$$

Proposition 1.3.1 Let A be a subset of $\mathbb{R}^n$.

- $\subseteq A$, and moreover, is the largest open set contained in A.
- A is open if and only if $A =$.

Definition 1.3.2 Closure

Let A be any subset of $\mathbb{R}^n$. The **closure of** A (also called the **closed hull**) is the set defined by

$$\overline{A} = \{x \in \mathbb{R}^n : \forall \delta > 0, B(x,\delta) \cap A \neq \varnothing\}.$$

Proposition 1.3.2

Let A be a subset of $\mathbb{R}^n$.

- $A \subseteq \overline{A}$, and moreover, $\overline{A}$ is the smallest closed set containing A.
- A is closed if and only if $A = \overline{A}$.

Definition 1.3.3 — Boundary. Let A be any subset of $\mathbb{R}^n$. The **boundary of** A (also called the **frontier**) is the set defined by

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

Note that ∂A is a closed set.

Definition 1.3.4 — Isolated Point. Let A be a non-empty subset of $\mathbb{R}^n$, and $a \in A$. The point a is said to be isolated in A if there exists a neighborhood $\mathscr{V}$ of a in $\mathbb{R}^n$ such that

$$\mathscr{V} \cap A = \{a\}.$$

Definition 1.3.5 — Accumulation Point. Let A be a non-empty subset of $\mathbb{R}^n$, and $a \in \mathbb{R}^n$. The point a is said to be an **accumulation point of** A if it is adherent to A without being isolated in A, in other words, if a is a point of closure of $A \setminus \{a\}$. The set of accumulation points of A, denoted A', is a closed set.

■ **Example 1.10** Let $A =]0,1] \cup \{2\}$. For the usual topology on $\mathbb{R}$, we have

$$\overline{A} = [0,1] \cup \{2\} \qquad \mathring{A} =]0,1[\qquad A' = [0,1] \qquad \partial A = \{0,1\}$$

2 is an isolated point in A, but 0 is not. ■

1.4 Numerical Sequences in $\mathbb{R}^n$

Definition 1.4.1 A extbfsequence of elements in $\mathbb{R}^n$ is any function from $\mathbb{N}$ to $\mathbb{R}^n$ that assigns each k to x_k. Such a function is denoted as $(x_k)_{k\in\mathbb{N}}$. The n coordinates of the element x_k in the canonical basis of $\mathbb{R}^n$ are written as $x_k = (x_{k,1},\dots,x_{k,n})$.

Definition 1.4.2 Let $A \subseteq \mathbb{R}^n$. The sequence $(x_k)_{k\in\mathbb{N}}$ is called a extbfsequence of elements in A if and only if

$$\forall k \in \mathbb{N}, \quad x_k \in A.$$

Definition 1.4.3 — Convergent Sequence. A sequence $(x_k)_{k\in\mathbb{N}}$ of elements in $\mathbb{R}^n$ is said to converge to a limit $l \in \mathbb{R}^n$ if and only if

$$\forall \varepsilon > 0, \exists k_\varepsilon \in \mathbb{N} : k \geq k_\varepsilon \implies \|x_k - l\| < \varepsilon.$$

This is denoted as

$$\lim_{k\to+\infty} x_k = l.$$

R If a sequence converges, its limit is unique.

Proposition 1.4.1 A sequence $(x_k)_{k\in\mathbb{N}}$ of elements in $\mathbb{R}^n$ converges to $l = (l_1,\dots,l_n) \in \mathbb{R}^n$ if and only if for all $1 \leq i \leq n$, the sequence of the i-th coordinate of x_k converges in $\mathbb{R}$ to the i-th coordinate of l, that is,

$$\lim_{k\to+\infty} x_k = l \iff \lim_{k\to+\infty} x_{k,i} = l_i, \quad \forall i \in \{1,\dots,n\}.$$

■ **Example 1.11** Consider in $\mathbb{R}^2$ the sequence $(x_k)_{k\in\mathbb{N}}$ defined by

$$x_k = \left(\frac{k-1}{2k+1}, \frac{1}{k+3}\right).$$

Since

$$\lim_{k\to+\infty} \frac{k-1}{2k+1} = \frac{1}{2} \quad \text{and} \quad \lim_{k\to+\infty} \frac{1}{k+3} = 0,$$

then $(x_k)_{k\in\mathbb{N}}$ is convergent and we have

$$\lim_{k\to+\infty} x_k = \left(\frac{1}{2}, 0\right).$$

■

Proposition 1.4.2 Characterization of Closed Sets via Sequences
A subset $A \subset \mathbb{R}^n$ is closed if and only if, for any convergent sequence $(x_k)_{k\in\mathbb{N}}$ of elements in A, we have

$$\lim_{k\to+\infty} x_k \in A.$$

Definition 1.4.4 — Bounded Sequence. A sequence $(x_k)_{k\in\mathbb{N}}$ of elements in $\mathbb{R}^n$ is bounded if and only if

$$\exists M > 0, \forall k \in \mathbb{N}, \|x_k\| \leq M.$$

Theorem 1.4.3 — Bolzano-Weierstrass. Every bounded sequence of elements in $\mathbb{R}^n$ has a convergent subsequence.

Definition 1.4.5 Cauchy Sequence A sequence $(x_k)_{k\in\mathbb{N}}$ of elements in $\mathbb{R}^n$ is called a Cauchy sequence if it satisfies the following property:

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N}, \forall (p,q) \in \mathbb{N}^2 : p \geq N_\varepsilon, \quad \text{and} \quad q \geq N_\varepsilon \implies \|x_p - x_q\| < \varepsilon.$$

Proposition 1.4.4
Every convergent sequence $(x_k)_{k\in\mathbb{N}}$ is a Cauchy sequence.

Theorem 1.4.5
Every Cauchy sequence in $\mathbb{R}^n$ converges. Consequently, the normed space $(\mathbb{R}^n, \|\cdot\|)$ is complete.

1.5 Compact Sets in $\mathbb{R}^n$

Definition 1.5.1
Let A be a subset of $\mathbb{R}^n$. The set A is called compact (or a compact set) if for every sequence $(x_k)_{k\in\mathbb{N}}$ of elements in A, there exists a subsequence $\left(x_{\varphi(k)}\right)_{k\in\mathbb{N}}$ that converges to an element of A.

By virtue of the Bolzano-Weierstrass theorem, we have the following result:

Proposition 1.5.1
A subset A of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.

Proposition 1.5.2

- Let A be a compact subset of $\mathbb{R}^n$. If F is a closed subset of A, then F is compact.
- The intersection of a closed set and a compact set is compact.
- Any arbitrary intersection of compact sets is compact.
- Any finite union of compact sets is compact.

■ **Example 1.12**

1. The closed unit ball of $\mathbb{R}^n$ is compact.
2. The set $A = [0,1]^3$ is a compact set from $\mathbb{R}^3$.

■

1.6 Exercises

Exercise 1.1
Determine whether the following sets are open or closed:

$$A = \{(x,y) \in \mathbb{R} \mid 0 < |x-1| < 1\} \qquad B = \{(x,y) \in \mathbb{R} \mid 0 \le x \le y\}$$
$$C = \{(x,y) \in \mathbb{R} \mid |x| < 1,\ |y| \le 1\} \qquad D = \{(x,y) \in \mathbb{R} \mid x \in \mathbb{Q} \text{ and } y \in \mathbb{Q}\}$$
$$E = \{(x,y) \in \mathbb{R} \mid x \notin \mathbb{Q} \text{ or } y \notin \mathbb{Q}\} \qquad F = \{(x,y) \in \mathbb{R} \mid x^2 + y^2 < 4\}.$$

■

Exercise 1.2
Draw, then determine the interior and closure of the following subsets of $\mathbb{R}^2$:

$$A = \{(x,y) \in \mathbb{R}^2 \mid x > 0\}$$
$$B = \{(x,y) \in \mathbb{R}^2 \mid xy = 1\}$$
$$C = \{(x,y) \in \mathbb{R}^2 \mid xy > 1\}$$
$$D = \{(x,y) \in \mathbb{R} \mid x^2 + y^2 \le 2\} \setminus \{(x,y) \in \mathbb{R} \mid (x-1)^2 + y^2 < 1\}.$$

■

Exercise 1.3
Let $\lambda > 0$. For any integer $n \ge 1$, denote B_n as the disk

$$B_n = \left\{ (x,y) \in \mathbb{R}^2 \mid \left(x - \frac{1}{n}\right)^2 + \left(y - \frac{1}{n}\right)^2 \le \frac{\lambda^2}{n^2} \right\}.$$

1. Under what condition on λ do we have $B_{n+1} \subset B_n$?
2. Let $B = \bigcup_{n\ge1} B_n$. Provide a necessary and sufficient condition on λ for B to be closed.

■

Exercise 1.4

Draw each of the following subsets of $\mathbb{R}^2$. Sketch their boundaries and interiors. Study whether they are open, closed, bounded.

1. $A = \{(x,y) \in \mathbb{R}^2 \mid 0 < \|k(x,y) - (1,3)\| < 2\}$.
2. $B = \{(x,y) \in \mathbb{R}^2 \mid y \leq x^3\}$.
3. $C = \{(x,y) \in \mathbb{R}^2 \mid |x| < 1, |y| \leq 2\}$.
4. $D = \{(x,y) \in \mathbb{R}^2 \mid |x| + |y| < 1\}$.
5. $E = \{(x,y) \in \mathbb{R}^2 \mid y < x^2, y < \frac{1}{x}, x > 0\}$.
6. $F = \{(x,y) \in \mathbb{R}^2 \mid xy \leq y + 1\}$.
7. $G = \{(x,y) \in \mathbb{R}^2 \mid (x-1)^2 + y^2 \leq 1, x \leq 1\}$.

■

Exercise 1.5

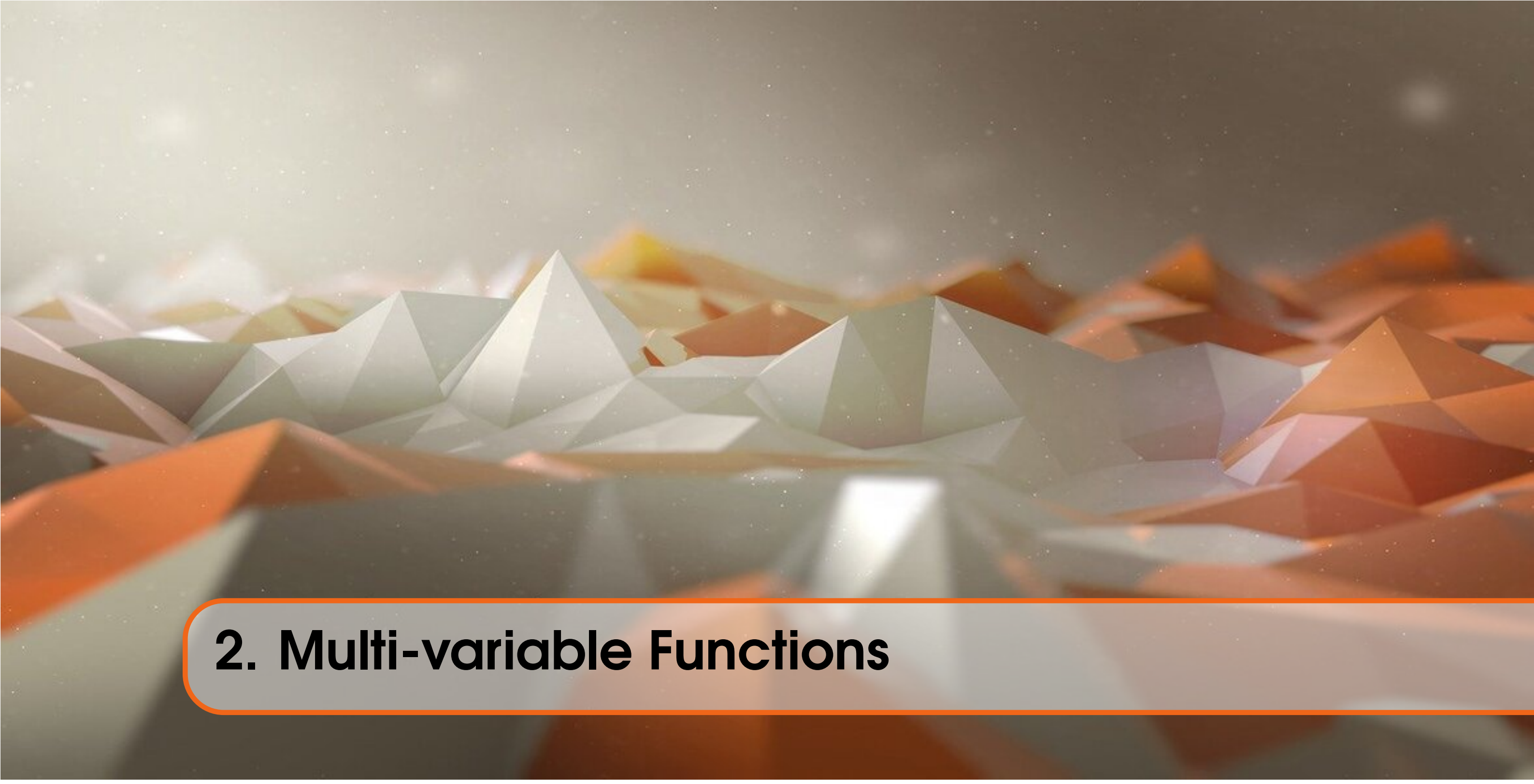
Let A be a subset of $\mathbb{R}^2$. Discuss the truth or falsehood of the following statements.

1. $\mathrm{Int}(A) = A - \mathrm{Fr}(A)$.
2. $\mathrm{Fr}(A) = \mathrm{Fr}(\mathbb{R}^2 - A) = \mathrm{Fr}(A^C)$.
3. $\mathrm{Fr}(A)$ is bounded.
4. A is closed $\Leftrightarrow A^C$ is open.
5. A is bounded $\Leftrightarrow A^C$ is unbounded.
6. A is closed $\Leftrightarrow \mathrm{Fr}(A) \not\subset A$.
7. A is open $\Leftrightarrow \mathrm{Fr}(A) \cap A = \emptyset$.

■

II

Functions of several variables



2. Multi-variable Functions

In many real-world applications, quantities depend on more than one variable. A multivariable function is a natural generalization of a function of a single variable. These functions arise in various fields such as physics, economics, and engineering.

A function of n variables is a rule that assigns a unique real number to each ordered n-tuple in a subset of $\mathbb{R}^n$. Formally, a function f of n real variables is defined as:

$$f : D \subseteq \mathbb{R}^n \to \mathbb{R}, \quad (x_1, x_2, \ldots, x_n) \mapsto f(x_1, x_2, \ldots, x_n).$$

The set D is called the domain of f, and its elements are points in $\mathbb{R}^n$.

We give here some examples.

- **Temperature distribution:** The temperature $T(x,y,z)$ at a point in space depends on three spatial coordinates. This function is used in thermodynamics and meteorology.
- **Economics:** The profit $P(x,y)$ of a company may depend on two factors, such as the quantity of two different products produced.
- **Physics:** The gravitational potential $V(x,y,z)$ at a point in space depends on the spatial coordinates, and is used in celestial mechanics.
- **Engineering:** The stress on a material can be modeled as a function of multiple spatial variables.

These functions form the foundation for concepts such as partial derivatives, gradients, and optimization, which will be explored in this chapter.

2.1 Definitions

Definition 2.1.1

Let E be a subset of $\mathbb{R}^n$. A *function* $f : E \to \mathbb{R}$ associates a single real number $f(x_1, \ldots, x_n)$ to each $(x_1, \ldots, x_n)$ in E.

■ Example 2.1

1. Distance from a point to the origin as a function of its coordinates:

$$\begin{array}{cccc} f & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & & (x,y) & \longmapsto & \sqrt{x^2+y^2}. \end{array}$$

2. Area of a rectangle as a function of its length and width:

$$\begin{array}{rccc} f & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & & (x,y) & \mapsto & xy. \end{array}$$

3. Surface area of a rectangular parallelepiped as a function of its three dimensions:

$$\begin{array}{rccc} f & : & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & & (x,y,z) & \mapsto & 2(xy+yz+xz). \end{array}$$

■

Definition 2.1.2

If an expression for $f(x_1,\ldots,x_n)$ is given first, then the *domain of definition* of f is the largest subset $D_f \subset \mathbb{R}^n$ such that, for each $(x_1,\ldots,x_n)$ in D_f, $f(x_1,\ldots,x_n)$ is well-defined. The function is then $f : D_f \to \mathbb{R}$.

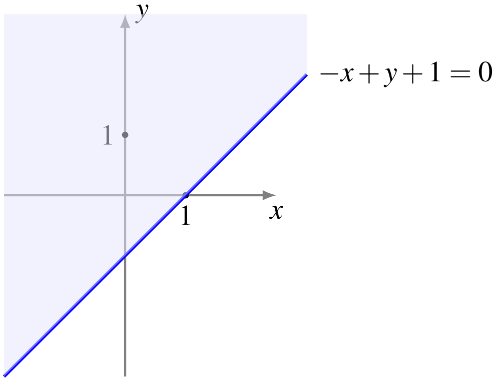
■ **Example 2.2**

1. $f(x,y) = \ln(1-x+y)$

   The quantity $1-x+y$ must be strictly positive in order to compute its logarithm. Thus:

$$D_f = \{(x,y) \in \mathbb{R}^2 \mid 1-x+y > 0\}$$

   To visualize this set, first plot the line with equation $1-x+y=0$. Then determine on which side of the line the inequality $1-x+y>0$ holds. Here, it is above the line.

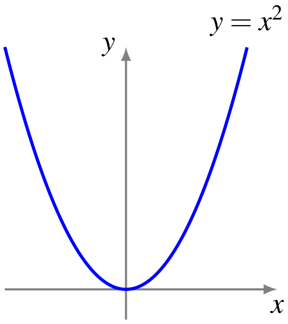


2. $f(x,y) = \dfrac{x+y}{x^2-y}$

   The function f is defined when the denominator in not zero. Then,

$$D_f = \{(x,y) \in \mathbb{R}^2 \mid x^2-y \neq 0\}$$

   The points in the domain are all the points in the plane except those on the parabola with equation $(y=x^2)$.

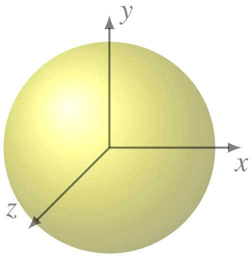


3. $f(x,y,z) = \dfrac{1}{\sqrt{4-x^2-y^2-z^2}}$

   The expression under the square root must be positive and must not be zero. Thus:

$$D_f = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 < 4\}$$

   In other words, these are all the points of the open ball centered at $(0,0,0)$ with radius 2.



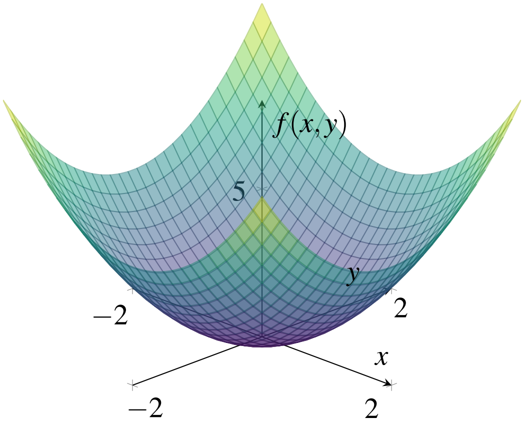
■

Definition 2.1.3
The *range* (image) of a function $f : E \to \mathbb{R}$ is the set of values $f(x_1,\ldots,x_n)$ taken by f as $(x_1,\ldots,x_n)$ varies over E:

$$R(f) = \left\{f(x_1,\ldots,x_n) \mid (x_1,\ldots,x_n) \in E\right\} \subset \mathbb{R}$$

■ Example 2.3

1. $f(x,y) = x^2 + y^2$ is defined on $\mathbb{R}^2$. The image of f is the half positive real line: $R(f) = \mathbb{R}^+$.

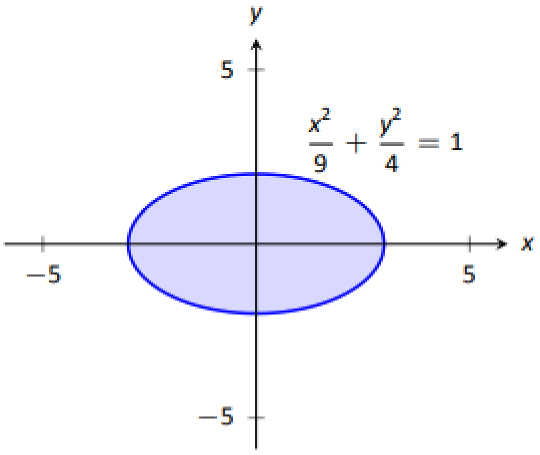


2. $f(x,y) = \ln(1-x+y)$
   The range (image) of f is the entire real line: $\text{Im } f = \mathbb{R}$.
   Indeed, we know that the natural logarithm is defined from $\mathbb{R}^+$ to $\mathbb{R}$.
3.
$$f(x,y) = \sqrt{1-\frac{x^2}{9}-\frac{y^2}{4}},$$

   we are going to find the domain and range of f.
   The domain is all pairs (x,y) allowable as input in f. Because of the square-root, we need (x,y) such that $0 \le 1-\frac{x^2}{9}-\frac{y^2}{4}$:

$$0 \le 1-\frac{x^2}{9}-\frac{y^2}{4} \iff \frac{x^2}{9}+\frac{y^2}{4} \le 1,$$

   which describes the interior of an ellipse. We can represent the domain D graphically with the figure; in set notation, we can write $D = \{(x,y) | \frac{x^2}{9}+\frac{y^2}{4} \le 1\}$.



The range is the set of all possible output values. The square-root ensures that all output is ≥ 0. Since the x and y terms are squared, then subtracted, inside the square-root, the largest output value comes at $x=0$, $y=0$: $f(0,0)=1$. Thus the range R is the interval $[0,1]$.

■

Definition 2.1.4
Let E be a subset of $\mathbb{R}^n$. A *vector-valued function* $f : E \to \mathbb{R}^m$ assigns a m-tuple of real numbers to each $(x_1,\ldots,x_n)$ in E. It is written as:

$$\begin{array}{cccl} f & : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^p \\ & & x & \longmapsto & (f_1(x), f_2(x), \ldots, f_m(x)) . \end{array}$$

■ **Example 2.4**

1. Surface area and volume of a rectangular parallelepiped as a function of its three dimensions:

$$\begin{array}{cccl} f & : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & & (x,y,z) & \longmapsto & \big(2(xy+yz+xz), xyz\big) . \end{array}$$

2. Polar coordinates of a point in the plane:

$$\begin{array}{cccl} f & : & \mathbb{R}_+ \times [0,2\pi[& \longrightarrow & \mathbb{R}^2 \\ & & (r,\theta) & \longmapsto & (r\cos\theta, r\sin\theta). \end{array}$$

■

2.1.1 Graph and Level Curves

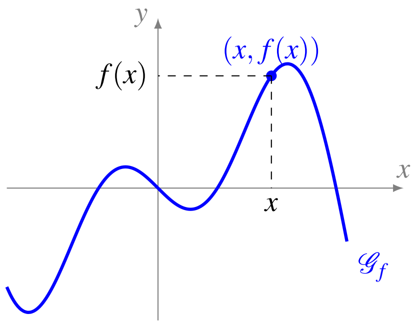
Definition 2.1.5
Let $f : D_f \subset \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables. The *graph* $\mathscr{G}_f$ of f is the subset of $\mathbb{R}^3$ consisting of the points with coordinates $(x,y,f(x,y))$ where (x,y) belongs to the domain. Thus, the graph is:

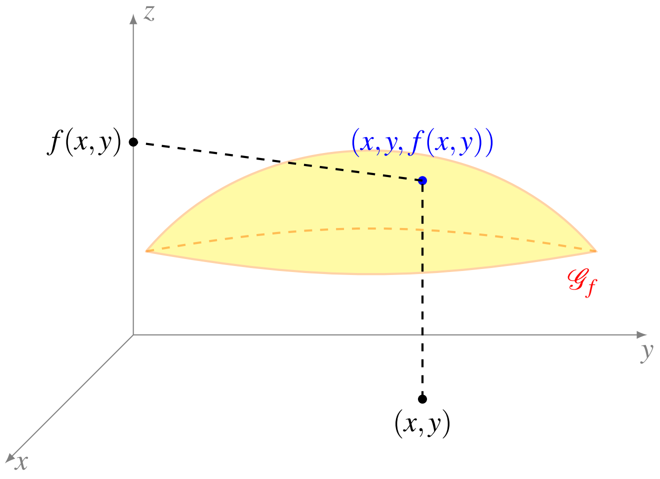
$$\mathscr{G}_f = \big\{(x,y,z) \in \mathbb{R}^3 \mid (x,y) \in D_f \text{ and } z = f(x,y)\big\}.$$

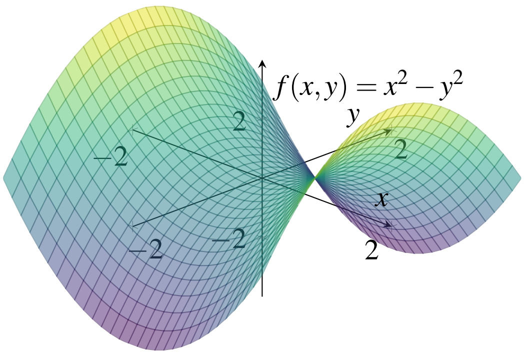
Graphically representing the graph is possible only for functions of one or two variables. For functions of one variable $f : D_f \subset \mathbb{R} \to \mathbb{R}$, the graph is given by:

$$\mathscr{G}_f = \big\{(x,y) \in \mathbb{R}^2 \mid x \in D_f \text{ and } y = f(x)\big\}.$$

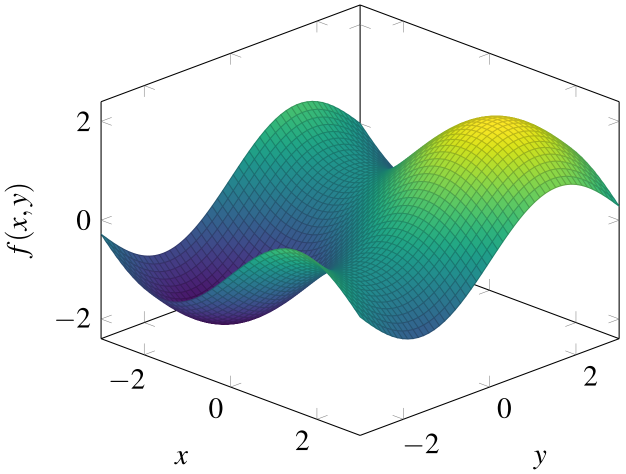
In the one-variable case (left), the graph is a curve; in the two-variable case, which we consider here, it is a surface.







Plot of $f(x,y) = \sin(x) + \sin(y)$



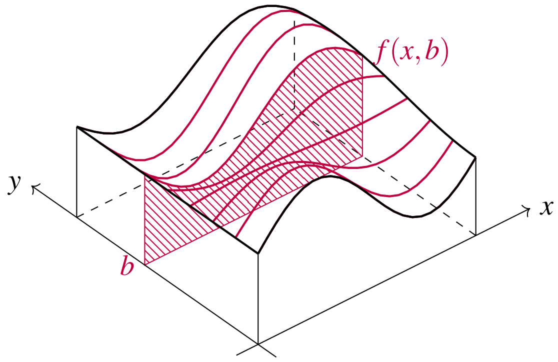
To plot the graph of a function of two variables, we divide the surface into sections.

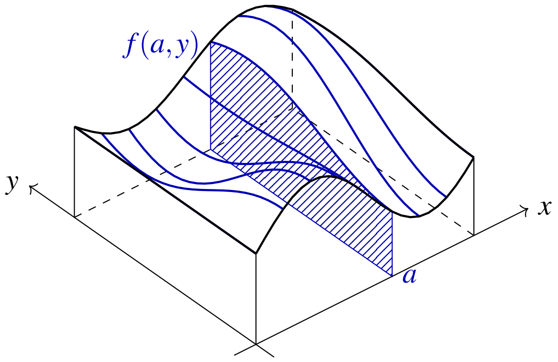
2.1.2 Partial functions (slices)

One way to proceed: for some values of a and b, plot the graphs of the partial functions

$$f_1 : x \mapsto f(x,b) \quad \text{and} \quad f_2 : y \mapsto f(a,y).$$

The first represents the intersection of the graph $\mathscr{G}_f$ with the plane $(y = b)$ (in purple) and the second represents the intersection of the graph with the plane $(x = a)$ (in blue).





2.1.3 Level Curves.

We are also interested in other curves traced on the surface: the level curves.

Definition 2.1.6
Let $f : D_f \subset \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables.

- The *level line* $z = c \in \mathbb{R}$ is
$$L_c = \{(x,y) \in D_f \mid f(x,y) = c\}.$$
- The *level curve* $z = c$ is the trace of $\mathscr{G}_f$ in the plane $(z = c)$:
$$\mathscr{G}_f \cap (z = c) = \{(x,y,c) \in \mathbb{R}^3 \mid f(x,y) = c\}.$$

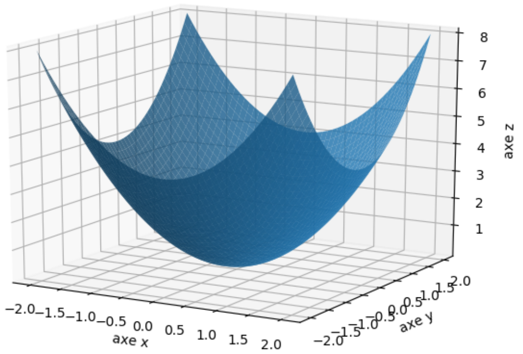
The level line c is a curve in the plane $\mathbb{R}^2$, while the level curve c is a curve in space $\mathbb{R}^3$. The level curve c is obtained by lifting the level line c to altitude c.

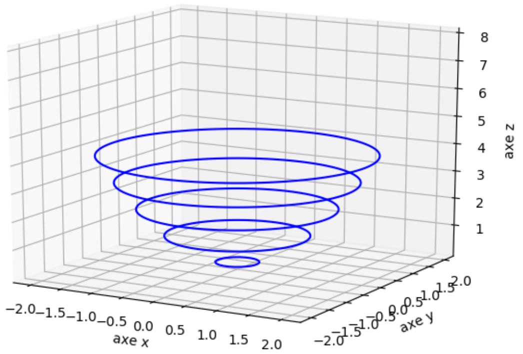
■ Example 2.5
Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x,y) = x^2 + y^2$.

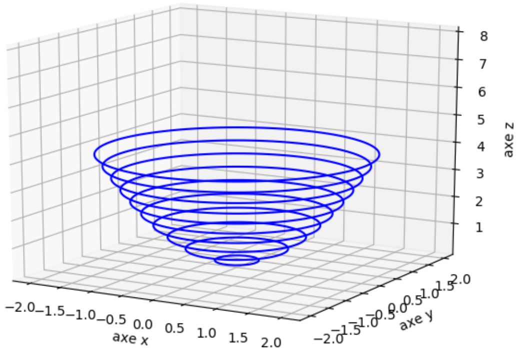
- If $c < 0$, the level line L_c is empty (no point has a negative altitude).
- If $c = 0$, the level line L_0 reduces to $\{(0,0)\}$.
- If $c > 0$, the level line L_c is the circle centered at $(0,0)$ with radius $\sqrt{c}$. Lifting L_c to altitude $z = c$, the level curve is the horizontal circle in space centered at $(0,0,c)$ with radius $\sqrt{c}$.

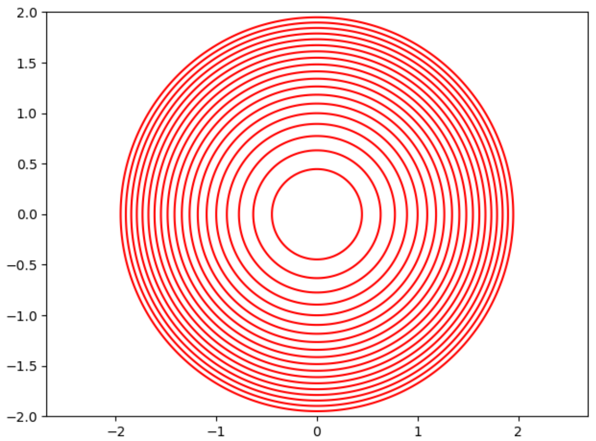
The graph is a superposition of horizontal circles in space centered at $(0,0,c)$ with radius $\sqrt{c}$ for $c \geq 0$.

Below: (a) the graph, called a revolution paraboloid, (b) five level curves, (c) ten level curves, (d) level lines in the plane.





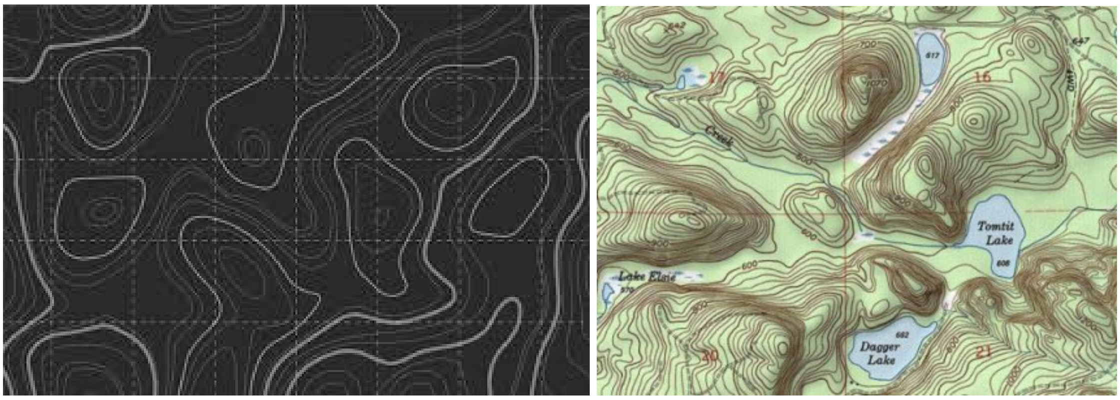




■

■ Example 2.6

On a topographic map, the level lines represent curves of equal altitude.



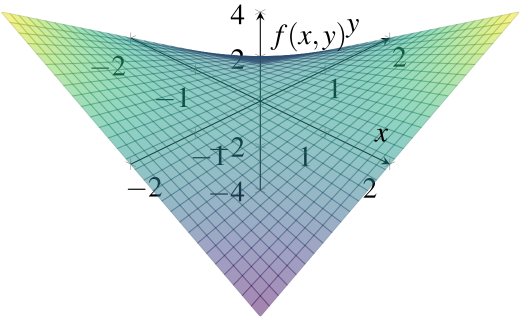
- When the level lines are widely spaced, the terrain is relatively flat; when they are close together, the terrain is steep.
- By definition, if you follow a level line, you remain at the same altitude!

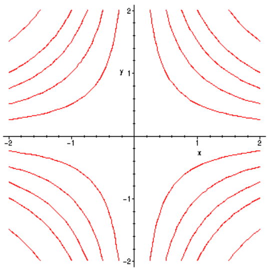
■

■ Example 2.7

In this example we find the graph and level lines of the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$(x,y) \mapsto z = xy$$





■

2.2 Exercises

Exercise 2.1

Find the domain of definition of the given function.

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

3. $f(x,y) = \frac{\ln(y-x)}{x}$ 4. $f(x,y) = \frac{1}{x} + \sqrt{y+4} - \sqrt{x+1}$

5.$f(x,y) = \ln(2x-3y+1)$ 6.$f(x,y,z) = \frac{1}{x^2+y^2+4z}$.

■

Exercise 2.2

Identify and represent the level curves (or contours) for the given function:

$$f_1(x,y) = y^2, \text{ with } k = -1 \text{ and } k = 1 \qquad f_2(x,y) = \frac{x^4+y^4}{8-x^2y^2} \text{ with } k = 2.$$

Represent the level lines of the following functions:

$$\textbf{1}.\ 2x-3y+z^2=1,\quad \textbf{2}.\ 4z+2y^2-x=0,\quad \textbf{3}.\ y^2=2x^2+z.$$

$$\textbf{4}.\ \ f(x,y)=x+y-1\quad \textbf{5}.\ \ f(x,y)=e^{y-x^2}\quad \textbf{6}.\ \ f(x,y)=\sin(xy)$$

■

3. Limits, Continuity

3.1 Limits

The notions of limit and continuity of functions of a single variable generalize to multiple variables without additional complexity: it is sufficient to replace the absolute value with the Euclidean norm.

3.1.1 Definition

Let f be a function $f : E \subset \mathbb{R}^n \to \mathbb{R}$ defined in a neighborhood of $x_0 \in \mathbb{R}^n$, possibly except at x_0.

Definition 3.1.1
The function f has the *limit* ℓ as x approaches x_0 if:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \qquad \forall x \in E \qquad 0 < \|x - x_0\| < \delta \implies |f(x) - \ell| < \varepsilon$$

We then write:

$$\lim_{x_0} f = \ell \qquad \text{or} \qquad \lim_{x \to x_0} f(x) = \ell \qquad \text{or} \qquad f(x) \underset{x \to x_0}{\longrightarrow} \ell.$$

Similarly, we define $\lim_{x \to x_0} f(x) = +\infty$ as:

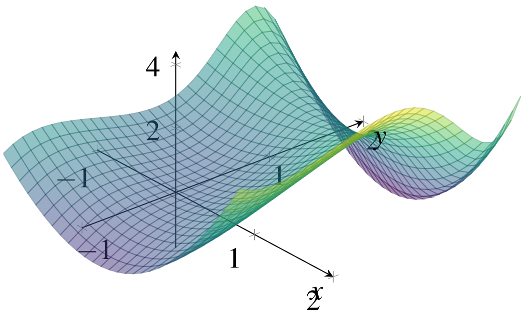
$$\forall A > 0 \quad \exists \delta > 0 \qquad \forall x \in E \qquad 0 < \|x - x_0\| < \delta \implies |f(x)| > A$$

R

- The notion of limit does not depend on the choice of norm.
- If it exists, the limit is unique.

■ Example 3.1
Let f be the function defined by $f(x,y) = x^2 + y\sin(x + y^2)$.

Let's Show that $f(x,y)$ tends to 0 as $(x,y) \to (0,0)$. Then, we are going to find an open set U containing the origin such that for all $(x,y) \in U$, we have $|f(x,y)| < \frac{1}{100}$.



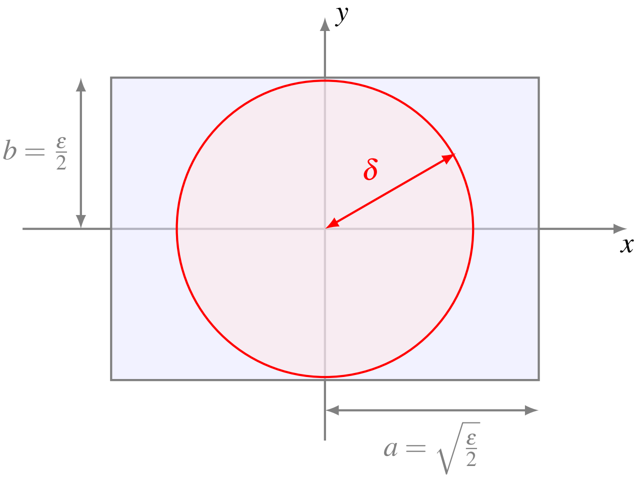
1. We bound $|f(x,y)|$ using the triangle inequality and $|\sin(t)| \leq 1$:

$$\left|f(x,y)\right| = \left|x^2 + y\sin(x+y^2)\right| \leq x^2 + |y|\left|\sin(x+y^2)\right| \leq x^2 + |y|$$

   Let $0 < \varepsilon < 1$. Define $a = \sqrt{\frac{\varepsilon}{2}}$ and $b = \frac{\varepsilon}{2}$. Then, for $x \in]-a,a[$, we have $x^2 < \frac{\varepsilon}{2}$ and for $y \in]-b,b[$, we have $|y| < \frac{\varepsilon}{2}$. For $(x,y) \in]-a,a[\times]-b,b[$, we then get:

$$\left|f(x,y)\right| \leq x^2 + |y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

   A suitable value for δ is therefore $\delta = \frac{\varepsilon}{2}$. Indeed, if $\|(x,y)\| < \delta$ then $|x| < \delta = \frac{\varepsilon}{2} \leq \sqrt{\frac{\varepsilon}{2}}$ and $|y| < \delta = \frac{\varepsilon}{2}$, so $|f(x,y)| < \varepsilon$. Conclusion: f has limit 0 as $(x,y) \to (0,0)$.



2. For $\varepsilon = \frac{1}{100}$, we have $a = \frac{1}{\sqrt{200}}$ and $b = \frac{1}{200}$. For each (x,y) in the open set $]-a,a[\times]-b,b[$, we have $|f(x,y)| < \frac{1}{100}$.

■

3.1.2 Operations on Limits

To compute limits, this definition is rarely used directly. Instead, general theorems on limits and bounding are applied. These are the same statements as for single-variable functions: there is no difficulty or novelty.

Proposition 3.1.1 — Operations on limits.
Let $f,g : \mathbb{R}^n \to \mathbb{R}$ be defined in a neighborhood of $x_0 \in \mathbb{R}^n$ such that f and g have limits at x_0. Then:

$$\lim_{x_0}(f+g) = \lim_{x_0} f + \lim_{x_0} g \qquad \lim_{x_0}(f\cdot g) = \lim_{x_0} f \times \lim_{x_0} g$$

And if g does not vanish in a neighborhood of x_0:

$$\lim_{x_0} \frac{1}{g} = \frac{1}{\lim_{x_0} g} \qquad \lim_{x_0} \frac{f}{g} = \frac{\lim_{x_0} f}{\lim_{x_0} g}$$

R

- The above results also hold for infinite limits using the usual conventions:

$$\ell+\infty=+\infty,\quad \ell-\infty=-\infty,\quad \frac{1}{0^+}=+\infty,\quad \frac{1}{0^-}=-\infty,\quad \frac{1}{\pm\infty}=0,$$

$$\ell\times\infty=\infty\ (\ell\neq 0),\ \infty\times\infty=\infty \text{ (with sign multiplication rules).}$$

- Indeterminate forms include: $+\infty-\infty$, $\frac{0}{0}$, $\frac{\infty}{\infty}$, $0\times\infty$, ∞^0, 1^∞ and 0^0.

Composition is also often useful:

- Let $f:\mathbb{R}^n\to\mathbb{R}$ be a function of multiple variables such that $\lim_{x\to x_0} f(x)=\ell$.
- Let $g:\mathbb{R}\to\mathbb{R}$ be a function of a single variable such that $\lim_{t\to\ell} g(t)=\ell'$.
- Then, the composite function $g\circ f:\mathbb{R}^n\to\mathbb{R}$ defined by $(g\circ f)(x)=g\big(f(x)\big)$ satisfies: $\lim_{x\to x_0}(g\circ f)(x)=\ell'$.

Using Example 3.1, and since $e^t\to 1$ as $t\to 0$, we deduce:

$$e^{x^2+y\sin(x+y^2)} \xrightarrow[(x,y)\to(0,0)]{} 1$$

There is also a *squeeze theorem*.

Theorem 3.1.2 — Bounding Theorem.
Let $f,g,h:\mathbb{R}^n\to\mathbb{R}$ be three functions defined in a neighborhood U of $x_0\in\mathbb{R}^n$.

- If, for all $x\in U$, we have $f(x)\leq h(x)\leq g(x)$,
- and if $\lim_{x_0} f=\lim_{x_0} g=\ell$,

then h has a limit at the point x_0 and $\lim\limits_{x_0} h=\ell$.

■ Example 3.2
Let h be defined by $h(x,y)=\cos(x+y^2)\big(x^2+y\sin(x+y^2)\big)$. We bound the absolute value of the cosine by 1:

$$\big|h(x,y)\big|\leq\big|x^2+y\sin(x+y^2)\big|.$$

We saw in Example 3.1 that the function $f(x,y)=x^2+y\sin(x+y^2)$ tends to 0 at $(0,0)$. Thus, by the squeeze theorem, $h(x,y)$ also tends to 0 as (x,y) approaches $(0,0)$. ■

3.1.3 Limit Along a Path

The uniqueness of the limit implies that, regardless of how we approach the point x_0, the limit value is always the same.

Proposition 3.1.3
Let $f:\mathbb{R}^n\to\mathbb{R}$ be a function defined in a neighborhood of $x_0\in\mathbb{R}^n$, possibly except at x_0.

1. If f has a limit ℓ at the point x_0, then the restriction of f along any curve passing through x_0 has a limit at x_0, and this limit is ℓ.
2. By contrapositive, if the restrictions of f along two different curves passing through x_0 have different limits at x_0, then f does not have a limit at x_0.

Let's detail this in the case of functions of two variables:

- A curve passing through the point $(x_0,y_0)\in\mathbb{R}^2$ is a continuous function $\gamma:\mathbb{R}\to\mathbb{R}^2$, $t\mapsto(x(t),y(t))$, such that $\gamma(0)=(x_0,y_0)$.
- The restriction of f along γ is the single-variable function $f\circ\gamma: t\mapsto f\big(x(t),y(t)\big)$.
- If f has a limit ℓ at (x_0,y_0), then the first part of the proposition states that $f\big(x(t),y(t)\big)\xrightarrow[t\to 0]{}\ell$.

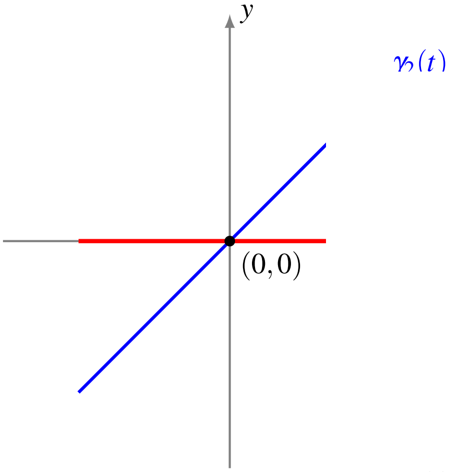
■ **Example 3.3**

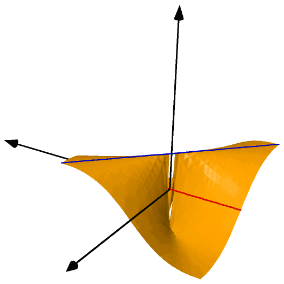
Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

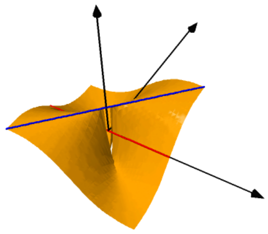
$$f(x,y) = \frac{xy}{x^2+y^2} \text{ if } (x,y) \neq (0,0) \quad \text{and} \quad f(0,0) = 0.$$

Does the function f have a limit at $(0,0)$?

- If we take the path $\gamma_1(t) = (t,0)$, then $(f \circ \gamma_1)(t) = f(t,0) = 0$. So, as $t \to 0$, $(f \circ \gamma_1)(t) \to 0$.
- If we take the path $\gamma_2(t) = (t,t)$, then $(f \circ \gamma_2)(t) = f(t,t) = \frac{t^2}{2t^2} = \frac{1}{2}$. So, as $t \to 0$, $(f \circ \gamma_2)(t) \to \frac{1}{2}$. Below, on the left figure, we see the two paths in the plane; on the two right figures, two different views of the values taken by f along these paths.







- If f had a limit ℓ, then for any path $\gamma(t)$ such that $\gamma(t) \to (0,0)$ as $t \to 0$, we would have $(f \circ \gamma)(t) \to \ell$. This would imply $\ell = 0$ and $\ell = \frac{1}{2}$, which contradicts the uniqueness of the limit. Thus, f does not have a limit at $(0,0)$.

■

Another possible formulation:

If $f : \mathbb{R}^n \to \mathbb{R}$ has a limit ℓ at $x_0 \in \mathbb{R}^n$, then for any sequence (u_n) of elements in $\mathbb{R}^n$ such that $u_n \to x_0$, we have $f(u_n) \to \ell$. For functions of two variables, this can be stated as: if f has a limit ℓ at (a,b), then for any sequence $(a_n, b_n) \to (a,b)$, we have $f(a_n, b_n) \to \ell$.

3.2 Continuity

Definition 3.2.1

1. $f : E \subset \mathbb{R}^n \to \mathbb{R}$ is continuous at $x_0 \in E$ if $\lim\limits_{x \to x_0} f(x) = f(x_0)$.
2. f is continuous on E if it is continuous at every point of E.

By the properties of limits, if f and g are two functions continuous at x_0, then:

- the function $f+g$ is continuous at x_0,
- similarly, fg and f/g (provided that $g(x) \neq 0$ in a neighborhood of x_0) are continuous at x_0,
- if $h : \mathbb{R} \to \mathbb{R}$ is continuous, then $h \circ f$ is continuous at x_0.

■ **Example 3.4**

- The functions defined by $(x,y) \mapsto x+y$, $(x,y) \mapsto xy$, and more generally, all polynomial functions in two variables x and y are continuous on $\mathbb{R}^2$ (for example, $(x,y) \mapsto x^2 + 3xy$). Similarly, all rational functions in two variables are continuous wherever they are defined.

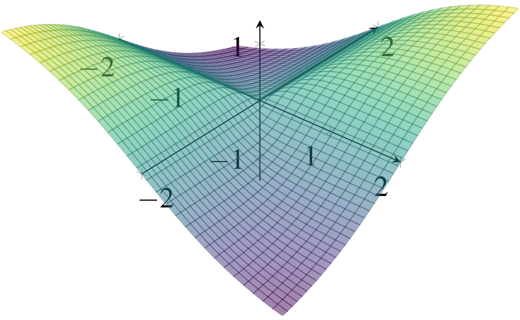
- Since the exponential function is continuous, the function $(x,y) \mapsto e^{xy}$ is continuous on $\mathbb{R}^2$.
- The function defined by $f(x,y) = \frac{1}{\sqrt{x^2+y^2}}$ is continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$.

Definition 3.2.2 — Extension by Continuity.
Let $f : E \subset \mathbb{R}^n \to \mathbb{R}$. Let x_0 be an accumulation point of E that does not belong to E. If $f(x)$ has a limit ℓ as $x \to x_0$, we can extend the domain of definition of f to $E \cup \{x_0\}$ by setting $f(x_0) = \ell$. The extended function is continuous at x_0. We say that we have obtained an **extension of f by continuity** at the point x_0.

Example 3.5
Let $f : \mathbb{R}^2 \setminus \{(0,0)\}$ be defined by:

$$f(x,y) = \frac{xy}{\sqrt{x^2+y^2}}.$$

In the figure below, the question is simply: is it possible to fill the hole in the middle of the surface by adding just one point? in that case we say that we extend f by continuity at $(0,0)$.



- **Limit at the origin.**
  We use the inequalities $|x| \le \sqrt{x^2+y^2}$ and $|y| \le \sqrt{x^2+y^2}$. Therefore:

$$|f(x,y)| = \frac{|x| \cdot |y|}{\sqrt{x^2+y^2}} \le \sqrt{x^2+y^2} \xrightarrow[(x,y)\to(0,0)]{} 0.$$

- **Extension.**
  To extend f at $(0,0)$, we choose as its value the obtained limit. We set $f(0,0) = 0$. (We still denote the extended function as $f : \mathbb{R}^2 \to \mathbb{R}$.)
- **Continuity.**
  By our choice of $f(0,0)$, f is continuous at $(0,0)$. Outside the origin, f is continuous as a sum, product, composition, and inverse of continuous functions. Conclusion: the extended function is continuous on all of $\mathbb{R}^2$.

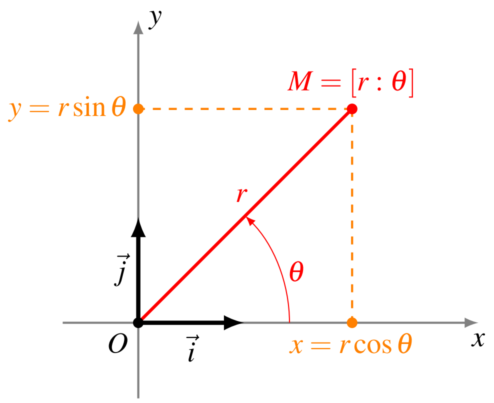
3.3 Polar Coordinates

Rather than locating a point in the plane $\mathbb{R}^2$ using its Cartesian coordinates (x,y), we can do so using its distance from the origin and the angle it forms with the horizontal axis: these are the polar coordinates.

3.3.1 Definition

Let M be a point in the plane $\mathbb{R}^2$. Let $O = (0,0)$ be the origin. Let $(O, \vec{i}, \vec{j})$ be a direct orthonormal frame.

- We denote by $r = \|\overrightarrow{OM}\|$ the distance from M to the origin.
- We denote by θ the angle between $\vec{i}$ and $\overrightarrow{OM}$.



The *polar coordinates* of the point M are written as $[r : \theta]$. In this course, r will always be positive. The angle is not uniquely determined; multiple choices are possible. To ensure uniqueness, we can restrict θ to the interval $[0, 2\pi[$, or $]-\pi, +\pi]$. Typically, we do not assign polar coordinates to the origin (as the angle would be meaningless).

From Polar to Cartesian Coordinates

The Cartesian coordinates (x, y) can be obtained from the polar coordinates $[r : \theta]$ using the formulas:

$$x = r\cos\theta \qquad \text{and} \qquad y = r\sin\theta.$$

In other words, we have defined a mapping:

$$]0, +\infty[\times[0, 2\pi[\to \mathbb{R}^2 \qquad (r, \theta) \mapsto (r\cos\theta, r\sin\theta).$$

From Cartesian to Polar Coordinates

The values of r and θ can be determined from (x, y) using the following formulas:

$$r = \sqrt{x^2 + y^2}$$

and, in the case where $x > 0$ and $y \geq 0$,

$$\theta = \arctan\left(\frac{y}{x}\right).$$

For points in other quadrants, we reduce the problem to the principal quadrant where $x > 0$ and $y \geq 0$.

3.3.2 Limit and Continuity

When considering functions $f : E \subset \mathbb{R}^2 \to \mathbb{R}$, it is sometimes easier to prove results about limits, continuity, etc., using polar coordinates.

Proposition 3.3.1

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function defined in a neighborhood of $(0,0) \in \mathbb{R}^2$, possibly except at $(0,0)$. If

$$\lim_{r \to 0} f(r\cos\theta, r\sin\theta) = \ell \in \mathbb{R}$$

exists independently of θ, meaning that there exists a function $\varepsilon(r) \xrightarrow[r \to 0]{} 0$ such that for all $r > 0$ and all θ, we have:

$$\left| f(r\cos\theta, r\sin\theta) - \ell \right| \leq \varepsilon(r),$$

then $\displaystyle\lim_{(x,y) \to (0,0)} f(x, y) = \ell$.

To clarify this proposition and explain the different practical cases of limits, here is the method. We express $f(x, y)$ in polar coordinates by computing $f(r\cos\theta, r\sin\theta)$.

1. If $\lim_{r\to 0} f(r\cos\theta, r\sin\theta)$ exists and does not depend on the variable θ, then this limit is the limit of f at the point $(0,0)$.
2. If $\lim_{r\to 0} f(r\cos\theta, r\sin\theta)$ does not exist (or the limit is not finite), then f does not have a finite limit at $(0,0)$.
3. If $\lim_{r\to 0} f(r\cos\theta, r\sin\theta) = \ell(\theta)$ depends on θ, then f does not have a limit at $(0,0)$. To justify this, we find two values θ_1 and θ_2 such that $\ell(\theta_1) \neq \ell(\theta_2)$.

Let's look at an example of each situation.

■ Example 3.6

1. $f(x,y) = \dfrac{x^2y}{x^2+y^2}$

$$f(r\cos\theta, r\sin\theta) = \frac{r^3\cos^2\theta\sin\theta}{r^2} = r\cos^2\theta\sin\theta$$

Since $\left|\cos^2\theta\sin\theta\right| \leq 1$, then $r|\cos^2\theta\sin\theta| \leq r$. This implies that $f(r\cos\theta, r\sin\theta) \underset{r\to 0}{\longrightarrow} 0$. The limit exists (independent of the values taken by θ), so the function f has a limit at $(0,0)$: $f(x,y) \xrightarrow[(x,y)\to(0,0)]{} 0$.

2. $f(x,y) = \dfrac{x}{x^2+y^3}$

$$f(r\cos\theta, r\sin\theta) = \frac{r\cos\theta}{r^2(\cos^2\theta + r\sin^3\theta} = \frac{1}{r}\frac{\cos\theta}{\cos^2\theta + r\sin^3\theta}$$

For a fixed θ such that $\sin\theta \neq 0$ (i.e., $\theta \neq 0 \pmod{\pi}$), as $r \to 0$, $f(r\cos\theta, r\sin\theta)$ does not have a finite limit. In particular, the function $(x,y) \mapsto f(x,y)$ does not have a finite limit at $(0,0)$.

3. $f(x,y) = \dfrac{x^2}{x^2+y^2}$

$$f(r\cos\theta, r\sin\theta) = \frac{r^2\cos^2\theta}{r^2} = \cos^2\theta$$

For fixed θ, the function $r \mapsto f(r\cos\theta, r\sin\theta)$ has a limit $\ell(\theta) = \cos^2\theta$ as $r \to 0$. However, this limit depends on the angle θ: if $\theta = 0$, then $\ell(\theta) = 1$; whereas if $\theta = \frac{\pi}{3}$, then $\ell(\theta) = \frac{1}{4}$. Since the limit depends on the angle, the function $(x,y) \mapsto f(x,y)$ does not have a limit at $(0,0)$.

■

R

Let $\ell \in \mathbb{R}$. Consider a function $f : \mathbb{R}^2 \to \mathbb{R}$ with the property that for every fixed angle θ, the limit $\lim\limits_{r\to 0} f(r\cos\theta, r\sin\theta) = \ell$ holds. Does this imply that f has ℓ as its limit at the origin $(0,0)$? The answer is **no**.

In other words, evaluating the limit of f along radial paths is insufficient to determine the limit of f at the origin.

3.4 Exercises

Exercise 3.1

Study the existence and possibly the value of the limit in $(0,0)$ for functions defined (on the largest domain of $\mathbb{R}^2$ possible) by

$$\begin{array}{lll} f_1(x,y) = \frac{x^2y^2}{x^2+y^2}, & f_2(x,y) = \frac{xy}{x^2+y^2}, & f_3(x,y) = \frac{xy}{x+y}, \\ f_4(x,y) = \frac{x^2-y^2}{x^2+y^2}, & f_5(x,y) = (x+y)\sin\left(\frac{1}{x^2+y^2}\right), & f_6(x,y) = \frac{x+y}{x^2+y^2}, \\ f_7(x,y) = \frac{1+x^2+y^2}{y}\sin(y), & f_8(x,y) = \frac{x^3+y^3}{x^2+y^2}, & f_9(x,y) = \frac{3x^2+xy}{\sqrt{x^2+y^2}}. \end{array}$$

■

Exercise 3.2

Do the following limits exist?

$$\lim_{(x,y)\to(1,1)} \frac{1}{x-y}, \quad \lim_{(x,y)\to(1,0)} \frac{y^3}{(x-1)^2+y^2}?$$

■

Exercise 3.3

Give the domain of definition of the following functions, then determine if they can be extended by continuity on $\mathbb{R}^2$:

$$f_1(x,y) = \frac{x^4+y^4}{x^2+y^2}, f_2(x,y) = \frac{y\sin(x+1)}{x^2-2x+1}, f_3(x,y) = \frac{xy-2y}{x^2+y^2-4x+4}.$$

■

Exercise 3.4

Study the continuity of the following functions:

$$f(x,y) = \begin{cases} \frac{x^2-y^2}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & \text{otherwise} \end{cases}$$

$$g(x,y) = \begin{cases} \frac{y^3}{(x-1)^2+y^2}, & (x,y) \neq (1,0) \\ 0, & \text{otherwise} \end{cases}$$

$$h(x,y) = \begin{cases} \frac{x\ln(1+x^2)}{y(x^2+y^2)}, & (x,y) \neq (0,0) \\ 0, & \text{otherwise} \end{cases}$$

$$k(x,y) = \begin{cases} \frac{6x^2y}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & \text{otherwise} \end{cases}$$

■

Exercise 3.5

Let $f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2}, & (x,y) \neq (0,0) \\ 0, & \text{otherwise} \end{cases}.$

Show that the restriction of f to any line passing through $(0,0)$ is continuous, but f is not continuous at the point $(0,0)$. ■

Exercise 3.6

Show that the function $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ defined by

$$f(x,y) = \frac{\sin(x^2) - \sin(y^2)}{x^2+y^2}$$

is not extendable by continuity in $(0,0)$. ■

Exercise 3.7

Let $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x,y) = \begin{cases} \frac{1}{2}(x^2+y^2) - 1, & \text{si } x^2+y^2 > 1 \\ -\frac{1}{2}, & \text{otherwise} \end{cases}.$$

Show that f is continuous. ■

Exercise 3.8
Extend the function f by continuity $f(x,y) = xy\ln(x^2+y^2)$ at the point $(0,0)$. ■

Exercise 3.9
Say whether

$$f(x,y) = xy - \frac{2y}{x^2+y^2-4x+4}$$

is extendable by continuity at the point $(2,0)$. ■

4. Differentiable Functions

The notion of differentiability is the generalization, to functions of several variables, of the concept of differentiability for functions of a single real variable. We begin by showing how to define and compute a partial derivative from the notion of the derivative of a function of a single variable.

4.1 First-order Partial Derivatives

Let f be a function defined on a neighborhood Ω of the point $a=(a_1,\dots,a_n)\in\mathbb{R}^n$, with values in $\mathbb{R}^p$, where $(n,p)\in(\mathbb{N}^*)^2$. Let $f_{i,a}: t\mapsto f(a_1,\dots,t,\dots,a_n)$ be the i^{th} partial function of f at a. Suppose that $f_{i,a}$ is differentiable at a_i. Then, the derivative of $f_{i,a}$ at a_i is:

$$f'_{i,a}(a_i)=\lim_{h\to 0}\frac{f(a_1,\dots,a_i+h,\dots,a_n)-f(a_1,\dots,a_n)}{h},$$

which is called the i^{th} partial derivative of f at a, and is denoted by $\frac{\partial f}{\partial x_i}(a)$.

Definition 4.1.1
Let $i\in\{1,\dots,n\}$. We say that the function $f:\Omega\subseteq\mathbb{R}^n\to\mathbb{R}^p$ has a partial derivative with respect to its i^{th} variable at the point a if the i^{th} partial function of f at a is differentiable at the point a_i. In this case, $f'_{i,a}(a_i)$ is called the partial derivative of f with respect to its i^{th} variable at a and is denoted:

$$\frac{\partial f}{\partial x_i}(a)=f'_{i,a}(a_i)=\lim_{h\to 0}\frac{f(a_1,\dots,a_i+h,\dots,a_n)-f(a_1,\dots,a_n)}{h}.$$

■ **Example 4.1**
Compute the partial derivatives of the function $f(x,y)=\ln(x+y^2)$ at the point $a=(1,0)$.

- $\frac{\partial f}{\partial x}(1,0)=\lim_{h\to 0}\frac{f(1+h,0)-f(1,0)}{h}=\lim_{h\to 0}\frac{\ln(1+h)-\ln(1)}{h}=\lim_{h\to 0}\frac{\ln(1+h)}{h}=1.$
- $\frac{\partial f}{\partial y}(1,0)=\lim_{k\to 0}\frac{f(1,k)-f(1,0)}{k}=\lim_{k\to 0}\frac{\ln(1+k^2)-\ln(1)}{k}=\lim_{k\to 0}\frac{\ln(1+k^2)}{k}=0.$

■

R The partial derivative of the function

$$\begin{aligned}&f:\Omega\subseteq\mathbb{R}^n\to\mathbb{R}^p\\&(x_1,\dots,x_n)\mapsto f(x_1,\dots,x_n)=(f_1(x_1,\dots,x_n),\dots,f_p(x_1,\dots,x_n)),\end{aligned}$$

with respect to the i^{th} variable at the point a is given by

$$\frac{\partial f}{\partial x_i}(a)=\begin{pmatrix}\frac{\partial f_1}{\partial x_i}(a)\\ \vdots\\ \frac{\partial f_p}{\partial x_i}(a)\end{pmatrix}.$$

R In practice, to determine a partial derivative of f, it suffices to differentiate the expression of f with respect to the considered variable, treating the other variables as constants.

■ Example 4.2

Consider the function $f(x,y,z)=xy^2+y^3+xz$. The partial derivatives at a point $a=(x,y,z)$ are:

$$\frac{\partial f}{\partial x}(x,y,z)=y^2+z,\quad \frac{\partial f}{\partial y}(x,y,z)=2xy+3y^2,\quad \frac{\partial f}{\partial z}(x,y,z)=x.$$

■

R A function can have partial derivatives at a point without being continuous at that point.

■ Example 4.3

The function defined on $\mathbb{R}^2$ by

$$f(x,y)=\begin{cases}\frac{xy}{x^2+y^2} & \text{if } (x,y)\neq(0,0),\\ 0 & \text{if } (x,y)=(0,0).\end{cases}$$

has partial derivatives at $(0,0)$:

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

But f is not continuous at $(0,0)$. ■

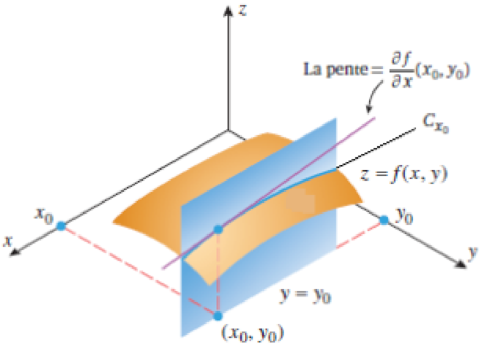
Proposition 4.1.1

The rules of differentiation are the same as for functions of one variable. Let Ω be an open subset of $\mathbb{R}^n$, $a\in\Omega$, and $i\in\{1,\dots,n\}$.

- If $f,g:\Omega\to\mathbb{R}^p$ have i^{th} partial derivatives at a, then for any $(\alpha,\beta)\in\mathbb{R}^2$:
$$\frac{\partial(\alpha f+\beta g)}{\partial x_i}(a)=\alpha\frac{\partial f}{\partial x_i}(a)+\beta\frac{\partial g}{\partial x_i}(a).$$
- Product rule:
$$\frac{\partial(f\cdot g)}{\partial x_i}(a)=\frac{\partial f}{\partial x_i}(a)g(a)+f(a)\frac{\partial g}{\partial x_i}(a).$$
- Quotient rule (if $g(a)\neq 0$):
$$\frac{\partial\left(\frac{f}{g}\right)}{\partial x_i}(a)=\frac{\frac{\partial f}{\partial x_i}(a)g(a)-f(a)\frac{\partial g}{\partial x_i}(a)}{(g(a))^2}.$$

4.1.1 Geometric Representation

Let S be the surface with the equation $z = f(x,y)$ and the point $M_0(x_0, y_0, z_0)$. The intersection of the surface S with the plane $y = y_0$ is a curve C_{x_0}. In this plane, C_{x_0} is the graph of $z = f_{1,x_0}(x) = f(x, y_0)$, which is the first partial function of f at $a = (x_0, y_0)$.
We know that $f'_{1,x_0}(x_0) = \frac{\partial f}{\partial x}(x_0, y_0)$ is the slope of the tangent to the curve C_{x_0} at M_0, as shown in the adjacent figure.



4.1.2 Partial Derivative Functions

Definition 4.1.2
Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^p$ be a function defined on a non-empty open set Ω and let $i \in \{1, \dots, n\}$. f is said to have a partial derivative with respect to x_i on Ω if and only if f has a partial derivative with respect to x_i at every point $a \in \Omega$. In this case, the partial derivative function of f with respect to x_i on Ω is defined as follows:

$$\frac{\partial f}{\partial x_i} : \Omega \subset \mathbb{R}^n \to \mathbb{R}^p$$

$$(x_1, \dots, x_n) \mapsto \frac{\partial f}{\partial x_i}(x_1, \dots, x_n).$$

■ Example 4.4
Consider the function defined on $\mathbb{R}^2$ by

$$f(x,y) = \ln(x^2 + y^2 + 1).$$

The partial derivatives of f exist at every point $(x,y) \in \mathbb{R}^2$, and we have

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

■

■ Example 4.5
Consider the function defined on $\mathbb{R}^2$ by

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

Partial derivatives at $(0,0)$: We saw in Example 3.4 that f has partial derivatives at the point $(0,0)$:

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

Partial derivatives on $\mathbb{R}^2 \setminus \{(0,0)\}$: f has a partial derivative with respect to x on $\mathbb{R}^2 \setminus \{(0,0)\}$ as a quotient of functions that have partial derivatives with respect to x, and the denominator is nonzero on this set. For all $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$:

$$\frac{\partial f}{\partial x}(x,y) = \frac{y^3 - x^2 y}{(x^2 + y^2)^2}, \quad \frac{\partial f}{\partial y}(x,y) = \frac{x^3 - x y^2}{(x^2 + y^2)^2}.$$

Summary: f has partial derivatives on $\mathbb{R}^2$ with respect to each of its two variables, and we have:

$$\frac{\partial f}{\partial x}(x,y)=\begin{cases}\frac{y^3-x^2y}{(x^2+y^2)^2} & \text{if } (x,y)\neq(0,0)\\ 0 & \text{if } (x,y)=(0,0)\end{cases}$$

$$\frac{\partial f}{\partial y}(x,y)=\begin{cases}\frac{x^3-xy^2}{(x^2+y^2)^2} & \text{if } (x,y)\neq(0,0)\\ 0 & \text{if } (x,y)=(0,0)\end{cases}$$

■

4.1.3 Directional Derivative

Now, we give a generalization of the notion of a partial derivative.

Definition 4.1.3
Let $f:\mathbb{R}^n\to\mathbb{R}$. Let $v\in\mathbb{R}^n$ be a nonzero vector. The *directional derivative* of f at $x_0\in\mathbb{R}^n$ in the direction of the vector v is defined, if it exists, as

$$D_vf(x_0)=\lim_{t\to0}\frac{f(x_0+tv)-f(x_0)}{t}.$$

For a function $f:\mathbb{R}^2\to\mathbb{R}$, the directional derivative at the point (x_0,y_0) in the direction of the vector $v=(h,k)$ is given by

$$D_vf(x_0,y_0)=\lim_{t\to0}\frac{f(x_0+th,y_0+tk)-f(x_0,y_0)}{t}.$$

■ **Example 4.6**
Let f be the function defined on $\mathbb{R}^2$ by

$$f(x,y)=\frac{x^3+y^3}{x^2+y^2}\quad\text{if } (x,y)\neq(0,0)\quad\text{and}\quad f(0,0)=0.$$

Study the existence of the directional derivative of f in the direction of a nonzero vector at the point $(0,0)$.

Solution.

For any nonzero vector $v=(h,k)$, we have:

$$\lim_{t\to0}\frac{f(0+th,0+tk)-f(0,0)}{t}=\lim_{t\to0}\frac{\frac{(th)^3+(tk)^3}{(th)^2+(tk)^2}-0}{t}=\frac{h^3+k^3}{h^2+k^2}.$$

Thus, f has a directional derivative in every direction at the point $(0,0)$, and when $v=(h,k)$, we get:

$$D_vf(0,0)=\frac{h^3+k^3}{h^2+k^2}.$$

■

In general, if the vector v is a vector of the standard basis, we recover a partial derivative. Let $f:\mathbb{R}^2\to\mathbb{R}$.

1. If $v=(1,0)$, we get $D_vf(x,y)=\frac{\partial f}{\partial x}(x,y)$.
2. If $v=(0,1)$, we get $D_vf(x,y)=\frac{\partial f}{\partial y}(x,y)$.

When f is differentiable (as discussed later in this chapter), we will have a simple and direct formula to compute $D_vf(x,y)$ using the partial derivatives. If f is differentiable and $v=(h,k)$, then:

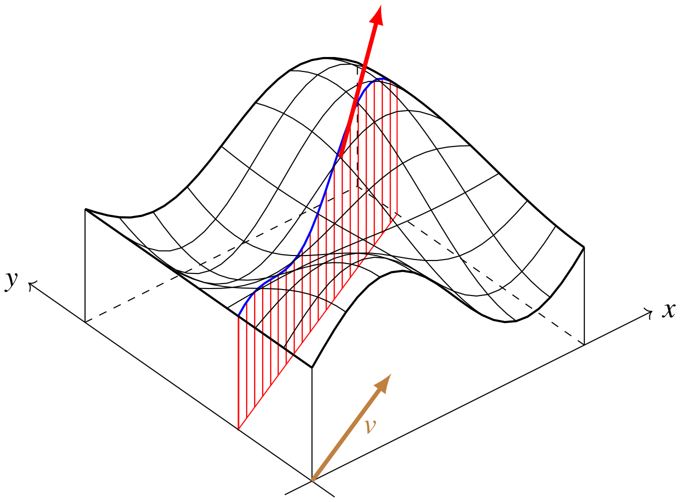
$$D_vf(x,y)=h\frac{\partial f}{\partial x}(x,y)+k\frac{\partial f}{\partial y}(x,y).$$

Geometric Interpretation.

For a function of one variable, the derivative at a point is the slope of the tangent line to the graph at that point (the graph is a curve). For a function of two variables $(x,y) \mapsto f(x,y)$, the partial derivatives indicate the slopes of the graph of f in certain directions (the graph is a surface). More precisely:

- $\frac{\partial f}{\partial x}(x_0,y_0)$ is the slope of the graph of f at (x_0,y_0) in the direction of the (Ox)-axis. Indeed, this slope is the slope of the tangent to the curve $z = f(x,y_0)$ and is given by the derivative of $x \mapsto f(x,y_0)$ at x_0, which is exactly $\frac{\partial f}{\partial x}(x_0,y_0)$.
- $\frac{\partial f}{\partial y}(x_0,y_0)$ is the slope of the graph of f at (x_0,y_0) in the direction of the (Oy)-axis.
- More generally, if v is a unit vector (i.e., with norm 1), then $D_v f(x_0,y_0)$ is the slope of the tangent line in the direction of v.

Below, the directional derivative $D_v f$ indicates the slope at a point in a slice in the direction of a vector v.



Exercise 4.1
Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x,y) = x^2y + x - y$. Find the directional derivative of f at $(0,0)$ along any non zero vector $v = (h,k)$. ■

4.2 Differential

The differential of a function generalizes the derivative to higher dimensions, capturing how a function changes in response to small variations in all its input variables. It combines all the partial derivatives into a single linear map, providing the best linear approximation of the function near a point.

4.2.1 Differentiability

For a function $f : \mathbb{R} \to \mathbb{R}$ of a single variable, another way to express that it is differentiable at x_0 is to check that there exists $\ell \in \mathbb{R}$ such that:

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

We denote this ℓ by $f'(x_0)$, so that:

$$f(x_0+h) \simeq f(x_0) + f'(x_0) \cdot h \quad \text{(for sufficiently small real } h\text{)}.$$

In other words, we approximate the mapping $h \mapsto f(x_0+h) - f(x_0)$ with a linear function $h \mapsto f'(x_0) \cdot h$.

We will extend this idea to higher dimensions.

Definition 4.2.1

Let $f:\mathbb{R}^n\to\mathbb{R}$. The function f is said to be *differentiable* at $x_0\in\mathbb{R}^n$ if there exists a linear map $\ell:\mathbb{R}^n\to\mathbb{R}$ such that:

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

The map ℓ is called the *differential* of f at x_0 and is denoted by $\mathrm{d}f(x_0)$.

In the case of single-variable functions, we have $\mathrm{d}f(x_0)=f'(x_0)$ (and $\mathrm{d}f(x_0)(h)=f'(x_0)\cdot h$). For functions of multiple variables, we'll soon see how to express the differential using partial derivatives. Note that $\mathrm{d}f(x_0)$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}$, so $\mathrm{d}f(x_0)(h)$ is a real number for each $h\in\mathbb{R}^n$.

Just like in the single-variable case, if a function is differentiable, then it is continuous.

Proposition 4.2.1

If f is differentiable at $x_0\in\mathbb{R}^n$, then f is continuous at x_0.

Proof. Let g be the function defined by $g(h)=\frac{f(x_0+h)-f(x_0)-\mathrm{d}f(x_0)(h)}{\|h\|}$. Then:

$$f(x_0+h)=f(x_0)+\mathrm{d}f(x_0)(h)+\|h\|g(h).$$

It is clear that $\mathrm{d}f(x_0)(h)$ and $\|h\|g(h)$ tend to 0 as h approaches the zero vector. Therefore, the limit of f at x_0 exists and equals $f(x_0)$, so f is continuous at x_0. ■

■ **Example 4.7**

If $\ell:\mathbb{R}^n\to\mathbb{R}$ is linear, then ℓ is differentiable, and its differential at every point is the function ℓ itself: for all $x_0\in\mathbb{R}^n$ and $h\in\mathbb{R}^n$,

$$\mathrm{d}\ell(x_0)(h)=\ell(h).$$

Indeed, by definition, the differential of a function at a point involves approximating the function with a linear map. We need to check the definition of differentiability:

$$\mathrm{d}\ell(x_0)(h)=\lim_{\|h\|\to 0}\frac{\ell(x_0+h)-\ell(x_0)-L(h)}{\|h\|}.$$

Since ℓ is linear, we know: $\ell(x_0+h)=\ell(x_0)+\ell(h)$. Substitute into the limit:

$$\mathrm{d}\ell(x_0)(h)=\lim_{\|h\|\to 0}\frac{\ell(x_0)+\ell(h)-\ell(x_0)-\ell(h)}{\|h\|}=0.$$

Thus, the remainder term vanishes, and the linear map itself is exactly the differential:

$$\mathrm{d}\ell(x_0)(h)=\ell(h).$$

■

4.2.2 Differential

Proposition 4.2.2

If $f:\mathbb{R}^n\to\mathbb{R}$ is differentiable at $x_0\in\mathbb{R}^n$, then its partial derivatives exist, and we have:

$$\mathrm{d}f(x_0)(h)=h_1\frac{\partial f}{\partial x_1}(x_0)+\cdots+h_n\frac{\partial f}{\partial x_n}(x_0)$$

where $h=(h_1,\ldots,h_n)$.

In particular, when it exists, the differential is unique.

For a function $f:\mathbb{R}^2\to\mathbb{R}$ differentiable at (x_0,y_0), the formula becomes:

$$\mathrm{d}f(x_0,y_0)(h,k)=h\frac{\partial f}{\partial x}(x_0,y_0)+k\frac{\partial f}{\partial y}(x_0,y_0).$$

Proof. Let's prove the formula for two variables. Suppose $f:\mathbb{R}^2\to\mathbb{R}$ is differentiable at $(x_0,y_0)\in\mathbb{R}^2$. Let $\ell(h,k)=ah+bk$ be its differential. By definition, as $\|(h,k)\|\to 0$, we have:

$$\frac{f(x_0+h,y_0+k)-f(x_0,y_0)-\ell(h,k)}{\|(h,k)\|}\longrightarrow 0.$$

For $(h,k)=(t,0)$ with $t>0$ and $t\to 0$, this gives:

$$\frac{f(x_0+t,y_0)-f(x_0,y_0)-t\ell(1,0)}{t}=\frac{f(x_0+t,y_0)-f(x_0,y_0)}{t}-\ell(1,0)\longrightarrow 0.$$

This exactly means:

$$\frac{\partial f}{\partial x}(x_0,y_0)=\ell(1,0)=a.$$

Similarly, for $(h,k)=(0,t)$, we find:

$$\frac{\partial f}{\partial y}(x_0,y_0)=\ell(0,1)=b.$$

Thus:

$$\ell(h,k)=h\frac{\partial f}{\partial x}(x_0,y_0)+k\frac{\partial f}{\partial y}(x_0,y_0).$$

■

When proving that a function is differentiable, we can use the fact that sums, products, inverses (of non-zero functions), and compositions of differentiable functions are differentiable. Otherwise, we return to the definition. For example, for $f:\mathbb{R}^2\to\mathbb{R}$:

1. First, compute the partial derivatives $\frac{\partial f}{\partial x}(x_0,y_0)$ and $\frac{\partial f}{\partial y}(x_0,y_0)$.
2. Write the candidate for the differential:

$$\ell(h,k)=h\frac{\partial f}{\partial x}(x_0,y_0)+k\frac{\partial f}{\partial y}(x_0,y_0).$$

3. Finally, prove the limit as $\|(h,k)\|\to 0$:

$$\frac{f(x_0+h,y_0+k)-f(x_0,y_0)-\ell(h,k)}{\|(h,k)\|}\longrightarrow 0.$$

■ **Example 4.8**

Let's study the differentiability at every point of the function f defined by:

$$f(x,y)=x-3y+\frac{x^4}{x^2+y^2}\quad\text{if } (x,y)\neq(0,0)\quad\text{and}\quad f(0,0)=0.$$

We have

- Outside of the point $(0,0)$, the function f is differentiable because it is a sum, product, and quotient of differentiable functions (since x^2+y^2 only vanishes at the origin).
- At the point $(0,0)$, we need to check differentiability manually.
  - Partial derivative with respect to x:

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

  - Partial derivative with respect to y:

$$\frac{\partial f}{\partial y}(0,0)=\lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k}=\lim_{k\to 0}\frac{-3k}{k}=-3.$$

- The candidate for the differential is therefore:

$$\ell(h,k) = h\frac{\partial f}{\partial x}(0,0) + k\frac{\partial f}{\partial y}(0,0) = h - 3k.$$

- Let's compute the limit:

$$0 \leq \frac{f(0+h,0+k) - f(0,0) - \ell(h,k)}{\sqrt{h^2+k^2}} = \frac{h^4}{(h^2+k^2)^{\frac{3}{2}}} \leq \frac{h^4}{|h|^3} = |h| \xrightarrow[(h,k)\to(0,0)]{} 0.$$

Thus, f is differentiable at the point $(0,0)$, and:

$$\mathrm{d}f(0,0)(h,k) = h - 3k.$$

■

4.2.3 Link with Partial Derivatives

Partial Derivatives.

From Proposition 4.2.2, we know that if $f : \mathbb{R}^2 \to \mathbb{R}$ is differentiable at (x_0, y_0), then:

$$\mathrm{d}f(x_0,y_0)(1,0) = \frac{\partial f}{\partial x}(x_0,y_0) \qquad \text{and} \qquad \mathrm{d}f(x_0,y_0)(0,1) = \frac{\partial f}{\partial y}(x_0,y_0).$$

In any dimension, for a function $f : \mathbb{R}^n \to \mathbb{R}$ that is differentiable at $x_0 \in \mathbb{R}^n$, and for e_i the i-th vector of the standard basis:

$$\mathrm{d}f(x_0)(e_i) = \frac{\partial f}{\partial x_i}(x_0).$$

Directional Derivative.

More generally, if $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $x_0 \in \mathbb{R}^n$, then:

$$\mathrm{d}f(x_0)(v) = D_v f(x_0).$$

For $f : \mathbb{R}^2 \to \mathbb{R}$, this means that if $v = (h,k)$, we have:

$$D_{(h,k)} f(x_0,y_0) = h\frac{\partial f}{\partial x}(x_0,y_0) + k\frac{\partial f}{\partial y}(x_0,y_0)$$

If f is not differentiable, this formula may not hold.

Gradient.

The gradient provides another way to encode the differential. The **gradient** of f at x_0 is the vector:

$$\mathrm{grad} f(x_0) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(x_0) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(x_0) \end{pmatrix}.$$

If f is differentiable at x_0, then:

$$\mathrm{d}f(x_0)(v) = \langle \mathrm{grad} f(x_0) \mid v \rangle,$$

where $\langle u \mid v \rangle$ denotes the dot product of u and v.

We will return to the gradient and its applications in detail in the chapter *"Gradient — Mean Value Theorem"*.

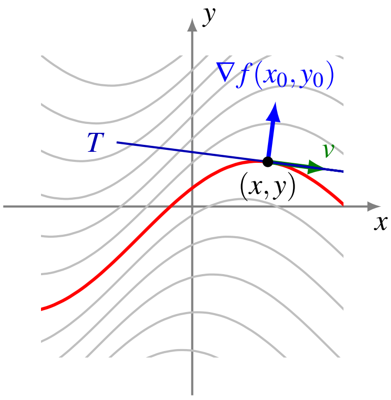
Hence, when f is differentiable, the differential, the directional derivative, and the gradient carry the same information and are connected by the formulas:

$$D_v f(x_0) = \mathrm{d}f(x_0)(v) = \langle \operatorname{grad} f(x_0) \mid v \rangle$$

Geometric Interpretation of the Gradient

Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be a function admitting partial derivatives at $a \in \Omega$.

- The gradient $\nabla f(a)$ is perpendicular to the level curve of f passing through the point a.
- The gradient $\nabla f(a)$ indicates the direction of the steepest ascent, i.e., the direction in which the variations of f increase the most rapidly in the neighborhood of a.



Exercise 4.2

Let f be the function defined by:

$$f(x,y) = \ln(1 + x + y^2).$$

1. Find the domain of definition U of f.
2. Compute the partial derivatives of f.
3. Show that f is differentiable on U.
4. Compute the gradient of f at $(0,1)$ and express the differential at this point.
5. Compute the directional derivative of f at $(0,1)$ in the direction of the vector $(2,1)$.

■

Solution.

1. The domain of definition is:
$$U = \left\{(x,y) \in \mathbb{R}^2 \mid 1 + x + y^2 > 0\right\}.$$
2. The partial derivatives are:
$$\frac{\partial f}{\partial x}(x,y) = \frac{1}{1+x+y^2} \qquad \frac{\partial f}{\partial y}(x,y) = \frac{2y}{1+x+y^2}.$$
3. The function f is differentiable on U as it is a sum, product, and composition of differentiable functions.
4. The gradient is obtained directly from the partial derivatives:
$$\operatorname{grad} f(0,1) = \begin{pmatrix} \frac{\partial f}{\partial x}(0,1) \\ \frac{\partial f}{\partial y}(0,1) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}.$$
The differential at this point, $\mathrm{d}f(0,1) : \mathbb{R}^2 \to \mathbb{R}$, is the linear map defined by:
$$\mathrm{d}f(0,1)(h,k) = \langle \operatorname{grad} f(0,1) \mid (h,k) \rangle = \frac{1}{2}h + k.$$
5. Since f is differentiable, the directional derivative is simply the linear combination of the partial derivatives:
$$D_{(2,1)} f(0,1) = 2 \times \frac{\partial f}{\partial x}(0,1) + 1 \times \frac{\partial f}{\partial y}(0,1) = 2.$$
Alternatively, we could compute it via the differential:
$$D_{(2,1)} f(0,1) = \mathrm{d}f(0,1)(2,1).$$

4.3 Class C^1 Functions

Definition 4.3.1
Let $\Omega \subseteq \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}^p$. We say that f is of class C^1 (continuously differentiable) on Ω if all partial derivatives of f exist and are continuous on Ω. We denote the set of class C^1 functions on Ω with values in $\mathbb{R}^p$ as $C^1(\Omega;\mathbb{R}^p)$, or simply $C^1(\Omega)$ when $p = 1$.

■ Example 4.9
The function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x,y) = xy^2 + 3x,$$

is of class C^1 on $\mathbb{R}^2$ because

$$\frac{\partial f}{\partial x}(x,y) = y^2 + 3, \quad \frac{\partial f}{\partial y}(x,y) = 2xy,$$

are continuous on $\mathbb{R}^2$. ■

Theorem 4.3.1
Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}^p$. If f is of class C^1 on Ω, then it is continuous on Ω, and we have the inclusion $C^1(\Omega;\mathbb{R}^p) \subset C^0(\Omega;\mathbb{R}^p)$.

■ Example 4.10
Consider the function defined on $\mathbb{R}^2$ by

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

We saw in Example 3.4 that f is not continuous at $(0,0)$, so it cannot be of class C^1 on $\mathbb{R}^2$. ■

Exercise 4.3
Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$f(x,y) = \begin{cases} \frac{xy^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

1. Is f continuous on $\mathbb{R}^2$?
2. Study the existence of the partial derivatives of f on $\mathbb{R}^2$ and determine them.
3. Is f of class $C^1(\mathbb{R}^2)$?

■

Solution

1. The function f is continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$ as a quotient of continuous functions on $\mathbb{R}^2 \setminus \{(0,0)\}$ whose denominator does not vanish on this set. For all $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, we have

$$|f(x,y)| = \left| \frac{xy^3}{x^2+y^2} \right| = \frac{|xy||y^2|}{x^2+y^2} \leq \frac{|xy|(x^2+y^2)}{x^2+y^2} = |xy|.$$

Since $\lim_{(x,y)\to(0,0)} |xy| = 0$, it follows that $\lim_{(x,y)\to(0,0)} f(x,y) = 0 = f(0,0)$. This shows that f is continuous at $(0,0)$. In summary, f is continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$ and at $(0,0)$, so f is continuous on $\mathbb{R}^2$.

2. **Partial derivatives on $\mathbb{R}^2 \setminus \{(0,0)\}$:** f has a partial derivative with respect to x on $\mathbb{R}^2 \setminus \{(0,0)\}$ as a quotient of functions that have partial derivatives with respect to their variables, and the denominator does not vanish on this set. Moreover, for all $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$ we have

$$\frac{\partial f}{\partial x}(x,y) = \frac{y^3(x^2+y^2) - 2x(xy^3)}{(x^2+y^2)^2} = \frac{y^5 - x^2y^3}{(x^2+y^2)^2}.$$

Similarly, for all $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$ we have

$$\frac{\partial f}{\partial y}(x,y) = \frac{xy^2(x^2+y^2) - 2y(xy^3)}{(x^2+y^2)^2} = \frac{xy^4+3xy^3}{(x^2+y^2)^2}.$$

Partial derivatives at $(0,0)$: We have

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

- $\frac{\partial f}{\partial y}(0,0) = \lim_{k\to 0} \frac{f(0,k)-f(0,0)}{k} = 0.$

In summary, f has partial derivatives with respect to each of its two variables on $\mathbb{R}^2$, and we have

$$\forall (x,y) \in \mathbb{R}^2 : \frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{y^5-x^2y^3}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

$$\forall (x,y) \in \mathbb{R}^2 : \frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{xy^4+3xy^3}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

3. The partial derivative functions $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$ as quotients of continuous functions on $\mathbb{R}^2 \setminus \{(0,0)\}$ whose denominator does not vanish on this set. Moreover, we have

$$\begin{aligned}\lim_{(x,y)\to(0,0)} \frac{\partial f}{\partial x}(x,y) &= \lim_{(x,y)\to(0,0)} \frac{y^5-x^2y^3}{(x^2+y^2)^2} \\ &= \lim_{r\to 0} \frac{r^5(\sin^5\theta - \cos^2\theta\sin^3\theta)}{r^4} = 0 = \frac{\partial f}{\partial x}(0,0).\end{aligned}$$

$$\begin{aligned}\lim_{(x,y)\to(0,0)} \frac{\partial f}{\partial y}(x,y) &= \lim_{(x,y)\to(0,0)} \frac{xy^4+3xy^3}{(x^2+y^2)^2} \\ &= \lim_{r\to 0} \frac{r^5(\cos\theta\sin^4\theta + 3\cos^3\theta\sin^2\theta)}{r^4} = 0 = \frac{\partial f}{\partial y}(0,0).\end{aligned}$$

This shows that the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are continuous at $(0,0)$, and therefore on $\mathbb{R}^2$. The function f is thus of class $C^1(\mathbb{R}^2)$.

Exercise 4.4

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$f(x,y) = \begin{cases} \frac{x^3-y^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

1. Is f continuous on $\mathbb{R}^2$?
2. Study the existence of the partial derivatives of f on $\mathbb{R}^2$ and determine them.
3. Is f of class $C^1(\mathbb{R}^2)$?

■

Solution.

1. The function f is continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$ as a quotient of continuous functions on $\mathbb{R}^2 \setminus \{(0,0)\}$ whose denominator does not vanish on this set. Moreover,

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} \frac{x^3-y^3}{x^2+y^2} = \lim_{r\to 0} \frac{r^3(\cos^3\theta - \sin^3\theta)}{r^2} = 0.$$

This shows that f is continuous at $(0,0)$. In summary, f is continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$ and at $(0,0)$; therefore, f is continuous on $\mathbb{R}^2$.

2. **Partial derivatives on $\mathbb{R}^2\setminus\{(0,0)\}$:** f has a partial derivative with respect to x on $\mathbb{R}^2\setminus\{(0,0)\}$ as a quotient of functions that have partial derivatives with respect to x, and the denominator does not vanish on this set. Moreover, for every $(x,y)\in\mathbb{R}^2\setminus\{(0,0)\}$ we have
$$\frac{\partial f}{\partial x}(x,y)=\frac{3x^2(x^2+y^2)-2x(x^3-y^3)}{(x^2+y^2)^2}=\frac{x^4+3x^2y^2+2xy^3}{(x^2+y^2)^2}.$$
Similarly, for every $(x,y)\in\mathbb{R}^2\setminus\{(0,0)\}$ we have
$$\frac{\partial f}{\partial y}(x,y)=\frac{y^4+3x^2y^2+2xy^3}{(x^2+y^2)^2}.$$
Partial derivatives at $(0,0)$:
   - $\frac{\partial f}{\partial x}(0,0)=\lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{h^3}{h}=h^2\to 0.$
   - $\frac{\partial f}{\partial y}(0,0)=\lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k}=\lim_{k\to 0}\frac{-k^3}{k}=-k^2\to 0.$

   In summary, f has partial derivatives with respect to each of its two variables on $\mathbb{R}^2$, and we have
$$\forall(x,y)\in\mathbb{R}^2:\frac{\partial f}{\partial x}(x,y)=\begin{cases}\frac{x^4+3x^2y^2+2xy^3}{(x^2+y^2)^2} & \text{if } (x,y)\neq(0,0),\\ 0 & \text{if } (x,y)=(0,0).\end{cases}$$
$$\forall(x,y)\in\mathbb{R}^2:\frac{\partial f}{\partial y}(x,y)=\begin{cases}\frac{y^4+3x^2y^2+2xy^3}{(x^2+y^2)^2} & \text{if } (x,y)\neq(0,0),\\ 0 & \text{if } (x,y)=(0,0).\end{cases}$$

3. **Study on $\mathbb{R}^2\setminus\{(0,0)\}$:** The partial derivative functions $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are continuous on $\mathbb{R}^2\setminus\{(0,0)\}$ as quotients of continuous functions on $\mathbb{R}^2\setminus\{(0,0)\}$ whose denominator does not vanish on this set. Hence, f is of class $C^1(\mathbb{R}^2\setminus\{(0,0)\})$.
Study at $(0,0)$: The partial derivatives are continuous at $(0,0)$ if and only if
$$\lim_{(x,y)\to(0,0)}\frac{\partial f}{\partial x}(x,y)=\frac{\partial f}{\partial x}(0,0)\quad\text{and}\quad\lim_{(x,y)\to(0,0)}\frac{\partial f}{\partial y}(x,y)=\frac{\partial f}{\partial y}(0,0).$$
Since
$$\lim_{(x,y)\to(0,0)}\frac{\partial f}{\partial x}(x,y)=0=\frac{\partial f}{\partial x}(0,0),$$
and
$$\lim_{(x,y)\to(0,0)}\frac{\partial f}{\partial y}(x,y)=0=\frac{\partial f}{\partial y}(0,0),$$
the partial derivatives are continuous at $(0,0)$. Therefore, the function f is of class $C^1(\mathbb{R}^2)$.

4.4 Jacobian Matrix

Definition 4.4.1
Let
$$f:\Omega\subset\mathbb{R}^n\to\mathbb{R}^p$$
$$(x_1,\dots,x_n)\mapsto f(x_1,\dots,x_n)=(f_1(x_1,\dots,x_n),f_2(x_1,\dots,x_n),\dots,f_p(x_1,\dots,x_n)),$$
admitting partial derivatives at $a\in\Omega$. The Jacobian matrix of f at a, denoted by $J_f(a)$, is the $p\times n$ matrix defined as follows:
$$J_f(a)=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a)\\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a)\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial f_p}{\partial x_1}(a) & \frac{\partial f_p}{\partial x_2}(a) & \cdots & \frac{\partial f_p}{\partial x_n}(a)\end{pmatrix}\in\mathbb{M}_{p,n}(\mathbb{R}).$$
In other words, the Jacobian matrix of f has as its columns the vectors $\frac{\partial f}{\partial x_i}$.

If $n = p$, the determinant of the matrix $J_f(a)$ is called the Jacobian determinant of f at a, and is written as:

$$\det J_f(a) = \frac{\partial(f_1, f_2, \ldots, f_p)}{\partial(x_1, x_2, \ldots, x_n)}.$$

Example 4.11
Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be a function defined by:

$$f(x,y,z) = \big(f_1(x,y,z), f_2(x,y,z)\big) = \big(xy^2, \sin(yz)\big).$$

For all $(x,y,z) \in \mathbb{R}^3$:

$$J_f(x,y,z) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x,y,z) & \frac{\partial f_1}{\partial y}(x,y,z) & \frac{\partial f_1}{\partial z}(x,y,z) \\ \frac{\partial f_2}{\partial x}(x,y,z) & \frac{\partial f_2}{\partial y}(x,y,z) & \frac{\partial f_2}{\partial z}(x,y,z) \end{pmatrix} = \begin{pmatrix} y^2 & 2xy & 0 \\ 0 & z\cos(yz) & y\cos(yz) \end{pmatrix}.$$

■

Example 4.12
Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be a function defined by:

$$f(r,\theta) = \big(f_1(r,\theta), f_2(r,\theta)\big) = \big(r\cos\theta, r\sin\theta\big).$$

For all $(r,\theta) \in \mathbb{R}^2$:

$$J_f(r,\theta) = \begin{pmatrix} \frac{\partial f_1}{\partial r}(r,\theta) & \frac{\partial f_1}{\partial \theta}(r,\theta) \\ \frac{\partial f_2}{\partial r}(r,\theta) & \frac{\partial f_2}{\partial \theta}(r,\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

Therefore, the Jacobian determinant of f is:

$$\det J_f(r,\theta) = \frac{\partial(f_1, f_2)}{\partial(r,\theta)} = r.$$

■

4.4.1 Differential

The theoretical counterpart of the Jacobian matrix is the differential associated with a function $F : \mathbb{R}^n \to \mathbb{R}^p$ at a point x. This section is more theoretical: for a first reading, it's enough to remember that the differential $\mathrm{d}F(x)$ is a linear map whose matrix (in the canonical basis) is the Jacobian matrix $J_F(x)$.

Let's break these concepts down in more detail. The notions of limit and continuity for $F : \mathbb{R}^n \to \mathbb{R}^p$ are similar to those for scalar functions $f : \mathbb{R}^n \to \mathbb{R}$, except that the absolute value in $\mathbb{R}$ is replaced by a norm in $\mathbb{R}^p$.

We'll now explore the differential of a vector-valued function. Let $F : \mathbb{R}^n \to \mathbb{R}^p$ with components $F = (f_1, \ldots, f_p)$, where each component is a function $f_j : \mathbb{R}^n \to \mathbb{R}$.

Definition 4.4.2

- $F : \mathbb{R}^n \to \mathbb{R}^p$ is said to be *differentiable* at a point $x \in \mathbb{R}^n$ if each component function $f_j : \mathbb{R}^n \to \mathbb{R}$ (for $j = 1, \ldots, p$) is differentiable at x. The differential of the component f_j at x is denoted by $\mathrm{d}f_j(x) : \mathbb{R}^n \to \mathbb{R}$.
- The *differential* of a differentiable vector-valued function $F : \mathbb{R}^n \to \mathbb{R}^p$ at $x \in \mathbb{R}^n$ is the linear map $\mathrm{d}F(x) : \mathbb{R}^n \to \mathbb{R}^p$ defined as:
$$\mathrm{d}F(x) = \big(\mathrm{d}f_1(x), \ldots, \mathrm{d}f_p(x)\big).$$

Important note: The differential $\mathrm{d}F(x)$ is a linear map, so it is a function (not just a vector). Evaluating this function gives a vector-valued expression:

$$\forall h \in \mathbb{R}^n, \qquad \mathrm{d}F(x)(h) = \big(\mathrm{d}f_1(x)(h), \ldots, \mathrm{d}f_p(x)(h)\big).$$

Proposition 4.4.1

Let $F : \mathbb{R}^n \to \mathbb{R}^p$ be differentiable at $x \in \mathbb{R}^n$. Then:

$$\mathrm{d}F(x)(h) = J_F(x) \times h$$

where $J_F(x)$ is the Jacobian matrix of F at x, for any $h \in \mathbb{R}^n$.

In other words, computing the differential at a point is the same as computing the Jacobian matrix at that point. This result comes from the expression of each differential $\mathrm{d}f_j(x)$ in terms of the partial derivatives $\frac{\partial f_j}{\partial x_i}$ for $i = 1, \dots, n$.

■ Example 4.13

Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by:

$$F(x,y) = (ye^{x^2}, x^2 - y).$$

Let's compute $\mathrm{d}F(x,y)(h,k)$ for arbitrary points (x,y) and vectors $(h,k) \in \mathbb{R}^2$.

- The Jacobian matrix of F is:

$$J_F(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x,y) & \frac{\partial f_1}{\partial y}(x,y) \\ \frac{\partial f_2}{\partial x}(x,y) & \frac{\partial f_2}{\partial y}(x,y) \end{pmatrix} = \begin{pmatrix} 2xye^{x^2} & e^{x^2} \\ 2x & -1 \end{pmatrix}.$$

- At point (x,y), for a vector $(h,k) \in \mathbb{R}^2$, the differential is:

$$\mathrm{d}F(x,y)(h,k) = J_F(x,y) \times \begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} (2xyh + k)e^{x^2} \\ 2xh - k \end{pmatrix}.$$

- For example, at the point $(x_0, y_0) = (1,1)$:

$$\mathrm{d}F(1,1)(h,k) = \big((2h+k)e, 2h - k\big).$$

■

R

- If the components of F are of class $\mathscr{C}^1$ (i.e., all partial derivatives exist and are continuous), then they are differentiable, and F itself is differentiable.
- If F is differentiable at x, then F is continuous at x.
- If $L : \mathbb{R}^n \to \mathbb{R}^p$ is a linear map, then its differential at every point is the map itself: that is, $\mathrm{d}L(x) = L$ for all $x \in \mathbb{R}^n$.

R There is an equivalent definition for the two concepts we've discussed.

- A function $F : \mathbb{R}^n \to \mathbb{R}^p$ is *differentiable* at a point $x \in \mathbb{R}^n$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^p$ such that:

$$\lim_{\|h\| \to 0} \frac{F(x+h) - F(x) - L(h)}{\|h\|} = 0.$$

- In this case, L is the *differential* of F at x, denoted by $\mathrm{d}F(x)$.

4.5 Operations on Differentiable Functions

The algebraic operations on differentiable functions — sums, products, quotients, and scalar multiplication — are the same as those previously seen for differentiable functions of a single real variable.

Theorem 4.5.1 — Linearity.
Let $f, g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^p$ be two functions differentiable at $a \in \Omega$ and let $\lambda, \mu \in \mathbb{R}$. Then the function $\lambda f + \mu g$ is differentiable at a, and we have:

$$d_a(\lambda f + \mu g) = \lambda d_a f + \mu d_a g.$$

In terms of Jacobians:

$$J_{\lambda f+\mu g}(a) = \lambda J_f(a) + \mu J_g(a).$$

Proof. For all $h \in \mathbb{R}^n$, we have:

$$f(a+h) - f(a) = d_a f(h) + \|h\|\varepsilon_1(h) \quad \text{and} \quad g(a+h) - g(a) = d_a g(h) + \|h\|\varepsilon_2(h),$$

where $\lim_{h\to 0} \varepsilon_1(h) = \lim_{h\to 0} \varepsilon_2(h) = 0$. Therefore:

$$\begin{aligned}(\lambda f + \mu g)(a+h) - (\lambda f + \mu g)(a) &= \lambda(f(a+h) - f(a)) + \mu(g(a+h) - g(a)) \\ &= \lambda d_a f(h) + \mu d_a g(h) + \|h\|(\lambda\varepsilon_1(h) + \mu\varepsilon_2(h)).\end{aligned}$$

Since $\lambda d_a f + \mu d_a g$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^p$, the function $\lambda f + \mu g$ is differentiable at a and:

$$d_a(\lambda f + \mu g) = \lambda d_a f + \mu d_a g.$$

■

■ Example 4.14
The function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by:

$$\varphi(x,y) = 2x - 3y + 4,$$

is differentiable at every point $a = (x,y) \in \mathbb{R}^2$ because we can write it as the sum of a linear function and a constant: - $f(x,y) = 2x - 3y$ (linear map), - $g(x,y) = 4$ (constant map).

For any $h = (h_1, h_2) \in \mathbb{R}^2$:

$$d_a f(h_1,h_2) = 2h_1 - 3h_2 \quad \text{and} \quad d_a g(h_1,h_2) = 0.$$

Thus:

$$d_a\varphi(h_1,h_2) = d_a f(h_1,h_2) + d_a g(h_1,h_2) = 2h_1 - 3h_2.$$

■

Theorem 4.5.2
Let $f, g : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be two functions differentiable at $a \in \Omega$. Then:

- The product $f \times g$ is differentiable at a and:

$$d_a(f \times g) = g(a) d_a f + f(a) d_a g.$$

- If $f(a) \neq 0$, the inverse $\frac{1}{f}$ is differentiable at a and:

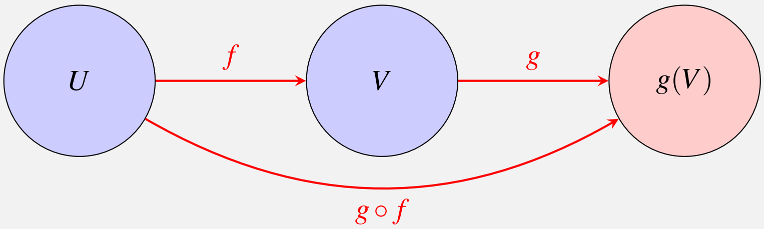
$$d_a\left(\frac{1}{f}\right) = -\frac{d_a f}{(f(a))^2}.$$

- If $g(a) \neq 0$, the quotient $\frac{f}{g}$ is differentiable at a and:

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

The **chain rule** (or composition rule) is a key result for differentiating functions of multiple variables.

Theorem 4.5.3 — Differentiating Composite Functions.
Let f be a function from an open set $U \subset \mathbb{R}^n$ to $\mathbb{R}^p$, and g a function from an open set $V \subset \mathbb{R}^p$ to $\mathbb{R}^m$, with $f(U) \subset V$.



If f is differentiable at a and g is differentiable at $f(a)$, then $g \circ f$ is differentiable at a and:

$$d(g \circ f)_a = d_{f(a)}g \circ d_a f.$$

In terms of Jacobians:

$$J_{g \circ f}(a) = J_g(f(a)) \times J_f(a). \tag{4.1}$$

It is very useful to break down formula (4.1) to understand how to apply it. Let:

$$\begin{aligned} f : U \subset \mathbb{R}^n &\longrightarrow \mathbb{R}^p \\ (x_1,\dots,x_n) &\longmapsto f(x_1,\dots,x_n) = (f_1(x_1,\dots,x_n),\dots,f_p(x_1,\dots,x_n)), \end{aligned}$$

and:

$$\begin{aligned} g : V \subset \mathbb{R}^p &\longrightarrow \mathbb{R}^m \\ (y_1,\dots,y_p) &\longmapsto g(y_1,\dots,y_p) = (g_1(y_1,\dots,y_p),\dots,g_m(y_1,\dots,y_p)). \end{aligned}$$

Since $J_{g \circ f}(a) = J_g(f(a)) \times J_f(a)$, we obtain:

$$\begin{pmatrix} \frac{\partial g_1}{\partial y_1}(f(a)) & \cdots & \frac{\partial g_1}{\partial y_p}(f(a)) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial y_1}(f(a)) & \cdots & \frac{\partial g_m}{\partial y_p}(f(a)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \cdots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix} = \begin{pmatrix} \frac{\partial (g\circ f)_1}{\partial x_1}(a) & \cdots & \frac{\partial (g\circ f)_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial (g\circ f)_m}{\partial x_1}(a) & \cdots & \frac{\partial (g\circ f)_m}{\partial x_n}(a) \end{pmatrix}.$$

Thus:

$$\frac{\partial (g \circ f)_1}{\partial x_1}(a) = \sum_{k=1}^{p} \frac{\partial g_1}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_1}(a).$$

And similarly:

$$\frac{\partial (g \circ f)_m}{\partial x_n}(a) = \sum_{k=1}^{p} \frac{\partial g_m}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_n}(a).$$

Thus, to compute the partial derivatives of composite functions or after a change of variables, we can use the component-wise version of the chain rule.

Corollary 4.5.4 — Chain Rule (Component Form).
For all $(i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}$:

$$\frac{\partial (g \circ f)_i}{\partial x_j}(a) = \sum_{k=1}^{p} \frac{\partial g_i}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a).$$

Exercise 4.5 — Example of Composition and Jacobians.

Consider the functions f and g defined by:

$$\begin{aligned} f &: \mathbb{R}^2 \to \mathbb{R}^3 \\ (x,y) &\mapsto (f_1(x,y), f_2(x,y), f_3(x,y)) = (x^2y, xy, xy^3), \end{aligned}$$

and:

$$\begin{aligned} g &: \mathbb{R}^3 \to \mathbb{R}^2 \\ (x,y,z) &\mapsto (g_1(x,y,z), g_2(x,y,z)) = (x+y+z, xyz). \end{aligned}$$

1. Compute:
    (a) The Jacobian matrix of f at $a = (x,y)$.
    (b) The Jacobian matrix of g at $f(a)$.
    (c) The Jacobian matrix of $g \circ f$ at a.
2. Verify the Jacobian of $g \circ f$ explicitly.

Solution. It's clear that $f \in C^1(\mathbb{R}^2)$ and $g \in C^1(\mathbb{R}^3)$.

(a) Jacobian of f at $a = (x,y)$:

$$J_f(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x,y) & \frac{\partial f_1}{\partial y}(x,y) \\ \frac{\partial f_2}{\partial x}(x,y) & \frac{\partial f_2}{\partial y}(x,y) \\ \frac{\partial f_3}{\partial x}(x,y) & \frac{\partial f_3}{\partial y}(x,y) \end{pmatrix} = \begin{pmatrix} 2xy & x^2 \\ y & x \\ y^3 & 3xy^2 \end{pmatrix}.$$

(b) Jacobian of g at $f(a)$:

$$J_g(x,y,z) = \begin{pmatrix} 1 & 1 & 1 \\ yz & xz & xy \end{pmatrix}.$$

Therefore, the Jacobian of g at $f(a)$ is:

$$J_g(f(a)) = \begin{pmatrix} 1 & 1 & 1 \\ x^2y^4 & x^3y^4 & x^3y^2 \end{pmatrix}.$$

(c) Jacobian of $g \circ f$ at a:

$$\begin{aligned} J_{g \circ f}(x,y) &= J_g(f(x,y)) \times J_f(x,y) \\ &= \begin{pmatrix} 1 & 1 & 1 \\ x^2y^4 & x^3y^4 & x^3y^2 \end{pmatrix} \times \begin{pmatrix} 2xy & x^2 \\ y & x \\ y^3 & 3xy^2 \end{pmatrix} \\ &= \begin{pmatrix} 2xy + y + y^3 & x^2 + x + 3xy^2 \\ 4x^3y^5 & 5x^4y^4 \end{pmatrix}. \end{aligned}$$

4.6 Exercises

Exercise 4.6

Justify the existence of partial derivatives of the following functions, and calculate them.

$$f(x,y) = e^x \cos y, \quad g(x,y) = (x^2+y^2)\ \cos(xy), \quad h(x,y) = \sqrt{1+x^2y^2}.$$

For the following functions, prove that they admit a derivative following any vector at $(0,0)$ without being continuous there.

$$f(x,y) = \begin{cases} y^2 \ln|x| & \text{si } x \neq 0 \\ 0 & \text{otherwise.} \end{cases}, \qquad g(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{othrwise.} \end{cases}$$

We define $f:\mathbb{R}^2\setminus\{(0,0)\}\to\mathbb{R}$ by

$$f(x,y)=\frac{x^2}{(x^2+y^2)^{3/4}}.$$

Justify that we can extend f in a continuous function on $\mathbb{R}^2$. Study the existence of partial derivatives at $(0,0)$ for this extension. ■

Exercise 4.7
Justify that the following functions are differentiable, and calculate their differential

$$(a)\ f(x,y)=e^{xy}(x+y)\quad (b)\ f(x,y,z)=xy+yz+zx\quad (c)\ f(x,y)=(y\sin x,\cos x).$$

■

Exercise 4.8
Let $f:\mathbb{R}^2\to\mathbb{R}$ defined by :

$$f(x,y)=\begin{cases}\frac{x^2-y^2}{x^2+y^2} & \text{if } (x,y)\neq(0,0)\\ 0 & \text{if } (x,y)=(0,0)\end{cases}$$

Is f continuous on $\mathbb{R}^2$? Is f of class C^1 on $\mathbb{R}^2$? Is f differentiable on$\mathbb{R}^2$? ■

Exercise 4.9

1. Prove that, for all real numbers (x,y) , $|xy|\leq x^2-xy+y^2$.
2. Let f the function from $\mathbb{R}^2$ in $\mathbb{R}$ defined by $f(0,0)=0$ and $f(x,y)=(x^py^q)/(x^2-xy+y^2)$ if $(x,y)\neq(0,0)$, where p and q are non zero natural integers . For which values of p and q this function is continuous?
3. Show that if $p+q=2$, then f is not differentiable.
4. Suppose that $p+q=3$, and f is differentiable at $(0,0)$. Justify that then there are two constants a and b such that $f(x,y)=ax+by+o(\|(x,y)\|)$. By studying partial applications $x\mapsto f(x,0)$ and $y\mapsto f(0,y)$, justify that $a=b=0$. Conclude, using $x\mapsto f(x,x)$, that f is not differentiable at $(0,0)$.

■

4.7 Higher-Order Partial Derivatives

Definition 4.7.1
Let $i\in\{1,\dots,n\}$, Ω be a non-empty open subset of $\mathbb{R}^n$, $a\in\Omega$, and $f:\Omega\to\mathbb{R}^p$ a function admitting a partial derivative $\frac{\partial f}{\partial x_i}$ on Ω. If the function $\frac{\partial f}{\partial x_i}:\Omega\to\mathbb{R}^p$ itself possesses a partial derivative with respect to x_j at a, where $j\in\{1,\dots,n\}$, we say that f is twice differentiable with respect to x_i and x_j at a. The second-order partial derivative is denoted as follows:

R

By induction, the partial derivatives of order k of f at point a are defined as the partial derivatives of the order $k-1$ derivatives.

The k-th order partial derivative of a function of the variables $x_1,x_2,\dots,x_n$ is obtained by differentiating k_1 times with respect to $x_1,\dots$, k_n times with respect to x_n, where $(k_1,\dots,k_n)\in\mathbb{N}^n$ with $k_1+\dots+k_n=k$. It is denoted as follows:

$$\frac{\partial^k f}{\partial x_1^{k_1}\partial x_2^{k_2}\dots\partial x_n^{k_n}}.$$

Definition 4.7.2 — Class C^k Functions.
Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}^p$.

- We say that f is of class C^k on Ω if all partial derivatives of f up to order k exist and are continuous on Ω. The set of class C^k functions on Ω with values in $\mathbb{R}^p$ is denoted by $C^k(\Omega;\mathbb{R}^p)$, and simply $C^k(\Omega)$ when $p = 1$.
- We say that f is of class C^∞ on Ω if f is of class C^k for all $k \in \mathbb{N}$:

$$C^\infty(\Omega;\mathbb{R}^p) = \bigcap_{k\in\mathbb{N}} C^k(\Omega;\mathbb{R}^p).$$

4.7.1 Second-Order Partial Derivatives

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function. The two partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are also functions from $\mathbb{R}^2$ to $\mathbb{R}$; suppose they are also differentiable. Then we can compute the partial derivatives of $\frac{\partial f}{\partial x}$:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \quad \text{and} \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

We can also compute the partial derivatives of $\frac{\partial f}{\partial y}$:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \quad \text{and} \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right).$$

We denote these partial derivatives as:

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y^2}.$$

These are functions from $\mathbb{R}^2$ to $\mathbb{R}$.

More generally, for $f : \mathbb{R}^n \to \mathbb{R}$, we denote $\frac{\partial f}{\partial x_i} : \mathbb{R}^n \to \mathbb{R}$ as the first-order partial derivatives ($1 \le i \le n$) and $\frac{\partial^2 f}{\partial x_j \partial x_i}$ as the second-order partial derivatives ($1 \le i, j \le n$).

■ **Example 4.15**
Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$f(x,y) = x^3y^2 + \cos x - \sin y.$$

- The first-order partial derivatives are:

$$\frac{\partial f}{\partial x}(x,y) = 3x^2y^2 - \sin x, \quad \frac{\partial f}{\partial y}(x,y) = 2x^3y - \cos y.$$

- The second-order partial derivatives are:

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 6xy^2 - \cos x, \quad \frac{\partial^2 f}{\partial x \partial y}(x,y) = 6x^2y,$$

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = 6x^2y, \quad \frac{\partial^2 f}{\partial y^2}(x,y) = 2x^3 + \sin y.$$

- The third-order partial derivatives are:

$$\frac{\partial^3 f}{\partial x^3}(x,y) = 6y^2 + \sin x, \quad \frac{\partial^3 f}{\partial y \partial x^2}(x,y) = 12xy = \frac{\partial^3 f}{\partial x^2 \partial y}(x,y),$$

$$\frac{\partial^3 f}{\partial y^3}(x,y) = \cos y, \quad \frac{\partial^3 f}{\partial x \partial y^2}(x,y) = 6x^2 = \frac{\partial^3 f}{\partial y^2 \partial x}(x,y).$$

Since all partial derivatives of f up to order 3 exist and are continuous on $\mathbb{R}^2$, we conclude that $f \in C^3(\mathbb{R}^2)$. We can even show that $f \in C^\infty(\mathbb{R}^2)$. ■

■ Example 4.16

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$f(x,y) = \begin{cases} \frac{xy^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

We calculate $\frac{\partial^2 f}{\partial x \partial y}(0,0)$, $\frac{\partial^2 f}{\partial y \partial x}(0,0)$, $\frac{\partial^2 f}{\partial x^2}(0,0)$, and $\frac{\partial^2 f}{\partial y^2}(0,0)$.
we know that:

$$\forall (x,y) \in \mathbb{R}^2 : \frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{y^5 - x^2y^3}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

$$\forall (x,y) \in \mathbb{R}^2 : \frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{xy^4 + 3x^3y^2}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

We find:

- $\frac{\partial^2 f}{\partial x^2}(0,0) = 0$.
- $\frac{\partial^2 f}{\partial y \partial x}(0,0) = 1$.
- $\frac{\partial^2 f}{\partial x \partial y}(0,0) = 0$.
- $\frac{\partial^2 f}{\partial y^2}(0,0) = 0$.

Hence, $\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$. ■

The following Schwarz's theorem provides a sufficient condition for the equality of mixed partial derivatives.

Theorem 4.7.1 — Schwarz's Theorem. Let Ω be a non-empty open subset of $\mathbb{R}^n$, $a \in \Omega$, and $i, j \in \{1, \ldots, n\}$. Let $f : \Omega \to \mathbb{R}^p$ be a function such that $\frac{\partial^2 f}{\partial x_i \partial x_j}$ and $\frac{\partial^2 f}{\partial x_j \partial x_i}$ exist on Ω and are continuous at the point a. Then:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

R

The preceding theorem can be extended to the case of partial derivatives of order $k \geq 3$ as follows:
If $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^p$ is of class C^k on an open set Ω, then the order in which any k-th order partial derivative is calculated does not matter. If σ is a permutation of $\{1, 2, \ldots, k\}$, then:

$$\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{i_{\sigma(1)}} \partial x_{i_{\sigma(2)}} \ldots \partial x_{i_{\sigma(k)}}} \quad \text{on } \Omega, \text{ with } 1 \leq i_1, \ldots, i_k \leq n.$$

■ Example 4.17 — Solving a Second-Order Partial Differential Equation. Let $c \neq 0$. Find the C^2 solutions to the following PDE:

$$(E) : c^2 \frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial t^2} = 0.$$

using the change of variables $u = x + ct$ and $v = x - ct$. Let's define the map:

$$\varphi(x,t) = (x + ct, x - ct) = (u, v).$$

If we let $f(x,t) = g(u,v)$ for some function g, we get:

$$f(x,t) = g(x + ct, x - ct).$$

In diagram form:

$$(x,t) \xrightarrow{\varphi} (u,v) \xrightarrow{g} g(u,v) = f(x,t).$$

Since g is C^2 if and only if f is C^2, we can apply the chain rule:

$$\frac{\partial f}{\partial x} = \frac{\partial g}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial x} = \frac{\partial g}{\partial u} + \frac{\partial g}{\partial v}.$$

$$\frac{\partial f}{\partial t} = \frac{\partial g}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial t} = c\frac{\partial g}{\partial u} - c\frac{\partial g}{\partial v}.$$

Differentiating again:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 g}{\partial u^2} + 2\frac{\partial^2 g}{\partial u \partial v} + \frac{\partial^2 g}{\partial v^2}.$$

$$\frac{\partial^2 f}{\partial t^2} = c^2\frac{\partial^2 g}{\partial u^2} - 2c^2\frac{\partial^2 g}{\partial u \partial v} + c^2\frac{\partial^2 g}{\partial v^2}.$$

Substitute into the original PDE:

$$c^2\frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial t^2} = c^2\frac{\partial^2 g}{\partial u^2} + 2c^2\frac{\partial^2 g}{\partial u \partial v} + c^2\frac{\partial^2 g}{\partial v^2} - \left(c^2\frac{\partial^2 g}{\partial u^2} - 2c^2\frac{\partial^2 g}{\partial u \partial v} + c^2\frac{\partial^2 g}{\partial v^2}\right).$$

Simplify:

$$= 4c^2\frac{\partial^2 g}{\partial u \partial v}.$$

Thus, the PDE reduces to:

$$\frac{\partial^2 g}{\partial u \partial v} = 0.$$

The general solution is then:

$$g(u,v) = A(u) + B(v),$$

where A and B are arbitrary C^2 functions. Returning to the original variables, we get the general solution of the PDE:

$$f(x,t) = A(x+ct) + B(x-ct).$$

Here, A and B represent traveling waves moving to the right and left, respectively, with speed c. ■

Exercise 4.10

Let f be the function defined on $\mathbb{R}^2$ by:

$$f(x,y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

1. Show that f is continuous on $\mathbb{R}^2$.
2. Show that f is of class C^1 on $\mathbb{R}^2$.
3. Calculate $\frac{\partial^2 f}{\partial x \partial y}(0,0)$ and $\frac{\partial^2 f}{\partial y \partial x}(0,0)$. Is f of class C^2 on $\mathbb{R}^2$?

■

Solution.

1. The function f is continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$ as a quotient of continuous functions on $\mathbb{R}^2 \setminus \{(0,0)\}$ whose denominator does not vanish on this set. Moreover:

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} \frac{xy(x^2-y^2)}{x^2+y^2} = \lim_{r\to 0} \frac{r^4 \cos\theta \sin\theta(\cos^2\theta - \sin^2\theta)}{r^2} = 0.$$

   This shows that f is continuous at $(0,0)$. In summary, f is continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$ and at $(0,0)$; therefore, f is continuous on $\mathbb{R}^2$.

2. **Partial derivatives on $\mathbb{R}^2 \setminus \{(0,0)\}$:** f has a partial derivative with respect to x on $\mathbb{R}^2 \setminus \{(0,0)\}$ as a quotient of functions that have partial derivatives with respect to x, and the denominator does not vanish. Moreover, for all $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$:

$$\frac{\partial f}{\partial x}(x,y) = \frac{x^4y + 4x^2y^3 - y^5}{(x^2+y^2)^2}.$$

Similarly, for all $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$:

$$\frac{\partial f}{\partial y}(x,y) = \frac{x^5 - 4x^3y^2 - xy^4}{(x^2+y^2)^2}.$$

Partial derivatives at $(0,0)$:
- $\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h} = 0.$
- $\frac{\partial f}{\partial y}(0,0) = \lim_{k\to 0} \frac{f(0,k)-f(0,0)}{k} = 0.$

In summary, f has partial derivatives with respect to x and y on $\mathbb{R}^2$, and:

$$\forall (x,y) \in \mathbb{R}^2 : \frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{x^4y+4x^2y^3-y^5}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

$$\forall (x,y) \in \mathbb{R}^2 : \frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{x^5-4x^3y^2-xy^4}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

3. **Second-order derivatives at $(0,0)$:**
- $\frac{\partial^2 f}{\partial y \partial x}(0,0) = -1.$
- $\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1.$

Since $\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$, by Schwarz's theorem, at least one of the mixed partial derivatives is not continuous at $(0,0)$. Therefore, f is not of class $C^2(\mathbb{R}^2)$.

4.7.2 Taylor's Formula

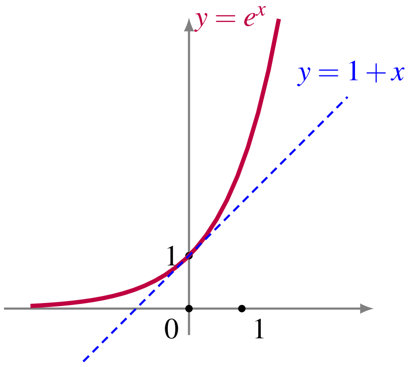
First, we recall it in case of one variable function. Let $f : \mathbb{R} \to \mathbb{R}$ be a function of one variable of class $\mathscr{C}^2$.

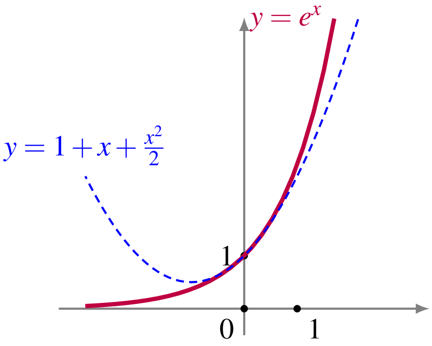
Theorem 4.7.2 — Taylor's Formula to Order 2. For all $x_0 \in \mathbb{R}$, we have

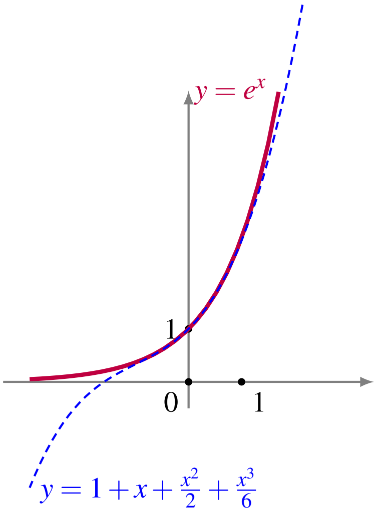
$$f(x_0+h) = f(x_0) + hf'(x_0) + \frac{h^2}{2} f''(x_0) + h^2\varepsilon(h)$$

where $\varepsilon(h) \to 0$ as $h \to 0$.

The first-order Taylor expansion, $f(x_0+h) \simeq f(x_0) + hf'(x_0)$, corresponds to approximating the graph of f by its tangent at x_0 (left figure below). The second-order Taylor expansion, $f(x_0+h) \simeq f(x_0) + hf'(x_0) + \frac{h^2}{2} f''(x_0)$, corresponds to an approximation by a parabola (right figure).







Choose x_0 such that $f'(x_0)=0$ and $f''(x_0)\neq 0$. Then, for h sufficiently small, the term $\frac{h^2}{2}f''(x_0)+h^2\varepsilon(h)$ has the same sign as $f''(x_0)$. If, for example, $f''(x_0)>0$, we deduce that $f(x_0+h)\geq f(x_0)$ (for h near 0) and thus f has a local minimum at x_0.

Now, the Taylor's formula for two variable functions.

4.7.3 Taylor's Formula to Order 1

We have already seen that one way to say that f is differentiable is to say that f admits a **first-order Taylor expansion**. For $f:\mathbb{R}^2\to\mathbb{R}$, at the point (x_0,y_0), if f is differentiable, then

$$f(x_0+h,y_0+k)=f(x_0,y_0)+h\frac{\partial f}{\partial x}(x_0,y_0)+k\frac{\partial f}{\partial y}(x_0,y_0)+o\left(\|(h,k)\|\right).$$

Knowing the values of f, $\frac{\partial f}{\partial x}$, and $\frac{\partial f}{\partial y}$ only at (x_0,y_0), we deduce an approximation of f at any (x,y) near (x_0,y_0).

The goal will be to improve this approximation with a second-order Taylor expansion.

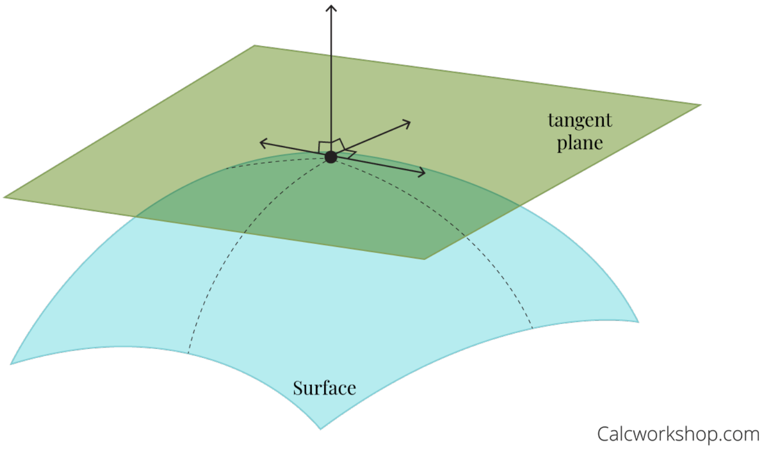
We recall the notation *little o*.

Notation. Let $g:\mathbb{R}^2\to\mathbb{R}$ be a function defined in a neighborhood of $(0,0)$. We say that g is **negligible** with respect to $\|(x,y)\|^n$ in a neighborhood of $(0,0)$ and write $g=o\left(\|(x,y)\|^n\right)$ if

$$\lim_{(x,y)\to(0,0)}\frac{g(x,y)}{\|(x,y)\|^n}=0.$$

Geometric Interpretation.

The tangent plane (in blue) to the graph of f at the point (x_0,y_0) is the plane that best approximates the graph of f (in red) around this point. To compute $f(x,y)$ when $(x,y)=(x_0+h,y_0+k)$ is near (x_0,y_0), we replace the exact value $z_{\text{exact}} = f(x,y)$ with its approximation $z_{\text{approx}} = f(x_0,y_0) + h\frac{\partial f}{\partial x}(x_0,y_0) + k\frac{\partial f}{\partial y}(x_0,y_0)$. The point (x,y,z_{exact}) is a point on the graph of f, while (x,y,z_{approx}) is a point on the tangent plane to the graph of f at (x_0,y_0).



4.7.4 Taylor's Formula to Order 2 (in 2 Variables)

Theorem 4.7.3

Let $f : U \to \mathbb{R}$ be a function of class $\mathscr{C}^2$ on an open set $U \subset \mathbb{R}^2$, and let $(x_0,y_0) \in U$. Then:

$$\begin{aligned} f(x_0+h,y_0+k) = f(x_0,y_0) \; &+ \; h\frac{\partial f}{\partial x}(x_0,y_0) + k\frac{\partial f}{\partial y}(x_0,y_0) \\ &+ \frac{1}{2}\left[h^2\frac{\partial^2 f}{\partial x^2}(x_0,y_0) + 2hk\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) + k^2\frac{\partial^2 f}{\partial y^2}(x_0,y_0)\right] \\ &+ o\left(\|(h,k)\|^2\right) \end{aligned}$$

- We also say that f admits a second-order Taylor expansion at the point (x_0,y_0).
- Be careful not to forget the factor $\frac{1}{2}$ in front of the second-degree terms, nor the factor 2 in front of the hk term.
- Recall that if we choose the Euclidean norm, then $\|(h,k)\|^2 = h^2 + k^2$, and $o\left(\|(h,k)\|^2\right)$ denotes a function equal to $\|(h,k)\|^2\varepsilon(h,k)$ with $\varepsilon(h,k) \xrightarrow[(h,k)\to(0,0)]{} 0$.

▪ Example 4.18

Let us study $f(x,y) = \exp(xy^2 - 2)$ around $(x_0,y_0) = (2,1)$.

- $f(2,1) = 1$.
- **First-order Taylor expansion.**

$$\frac{\partial f}{\partial x}(x,y) = y^2\exp(xy^2-2) \quad \frac{\partial f}{\partial y}(x,y) = 2xy\exp(xy^2-2) \quad \text{so} \quad \frac{\partial f}{\partial x}(2,1) = 1 \quad \frac{\partial f}{\partial y}(2,1) = 4.$$

  Thus:

$$f(2+h,1+k) = 1 + h + 4k + o\left(\|(h,k)\|\right).$$

- **Second-order Taylor expansion.**

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x,y) &= y^4\exp(xy^2-2), \\ \frac{\partial^2 f}{\partial x \partial y}(x,y) &= (2y+2xy^3)\exp(xy^2-2), \\ \frac{\partial^2 f}{\partial y^2}(x,y) &= (2x+4x^2y^2)\exp(xy^2-2). \end{aligned}$$

Therefore:

$$\frac{\partial^2 f}{\partial x^2}(2,1) = 1, \quad \frac{\partial^2 f}{\partial x \partial y}(2,1) = 6, \quad \frac{\partial^2 f}{\partial y^2}(2,1) = 20.$$

Thus:

$$f(2+h,1+k) \; = \; 1 \quad + \quad h+4k \quad + \quad \frac{1}{2}\left(h^2+2\cdot 6\cdot hk+20k^2\right) \quad + \quad o\left(\|(h,k)\|^2\right).$$

■

Geometric Interpretation. A first-order Taylor expansion corresponds to approximating the graph of f by a tangent plane with equation $z = a+bx+cy$. A second-order Taylor expansion corresponds to an approximation by a quadratic surface, i.e., a surface defined by a second-degree equation:

$$z = a+bx+cy+dx^2+2exy+fy^2.$$

This approximation is better but remains valid only around the considered point (x_0,y_0).

Now, we are going to express the Taylor expansion of a function with two variables in terms of vectors and matrices, which provides a compact and powerful way to understand the local behavior of functions.

The Jacobian matrix, in this context, is a row vector:

$$J_f(x_0,y_0) \times \begin{pmatrix} h \\ k \end{pmatrix} = h\frac{\partial f}{\partial x}(x_0,y_0) + k\frac{\partial f}{\partial y}(x_0,y_0).$$

Remark: We could alternatively use the gradient of the function, which is a column vector:

$$\mathrm{d}f(x_0,y_0)(h,k) = J_f(x_0,y_0) \times \begin{pmatrix} h \\ k \end{pmatrix} = \langle \nabla f(x_0,y_0) \mid \begin{pmatrix} h \\ k \end{pmatrix} \rangle.$$

We can verify that the second-degree terms are expressed using the Hessian matrix:

$$\frac{1}{2}(h,k)\; H_f(x_0,y_0) \begin{pmatrix} h \\ k \end{pmatrix} = \frac{1}{2}\left[h^2\frac{\partial^2 f}{\partial x^2}(x_0,y_0) + 2hk\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) + k^2\frac{\partial^2 f}{\partial y^2}(x_0,y_0)\right],$$

where $(h,k) = \begin{pmatrix} h \\ k \end{pmatrix}^\top$ is a row vector.

More generally, for a function $f : \mathbb{R}^n \to \mathbb{R}$, if we let $x_0 = (x_1,\ldots,x_n)$ and $h = (h_1,\ldots,h_n)$ (considered as column vectors), the Taylor expansion can be written as:

$$f(x_0+h) = f(x_0) + J_f(x_0)h + \frac{1}{2}h^\top H_f(x_0)h + o(\|h\|^2),$$

where $h^\top$ denotes the transpose of the vector h.

This formula is often encountered in the following form:

$$f(x) = f(x_0) + J_f(x_0)(x-x_0) + \frac{1}{2}(x-x_0)^\top H_f(x_0)(x-x_0) + o(\|x-x_0\|^2),$$

where we perform the change of variables $x = x_0+h$ (so $h = x-x_0$).

■ **Example 4.19**

Let $f(x,y,z) = ye^{x^2} + xz^2$. We compute the second-order Taylor expansion at the point $(x_0,y_0,z_0) = (1,2,3)$, where:

$$f(x_0,y_0,z_0) = 2e+9.$$

The Jacobian is:

$$J_f(x,y,z) = \left(2xye^{x^2} + z^2,\ e^{x^2},\ 2xz\right),$$

so at the point $(1,2,3)$:

$$J_f(x_0,y_0,z_0) = (4e+9,\ e,\ 6).$$

The Hessian matrix is:

$$H_f(x,y,z) = \begin{pmatrix} (4x^2y+2y)e^{x^2} & 2xe^{x^2} & 2z \\ 2xe^{x^2} & 0 & 0 \\ 2z & 0 & 2x \end{pmatrix},$$

which evaluates at $(1,2,3)$ to:

$$H_f(x_0,y_0,z_0) = \begin{pmatrix} 12e & 2e & 6 \\ 2e & 0 & 0 \\ 6 & 0 & 2 \end{pmatrix}.$$

Thus, the second-order Taylor expansion becomes:

$$\begin{aligned} f(x_0+h_x, y_0+h_y, z_0+h_z) = 2e+9 & \\ &+ (4e+9)h_x + eh_y + 6h_z \\ &+ \frac{1}{2}\left[12eh_x^2 + 4eh_xh_y + 12h_xh_z + 2h_z^2\right] \\ &+ o(h_x^2+h_y^2+h_z^2). \end{aligned}$$

■

Exercise 4.11
Let $f(x,y) = x^3 + 2xy - y^2$. Compute the Taylor expansion of f up to order 1 at the point $(x_0,y_0) = (1,2)$. Deduce the equation of the tangent plane to the graph of f at (x_0,y_0).

Compute the Taylor expansion of f up to order 2 at the point (x_0,y_0). Deduce the equation of the quadratic surface that best approximates the graph of f at (x_0,y_0).

■

4.8 Exercises

Exercise 4.12
Justify that the following functions are differentiable and calculate their Jacobian matrices.

1. $f(x,y,z) = \left(\frac{1}{2}(x^2-z^2), \sin x \sin y\right).$
2. $f(x,y) = \left(xy, \frac{1}{2}x^2 + y, \ln(1+x^2)\right).$

■

Exercise 4.13
Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$\begin{aligned} f(x,y) &= \frac{x^2y+3y^3}{x^2+y^2} \text{ for } (x,y) \neq (0,0), \\ f(0,0) &= 0. \end{aligned}$$

1. Is the function f continuous at $(0,0)$? Justify the answer.
2. Does the function f have partial derivatives with respect to x and y at $(0,0)$? Provide the value(s) if applicable and justify the answer.
3. Is the function f differentiable at $(0,0)$? Justify the answer.

4. Determine the partial derivatives of f at a point $(x_0,y_0) \neq (0,0)$.
5. Determine the equation of the tangent plane to the graph of f at the point $(1,1,2)$.
6. Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be the function defined by $F(x,y) = (f(x,y), f(y,x))$. Determine the Jacobian matrix of F at the point $(1,1)$. Does the function F have a local inverse in the vicinity of the point $(2,2)$?

■

Exercise 4.14

Consider the functions $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ and $g\colon \mathbb{R}^3 \longrightarrow \mathbb{R}$ defined by

$$f(x,y) = (\sin(xy), y\cos x, xy\sin(xy)\exp(y^2)), \quad g(u,v,w) = uvw.$$

1. Explicitly calculate $g \circ f$.
2. Using the expression found in (1), compute the partial derivatives of $g \circ f$.
3. Determine the Jacobian matrices $J_f(x,y)$ and $J_g(u,v,w)$ for f and g.
4. Verify the result in (2) using an appropriate matrix product of Jacobian matrices.

■

Exercise 4.15

Compute the second-order partial derivatives of the following functions:

1. $f(x,y) = e^{xy}$ 2. $f(x,y) = x^2(x+y)$

■

Exercise 4.16

For $(x,y) \neq (0,0)$, define

$$f(x,y) = xy\frac{x^2-y^2}{x^2+y^2}.$$

1. Does f admit a continuous extension to $\mathbb{R}^2$?
2. Does f admit a C^1 extension to $\mathbb{R}^2$?
3. Does f admit a C^2 extension to $\mathbb{R}^2$?

■

Exercise 4.17

Let f be a C^1 function from $\mathbb{R}^2$ to $\mathbb{R}$, and let $r \in \mathbb{R}$. We say that f is homogeneous of degree r if

$$\forall (x,y) \in \mathbb{R}^2,\ \forall t > 0,\ f(tx,ty) = t^r f(x,y).$$

1. Show that if f is homogeneous of degree r, then its partial derivatives are homogeneous of degree $r-1$.
2. Show that f is homogeneous of degree r if and only if:
$$\forall (x,y) \in \mathbb{R}^2,\ x\frac{\partial f}{\partial x}(x,y) + y\frac{\partial f}{\partial y}(x,y) = rf(x,y).$$
3. Assume that f is of class C^2. Show that:
$$x^2\frac{\partial^2 f}{\partial x^2}(x,y) + 2xy\frac{\partial^2 f}{\partial x \partial y}(x,y) + y^2\frac{\partial^2 f}{\partial y^2}(x,y) = r(r-1)f(x,y).$$

■

Exercise 4.18

A function $f : U \to \mathbb{R}$ of class C^2, defined on an open set U of $\mathbb{R}^2$, is called harmonic if its Laplacian is zero, i.e., if

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0.$$

In what follows, let f be a harmonic function.

1. Assume that f is of class C^3. Show that $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, and $x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y}$ are harmonic.
2. Now assume that f is defined on $\mathbb{R}^2 \setminus \{(0,0)\}$ and is radial, i.e., there exists a C^2 function $\varphi : \mathbb{R}^* \to \mathbb{R}$ such that $f(x,y) = \varphi(x^2+y^2)$.
3. Show that φ' is a solution to a first-order linear differential equation. Deduce all radial harmonic functions.

■

Exercise 4.19

Compute the following differentials without calculating partial derivatives, using the properties of differentials of sums, products, and compositions:

$$\text{(a) } \mathrm{d}\left(\ln(xy)\right) \quad \text{(b) } \mathrm{d}\left(xyz(1+\sinh(yz))\right) \quad \text{(c) } \mathrm{d}\left(\sin(x^2y)e^{x-y}\right)$$

■

Exercise 4.20

Compute the Hessian matrices of the functions f defined by the following expressions on their natural domain of definition:

$$\sin(xyz), \quad \sin^2(y/x).$$

■

Exercise 4.21

Let $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ be a C^2 function, and let r and θ be the standard polar coordinates in the plane such that the mapping

$$]0,+\infty[\times]0,2\pi[\to \mathbb{R}^2 \setminus \{(0,0)\}, \quad (r,\theta) \mapsto (x,y) = (r\cos\theta, r\sin\theta),$$

is a change of variables. Let F be the function defined by

$$F(r,\theta) = f(r\cos\theta, r\sin\theta).$$

This is the "expression of f in polar coordinates." Show that

$$\frac{\partial^2 f}{\partial x^2}(x,y) + \frac{\partial^2 f}{\partial y^2}(x,y) = \frac{\partial^2 F}{\partial r^2}(r,\theta) + \frac{1}{r}\frac{\partial F}{\partial r}(r,\theta) + \frac{1}{r^2}\frac{\partial^2 F}{\partial \theta^2}(r,\theta). \tag{4.2}$$

■

5. Extremum

In this chapter, we explore the methods and concepts used to determine the extrema of a function — that is, its maximum or minimum values. Understanding extrema is a fundamental aspect of mathematical optimization, with applications across physics, economics, machine learning, and more.

We begin by investigating the conditions under which a function has local extrema, which are points where the function reaches a maximum or minimum within a small neighborhood. It is important to note that local extrema do not necessarily correspond to the highest or lowest values of the function globally.

Throughout this chapter, we will:

- Learn the necessary and sufficient conditions for local extrema, using tools like the gradient and the Hessian matrix.
- Study how critical points arise, and understand their classification as local maxima, minima, or saddle points through second-order conditions.
- Extend these ideas to constrained optimization, where the function's extrema are restricted by one or more constraints. We will introduce powerful techniques like the method of Lagrange multipliers to handle these scenarios.

To build a solid foundation, we will first revisit the case of functions of a single variable. This simpler context will help us intuitively grasp the behavior of functions and their critical points, before generalizing to higher dimensions.

By the end of the chapter, you will have a thorough understanding of how to identify and classify extrema, both with and without constraints, in multidimensional settings.

5.1 Minimum, Maximum, Critical Point

We first recall the extremum in one variable function. Let $f : \mathbb{R} \to \mathbb{R}$ be a function of one variable.

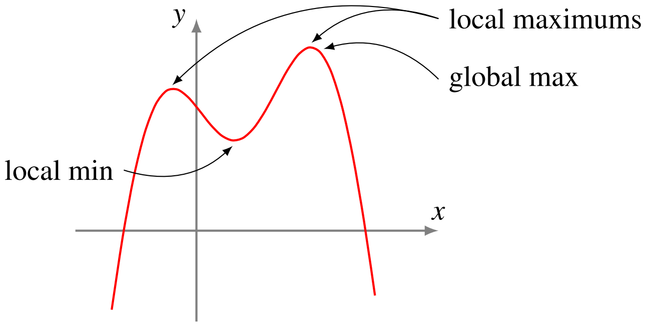
- f has a **local maximum** at $x_0 \in \mathbb{R}$ if there exists an open interval I containing x_0 such that:

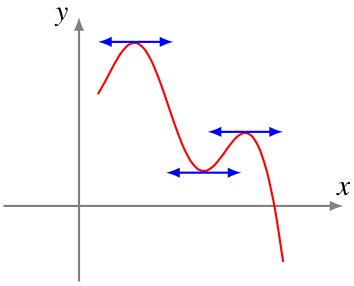
$$\text{for all } x \in I, \quad f(x) \leq f(x_0).$$

- f has a **local minimum** at $x_0 \in \mathbb{R}$ if there exists an open interval I containing x_0 such that:

$$\text{for all } x \in I, \quad f(x) \geq f(x_0).$$

- f has a **local extremum** at $x_0 \in \mathbb{R}$ if f has a local maximum or a local minimum at this point.
- f has a **critical point** at $x_0 \in \mathbb{R}$ if $f'(x_0) = 0$. Geometrically, this is a point with a horizontal tangent.
- Proposition: If f is differentiable and has a local minimum or local maximum at x_0, then $f'(x_0) = 0$. In other words, if x_0 is a local extremum, then it is a critical point.
- The converse is not always true. For example, for $f : x \mapsto x^3$, the point $x_0 = 0$ is a critical point, but it is neither a local maximum nor a local minimum (it is a **point of inflection**).





On the left figure: examples of a local minimum, local maximum, and global maximum; there is no global minimum on $\mathbb{R}$. On the right figure: a local extremum is necessarily a critical point.

5.1.1 Characterization of Minima and Maxima

The practical search for local extrema for a function of one variable proceeds as follows:

1. Find the critical points given by the equation $f'(x) = 0$.
2. For each critical point x_0, compute the second derivative:
    - If $f''(x_0) > 0$, then f has a local minimum at x_0.
    - If $f''(x_0) < 0$, then f has a local maximum at x_0.
    - If $f''(x_0) = 0$, further study is required.

When $f : [a,b] \to \mathbb{R}$ is defined on a compact interval, we must also study the behavior of f at a and b (i.e., at the boundary of the domain). Since the domain is compact, we are guaranteed the existence of global extrema.

5.2 Second-Order Partial Derivatives

5.2.1 Hessian Matrix

The Jacobian matrix is the matrix of partial derivatives. The Hessian matrix is the matrix of second-order partial derivatives.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of n variables. The **Hessian matrix** of f at $x = (x_1, \dots, x_n)$ is the $n \times n$ matrix:

$$H_f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{1 \le i,j \le n}$$

For a function of class $\mathscr{C}^2$, by Schwarz's theorem, this is a **symmetric matrix**.

In the case of a function of two variables:

$$H_f(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{pmatrix}$$

For three variables, the Hessian matrix (evaluated at (x,y,z)) is:

$$H_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix}.$$

■ Example 5.1

Let us compute the Hessian matrix of $f(x,y) = xy^2 + x^4 - y^4$.

First, compute

$$\frac{\partial f}{\partial x}(x,y) = 4x^3 + y^2 \qquad \frac{\partial f}{\partial y}(x,y) = 2xy - 4y^3.$$

Thus,

$$H_f(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{pmatrix} = \begin{pmatrix} 12x^2 & 2y \\ 2y & 2x - 12y^2 \end{pmatrix}.$$

■

Exercise 5.1

1. Let $f(x,y) = x^3 + 5x^2y - y^2$. Compute the first-order partial derivatives of f. Compute the second-order partial derivatives of f. Verify the validity of Schwarz's theorem. Compute the Hessian matrix of f. Compute the third-order partial derivatives of f.
2. Let $f(x,y) = xe^{x^2-y^2}$. Compute the first-order and second-order partial derivatives of f. Compute the Hessian matrix of f.
3. Let $f(x,y,z) = xy^2 \ln(z)$. Determine the domain of definition of f. Compute the first-order and second-order partial derivatives as well as the Hessian matrix of f.

■

5.3 Second-Order Differential

This section is highly theoretical and may be skipped on a first reading.

5.3.1 Second-Order Differential

Let $f : U \to \mathbb{R}$ be a function defined on an open set $U \subset \mathbb{R}^n$.

Definition 5.3.1

We say that f is **twice differentiable at** $x_0 \in U$:

- if it is differentiable in an open neighborhood V_{x_0} of x_0,
- and if its differential $\mathrm{d}f : V_{x_0} \to \mathscr{L}(\mathbb{R}^n, \mathbb{R})$ is differentiable at x_0.

We say that f is **twice differentiable on** U if it is twice differentiable at every point of U.

We denote by $\mathscr{L}(\mathbb{R}^n, \mathbb{R})$ the set of linear maps from $\mathbb{R}^n$ to $\mathbb{R}$. More generally, $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^p)$ refers to the set of linear maps from $\mathbb{R}^n$ to $\mathbb{R}^p$.

By definition, the differential of $\mathrm{d}f$ at point x, written $\mathrm{d}(\mathrm{d}f)(x)$, is a linear map from $\mathbb{R}^n$ to $\mathscr{L}(\mathbb{R}^n, \mathbb{R})$. In other words:

$$\mathrm{d}f : U \to \mathscr{L}(\mathbb{R}^n, \mathbb{R}) \quad \text{and} \quad \mathrm{d}(\mathrm{d}f) : U \to \mathscr{L}(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n, \mathbb{R})).$$

However, it can naturally be identified with a bilinear map from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$ thanks to the following proposition:

$$\mathscr{L}(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n, \mathbb{R})) \simeq \mathscr{L}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}).$$

The identification is as follows: If $L \in \mathscr{L}(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n, \mathbb{R}))$, then for $h \in \mathbb{R}^n$, $L(h) \in \mathscr{L}(\mathbb{R}^n, \mathbb{R})$, and for $k \in \mathbb{R}^n$:

$$k \mapsto L(h)(k) \in \mathbb{R}.$$

The map $(h,k) \mapsto L(h)(k)$ (denoted as $L(h,k)$) is linear in both h and k, so it is bilinear in (h,k).

Definition 5.3.2
The **second-order differential** of a twice-differentiable function $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is the map:

$$\mathrm{d}^2 f : U \to \mathscr{L}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$$

defined by:

$$\mathrm{d}^2 f(x)(h,k) = \mathrm{d}(\mathrm{d}f)(x)(h)(k) \quad \text{for all } (h,k) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Let's now compute the second-order differentials for two common types of functions: affine functions and quadratic functions.

- An **affine function** $f : x \mapsto \ell(x) + b$ with $\ell \in \mathscr{L}(\mathbb{R}^n, \mathbb{R})$ and $b \in \mathbb{R}$ is twice differentiable, and its second-order differential is identically zero. This generalizes the fact that the second derivative of a linear function in one variable is zero.
- A **quadratic function** $f : x \mapsto \phi(x,x)$ with $\phi \in \mathscr{L}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ is twice differentiable, and its second-order differential is constant — and even equals 2ϕ if ϕ is symmetric. This extends the result that the second derivative of a quadratic function in one variable is constant.

5.3.2 Schwarz's Theorem — Hessian Matrix

Schwarz's theorem on the equality of mixed partial derivatives can be restated as follows:

Theorem 5.3.1 — Schwarz's Theorem.
Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathscr{C}^2$. Then, for every point $x \in U$, the second-order differential $\mathrm{d}^2 f(x)$ is a **symmetric bilinear map**. In other words, for all $(h,k) \in \mathbb{R}^n \times \mathbb{R}^n$:

$$\mathrm{d}^2 f(x)(h,k) = \mathrm{d}^2 f(x)(k,h).$$

Let $f : U \to \mathbb{R}$ be a twice-differentiable function on an open set $U \subset \mathbb{R}^n$. Let $(e_1, \dots, e_n)$ be the canonical basis of $\mathbb{R}^n$. Then, for every point $x \in U$ and for all indices $i, j \in \{1, \dots, n\}$:

$$\mathrm{d}^2 f(x)(e_i, e_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

The Hessian matrix of f at point x is the matrix of second-order partial derivatives:

$$H_f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{1 \le i,j \le n}.$$

By bilinearity, if h and k are two vectors in $\mathbb{R}^n$ with coordinates $(h_1, \dots, h_n)$ and $(k_1, \dots, k_n)$ (considered as column vectors), we have:

$$\mathrm{d}^2 f(x)(h,k) = h^\top \cdot H_f(x) \cdot k = \sum_{i=1}^{n} \sum_{j=1}^{n} h_i k_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

In other words, the Hessian matrix $H_f(x)$ represents the bilinear form $\mathrm{d}^2 f(x)$ with respect to the canonical basis of $\mathbb{R}^n$. Moreover, Schwarz's theorem ensures that the Hessian matrix is symmetric if f is of class $\mathscr{C}^2$ on U.

5.3.3 Taylor Expansion

Let $f : U \to \mathbb{R}$ be a differentiable function on an open set $U \subset \mathbb{R}^n$. The first-order Taylor formula is:

$$f(x+h) = f(x) + \mathrm{d}f(x)(h) + o(\|h\|),$$

where we recall that $\mathrm{d}f(x)(h) = J_f(x) \cdot h$.

For a twice-differentiable function, the **second-order Taylor formula** (or second-order approximation) is given by:

$$f(x+h) = f(x) + \mathrm{d}f(x)(h) + \frac{1}{2}\mathrm{d}^2 f(x)(h,h) + o(\|h\|^2).$$

More generally, by iterating the differentials, we get:

Theorem 5.3.2 — Taylor Formula of Order p.
If U is an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ is p-times differentiable at $x \in U$, then:

$$f(x+h) = f(x) + \sum_{k=1}^{p} \frac{1}{k!}\mathrm{d}^k f(x)(h^{[k]}) + o(\|h\|^p),$$

where we denote $h^{[k]} = (h, h, \dots, h) \in (\mathbb{R}^n)^k$.

5.4 Minimum and Maximum: The Case of Two Variables

5.4.1 Local Maximum and Minimum

Let $f : U \to \mathbb{R}$ be a function of two variables, where U is an open subset of $\mathbb{R}^2$.

Definition 5.4.1
We say that f has a **local maximum** (resp. **local minimum**) at $(x_0, y_0) \in U$ if there exists an open disk $D \subset U$, centered at (x_0, y_0), such that:

$$\forall (x,y) \in D, \quad f(x,y) \leq f(x_0,y_0) \quad (\text{resp. } f(x,y) \geq f(x_0,y_0)).$$

We say that f has a **local extremum** at (x_0, y_0) if it has either a local maximum or a local minimum at that point.

5.4.2 Critical Points

Assume f is of class $\mathscr{C}^2$ on an open set U, meaning its partial derivatives up to order 2 exist and are continuous.

Proposition 5.4.1
If f has a local extremum at (x_0, y_0) in the open set U, then:

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad \text{and} \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

Proof. The function of one variable $x \mapsto f(x, y_0)$ has a local extremum at x_0 in an open interval of $\mathbb{R}$. Therefore, its derivative, which is precisely $\frac{\partial f}{\partial x}(x, y_0)$, must vanish at x_0. The same argument applies to the function $y \mapsto f(x_0, y)$. ■

In other words, if f has a local minimum or maximum at a point, then its gradient vanishes at that point. The points in U where the gradient vanishes are called **critical points** of f. The previous result tells us that the local extrema of a function can only occur at critical points, although the converse is false.

By definition, a critical point that is neither a local maximum nor a local minimum is called a **saddle point**.

5.4.3 Fundamental Examples

■ Example 5.2
The function $f(x,y) = x^2 + y^2$ is an example of a local minimum at the point $(0,0)$. ■

■ Example 5.3
The function $f(x,y) = x^2 - y^2$ is an example of a saddle point at $(0,0)$. The function decreases in the y-direction but increases in the x-direction. ■

5.4.4 Characterizing Minimum and Maximum Points

For a function $f : \mathbb{R}^2 \to \mathbb{R}$, we use Monge's notation, which provides a simple criterion to detect local minima or maxima.

Theorem 5.4.2 — Monge's Criterion.
Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function of class $\mathscr{C}^2$, and let (x_0, y_0) be a critical point of f. Define:

$$r = \frac{\partial^2 f}{\partial x^2}(x_0, y_0), \quad s = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0), \quad t = \frac{\partial^2 f}{\partial y^2}(x_0, y_0).$$

Then:

- If $rt - s^2 > 0$ and $r > 0$, then (x_0, y_0) is a local minimum of f.
- If $rt - s^2 > 0$ and $r < 0$, then (x_0, y_0) is a local maximum of f.
- If $rt - s^2 < 0$, then (x_0, y_0) is neither a local minimum nor a local maximum: it is a saddle point.
- If $rt - s^2 = 0$, we cannot conclude directly; a deeper analysis is required.

Note: The quantity $rt - s^2$ is the determinant of the Hessian matrix at (x_0, y_0):

$$H_f(x_0, y_0) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}.$$

■ Example 5.4
Let's revisit the three fundamental examples to verify this criterion.

1. Let $f(x,y) = x^2 + y^2$. The unique critical point is $(0,0)$. Compute the second derivatives:

$$r = \frac{\partial^2 f}{\partial x^2}(0,0) = 2, \quad s = \frac{\partial^2 f}{\partial x \partial y}(0,0) = 0, \quad t = \frac{\partial^2 f}{\partial y^2}(0,0) = 2.$$

   Thus, $rt - s^2 = 4$ with $r > 0$, so $(0,0)$ is indeed a local minimum of f.
2. **Exercise:** Do the same analysis for $f(x,y) = -x^2 - y^2$ to check for a local maximum.
3. Let $f(x,y) = x^2 - y^2$. The only critical point is $(0,0)$. Compute the Hessian matrix:

$$H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

   Here, $r = 2$, $s = 0$, $t = -2$. Therefore, $rt - s^2 = -4 < 0$, confirming that $(0,0)$ is a saddle point.

■

5.4.5 Other Examples

■ Example 5.5
Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by:

$$f(x,y) = x^3 + y^3 - 3xy.$$

- **First partial derivatives:**

$$\frac{\partial f}{\partial x}(x,y) = 3x^2 - 3y, \quad \frac{\partial f}{\partial y}(x,y) = 3y^2 - 3x.$$

- **Critical points:** The critical points are the solutions to:

$$3x^2 - 3y = 0, \quad 3y^2 - 3x = 0.$$

  This gives the system: $x^2 = y$ and $y^2 = x$, implying $x, y \geq 0$. Solving this yields the points: $(0,0)$ and $(1,1)$.

- **Second partial derivatives:**

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 6x, \quad \frac{\partial^2 f}{\partial x \partial y}(x,y) = -3, \quad \frac{\partial^2 f}{\partial y^2}(x,y) = 6y.$$

- **Analysis at** $(0,0)$**:** The Hessian matrix at $(0,0)$ is:

$$H_f(0,0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}.$$

  Thus, $r = 0$, $s = -3$, $t = 0$, and:

$$rt - s^2 = -9 < 0.$$

  So $(0,0)$ is a saddle point.

- **Analysis at** $(1,1)$**:** The Hessian matrix at $(1,1)$ is:

$$H_f(1,1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}.$$

  Here, $r = 6$, $s = -3$, $t = 6$, and:

$$rt - s^2 = 36 - 9 = 27 > 0.$$

  Since $r > 0$, $(1,1)$ is a local minimum (not a global minimum).

■

Here's an example where the criterion fails to conclude, and we need to continue the analysis manually.

■ Example 5.6

Let $f(x,y) = 2x^3 - y^4 - 3x^2$. We find two critical points: $(0,0)$ and $(1,0)$. The Hessian matrices are:

$$H_f(0,0) = \begin{pmatrix} -6 & 0 \\ 0 & 0 \end{pmatrix}, \quad H_f(1,0) = \begin{pmatrix} 6 & 0 \\ 0 & 0 \end{pmatrix}.$$

Since the determinant $rt - s^2 = 0$, we cannot conclude directly. Let's study each case manually.

- At $(0,0)$: Write $f(x,y) = x^2(2x-3) - y^4$. For $|x| \leq 1$, $2x - 3 \leq 0$, so:

$$f(x,y) \leq 0.$$

  Since $f(0,0) = 0$, there is a local maximum at $(0,0)$.

- At $(1,0)$: For points $(1,y)$:

$$f(1,y) = -1 - y^4 \leq -1 = f(1,0).$$

  For points $(x,0)$ near $(1,0)$:

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

  So $(1,0)$ is a saddle point.

■

5.4.6 Constrained Extrema in the Plane

Let's start with functions of two variables.

Problem: Find the maximum or minimum of a function $f(x,y)$ subject to the constraint $g(x,y) = 0$.

Example: Find the minimum of $f(x,y) = x^2 + 2y^2$ with the constraint $y = 2x - 1$ (i.e., $g(x,y) = y - 2x + 1$).

Clearly, the global minimum of f is 0, reached at $(0,0)$, but this point does not satisfy the constraint $g(x,y) = 0$.

Geometrically, we want to find the minimum of $f(x,y)$ restricted to points (x,y) that lie on the line $y = 2x - 1$.

Finding the minimum of f under the constraint means identifying the point on the curve with the lowest altitude.

Theorem 5.4.3 — Constrained Extrema — Two Variables.
Let $f, g : U \to \mathbb{R}$ be functions of class $\mathscr{C}^1$ on an open set $U \subset \mathbb{R}^2$. Suppose $(x_0, y_0) \in U$ is a constrained extremum of f subject to the condition $g(x,y) = 0$ and that $\nabla g(x_0, y_0) \neq (0,0)$. Then there exists a real number λ (called the **Lagrange multiplier**) such that:

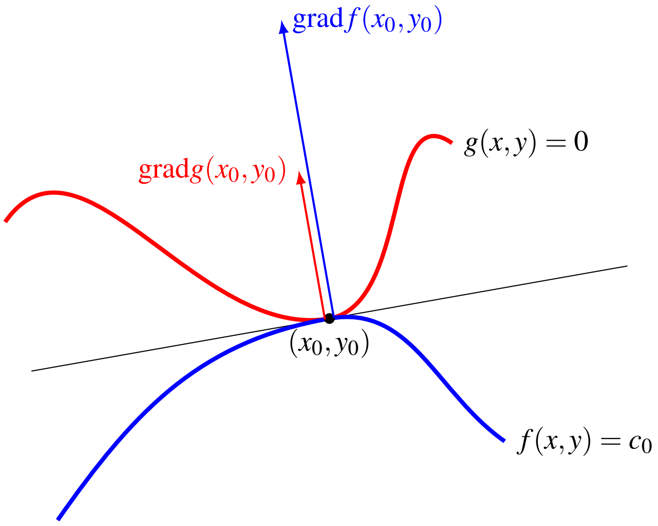
$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0).$$

In other words, the system becomes:

$$\begin{cases} g(x_0, y_0) = 0 \\ \dfrac{\partial f}{\partial x}(x_0, y_0) = \lambda \dfrac{\partial g}{\partial x}(x_0, y_0) \\ \dfrac{\partial f}{\partial y}(x_0, y_0) = \lambda \dfrac{\partial g}{\partial y}(x_0, y_0). \end{cases}$$

Note: We can also consider the constraint $g(x,y) = c$ for a constant c. This reduces to the previous case by defining $g(x,y) - c = 0$.

Geometric Interpretation: At a constrained extremum (x_0, y_0), the gradients of f and g are collinear. This means that the level curve $f(x,y) = c_0$ (with $c_0 = f(x_0, y_0)$) and the constraint curve $g(x,y) = 0$ share the same tangent line at (x_0, y_0). In other words, the curves are tangent to each other at the extremum.



Remark on the Equations: In dimension n, the constrained extremum theorem leads to a system of n equations with n unknowns. However, these equations are usually nonlinear and can be difficult (or impossible) to solve, even in simple cases. That's why we focus on relatively simple functions that allow for explicit solutions.

■ **Example 5.7**

Let's revisit the previous example: find the minimum of $f(x,y) = x^2 + 2y^2$ subject to the constraint $y = 2x - 1$.

Geometrically, we want the point on the line $y = 2x - 1$ that minimizes f.

Set $g(x,y) = y - 2x + 1$. Compute the gradients:

$$\nabla f(x,y) = \begin{pmatrix} 2x \\ 4y \end{pmatrix}, \quad \nabla g(x,y) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

At a constrained extremum, the gradients are collinear: $\nabla f(x,y) = \lambda \nabla g(x,y)$. This yields two equations:

$$2x = -2\lambda \tag{5.1}$$

$$4y = \lambda \tag{5.2}$$

along with the constraint:

$$y = 2x - 1. \tag{5.3}$$

From (5.1), we get $x = -\lambda$. From (5.2), we get $y = \frac{\lambda}{4}$. Substitute into the constraint (5.3):

$$\frac{\lambda}{4} = -2\lambda - 1.$$

Solving this, we find:

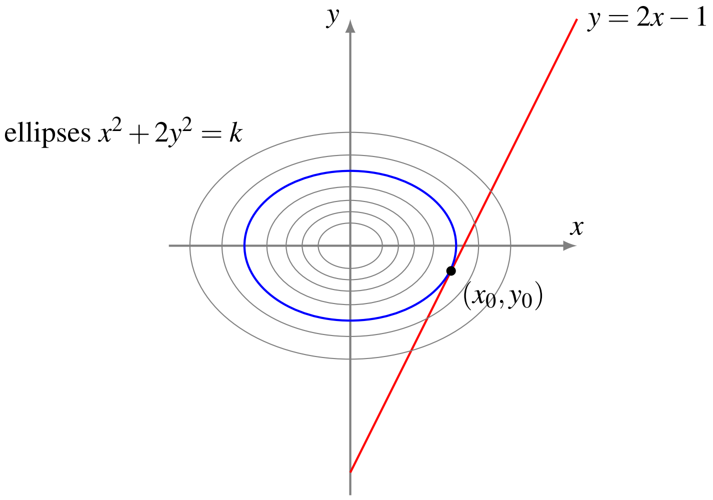
$$\lambda_0 = -\frac{4}{9}.$$

Thus, the constrained minimum is at:

$$(x_0, y_0) = \left(\frac{4}{9}, -\frac{1}{9}\right).$$

Since $f(x,y) \to \infty$ as $\|(x,y)\| \to \infty$ on the constraint line, f must attain a minimum. Conclusion: f has a unique minimum on $g(x,y) = 0$, reached at (x_0, y_0).

Geometrically, the level curves of $f(x,y)$ are ellipses. The minimum occurs where the smallest ellipse is tangent to the constraint line.



■

■ Example 5.8
What is the point on the parabola defined by $y = -x^2$ that is closest to the point $(2,1)$?

We want to minimize the distance between a point on the parabola and $(2,1)$. We define the objective function as the squared distance:

$$f(x,y) = (x-2)^2 + (y-1)^2.$$

The constraint is given by the parabola's equation:

$$g(x,y) = y + x^2 = 0.$$

Compute the gradients:

$$\nabla f(x,y) = \begin{pmatrix} 2(x-2) \\ 2(y-1) \end{pmatrix}, \quad \nabla g(x,y) = \begin{pmatrix} 2x \\ 1 \end{pmatrix}.$$

At a constrained extremum, the gradients must be collinear: $\nabla f(x,y) = \lambda \nabla g(x,y)$. This gives the system:

$$2(x-2) = 2\lambda x \tag{5.1}$$
$$2(y-1) = \lambda \tag{5.2}$$

along with the constraint:

$$y + x^2 = 0. \tag{5.3}$$

From equation (5.1), we get:

$$x = \frac{2}{1-\lambda}.$$

From equation (5.2), we find:

$$y = \frac{\lambda}{2} + 1.$$

Substitute these into the constraint equation (5.3):

$$\frac{\lambda}{2} + 1 + \frac{4}{(1-\lambda)^2} = 0.$$

Rearranging, we get a cubic equation:

$$(\lambda+2)(\lambda-1)^2 + 8 = 0.$$

Expand and simplify:

$$\lambda^3 - 3\lambda + 10 = 0.$$

This cubic equation has only one real solution (not easy to find by hand):

$$\lambda_0 = \frac{-1 - \left(\sqrt[3]{5-2\sqrt{6}}\right)^2}{\sqrt[3]{5-2\sqrt{6}}}.$$

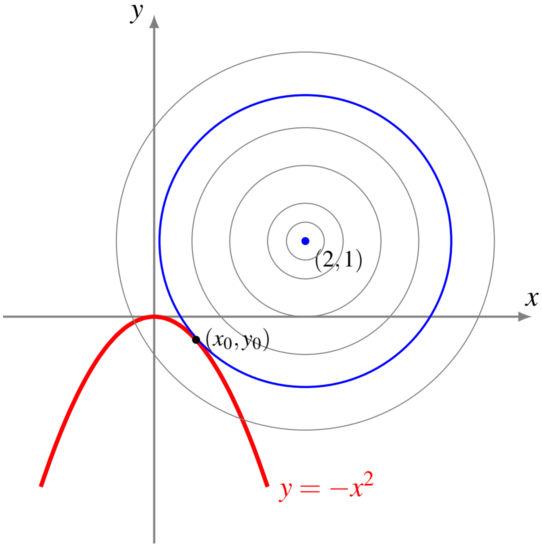
Thus, the point on the parabola closest to $(2,1)$ is:

$$(x_0, y_0) = \left(\frac{2}{1-\lambda_0}, \frac{\lambda_0}{2} + 1\right).$$

Numerical approximations give:

$$\lambda_0 \approx -2.613, \quad (x_0, y_0) \approx (0.554, -0.306).$$

Geometric Interpretation: We grow a circle centered at $(2,1)$ until it just touches the parabola. The contact point is the point on the parabola closest to $(2,1)$. If we kept enlarging the circle, we might encounter other constrained extrema (though not in this case).



■

5.5 Implicit function theorem

5.5.1 Definition

Bijective functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ that are of class $\mathscr{C}^1$ (continuously differentiable) and whose inverse is also of class $\mathscr{C}^1$ are used as variable changes. These functions are called **diffeomorphisms**.

Definition 5.5.1 Let U and V be open subsets of $\mathbb{R}^n$, and let $\Phi : U \to V$. We say that Φ is a **diffeomorphism** if:

1. Φ is a bijection from U onto V,
2. Φ is of class $\mathscr{C}^1$ on U, meaning it is differentiable on U with a continuous differential,
3. and Φ^{-1} is of class $\mathscr{C}^1$ on V.

■ Example 5.9

1. The function $f : (0,+\infty) \to (0,+\infty)$ defined by $f(x) = x^3$ is a diffeomorphism. Its inverse is the function $f^{-1} : \mathbb{R} \to \mathbb{R}$ given by:

   $$f^{-1}(x) = x^{\frac{1}{3}}.$$

   Both f and f^{-1} are of class $\mathscr{C}^1$.
2. A fundamental example that we will explore in this chapter is the transition to polar coordinates. This replaces Cartesian coordinates (x,y) with polar coordinates (r,θ). To obtain a diffeomorphism, we must exclude certain parts of the plane. Let:

   $$U = \{(r,\theta) \in \mathbb{R}^2 \mid r > 0 \text{ and } 0 < \theta < 2\pi\}, \quad V = \mathbb{R}^2 \setminus (\mathbb{R}_+ \times \{0\}).$$

   Then, the function $\Phi : U \to V$ defined by:

   $$\Phi(r,\theta) = (r\cos\theta, r\sin\theta)$$

is a diffeomorphism. Finding the inverse amounts to expressing the polar coordinates (r,θ) in terms of the Cartesian coordinates (x,y).

■

From the theorem giving the Jacobian matrix of a composition, it follows that the Jacobian matrices of Φ and Φ^{-1} are inverses of each other.

Let U and V be two open subsets of $\mathbb{R}^n$, and let $\Phi : U \to V$. Let $x \in U$ and $y = \Phi(x) \in V$. Let $J_\Phi(x)$ denote the Jacobian matrix of Φ at x, and $J_{\Phi^{-1}}(y)$ denote the Jacobian matrix of Φ^{-1} at y. These are square $n \times n$ matrices.

Proposition 5.5.1

$$J_{\Phi^{-1}}(y) = (J_\Phi(x))^{-1}.$$

As mentioned: the Jacobian matrix of the inverse is the inverse of the Jacobian matrix. Let $F : \mathbb{R}^2 \to \mathbb{R}$. Consider the level curve:

$$\mathscr{C} : \big(F(x,y) = 0\big).$$

We say that the function $y = \varphi(x)$ is **implicitly defined** by the equation $F(x,y) = 0$ if $F(x,\varphi(x)) = 0$, meaning that the point $(x,\varphi(x))$ lies on the curve $\mathscr{C}$. In this case, $y = \varphi(x)$ is called an **implicit function** of the equation $F(x,y) = 0$.

Theorem 5.5.2 — Implicit Function Theorem.
Let $F : \mathbb{R}^2 \to \mathbb{R}$ be a function of class $\mathscr{C}^1$, and let (x_0,y_0) be a point such that:

$$F(x_0,y_0) = 0.$$

If:

$$\frac{\partial F}{\partial y}(x_0,y_0) \neq 0$$

then:

1. There exists a function $\varphi : I \to J$ of class $\mathscr{C}^1$ that defines an implicit function $y = \varphi(x)$, where I is an open interval containing x_0, and J is an open interval containing y_0 with $y_0 = \varphi(x_0)$. More precisely, for all $(x,y) \in I \times J$:

$$F(x,y) = 0 \iff y = \varphi(x)$$

In particular:

$$\text{For all } x \in I, \quad F(x,\varphi(x)) = 0$$

2. Moreover, for all $x \in I$, we have $\frac{\partial F}{\partial y}(x,\varphi(x)) \neq 0$, and the derivative of φ is given by:

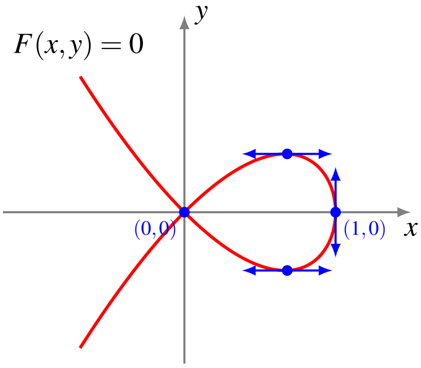
$$\varphi'(x) = \frac{-\frac{\partial F}{\partial x}(x,\varphi(x))}{\frac{\partial F}{\partial y}(x,\varphi(x))}$$

■ Example 5.10
Let

$$F(x,y) = x(x^2+y^2) - (x^2-y^2).$$

Curve $(F(x,y) = 0)$**.**

Let us begin by plotting the curve $\mathscr{C}$ defined by the equation $F(x,y) = 0$, which will help us anticipate the calculations. The curve forms a loop. We note that there is a vertical tangent at the point $(1,0)$. The point $(0,0)$ is a double point, and we will see that both partial derivatives vanish there simultaneously: it is a singular point. Outside these two points, we can locally express the curve $\mathscr{C}$ as the graph of a function of one variable $y = \varphi(x)$.



There are two points P_1 and P_2 where the tangent is horizontal. Outside the three points P_1, P_2, and $(0,0)$, we could also locally express the curve $\mathscr{C}$ as the graph of a function of one variable $x = \tilde{\varphi}(y)$.

Partial Derivatives.

We have $F(x,y) = x^3 + xy^2 - x^2 + y^2$. Therefore,

$$\frac{\partial F}{\partial x}(x,y) = 3x^2 + y^2 - 2x \qquad \text{and} \qquad \frac{\partial F}{\partial y}(x,y) = 2xy + 2y.$$

Vertical Tangency.

To find the points of vertical tangency, we first solve the equation $\frac{\partial F}{\partial y}(x,y) = 0$:

$$\frac{\partial F}{\partial y}(x,y) = 0 \iff 2y(x+1) = 0 \iff y = 0 \text{ or } x = -1.$$

Additionally, we seek a point $(x,y) \in \mathscr{C}$.

Case $y = 0$. Then, since we also require $F(x,y) = 0$, we have $x^3 - x^2 = 0$, so $x = 0$ or $x = 1$. Thus, there are two solutions: $(0,0)$ and $(1,0)$.

Case $x = -1$. We also require $F(x,y) = 0$, but the equation $F(-1,y) = 0$ is equivalent to

$$-2 + y^2 - y^2 = 0,$$

which has no solution.

Conclusion:

- $(1,0) \in \mathscr{C}$ is a point of vertical tangency: the derivative $\frac{\partial F}{\partial y}$ vanishes there, but $\frac{\partial F}{\partial x}$ does not (verify the calculation).
- $(0,0) \in \mathscr{C}$ is a singular point: both partial derivatives vanish there.

■

5.6 Exercises

Exercise 5.2

Determine the local extrema of the following functions $f : \mathbb{R}^2 \to \mathbb{R}$:

1. $f(x,y) = x^2 + xy + y^2 - 3x - 6y$
2. $f(x,y) = x^2 + 2y^2 - 2xy - 2y + 1$
3. $f(x,y) = x^3 + y^3$

■

Exercise 5.3
I) Let f be a function of class C^2 defined on $\mathbb{R}^2$ by:

$$f(x,y) = (x^2+y^2)e^{-x}$$

1. Find the local extrema of f on $\mathbb{R}^2$.
2. Show that f has a global minimum on $\mathbb{R}^2$ and does not have a global maximum.

II) Consider the function f defined on $\mathbb{R}^2$ by:

$$f(x,y) = x^2 - \cos(y) + \frac{1}{2}xy.$$

Find the critical points of f and discuss their nature. ■

Exercise 5.4
Consider the function h defined on $]-1,+\infty[^2$ by:

$$h(x,y) = x\sqrt{y+1} + y\sqrt{x+1}.$$

Find the critical points of h and determine their nature. ■

Exercise 5.5
Consider the function f defined on $\mathbb{R}^2$ by: $f(x,y) = x^2 + \frac{1}{3}y^3 + xy$.

1. Give the Taylor expansion of f at $(0,0)$ to order 2.
2. Find the critical points of f and determine their nature.
3. Does the function f have a global minimum or maximum?

■

Exercise 5.6
For each of the following examples, prove that f has a maximum on K, and determine this maximum.

1. $f(x,y) = xy(1-x-y)$ and $K = \{(x,y) \in \mathbb{R}^2; x,y \geq 0, x+y \leq 1\}$.
2. $f(x,y) = x - y + x^3 + y^3$ and $K = [0,1] \times [0,1]$.

■

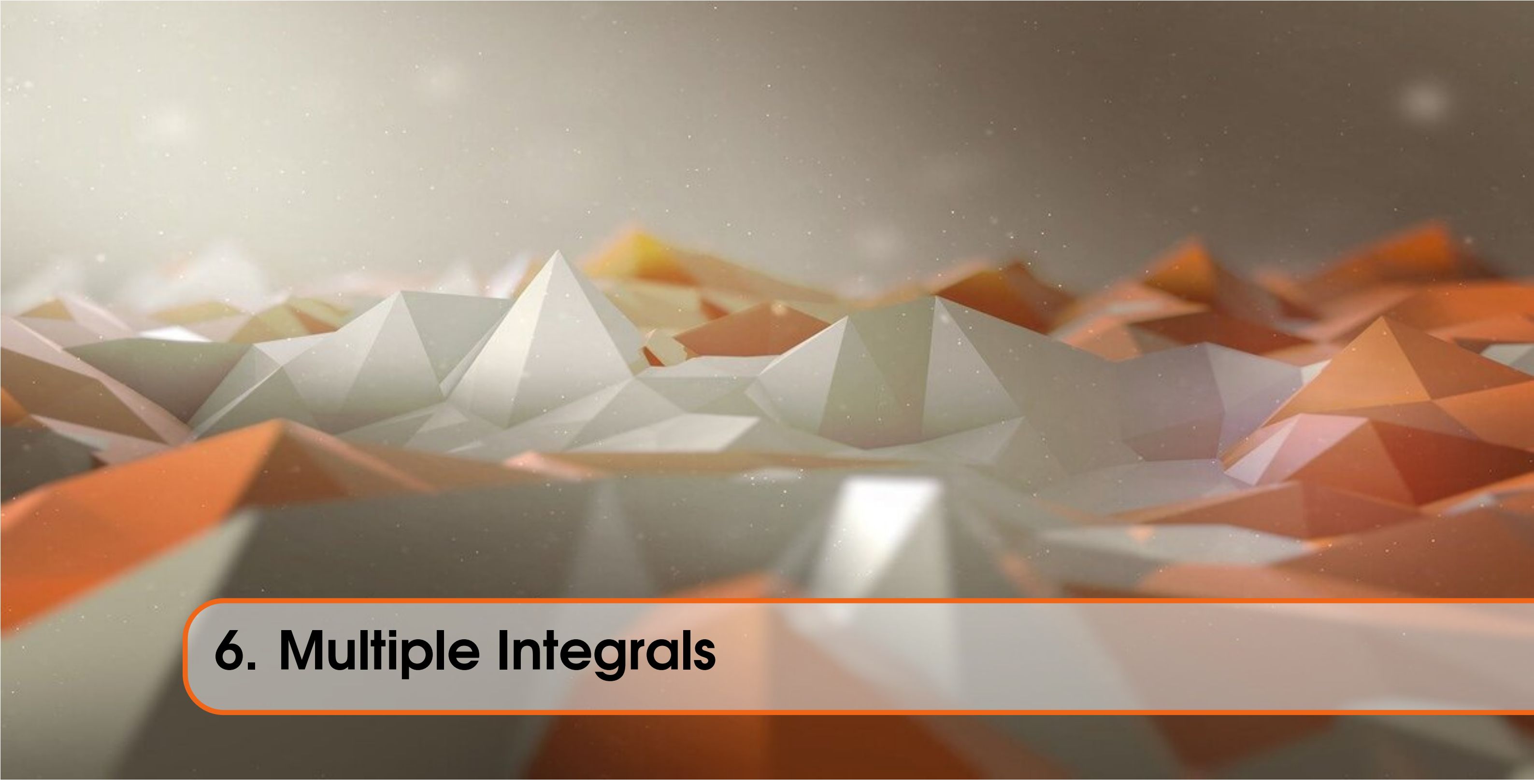
Exercise 5.7
Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function defined by $f(x,y) = (x^2+y^2)^x$ for $(x,y) \neq (0,0)$ and $f(0,0) = 1$.

1. Is the function f continuous at $(0,0)$?
2. Determine the partial derivatives of f at any point other than the origin.
3. Does the function f have partial derivatives with respect to x and y at $(0,0)$?

■

III

Multiple Integrals



6. Multiple Integrals

Multiple integrals extend the concept of integration to functions of several variables. While single integrals compute the area under a curve in one dimension, multiple integrals allow us to compute volumes, masses, and other quantities in higher dimensions. The most common types of multiple integrals are double integrals (for functions of two variables) and triple integrals (for functions of three variables).

For a function $f(x,y)$ defined over a region D in the xy-plane, the double integral is denoted by:

$$\iint_D f(x,y)\,dx\,dy.$$

This integral represents the volume under the surface $z = f(x,y)$ and above the region D. To compute a double integral, we often use iterated integrals, integrating first with respect to one variable and then the other.

For a function $f(x,y,z)$ defined over a region E in three-dimensional space, the triple integral is denoted by:

$$\iiint_E f(x,y,z)\,dx\,dy\,dz.$$

This integral can represent the mass of a solid with variable density $f(x,y,z)$. Similar to double integrals, triple integrals are often computed using iterated integration.

Multiple integrals have numerous applications in physics, engineering, and mathematics, including:

- Calculating volumes of solids.
- Finding the center of mass of an object.
- Computing moments of inertia.
- Solving problems in probability and statistics.

In some cases, it is easier to compute multiple integrals by changing variables. For example, switching to polar coordinates for double integrals or cylindrical/spherical coordinates for triple integrals can simplify the integration process. The Jacobian determinant is used to account for the change in area or volume elements during such transformations.

Multiple integrals are a powerful tool in multivariable calculus, enabling us to solve complex problems in higher dimensions.

6.1 Double integrals

6.1.1 Integration of Functions of Two Variables

Let $f : [a,b] \times [c,d] \to \mathbb{R}$. We denote R as the rectangle $[a,b] \times [c,d] \subset \mathbb{R}^2$. We will define the integral of a function over R in two steps: first, in the case where the function is piecewise constant, and then by approximating the given function (if possible) by piecewise constant functions.

a) f is piecewise constant.
There exists a partition of $[a,b] \times [c,d]$:
$a = s_0 < s_1 < s_2 < \cdots < s_m = b$
$c = t_0 < t_1 < \cdots < t_n = d$
such that f is constant inside each rectangle $]s_i, s_{i+1}[\times]t_j, t_{j+1}[$ (where it equals C_{ij}).
We define $\iint_R f(x,y)\,dx\,dy = \sum_{i,j} C_{ij}(s_{i+1} - s_i)(t_{j+1} - t_j)$.
In particular, we note that the value of the integral does not depend on the values of f on the boundaries of the small rectangles.

b) f is bounded on $R = [a,b] \times [c,d]$.
We approximate f by piecewise constant functions.

Definition 6.1.1
f is **integrable** on R if there exists a unique number I such that for all piecewise constant functions $u(x,y)$ and $v(x,y)$, with $u(x,y) \le f(x,y) \le v(x,y)$, we have:

$$\iint_R u(x,y)\,dx\,dy \le I \le \iint_R v(x,y)\,dx\,dy$$

and if, for every $\varepsilon > 0$, there exist piecewise constant functions u_ε and v_ε such that:

$$u_\varepsilon(x,y) \le f(x,y) \le v_\varepsilon(x,y)$$

and

$$0 \le \int_R v_\varepsilon(x,y)\,dx\,dy - \int_R u_\varepsilon(x,y)\,dx\,dy < \varepsilon.$$

Notation: I is called the integral of f over R and is denoted by $\iint_R f(x,y)\,dx\,dy$.

Theorem 6.1.1

1. If f is continuous on R, then f is integrable.
2. If f is positive on R, then $\iint_R f(x,y)\,dx\,dy$ is the volume under the graph of f above R.

Proposition 6.1.2 — Properties of the Double Integral.

1. $\iint_R (\alpha f + \beta g)(x,y)\,dx\,dy = \alpha \iint_R f(x,y)\,dx\,dy + \beta \iint_R g(x,y)\,dx\,dy.$
2. If $R = R_1 \cup R_2$ with $R_1 \cap R_2 = \emptyset$, then:

$$\iint_R f(x,y)\,dx\,dy = \iint_{R_1} f(x,y)\,dx\,dy + \iint_{R_2} f(x,y)\,dx\,dy.$$

Theorem 6.1.3
If f is continuous on R, then $\iint_R f(x,y)\,dx\,dy$ exists.

6.1.2 Calculation of Double Integrals

Theorem 6.1.4 — Fubini.
If f is continuous on $R=[a,b]\times[c,d]$, then:

$$\iint_R f(x,y)\,dx\,dy=\int_a^b\left(\int_c^d f(x,y)\,dy\right)\,dx=\int_c^d\left(\int_a^b f(x,y)\,dx\right)\,dy.$$

■ **Example 6.1**

(1) $\displaystyle\iint_R (x^2+y^2)\,dx\,dy$ with $R=[0,1]\times[0,1]$.

(2) $\displaystyle\iint_R (1+x+y)\,dx\,dy$ with $R=[0,1]\times[0,1]$.

■

Corollary 6.1.5
If the function f is the product of two single-variable functions g and h, i.e., $f(x,y)=g(x)h(y)$, then

$$\iint_R f(x,y)\,dx\,dy=\left(\int_a^b g(x)\,dx\right)\left(\int_c^d h(y)\,dy\right).$$

■ **Example 6.2** $\displaystyle\iint_{[0,1]\times[0,1]} e^{x+y}\,dx\,dy.$ ■

6.1.3 Integration over Bounded Regions in $\mathbb{R}^2$

Let $f:D\to\mathbb{R}$ where $D\subset\mathbb{R}^2$ is non-rectangular.
We consider a rectangle R such that $D\subset R$ and define $\overline{f}$ on R as:

$$\overline{f}(x,y)=\begin{cases} f(x,y) & \text{if } (x,y)\in D,\\ 0 & \text{if } (x,y)\notin D.\end{cases}$$

With the previous notation, we define, if it makes sense:

$$\iint_D f(x,y)\,dx\,dy=\iint_R \overline{f}(x,y)\,dx\,dy.$$

We can reduce the problem to two types of domains D:

Type 1: $D=\{(x,y)\mid a\le x\le b,\ g_1(x)\le y\le g_2(x)\}$ where g_1 and g_2 are continuous.

Type 2: $D=\{(x,y)\mid c\le y\le d,\ h_1(y)\le x\le h_2(y)\}$.

Theorem 6.1.6 — Fubini.

a) If f is continuous on D of Type 1, then f is integrable and we have:

$$\iint_D f(x,y)\,dx\,dy=\int_a^b\left(\int_{g_1(x)}^{g_2(x)} f(x,y)\,dy\right)\,dx.$$

b) If f is continuous on D of Type 2, then f is integrable and we have:

$$\iint_D f(x,y)\,dx\,dy=\int_c^d\left(\int_{h_1(y)}^{h_2(y)} f(x,y)\,dx\right)\,dy.$$

■ **Example 6.3**

(1) $\displaystyle\iint_D (x+2y)\,dx\,dy$: D is the region between the two parabolas $y=2x^2$ and $y=1+x^2$.

(2) $\iint_D e^{x^2}\,dx\,dy$ over the triangle $D = \{(x,y) \mid 0 \le x \le 1,\ 0 \le y \le x\}$.

(3) The choice of integrating first with respect to x or y can lead to longer or shorter calculations. For example, with $\iint_D xy\,dx\,dy$ where D is the trapezoid bounded by $y=0$, $y=1$, and the lines $y = 2-x$ and $y = 1 + x/2$.

■

Definition 6.1.2

If D is a bounded domain, the **area of** D is defined as: $\operatorname{area}(D) = \iint_D 1\,dx\,dy$.

6.1.4 Double Integrals and Change of Variables

Recall for a single variable:

$$\int_a^b f(x)\,dx = \int_c^d f(g(t))g'(t)\,dt$$

where g is a bijection from $[c,d]$ to $[a,b]$.

Proof. Let F be a primitive of f.

On one hand,

$$F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} F'(t)\,dt = \int_{g(a)}^{g(b)} f(t)\,dt,$$

and on the other hand,

$$F(g(b)) - F(g(a)) = \int_a^b (F \circ g)'(s)\,ds = \int_a^b F'(g(s))g'(s)\,ds = \int_a^b f(g(s))g'(s)\,ds.$$

■

Theorem 6.1.7

If $G(u,v) = (x(u,v), y(u,v))$, then:

$$\iint_{G(S)} f(x,y)\,dx\,dy = \iint_S f(x(u,v), y(u,v))\,|\mathrm{J}(G(u,v))|\,du\,dv$$

where $\mathrm{J}(G(u,v)) = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}$.

R

It is not very surprising to find the determinant here. For example, the absolute value of the determinant of a 2 × 2 matrix calculates the area of the parallelogram spanned by the column vectors.

Case of Polar Coordinates

$x = x(r,\theta) = r\cos\theta$

$y = y(r,\theta) = r\sin\theta$

$$J(G) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Thus,

$$\iint_{R=G(S)} f(x,y)\,dx\,dy = \iint_S f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta.$$

■ **Example 6.4**

1. Calculation of the area of a disk.
2. Calculation of the area inside an ellipse.
3. $\int_{-\infty}^{+\infty} e^{-x^2}\,dx = \sqrt{\pi}$.
4. $\iint_D (x-y)^2\,dx\,dy$ where D is the portion of the unit disk between the x-axis and the half-line $y = x$.

■

6.2 Triple integrals

Triple integrals are defined and computed in a manner entirely similar to integrals for functions of two variables (and more). We will look at a few examples...

$\iiint_R f(x,y,z)\,dx\,dy\,dz$

Example: $\iiint_D (x+y+z)^2\,dx\,dy\,dz = \frac{1}{10}$, where D is the domain bounded by the planes $x = 0$, $y = 0$, $z = 0$, and $x+y+z = 1$.

Change of Variables in 3D. Case of spherical coordinates.

$$\iiint_{R=G(S)} f(x,y,z)\,dx\,dy\,dz = \iiint_S f(r\cos\theta\sin\phi, r\sin\theta\sin\phi, r\cos\phi)\, r^2\sin\phi\, dr\,d\theta\,d\phi.$$

Example: $\iiint_D (x^2+y^2+z^2)\,dx\,dy\,dz = \frac{4\pi}{5}$, where D is the ball centered at the origin with radius 1.

6.2.1 Change of Variables in Triple Integrals

Let $T : (u,v,w) \mapsto (x(u,v,w), y(u,v,w), z(u,v,w))$ be a transformation from a region S in the uvw-space to a region E in the xyz-space. The Jacobian is:

$$J(u,v,w) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}.$$

The change of variables formula is:

$$\iiint_E f(x,y,z)\,dx\,dy\,dz = \iiint_S f(x(u,v,w), y(u,v,w), z(u,v,w))\,|J(u,v,w)|\,du\,dv\,dw.$$

Cylindrical Coordinates

In cylindrical coordinates, $x = r\cos\theta$, $y = r\sin\theta$, and $z = z$. The Jacobian is $J(r,\theta,z) = r$, and the change of variables formula becomes:

$$\iiint_E f(x,y,z)\,dx\,dy\,dz = \iiint_S f(r\cos\theta, r\sin\theta, z)\, r\,dr\,d\theta\,dz.$$

Spherical Coordinates

In spherical coordinates, $x = r\sin\phi\cos\theta$, $y = r\sin\phi\sin\theta$, and $z = r\cos\phi$. The Jacobian is $J(r,\theta,\phi) = r^2\sin\phi$, and the change of variables formula becomes:

$$\iiint_E f(x,y,z)\,dx\,dy\,dz = \iiint_S f(r\sin\phi\cos\theta, r\sin\phi\sin\theta, r\cos\phi)\, r^2\sin\phi\, dr\,d\theta\,d\phi.$$

Applications: Calculation of Volumes and Surfaces

- **Volume**: The volume of a region $D \subset \mathbb{R}^3$ is given by:

$$V = \iiint_D 1\, dx\, dy\, dz.$$

- **Surface Area**: The surface area of a surface S defined by $z = f(x,y)$ over a region D is:

$$A = \iint_D \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\, dx\, dy.$$

6.2.2 Some Classical Calculations

Area of a Disk

The area of a disk can be found using integral calculus. By integrating slice by slice:

$$\begin{aligned} \text{Area} &= 4\int_0^R \sqrt{R^2 - x^2}\, dx \\ &= 4\int_0^{\pi/2} R\cos(t)\cdot R\cos(t)\, dt \\ &= 4R^2 \int_0^{\pi/2} \frac{1+\cos(2t)}{2}\, dt \\ &= \pi R^2, \end{aligned}$$

computed using the change of variable $x = R\sin(t)$.

Alternatively (and more simply when the Jacobian for polar coordinates is known), we can use polar coordinates:

$$\begin{aligned} \text{Area} &= \int_0^R \int_0^{2\pi} r\, dr\, d\theta \\ &= 2\pi \cdot \frac{R^2}{2} \\ &= \pi R^2. \end{aligned}$$

Volume of a Ball

The volume of a ball can be calculated using a change of variables to spherical coordinates.

$$\iiint_{B(0,R)} dx\, dy\, dz = \iiint_{[0,R]\times[0,2\pi]\times[0,\pi]} r^2 \sin\phi\, dr\, d\theta\, d\phi = \int_0^R r^2\, dr \int_0^{2\pi} d\theta \int_0^{\pi} \sin\phi\, d\phi,$$

and thus the volume is $\frac{4\pi R^3}{3}$.

Volume of a Pyramid

Consider a pyramid P with a square base of side length a and height h. Place the pyramid on the plane $z = 0$, centered on the z-axis. The side length of the square C_z at height z is $a(1 - z/h)$. Thus, by Fubini's theorem,

$$\iiint_P dx\, dy\, dz = \int_0^h \left(\iint_{C_z} dx\, dy\right) dz = \int_0^h a^2(1 - z/h)^2\, dz = \frac{a^2 h}{3} = \frac{Sh}{3},$$

where S is the area of the base.

Solids of Revolution

Let f be a non-negative function defined on an interval $[a,b]$. Consider the region of space defined as follows:

$$V = \{(x,y,z) \mid x \in [a,b], \sqrt{y^2+z^2} \leq f(x)\}.$$

This is the solid obtained by rotating the graph of f around the x-axis. The volume of V is given by the triple integral

$$\iiint_V dx\,dy\,dz.$$

By integrating slice by slice (first in y and z, then in x), we obtain:

$$\iiint_V dx\,dy\,dz = \int_a^b \left(\iint_{\{(y,z)\mid \sqrt{y^2+z^2}\leq f(x)\}} dy\,dz \right) dx = \int_a^b \pi f(x)^2\,dx.$$

Thus, the calculation of the triple integral reduces to a single integral.

■ **Example 6.5**

(1) Case of a cone: take f as the function defined on $[0,h]$ by $f(x) = ax$ for $a > 0$. Here, the base area of the cone is $S = \pi(ah)^2$, and thus the volume of the cone is

$$\int_0^h \pi(ax)^2\,dx = \frac{\pi a^2 h^3}{3} = \frac{Sh}{3}.$$

(2) Case of a cylinder with radius R and height h:

$$\int_0^h \pi R^2\,dx = \pi R^2 h = Sh,$$

where S is the area of the base.

■

6.3 Exercises

Exercise 6.1

Let D be the domain: $D = \left\{(x,y) \in \mathbb{R}^2;\ x \geq 0,\ y \geq 0,\ x+y \leq 1\right\}$. Compute $\iint_D f(x,y)\,dx\,dy$ in the following cases:

1. $f(x,y) = x^2+y^2$ **2**. $f(x,y) = xy(x+y)$.

■

Exercise 6.2

Compute the following double integral $\iint_D f(x,y)\,dx\,dy$, where

1. $f(x,y) = x$ and $D = \left\{(x,y) \in \mathbb{R}^2;\ y \geq 0,\ x-y+1 \geq 0,\ x+2y-4 \leq 0\right\}$.
2. $f(x,y) = x+y$ and $D = \left\{(x,y) \in \mathbb{R}^2;\ 0 \leq x \leq 1,\ x^2 \leq y \leq x\right\}$.
3. $f(x,y) = \cos(xy)$ and $D = \left\{(x,y) \in \mathbb{R}^2;\ 1 \leq x \leq 2,\ 0 \leq xy \leq \frac{\pi}{2}\right\}$.
4. $f(x,y) = xy$ and $D = \left\{(x,y) \in \mathbb{R}^2;\ x \geq 0,\ y \geq 0,\ xy+x+y \leq 1\right\}$.
5. $f(x,y) = \frac{1}{(x+y)^3}$ and $D = \left\{(x,y) \in \mathbb{R}^2;\ 1 < x < 3,\ y > 2,\ x+y < 5\right\}$.

■

Exercise 6.3

Let D be the domain:

$$D = \left\{(x,y) \in \mathbb{R}^2;\ -1 \leq x \leq 1 \text{ and } x^2 \leq y \leq 4-x^3\right\}.$$

Compute the area of D. ■

Exercise 6.4
Let D be the domain:

$$D=\left\{(x,y)\in\mathbb{R}^2;\ 0\leq x\leq 1,\ 0\leq y\leq 1,\ x^2+y^2\geq 1\right\}.$$

Compute $\iint_D \frac{xy}{1+x^2+y^2}\,dx\,dy$. ■

Exercise 6.5
Compute $\iiint_D f(x,y,z)\,dx\,dy\,dz$ for:

1. $f(x,y,z)=\cos x$ and $D=\left\{(x,y,z)\in\mathbb{R}^3;\ x^2+y^2+z^2<1\right\}$.
2. $f(x,y,z)=\frac{z}{\sqrt{x^2+y^2}}$ and $D=\left\{(x,y,z)\in\mathbb{R}^3;\ x^2+y^2\leq a^2 \text{ and } 0<z<a\right\}$.

■

Exercise 6.6

1. Compute the double integral

$$\iint_\Delta \frac{1}{1+x^2+y^2}\,dx\,dy$$

where $\Delta=\left\{(x,y)\in\mathbb{R}^2;\ 0\leq x\leq 1,\ 0\leq y\leq 1,\ 0<x^2+y^2\leq 1\right\}$.
2. Let $D=\{(x,y)\in\mathbb{R}^2:\ x<y<2x,\ x<y^2<2x\}$. Compute

$$\iint_D \frac{y}{x}\,dx\,dy$$

using the change of variables $u=x/y$ and $v=y^2/x$.

■

Exercise 6.7
Let $D=\left\{(x,y)\in\mathbb{R}^2;\ x\geq 0,\ y\geq 0,\ \frac{x^2}{a^2}+\frac{y^2}{b^2}\leq 1\right\}$. Compute the integral:

$$J=\iint_D (2x^3-y)\,dx\,dy.$$

■

Exercise 6.8
Let B be the unit ball, and $a>1$. Compute:

$$\iiint_B \frac{dx\,dy\,dz}{\sqrt{x^2+y^2+(z-a)^2}}.$$

■

Exercise 6.9
Determine the volume inside the ellipsoid defined by:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1,$$

where a, b, and c are strictly positive real numbers. (Perform the change of variables $x=au$, $y=bv$, $z=cw$.) ■

Bibliography

[1] Apostol, T. M. *Calculus, Vol. 2: Multi-Variable Calculus and Linear Algebra*. Wiley, 1969.

[2] A. Bodin, Exo7, *Cours et exercices de mathématiques en ligne*. Available at: `https://exo7.emath.fr/`

[3] Hubbard, J., & Hubbard, B. *Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach*. Matrix Editions, 2009.

[4] Marsden, J. E., & Tromba, A. J. *Vector Calculus*. W. H. Freeman, 2012.

[5] Spivak, M. *Calculus on Manifolds*. CRC Press, 1965.

[6] Stewart, J. *Calculus: Early Transcendentals*. Cengage Learning, 2015.