MAT334 - Basic notions of calculus
Lecture notes (draft)

Estanislao Herscovich [1]

[1]Institut Fourier, Université Grenoble Alpes, France.

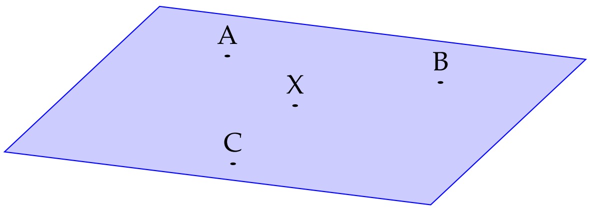
Contents

Chapter 1

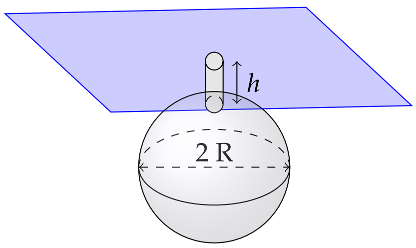
Introduction

The aim of these lectures is to learn tools that allow us to solve the following two kinds of problems.

Problem 1 (optimisation). *Suppose that there are lashing lugs on three fixed points A, B and C on the ground, and assume that one intends to fix a cargo of logs inside a bag by means of 3 ropes going from the top of the bag to each corresponding lashing lug. In which position X should we place the bag of logs to use the minimum amount possible of rope?*



Problem 2 (volumes). *Suppose we store petrol inside an underground cistern of spheric form and radius R whose center is buried $R + h$ under the level of the ground. The only communication between the tank and the exterior is via a small pipe of length h going vertically from the top of the spheric part to the ground. The only way to measure the amount of petrol inside the cistern is to use a stick to measure the level ℓ of petrol. Determine the volume of petrol inside the cistern as a function of the measured level $\ell \in [0, 2R]$.*



In order to be able to solve them we will need to recall some basic objects and facts concerning calculus.

Chapter 2

Vectors in the Euclidean space

The proof of all the results in this chapter are standard and left as an exercise to the reader.

2.1 Basic notions

We recall that $\mathbb{Z}_{>0}$ denotes the set of (strictly) positive integers and $\mathbb{R}$ is the set of real numbers. Open intervals will be denoted in the form $]a, b[$, where $a < b$ are real numbers. Given $n \in \mathbb{Z}_{>0}$, $\mathbb{R}^n$ is the set of n-tuples of real numbers, *i.e.* elements of the form $(a_1, \dots, a_n)$, where $a_i \in \mathbb{R}$. We will be mainly concerned with $n = 1, 2, 3$.

Definition 2.1.1. *Given $n \in \mathbb{Z}_{>0}$, a **vector (of size** n**)** is an element of $\mathbb{R}^n$. A vector of size 1 is called a **scalar**. We shall usually omit the size of the vector if it is clear from the context.*

Notation 2.1.2. *A vector of the form $(a_1, \dots, a_n)$ is usually denoted by $\vec{a}$, $\bar{a}$ or $\mathbf{a}$. In this manuscript we shall mainly utilize the first one. We shall typically denote the **zero vector** $(0, \dots, 0)$ of size n by $\vec{0}$. The vectors of the canonical basis of $\mathbb{R}^n$ are typically denoted by $\{\vec{e}_1, \dots, \vec{e}_n\}$, even though if $n = 3$, one usually writes them instead as $\{\vec{i}, \vec{j}, \vec{k}\}$.*

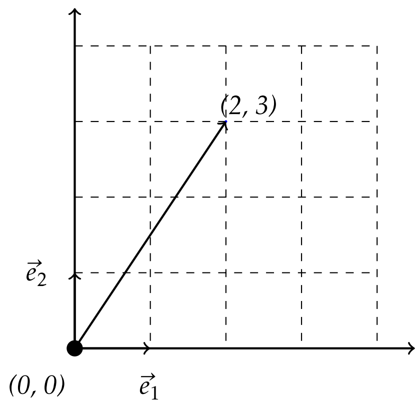
Definition 2.1.3. *We recall the following operations on the set of vectors of a fixed size:*

(i) ***(addition)*** *given $\vec{a} = (a_1, \dots, a_n)$ and $\vec{b} = (b_1, \dots, b_n)$, the **sum** $\vec{a} + \vec{b}$ is defined as $(a_1 + b_1, \dots, a_n + b_n)$;*

(ii) ***(external product)*** *given $\vec{a} = (a_1, \dots, a_n)$ and a scalar λ, the **product** $\lambda\vec{a}$ is defined as $(\lambda a_1, \dots, \lambda a_n)$.*

Notation 2.1.4. *Given two points $\vec{a}$ and $\vec{b}$ in $\mathbb{R}^n$, the **vector** from $\vec{a}$ to $\vec{b}$ is the difference $\vec{b} - \vec{a}$. It is typically denoted by $\vec{ab}$.*

Interpretation 2.1.5. *A vector $\vec{a}$ of the form $(a_1, \dots, a_n)$ is represented by an arrow from the origin to the point $(a_1, \dots, a_n)$. In this sense, **vectors are identified with points**. For instance, if $n = 2$, the vector $(2, 3)$ is depicted as*



In many graphic representations of vectors, specially when dealing with forces acting on a body or the tangential vector of a curve at a point, one often translates the vector from the origin to other point of application. However, ***the "new point of application" is not part of the information conveyed by the vector itself****, and must be regarded thus as further data.*

Definition 2.1.6. *A* ***smooth parameterized curve*** *in* $\mathbb{R}^n$ *is a function* $\vec{\alpha} :]a,b[\to \mathbb{R}^n$ *defined on a nonempty open interval satisfying that* $\alpha_i :]a,b[\to \mathbb{R}$ *is* C^∞*, for all* $i = 1,\dots,n$*, where we write* $\vec{\alpha}(t) = (\alpha_1(t),\dots,\alpha_n(t))$*, for all* $t \in]a,b[$*. The* ***velocity*** $\vec{\alpha}'$ *of* $\vec{\alpha}$ *is the smooth parameterized curve given by* $\vec{\alpha}'(t) = (\alpha_1'(t),\dots,\alpha_n'(t))$*, for all* $t \in]a,b[$*, where* α_i' *denotes the derivative of* α_i*. The* ***acceleration*** $\vec{\alpha}''$ *of* $\vec{\alpha}$ *is the velocity of* $\vec{\alpha}'$*. We say that the parameterized curve* $\vec{\alpha}$ *is* ***regular*** *if* $\vec{\alpha}'(t)$ *is nonzero for all* $t \in]a,b[$*.*

Example 2.1.7. *The map* $\vec{\alpha} :]0,1[\to \mathbb{R}^2$ *given by*

$$\vec{\alpha}(t) = \Big(\underbrace{\cos(2\pi t)}_{\alpha_1(t)}, \underbrace{\sin(2\pi t)}_{\alpha_2(t)}\Big)$$

is a smooth parameterized curved in the Euclidean plane, whose image is $S^1 \setminus \{(1,0)\}$*. The velocity of* $\vec{\alpha}$ *is* $\vec{\alpha}'(t) = 2\pi(-\sin(2\pi t), \cos(2\pi t))$ *and the acceleration is* $\vec{\alpha}'' = -4\pi^2\vec{\alpha}$*. Note that* $\vec{\alpha}$ *is regular in this example.*

Fact 2.1.8. *Let* $\vec{\alpha} :]a,b[\to \mathbb{R}^n$ *be a regular smooth parameterized curve. The* ***tangent line*** *to* $\vec{\alpha}$ *at the point* $\vec{\alpha}(t_0)$ *(where* $t_0 \in]a,b[$*) is given by the parameterized straight line* $\vec{\beta} : \mathbb{R} \to \mathbb{R}^n$ *defined as*

$$\beta(t) = \vec{\alpha}(t_0) + t\vec{\alpha}'(t_0).$$

Note that the regularity hypothesis is needed to assure that the direction vector $\vec{\alpha}'(t_0)$ *of the straight line is nonzero.*

2.2 Scalar product

2.2.1 Basic definitions and properties

Definition 2.2.1. *Given two vectors* $\vec{a} = (a_1,\dots,a_n)$ *and* $\vec{b} = (b_1,\dots,b_n)$ *of the same size, the* ***(Euclidean) scalar product*** $\vec{a}\cdot\vec{b}$ *is defined as the scalar* $a_1b_1 + \cdots + a_nb_n$*. Note that* $\vec{a}\cdot\vec{a} \geq 0$*, since it is a sum of squares of real numbers. The* ***(Euclidean) norm*** $||\vec{a}||$ *of* $\vec{a} = (a_1,\dots,a_n)$ *is given by* $\sqrt{\vec{a}\cdot\vec{a}}$*. The* ***distance*** *between two points* $\vec{a}$ *and* $\vec{b}$ *is defined as* $||\vec{a} - \vec{b}||$*.*

Proposition 2.2.2 (Cauchy-Schwarz inequality). *Given two vectors $\vec{a}$ and $\vec{b}$ of the same size, then*

$$|\vec{a} \cdot \vec{b}| \leq ||\vec{a}|| \, ||\vec{b}||.$$

The equality holds if and only if $\vec{a}$ and $\vec{b}$ are linearly dependent.

Definition 2.2.3. *Given two nonzero vectors $\vec{a}$ and $\vec{b}$ of the same size, the* ***angle θ determined by them*** *(and measured in radian units) is the only real number in the interval $[0, \pi]$ satisfying that*

$$\cos(\theta) = \frac{\vec{a} \cdot \vec{b}}{||\vec{a}|| \, ||\vec{b}||}.$$

Since the restriction of the cosine function to the interval $[0, \pi]$ is injective and its image is the interval $[-1, 1]$, the previous notion is well-defined.

Properties 2.2.4. *Let $\vec{a}$, $\vec{b}$ and $\vec{c}$ be three vectors of the same size, and λ a scalar.*

(i) $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$;

(ii) $\vec{a} \cdot (\vec{b} + \lambda \vec{c}) = \vec{a} \cdot \vec{b} + \lambda \vec{a} \cdot \vec{c}$;

(iii) $\vec{a} = \vec{0}$ *if and only if* $||\vec{a}|| = 0$;

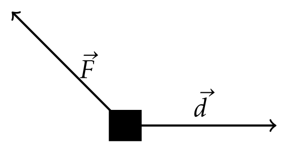
(iv) if $\vec{a}$ and $\vec{b}$ are nonzero, then the angle they determine is $\pi/2$ if and only if $\vec{a} \cdot \vec{b} = 0$. We say in this case that they are ***orthogonal****.*

2.2.2 Applications of the scalar product

The scalar product of vectors naturally appears in many situations.

2.2.2.1 Physics: Work of a force

Consider a constant force $\vec{F}$ acting on a body that moves from a position $\vec{a}$ to $\vec{a} + \vec{d}$, as depicted below.



The **work** done by the force is the scalar product $\vec{F} \cdot \vec{d}$.

2.2.2.2 Equation of a hyperplane

The **hyperplane in the space $\mathbb{R}^n$ containing a point $\vec{a}$ and perpendicular to a nonzero vector $\vec{n}$** is the set

$$\left\{ \vec{x} \in \mathbb{R}^n : (\vec{x} - \vec{a}) \cdot \vec{n} = 0 \right\}. \tag{2.2.1}$$

If $n = 3$, the previous set is an honest plane in $\mathbb{R}^3$.

2.2.2.3 Distance from a point to a hyperplane

The **distance** $d(\vec{b},\Pi)$ from a point $\vec{b}$ in the space $\mathbb{R}^n$ to a hyperplane Π is defined as the minimum of the set

$$\left\{\|\vec{b}-\vec{x}\| : \vec{x} \in \Pi\right\}.$$

Suppose now that Π is the hyperplaneplane given by (2.2.1). Using Proposition 2.2.2, prove that the distance $d(\vec{b},\Pi)$ is

$$\frac{|(\vec{b}-\vec{a})\cdot\vec{n}|}{\|\vec{n}\|}. \tag{2.2.2}$$

2.3 Vector product

2.3.1 Basic definitions and properties

The next definitions only concerns vectors of size 3.

Definition 2.3.1. *Given two vectors $\vec{a} = (a_1, a_2, a_3)$ and $\vec{b} = (b_1, b_2, b_3)$ of size 3, the **vector (or cross) product** $\vec{a}\times\vec{b}$ is defined as the vector $(a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)$.*

Proposition 2.3.2. *Given two nonzero vectors $\vec{a}$ and $\vec{b}$ of size 3, and let θ denote the angle determined by them given in Definition 2.2.3, then*

$$|\vec{a}\times\vec{b}| = \|\vec{a}\|\,\|\vec{b}\|\sin(\theta).$$

Properties 2.3.3. *Let $\vec{a}$, $\vec{b}$ and $\vec{c}$ be three vectors of size 3, and λ a scalar.*

(i) $\vec{a}\times\vec{b} = -\vec{b}\times\vec{a}$;

(ii) $\vec{a}\times(\vec{b}+\lambda\vec{c}) = \vec{a}\times\vec{b} + \lambda\vec{a}\times\vec{c}$;

(iii) $\vec{a}\times\vec{b} = \vec{0}$ *if and only if* $\vec{a}$ *and* $\vec{b}$ *are linearly dependent.*

2.3.2 Applications

The cross product of vectors naturally appears in the following situations.

2.3.2.1 Physics : Movement of a rigid body

Let $\{{}^1\vec{\alpha},\dots,{}^N\vec{\alpha}\}$ be a finite collection of smooth parameterized curves describing the movement of material points, where $N \geq 3$, and ${}^i\vec{\alpha} :]a,b[\to \mathbb{R}^3$ for all $i = 1,\dots,N$. It is said to be a *rigid object* if the distance $\|{}^i\vec{\alpha}(t) - {}^j\vec{\alpha}(t)\|$ is a constant function of $t \in]a,b[$, for all $i,j = 1,\dots,N$. By a theorem by L. Euler and M. Chasles (see [1,2]), there exists a smooth map $\vec{\omega} :]a,b[\to \mathbb{R}^3$ such that

$${}^i\vec{\alpha}'(t) - {}^j\vec{\alpha}'(t) = \vec{\omega}(t)\times\left({}^i\vec{\alpha}(t) - {}^j\vec{\alpha}(t)\right),$$

for all $i,j = 1,\dots,N$. Note that $\vec{\omega}(t)$ does not depend on the indices i and j, which is precisely the content of the mentioned theorem. The vector $\vec{\omega}(t)$ is called the **angular velocity** of the rigid body at time t.

2.3.2.2 Equation of a line in the three dimensional space

The equation of a line in $\mathbb{R}^3$ containing a point $\vec{a}$ and parallel to a nonzero vector $\vec{v}$ is the set

$$\left\{\vec{x} \in \mathbb{R}^3 : (\vec{x} - \vec{a}) \times \vec{v} = \vec{0}\right\}. \tag{2.3.1}$$

2.3.2.3 Distance from a point to a line in the three dimensional space

The **distance** $d(\vec{b}, \Lambda)$ from a point $\vec{b}$ in the space $\mathbb{R}^3$ to a line Λ is defined as the minimum of the set

$$\left\{\|\vec{b} - \vec{x}\| : \vec{x} \in \Lambda\right\}.$$

Suppose now that Λ is the line given by (2.3.1). Using Proposition 2.2.2, prove that the distance $d(\vec{b}, \Lambda)$ is

$$\frac{\|(\vec{b} - \vec{a}) \times \vec{v}\|}{\|\vec{v}\|}. \tag{2.3.2}$$

2.4 Other coordinates systems for the plane and the three dimensional space

2.4.1 Polar coordinates in the plane

Let $\vec{x} = (x_1, x_2) \in \mathbb{R}^2$ be a vector. In the case of the plane, the most widely used notation for the coordinates of a vector are (x, y), but we prefer to keep the subscripts for clarity. Define the mapping $\mathscr{P} : \mathbb{R}_{>0} \times [0, 2\pi[\to \mathbb{R}^2 \setminus \{(0,0)\}$ by

$$\mathscr{P}(r, \theta) = \Big(r\cos(\theta), r\sin(\theta)\Big), \tag{2.4.1}$$

where $r \in \mathbb{R}_{>0}$ and $\theta \in [0, 2\pi[$.

Proposition 2.4.1. *The map $\mathscr{P} : \mathbb{R}_{>0} \times [0, 2\pi[\to \mathbb{R}^2 \setminus \{(0,0)\}$ given in* (2.4.1) *is bijective. The inverse map sends a nonzero vector $\vec{x} = (x_1, x_2) \in \mathbb{R}^2$ to*

$$\Big(\sqrt{x_1^2 + x_2^2}, \operatorname{atan2}(x_1, x_2)\Big),$$

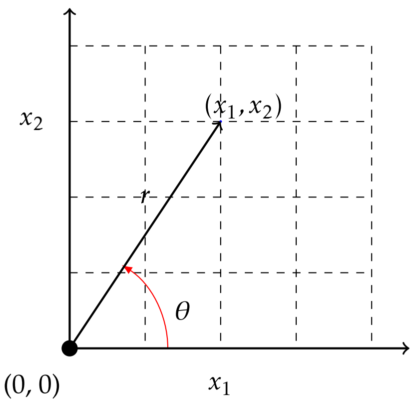
where $\operatorname{atan2} : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ *is a common variation on the arctangent function* $\arctan : \mathbb{R} \to]-\pi/2, \pi/2[$ *defined as*

$$\operatorname{atan2}(x_1, x_2) = \begin{cases} \arctan(\frac{x_2}{x_1}), & \text{if } x_1 > 0 \text{ and } x_2 \geq 0, \\ \arctan(\frac{x_2}{x_1}) + 2\pi, & \text{if } x_1 > 0 \text{ and } x_2 < 0, \\ \arctan(\frac{x_2}{x_1}) + \pi, & \text{if } x_1 < 0, \\ \frac{\pi}{2}, & \text{if } x_1 = 0 \text{ and } x_2 > 0, \\ -\frac{\pi}{2}, & \text{if } x_1 = 0 \text{ and } x_2 < 0. \end{cases} \tag{2.4.2}$$

Given any vector $\vec{x} = (x_1, x_2) \in \mathbb{R}^2$, we say that x_1 and x_2 are the **Cartesian coordinates** of the point. Analogously, the two components of $\mathscr{P}^{-1}(\vec{x})$ are called the **polar coordinates** of the point $\vec{x}$. The first component is

$$\sqrt{x_1^2 + x_2^2}$$

and it is called the **radial coordinate** or **radius**, and the second one is $\operatorname{atan2}(x_1, x_2)$ and it is called the **angular coordinate**, **polar angle**, or **azimuth**. The previous result allows us to give a different description of a point or a set of points, which may be simpler in the new coordinates. Graphically, the polar coordinates (r, θ) of a nonzero vector $\vec{x} = (x_1, x_2)$ are given as follows.



Remark 2.4.2. *Note that there is no canonical manner of defining the polar coordinates of the origin: even though the radial coordinate could be chosen to be zero, there is no reasonable choice of polar angle for* $(0,0)$.

Example 2.4.3. *A circle of center* $\vec{a} = (a_1, a_2)$ *and radius* $\rho > 0$ *is defined as the set of points*

$$\left\{\vec{x} = (x_1, x_2) \in \mathbb{R}^2 : \|\vec{x} - \vec{a}\| = \rho\right\}.$$

If $\vec{a} = \vec{0}$, *then the equation of the previous circle in polar coordinates is*

$$\left\{(r, \theta) \in \mathbb{R}_{>0} \times [0, 2\pi[\, : r = \rho\right\},$$

which is particularly simple.

Suppose that $\vec{a}$ *is nonzero and that its polar coordinates are* (r_0, γ), *with* $r_0 \neq \rho$. *Then, the equation of the previous circle in polar coordinates is*

$$\left\{(r, \theta) \in \mathbb{R}_{>0} \times [0, 2\pi[\, : r^2 - 2rr_0\cos(\theta - \gamma) + r_0^2 = \rho^2\right\}.$$

2.4.2 Cylindrical coordinates in the space

Cylindrical coordinates are the obvious extension of the polar coordinates to the space and can be described as follows. Let $\vec{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ be a vector. As before, in the case of the space the most widely used notation for the coordinates of a vector are (x, y, z), but we keep the subscripts for coherence.

The next result is a direct consequence of Proposition 2.4.1.

Proposition 2.4.4. *The map* $\mathscr{C} : \mathbb{R}_{>0} \times [0, 2\pi[\times \mathbb{R} \to \mathbb{R}^3 \setminus (\{(0,0)\} \times \mathbb{R})$ *given by sending* (r, θ, z) *to* $(\mathscr{P}(r, \theta), z)$ *is bijective, where the map* $\mathscr{P}$ *is defined in* (2.4.1). *The inverse map sends a vector* $\vec{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \setminus (\{(0,0)\} \times \mathbb{R})$ *to*

$$\left(\sqrt{x_1^2 + x_2^2}, \operatorname{atan2}(x_1, x_2), x_3\right),$$

where atan2 *is given in* (2.4.2).

Given any vector $\vec{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$, we say that x_1, x_2 and x_3 are the **Cartesian coordinates** of the point. If the vector $\vec{x}$ satisfies that $x_1^2 + x_2^2 > 0$, the three components of $\mathscr{C}^{-1}(\vec{x})$ are called the **cylindrical coordinates** of $\vec{x}$. The two first components of $\mathscr{C}^{-1}(\vec{x})$ are just the polar coordinates of (x_1, x_2) and are called as before. The third component is just x_3 and it is called **height** or **altitude** (if the reference plane is considered horizontal), **longitudinal position**, or **axial position**.

Remark 2.4.5. *Remark 2.4.2 tells us that there is no canonical manner of defining the cylindrical coordinates of a point of the form* $(0, 0, x_3)$: *although the radius and the height could be however chosen to be zero and* x_3, *respectively, there is no sensible choice of polar angle.*

2.4.3 Spherical coordinates in the space

There is another way of extending polar coordinates to the space by mimicking the geographical coordinates. Indeed, define the mapping $\mathscr{S} : \mathbb{R}_{>0} \times [0, 2\pi[\times]0, \pi[\to \mathbb{R}^3 \setminus (\{(0,0)\} \times \mathbb{R})$ by

$$\mathscr{S}(r, \phi, \theta) = \Big(r\cos(\phi)\sin(\theta), r\sin(\phi)\sin(\theta), r\cos(\theta) \Big), \tag{2.4.3}$$

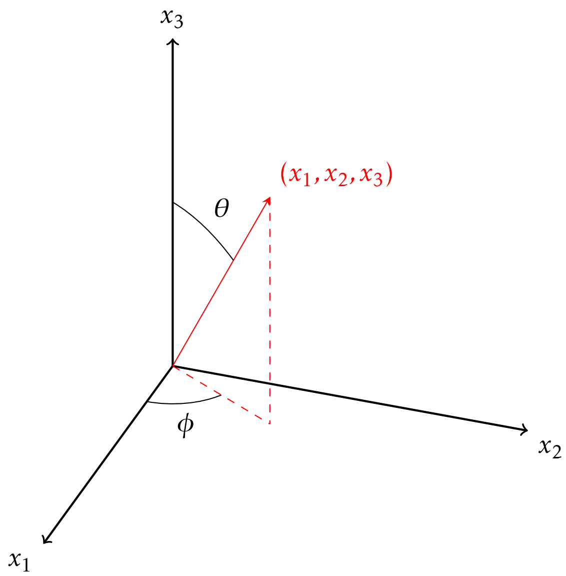
where $r \in \mathbb{R}_{>0}$, $\phi \in [0, 2\pi[$, and $\theta \in]0, \pi[$.

Proposition 2.4.6. *The map* $\mathscr{S} : \mathbb{R}_{>0} \times [0, 2\pi[\times]0, \pi[\to \mathbb{R}^3 \setminus (\{(0,0)\} \times \mathbb{R})$ *given in* (2.4.3) *is bijective. The inverse map sends a nonzero vector* $\vec{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ *to*

$$\left(\sqrt{x_1^2 + x_2^2 + x_3^2}, \operatorname{atan2}(x_1, x_2), \arccos\left(\frac{x_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \right) \right),$$

where atan2 *was defined in* (2.4.2).

If the vector $\vec{x}$ satisfies that $x_1^2 + x_2^2 > 0$, the three components of $\mathscr{S}^{-1}(\vec{x})$ are called the **spherical coordinates** of $\vec{x}$. The first component of $\mathscr{S}^{-1}(\vec{x})$ is called the **radial coordinate** or **radius**, the second component is called the **polar angle**, **co-latitude**, **zenith angle**, **normal angle**, or **inclination angle**, and the third one is called the **azimuthal angle** or **longitude**. The polar coordinates (r, θ, ϕ) of a nonzero vector $\vec{x} = (x_1, x_2, x_3)$ are given as follows.



Remark 2.4.7. *As in the case of cylindrical coordinates, there is no canonical manner of defining the spherical coordinates of a point of the form* $(0, 0, x_3)$. *We see that the radial coordinate of such point could be chosen to be* $|x_3|$. *Moreover, if* $x_3 \neq 0$, *the co-latitude can be reasonably defined to be* $(1 - \operatorname{sgn}(x_3))\pi/2$, *where* $\operatorname{sgn}(t)$ *is the sign function*[1]. *However, there is no reasonable choice of polar angle for any point of the form* $(0, 0, x_3)$, *and moreover, the origin has no sensible definition of co-latitude either.*

[1] The **sign function** $\operatorname{sgn} : \mathbb{R} \to \mathbb{R}$ sends any (strictly) positive t to 1, 0 to 0, and any negative t to -1.

Chapter 3

Scalar and vector maps

3.1 Basic definitions

Definition 3.1.1. *Consider $n, m \in \mathbb{Z}_{>0}$. Let $U \subseteq \mathbb{R}^n$ be an open subset*[1]*, let $f : U \to \mathbb{R}^m$ be a map. We write $f(\vec{x}) = (f_1(\vec{x}), \dots, f_m(\vec{x}))$, for all $\vec{x} \in U$. It is generally called a* ***field****, or* ***vector map****. If $m = 1$ we say that f is a* ***scalar map*** *(or* ***scalar field****).*

Definition 3.1.2. *Consider $n \in \mathbb{Z}_{>0}$ and let $U \subseteq \mathbb{R}^n$ be an open subset. A* ***vector field*** *is a vector map of the form $f : U \to \mathbb{R}^n$.*

Applications 3.1.3. *Scalar and vector maps are typical in physics. An usual example of scalar field is given by associating a temperature to every point of a certain region of the space. Other examples of scalar fields are given by the density function associating a density of mass to each point of a region, the density charge, etc.*

The most common example of vector field in mechanics is the gravitational field produced by a distribution of mass expressed by Newton's law. Similar examples are the electric (resp., magnetic) field produced by a distribution of (resp., steadily moving) charges expressed by the Coulomb (resp., Biot-Savart) law.

Definition 3.1.4. *Consider $n \in \mathbb{Z}_{>0}$. Let $U \subseteq \mathbb{R}^n$ be an open subset and let $f : U \to \mathbb{R}^m$ be a vector map. Given $\vec{c} \in \mathbb{R}^m$, the set*

$$L_c = \{\vec{x} \in U : f(\vec{x}) = \vec{c}\} \subseteq \mathbb{R}^n$$

is called the ***level set*** *of f level $\vec{c}$. In case $m = 1$ and $n = 2$ it is also called the* ***level curve*** *and for $m = 1$ and $n = 3$ the* ***level surface****.*

Example 3.1.5. *Let $U = \mathbb{R}^2 \to \mathbb{R}$ given by $f(x_1, x_2) = x_1^2 + x_2^2$. Then, $L_c = \emptyset$ if $c < 0$, $L_0 = \{(0,0)\}$, and L_c is the circle of radius $\sqrt{c}$ if $c > 0$.*

3.2 Applications: Thermodynamics

There is one description of (phenomenological) thermodynamics, as formulated by A. Neumaier (see [5]), which is very simple to introduce. We first recall that a **standard thermodynamic system** composed of $J \in \mathbb{Z}_{>0}$ species is any homogeneous isotropic material system which consists of J pure substances in the absence of chemical reactions and external forces, and it is very large in microscopic units.

The **space of intensive thermodynamic variables** $\mathscr{I}$ is just $\mathbb{R}_{>0} \times \mathbb{R}^{J+1}$. The elements of the space of intensive thermodynamic variables are denoted by $(T, P, \vec{\mu})$,

[1] A subset $U \subseteq \mathbb{R}^n$ is said to be **open** if given any $\vec{x} = (x_1, \dots, x_n) \in U$ there is $\epsilon > 0$ (depending on $\vec{x}$) such that $[x_1 - \epsilon, x_1 + \epsilon] \times \cdots \times [x_n - \epsilon, x_n + \epsilon] \subseteq U$.

where T is called the **temperature**, P is the **pressure**, and the last J arguments are the **chemical potentials** $\vec{\mu} \in \mathbb{R}^J$. These $J+2$ physical entities are called **intensive thermodynamic variables**. Note that $\mathscr{I}$ is a convex set[2]. A **state function** is a map $\Delta : \mathscr{I} \to \mathbb{R}$ such that Δ is **convex**[3].

The **space of extensive thermodynamic variables** $\mathscr{E}$ is just $\mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0} \times \mathbb{R}^J_{\geq 0}$. The elements of the space of extensive thermodynamic variables are denoted by $(U, S, V, \vec{n})$, where U is called the **internal energy**, S is the **entropy**, V is the **volume** and the last J arguments are the **number of moles** $\vec{n} \in \mathbb{R}^J_{\geq 0}$ **of the** J **species**. These $J+3$ physical entities are called **extensive thermodynamic variables**. The extensive and intensive thermodynamic variables form the **basic thermodynamic variables**.

Given a state function $\Delta : \mathscr{I} \to \mathbb{R}$, denote $\mathscr{I}_\Delta$ the level set of Δ of level zero, *i.e.*

$$\mathscr{I}_\Delta = \Big\{(T, P, \vec{\mu}) \in \mathscr{I} : \Delta(T, P, \vec{\mu}) = 0\Big\}. \tag{3.2.1}$$

Define now

$$\mathscr{E}_\Delta = \Big\{(U, S, V, \vec{n}) \in \mathscr{E} : U - TS + PV - \vec{\mu} \cdot \vec{n} \geq 0, \text{ for all } (T, P, \vec{\mu}) \in \mathscr{I}_\Delta\Big\}. \tag{3.2.2}$$

The (unique) axioms of thermodynamics given by Neumaier say the following.

(i) The extensive thermodynamic variables of a standard thermodynamic system formed by two subsystems is the sum of the corresponding extensive variables of each subsystem.

(ii) To every standard thermodynamic system there is an associated state function Δ such that the set of possible physical values of the intensive thermodynamic variables, called the **state space** associated to Δ, is given by the level set $\mathscr{I}_\Delta$ of level zero (see (3.2.1)). The state function of a standard thermodynamic system formed by two subsystems having the same state function Δ is also Δ.

(iii) Given a standard thermodynamic system with state function Δ, the set of possible values for the extensive variables $(U, S, P, \vec{n})$ is included in $\mathscr{E}_\Delta$ defined in (3.2.2), *i.e.* if $(U, S, P, \vec{n})$ is a collection of physical values for the extensive thermodynamic variables, then the following inequality

$$U - TS + PV - \vec{\mu} \cdot \vec{n} \geq 0$$

holds for all $(T, P, \vec{\mu})$ in the state space associated to Δ.

(iv) An **equilibrium state** is given by a vector in the space of intensive thermodynamic variables and one in the space of extensive thermodynamic variables such that the inequality of the previous item is an equality. In other words, the set of equilibrium states is given the level set of level $(0, 0)$ of the vector map $\mathscr{I} \times \mathscr{E} \to \mathbb{R}^2$ sending $(T, P, \vec{\mu}, U, S, P, \vec{n})$ to $(\Delta(T, P, \vec{\mu}), U - TS + PV - \vec{\mu} \cdot \vec{n})$. We will denote it by $\mathcal{E}q_\Delta$.

Remark 3.2.1. *Note that the there is no uniqueness of equilibrium states for prescribed values of the intensive variables. Indeed, if* $({}^0T, {}^0P, {}^0\vec{\mu}, {}^0U, {}^0S, {}^0V, {}^0\vec{n})$ *is an equilibrium state and* $\omega > 0$ *is a positive real number, then* $({}^0T, {}^0P, {}^0\vec{\mu}, \omega.{}^0U, \omega.{}^0S, \omega.{}^0V, \omega.{}^0\vec{n})$ *is also an equilibrium state. This is coherent with item (i) of the previous list of axioms, for the factor* ω *indicates how the size of the system has changed.*

[2] We recall that a set $C \subseteq \mathbb{R}^n$ is called **convex** if $t\vec{x} + (1-t)\vec{y} \in C$ for all $t \in [0,1]$ and all $\vec{x}, \vec{y} \in C$.

[3] A map $f : C \to \mathbb{R}$ defined on a convex set C is called **convex** if $f(t\vec{x} + (1-t)\vec{y}) \leq tf(\vec{x}) + (1-t)f(\vec{y})$ for all $t \in [0,1]$, and all $\vec{x}, \vec{y} \in C$.

Remark 3.2.2. *Note that the previous axioms do not explain when a concrete system is in equilibrium. However, suppose we have an equilibrium state* $({}^0T, {}^0P, {}^0\vec{\mu}, {}^0U, {}^0S, {}^0V, {}^0\vec{n})$*, and consider the set of points*

$$\mathscr{B}_{{}^0U,{}^0S,{}^0V,{}^0\vec{n}} = \Big\{(T, P, \vec{\mu}, U, S, V, \vec{n}) \in \mathscr{I}_\Delta \times \mathscr{E}_\Delta : (U, S, V, \vec{N}) = ({}^0U, {}^0S, {}^0V, {}^0\vec{n})\Big\}.$$

Then, the point $({}^0T, {}^0P, {}^0\vec{\mu}, {}^0U, {}^0S, {}^0V, {}^0\vec{n})$ *belongs to* $\mathscr{B}_{{}^0U,{}^0S,{}^0V,{}^0\vec{n}}$ *and it is precisely a maximum for the (nonnegative) function*

$$(T, P, \vec{\mu}, {}^0U, {}^0S, {}^0V, {}^0\vec{n}) \mapsto {}^0U - T.{}^0S + P.{}^0V - \vec{\mu} \cdot {}^0\vec{n}$$

defined on $\mathscr{B}_{{}^0U,{}^0S,{}^0P,{}^0\vec{n}}$*. In other words, in this formulation of thermodynamics, given the values of the extensive thermodynamic variables,* ***finding the equilibrium state corresponds exactly to solving an optimisation problem****. We will see how to solve this kind of problems in the sequel.*

Example 3.2.3. *An* ***ideal gas*** *is characterized by the state function*

$$\Delta(T, P, \mu) = \sum_{j=1}^{J} P_j(T) e^{\frac{\mu_j}{RT}} - P,$$

where R *is the* ***gas constant****, and* P_j *are (strictly) positive functions of the temperature. The state space describing all the possible physical values for the intensive variables is thus given by its level set of level zero.*

Chapter 4

Continuity

Definition 4.0.1. *Consider $n, m \in \mathbb{Z}_{>0}$. Let $U \subseteq \mathbb{R}^n$ be an open subset, let $f : U \to \mathbb{R}^m$ be a map and $\vec{a} \in U$ be a point of U. As usual, write $f(\vec{x}) = (f_1(\vec{x}), \dots, f_m(\vec{x}))$, for all $\vec{x} \in U$. We say that f is **continuous at** $\vec{a}$ if the limit*

$$\lim_{\vec{x} \to \vec{a}} f(\vec{x}) \tag{4.0.1}$$

exists (i.e. it gives an element of $\mathbb{R}^m$) and it coincides with $f(\vec{a})$. We recall that the limit (4.0.1) *exists if and only if the limit*

$$\lim_{\vec{x} \to \vec{a}} f_i(\vec{x}) \tag{4.0.2}$$

*exists (i.e. it gives a real number) for all $i = 1, \dots, m$. We say that f is **continuous** (on U) if it is continuous at every point $\vec{a} \in U$.*

Example 4.0.2. *Let $U = \mathbb{R}^n$, and $f : U \to \mathbb{R}$ given by any **polynomial function**, i.e. a mapping of the form*

$$f(\vec{x}) = \sum_{\bar{i} \in \mathbb{N}^n} c_{\bar{i}} x_1^{i_1} \dots x_n^{i_n}$$

*where $\vec{x} = (x_1, \dots, x_n)$, $\bar{i} = (i_1, \dots, i_n)$, and $c_{\bar{i}}$ are real numbers that are almost all zero. We say that f is **nonzero** if there exists at least one nonzero coefficient $c_{\bar{i}}$. The **degree** of a nonzero polynomial function is the maximum of the set*

$$\{i_1 + \dots + i_n : c_{\bar{i}} \neq 0\}.$$

It is easy to prove that any polynomial function is continuous on $\mathbb{R}^n$. It is also a direct consequence of items (i) and (ii) of Lemma 4.0.4.

Remark 4.0.3. *Let $U \subseteq \mathbb{R}^n$ be an open subset, and $f : U \to \mathbb{R}^m$ be a map. We write as usual $f(\vec{x}) = (f_1(\vec{x}), \dots, f_m(\vec{x}))$, for all $\vec{x} \in U$. Then f is continuous on U if and only if $f_i : U \to \mathbb{R}$ is continuous on U for all $i = 1, \dots, m$.*

We now give some properties of continuous mappings (for a proof see [3], Section VII.§1).

Lemma 4.0.4. *Let $U \subseteq \mathbb{R}^n$ be an open subset, $\lambda \in \mathbb{R}$ and $f, g : U \to \mathbb{R}^m$ be two continuous mappings. Then,*

*(i) the **linear combination** $f + \lambda g : U \to \mathbb{R}^m$ given by $(f + \lambda g)(\vec{x}) = f(\vec{x}) + \lambda g(\vec{x})$, for all $\vec{x} \in U$, is continuous;*

*(ii) the **pointwise scalar product** $f \cdot g : U \to \mathbb{R}^m$ given by $(f \cdot g)(\vec{x}) = f(\vec{x}) \cdot g(\vec{x})$, for all $\vec{x} \in U$, is continuous;*

(iii) if $m = 1$ and $f(\vec{x}) \neq 0$ for all $\vec{x} \in U$, the ***multiplicative inverse*** *$1/f : U \to \mathbb{R}$ given by $(1/f)(\vec{x}) = 1/f(\vec{x})$, for all $\vec{x} \in U$, is continuous;*

(iv) suppose that the image of f is included in an open set $V \subseteq \mathbb{R}^m$ and let $h : V \to \mathbb{R}^p$ be a continuous map. Then, the ***composition*** *$h \circ f : U \to \mathbb{R}^p$ given by $(h \circ f)(\vec{x}) = h(f(\vec{x}))$, for all $\vec{x} \in U$, is also continuous.*

Note that (iii) is a consequence of (iv).

Example 4.0.5. *Let $U = \mathbb{R}^2 \setminus \{(0,0)\}$. By item (iv), the function $f : U \to \mathbb{R}$ given by $f(x_1, x_2) = \ln(x_1^2 + x_2^2)$ is continuous on U.*

We give the following two interesting results. The first one will be handy when we deal with optimisation of functions. The second one will only be used as an auxiliary result to study the rule of differentiation under the integral sign.

For a proof of the following result, see [3], Thm. VIII.§1.5 and VIII.2.2.

Proposition 4.0.6. *Let R be a bounded and closed subset[1] of $\mathbb{R}^n$. Any continuous function $f : R \to \mathbb{R}$ has a maximum and a minimum in R, i.e. there exist $\vec{a}$ and $\vec{b}$ in R such that $f(\vec{a}) \leq f(\vec{x}) \leq f(\vec{b})$ for all $\vec{x} \in R$.*

For a proof of the following result, see [3], Thm. II.§4.6.

Proposition 4.0.7. *Let R be a bounded and closed subset of $\mathbb{R}^n$. Any continuous function $f : R \to \mathbb{R}$ is* ***uniformly continuous*** *on R, i.e. for every $\epsilon > 0$ there exists $\delta > 0$ such that for all $\vec{x}, \vec{y}$, the condition $\|\vec{x} - \vec{y}\| \leq \delta$ implies $|f(\vec{x}) - f(\vec{y})| \leq \epsilon$.*

[1] A subset of $\mathbb{R}^n$ is said to be **closed** if its complement is open. It is **bounded** if is included in a product of intervals, each of which is of finite length.

Chapter 5

Differentiability

5.1 Basic definitions and properties

Definition 5.1.1. *Consider $n \in \mathbb{Z}_{>0}$. Let $U \subseteq \mathbb{R}^n$ be an open subset, let $f : U \to \mathbb{R}$ be a map and $\vec{a} \in U$ be a point of U. Given a nonzero vector $\vec{v} \in \mathbb{R}^n$, the **directional derivative** of f at $\vec{a}$ along $\vec{v}$ is defined as the limit*

$$\frac{\partial f}{\partial \vec{v}}(\vec{a}) = \lim_{h \to 0} \frac{f(\vec{a}+h\vec{v}) - f(\vec{a})}{h}, \tag{5.1.1}$$

*provided it exists (i.e. it gives an element of $\mathbb{R}$). Given $i = 1,\dots,n$, the **i-th partial derivative***

$$\frac{\partial f}{\partial x_i}(\vec{a})$$

of f at $\vec{a}$ is defined as the directional derivative of f at $\vec{a}$ along the vector $\vec{e}_i$. It is also denoted by $D_i f(\vec{a})$.

We say that $f : U \to \mathbb{R}$ is C^1 on U if all the partial derivatives of f exist at every point of U, and the functions $D_i f : U \to \mathbb{R}$ are continuous on U for all $i = 1,\dots,n$. More generally, a function $f : U \to \mathbb{R}^m$, for which we write $f(\vec{x}) = (f_1(\vec{x}),\dots,f_m(\vec{x}))$ for all $\vec{x} \in U$, is said to be C^1 on U if the component function $f_j : U \to \mathbb{R}$ is C^1 on U for all $j = 1,\dots,m$.

Definition 5.1.2. *Consider a C^1 map $f : U \to \mathbb{R}^m$. We write $f(\vec{x}) = (f_1(\vec{x}),\dots,f_m(\vec{x}))$, for all $\vec{x} \in U$. The **Jacobian** $J_f(\vec{a})$ of f at a point $\vec{a}$ in U is defined as the $m \times n$ matrix*

$$J_f(\vec{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{a}) \\ \vdots & \dots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{a}) & \dots & \frac{\partial f_m}{\partial x_n}(\vec{a}) \end{pmatrix}.$$

It is also denoted by $Df(\vec{a})$. In case $m = 1$, i.e. we have $f : U \to \mathbb{R}$, the $1 \times n$ matrix $J_f(\vec{a})$ is only a vector of size n of the form

$$\left(\frac{\partial f}{\partial x_1}(\vec{a}),\dots,\frac{\partial f}{\partial x_n}(\vec{a})\right)$$

*and it is called the **gradient** of f at $\vec{a}$. It is usually denoted by $\nabla f(\vec{a})$.*

The gradient appears naturally in physics. We give a rather simple example of it.

Example 5.1.3. ***(Heat flux)*** *Suppose we have a mass distribution inside a region $U \subseteq \mathbb{R}^3$, and let $T : U \to \mathbb{R}$ be the scalar field indicating the temperature at each point of the material. Then the* ***heat flux*** *$\vec{q}$ is a vector field on U given by $\vec{q}(\vec{x}) = -K.\nabla T(\vec{x})$, for all $\vec{x} \in U$, where K is a constant, called the* ***thermal conductivity*** *of the material.*

The following operations were recalled in the list of properties in Lemma 4.0.4. The proof of the first three items is a direct consequence of the mentioned result, and for the last one (but also for the first three) the reader may see [3], Section XVII.§2.

Lemma 5.1.4. *Let $U \subseteq \mathbb{R}^n$ be an open subset, $\lambda \in \mathbb{R}$ and $f, g : U \to \mathbb{R}^m$ be two C^1 mappings.*

(i) the ***linear combination*** *$f + \lambda g : U \to \mathbb{R}^m$ is C^1, and $J_{f+\lambda g}(\vec{a}) = J_f(\vec{a}) + \lambda J_g(\vec{a})$;*

(ii) the ***pointwise scalar product*** *$f \cdot g : U \to \mathbb{R}^m$ is C^1, and $\nabla(f \cdot g)(\vec{a}) = g(\vec{a})J_f(\vec{a}) + f(\vec{a})J_g(\vec{a})$, where we identify a vector size m with a $1 \times m$ matrix and we are using the matrix product;*

(iii) if $m = 1$ and $f(\vec{x}) \neq 0$ for all $\vec{x} \in U$, the ***multiplicative inverse*** *$1/f : U \to \mathbb{R}$ is C^1, and $\nabla(1/f)(\vec{a}) = (-1/f(\vec{a})^2)\nabla f(\vec{a})$;*

(iv) suppose that the image of f is included in an open set $V \subseteq \mathbb{R}^m$ and let $h : V \to \mathbb{R}^p$ be a C^1 map. Then, the ***composition*** *$h \circ f : U \to \mathbb{R}^p$ is also C^1 and $J_{h\circ f}(\vec{a}) = J_h(f(\vec{a})).J_f(\vec{a})$. The expression of the Jacobian of a composition given previously is called the* ***chain rule****, and it is one of the most useful identities of elementary differential calculus.*

The following result is a direct consequence of Lemma 5.1.4, (iv), (for another proof see [3], Thm. XV.§2.1).

Proposition 5.1.5. *Let $f : U \to \mathbb{R}^m$ be a C^1 map. Then f is continuous on U. Moreover,*

$$\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{a}+\vec{h}) - f(\vec{a}) - J_f(\vec{a}).\vec{h}^t}{\|\vec{h}\|} = \vec{0},$$

for all $\vec{a} \in U$.

Proof. Since U is open, there exists $\vec{h} \in \mathbb{R}^n$ such that $\vec{a} + t\vec{h} \in U$, for all $t \in [-1, 2]$. Consider the smooth parameterized curve $\vec{\alpha} :]-1, 2[\to U$ given by $\vec{\alpha}(t) = \vec{a} + t\vec{h}$. Since $\vec{\alpha}$ is C^∞, *a fortiori* it is C^1, so the composition with the C^1 map f is also C^1. Define $\vec{\beta} :]-1, 2[\to \mathbb{R}^m$ as the composition $f \circ \vec{\alpha}$. The chain rule tells us that $\vec{\beta}$ is C^1. The theorem of Barrow for functions of one variables for each component of $\vec{\beta}$ implies that

$$\vec{\beta}(1) - \vec{\beta}(0) = \int_0^1 \vec{\beta}'(t)dt.$$

By the chain rule we see that $\vec{\beta}'(t) = J_f(\vec{a} + t\vec{h}).\vec{h}^t$, so

$$\begin{aligned}
\frac{f(\vec{a}+\vec{h}) - f(\vec{a}) - J_f(\vec{a}).\vec{h}^t}{\|\vec{h}\|} &= \int_0^1 J_f(\vec{a}+t\vec{h}).\Big(\frac{\vec{h}}{\|\vec{h}\|}\Big)^t dt - \int_0^1 J_f(\vec{a}).\Big(\frac{\vec{h}}{\|\vec{h}\|}\Big)^t dt \\
&= \int_0^1 \bigg(J_f(\vec{a}+t\vec{h}).\Big(\frac{\vec{h}}{\|\vec{h}\|}\Big)^t - J_f(\vec{a}).\Big(\frac{\vec{h}}{\|\vec{h}\|}\Big)^t\bigg) dt \\
&= \underbrace{\Big(\int_0^1 \big(J_f(\vec{a}+t\vec{h}) - J_f(\vec{a})\big)dt\Big)}_{\spadesuit_{\vec{h}}}.\Big(\frac{\vec{h}}{\|\vec{h}\|}\Big)^t,
\end{aligned}$$

where the integral in the last member indicates that we consider the integral componentwise. Hence, it suffices to show that $\spadesuit_{\vec{h}}$ goes to the zero matrix as $\vec{h} \to \vec{0}$. This follows directly from the continuity of the partial derivatives of f. □

A direct consequence of the important chain rule mentioned before is the following.

Corollary 5.1.6. *Let $U \subseteq \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}$ be a C^1 map. Given any nonzero vector $\vec{v} \in \mathbb{R}^n$ and any $\vec{a} \in U$, the directional derivative of f at $\vec{a}$ along $\vec{v}$ exists and it is given by*

$$\frac{\partial f}{\partial \vec{v}}(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{v}.$$

Proof. Since U is open, there exists $\varepsilon > 0$ such that $\vec{a} + t\vec{v} \in U$, for all $t \in [-\varepsilon, \varepsilon]$. Consider the smooth parameterized curve $\vec{\alpha} :]-\varepsilon, \varepsilon[\to U$ given by $\vec{\alpha}(t) = \vec{a} + t\vec{v}$. Since $\vec{\alpha}$ is C^∞, *a fortiori* it is C^1, so the composition with the C^1 map f is also C^1. Now, the chain rule tells us that the derivative of the composition map $f \circ \vec{\alpha} :]-\varepsilon, \varepsilon[\to \mathbb{R}$ at 0 is given by $\nabla f(\vec{\alpha}(0)) \cdot \vec{\alpha}'(0) = \nabla f(\vec{a}) \cdot \vec{v}$, because $\vec{\alpha}(0) = \vec{a}$ and $\vec{\alpha}'(0) = \vec{v}$. On the other hand, the Newton's incremental quotient definition of the derivative of the real function $f \circ \vec{\alpha}$ is precisely the directional derivative of f at $\vec{a}$ along $\vec{v}$. The corollary thus follows. □

The previous result tells us that we do not need to use the definition of directional derivative if we are dealing with a map of class C^1, since it can be computed by means of the gradient. We give an example of this situation.

Example 5.1.7. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x_1, x_2) = x_1^2 + x_2^3$, and let $\vec{v} = (1, 2)$. We want to show that the directional derivative of f at $\vec{a} = (-1, 3)$ along $\vec{v}$ exists and find its value. We first note that f is a polynomial function, so it is C^1. The partial derivatives are given by*

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 2x_1, \qquad \text{and} \qquad \frac{\partial f}{\partial x_2}(x_1, x_2) = 3x_2^2.$$

The gradient of f at $(-1, 3)$ is thus $\nabla f(-1, 3) = (-2, 27)$. By the previous corollary, we have that

$$\frac{\partial f}{\partial \vec{v}}(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{v} = (-2, 27) \cdot (1, 2) = 52.$$

Corollary 5.1.8. *Let $U, V \subseteq \mathbb{R}^n$ be two open subsets and $f : U \to V$ be a bijective C^1 map. Assume that the inverse f^{-1} of f is also C^1. Let $\vec{a} \in U$ and $\vec{b} = f(a) \in V$. Then*

$$J_{f^{-1}}(\vec{b}) = J_f(\vec{a})^{-1},$$

where the last inversion indicates that of the Jacobian matrix of f at $\vec{a}$.

Proof. We first note that the Jacobian of the identity map at any point is the identity $n \times n$ matrix Id_n. Hence, if we consider the Jacobian of the following compositions $f^{-1} \circ f = \mathrm{id}_U$ and $f \circ f^{-1} = \mathrm{id}_V$ at $\vec{a}$ and at $\vec{b}$, respectively, and we apply the chain rule, we obtain that

$$J_{f^{-1}}(\vec{b}).J_f(\vec{a}) = \mathrm{Id}_n = J_f(\vec{a}).J_{f^{-1}}(\vec{b}).$$

These identities imply that $J_{f^{-1}}(\vec{b})$ is the inverse matrix of $J_f(\vec{a})$ and the corollary follows. □

5.2 Applications: Some geometry

Let $U \subseteq \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}$ be a C^1 map. We recall that the level set L_c of f of level $c \in \mathbb{R}$ is the set

$$L_c = \{\vec{x} \in U : f(\vec{x}) = c\}.$$

Consider $\vec{a} \in L_c$ such that $\nabla f(\vec{a})$ is a nonzero vector. Given $\epsilon > 0$, let $\vec{\alpha} :]-\epsilon, \epsilon[\to L_c$ be any smooth parameterized curve such that $\alpha(0) = \vec{a}$. The vector $\vec{\alpha}'(0)$ is by definition tangent to the level set L_c. Since the function $f \circ \vec{\alpha}$ is constant (of value c), its derivative vanishes. The chain rule says thus that

$$0 = (f \circ \vec{\alpha})'(t) = \nabla f(\vec{\alpha}(t)) \cdot \vec{\alpha}'(t),$$

for all $t \in]-\epsilon, \epsilon[$. In particular, for $t = 0$ we see that

$$0 = \nabla f(\vec{a}) \cdot \vec{\alpha}'(0),$$

for any such curve $\vec{\alpha}$. Conversely, the Implicit Function Theorem implies that one can show that given any vector $\vec{w}$ such that $0 = \nabla f(\vec{a}) \cdot \vec{w}$, there is one smooth parameterized curve $\vec{\alpha} :]-\epsilon, \epsilon[\to L_c$ (for $\epsilon > 0$ sufficiently small) such that $\alpha(0) = \vec{a}$ and $\vec{w} = \vec{\alpha}'(0)$ (see [3], Cor. XVIII.§4.7). This gives the following result, which is almost a slogan.

Fact 5.2.1. *Let $U \subseteq \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}$ be a C^1 map. Consider $\vec{a} \in L_c$ in the level set of f of level $c \in \mathbb{R}$ such that $\nabla f(\vec{a})$ is a nonzero vector. Then the gradient vector $\nabla f(\vec{a})$ is orthogonal to the level set L_c at the point $\vec{a}$.*

5.3 Applications: The differentials of some changes of variables

In this section we will apply Corollary 5.1.8 to compute the Jacobian of some changes of variables we have considered in Section 2.4. Before doing so, let us define more precisely what a change of variables means.

Definition 5.3.1. *A **change of coordinates** (or **variables**) of class C^1 is a bijective map $\mathscr{T} : S \to T$, for $S, T \subseteq \mathbb{R}^n$, such that there exist open sets $U, V \subseteq \mathbb{R}^n$ and C^1 maps $\tilde{\mathscr{T}} : U \to V$ and $\tilde{\mathscr{U}} : V \to U$ satisfying that $S \subseteq U$, $T \subseteq V$, $\tilde{\mathscr{T}}(\vec{x}) = \mathscr{T}(\vec{x})$, for all $\vec{x} \in S$, and $\tilde{\mathscr{U}}(\vec{y}) = \mathscr{T}^{-1}(\vec{y})$, for all $\vec{y} \in T$. We will write in this case $J_{\mathscr{T}}(\vec{x})$ instead of $J_{\tilde{\mathscr{T}}}(\vec{x})$ for all $\vec{x} \in S$.*

Note that if $\mathscr{T}$ as above is a change of coordinates, then $\mathscr{T}^{-1}$ also. Moreover, it is easy to check that the maps $\mathscr{P}$, $\mathscr{C}$, and $\mathscr{S}$ of Section 2.4 satisfy this definition.

5.3.1 Polar coordinates

Recall the map $\mathscr{P} : \mathbb{R}_{>0} \times [0, 2\pi[\to \mathbb{R}^2 \setminus \{(0,0)\}$ given by

$$\mathscr{P}(r, \theta) = \big(r\cos(\theta), r\sin(\theta)\big), \tag{5.3.1}$$

where $r \in \mathbb{R}_{>0}$ and $\theta \in [0, 2\pi[$. Since each component of $\mathscr{P}$ is C^1, for they are products of maps of class C^1, we see that $\mathscr{P}$ is C^1. The Jacobian is given by

$$J_{\mathscr{P}}(r, \theta) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}.$$

The computation of the Jacobian matrix of $\mathscr{P}^{-1}$ would imply in principle that we need to calculate the derivative of $\mathrm{atan2}$ given in (2.4.2), which is rather tedious. On the other hand, Corollary 5.1.8 tells us that the Jacobian of $\mathscr{P}^{-1}$ can also be computed as the inverse of the matrix $J_{\mathscr{P}}(r,\theta)$, which in this case is simpler. More precisely,

$$J_{\mathscr{P}^{-1}}(r\cos(\theta), r\sin(\theta)) = J_{\mathscr{P}}(r,\theta)^{-1} = \frac{1}{r}\begin{pmatrix} r\cos(\theta) & r\sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

5.3.2 Cylindrical coordinates

This case follows immediately from the previous one. More precisely, the map $\mathscr{C} : \mathbb{R}_{>0} \times [0,2\pi[\times \mathbb{R} \to \mathbb{R}^3 \setminus (\{(0,0)\} \times \mathbb{R})$ given by sending (r,θ,z) to $(\mathscr{P}(r,\theta),z)$ is bijective, where the map $\mathscr{P}$ is defined in (5.3.1), is clearly C^1, because each of its components is so. The Jacobian is given by

$$J_{\mathscr{C}}(r,\theta,z) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) & 0 \\ \sin(\theta) & r\cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

By Corollary 5.1.8, the Jacobian of $\mathscr{C}^{-1}$ can be computed as the inverse of the matrix $J_{\mathscr{C}}(r,\theta,z)$, which in this case is simpler. Indeed,

$$J_{\mathscr{C}^{-1}}(r\cos(\theta), r\sin(\theta), z) = J_{\mathscr{C}}(r,\theta,z)^{-1} = \frac{1}{r}\begin{pmatrix} r\cos(\theta) & r\sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

5.3.3 Spherical coordinates

Recall the mapping $\mathscr{S} : \mathbb{R}_{>0} \times [0,2\pi[\times]0,\pi[\to \mathbb{R}^3 \setminus (\{(0,0)\} \times \mathbb{R})$ defined as

$$\mathscr{S}(r,\phi,\theta) = \Big(r\cos(\theta)\sin(\phi), r\sin(\theta)\sin(\phi), r\cos(\phi)\Big), \tag{5.3.2}$$

where $r \in \mathbb{R}_{>0}$, $\phi \in [0,2\pi[$, and $\theta \in]0,\pi[$. It is clear that $\mathscr{S}$ is C^1, for each of its components is a product of C^1 functions. The Jacobian is given by

$$J_{\mathscr{S}}(r,\theta,z) = \begin{pmatrix} \sin(\phi)\cos(\theta) & r\cos(\phi)\cos(\theta) & -r\sin(\theta)\sin(\phi) \\ \sin(\phi)\sin(\theta) & r\cos(\phi)\sin(\theta) & r\sin(\phi)\cos(\theta) \\ \cos(\phi) & -r\sin(\phi) & 0 \end{pmatrix}.$$

Hence, Corollary 5.1.8 gives us that

$$\begin{aligned} &J_{\mathscr{S}^{-1}}\Big(r\cos(\theta)\sin(\phi), r\sin(\theta)\sin(\phi), r\cos(\phi)\Big) \\ &= J_{\mathscr{S}}(r,\phi,\theta)^{-1} = \frac{1}{r}\begin{pmatrix} r\sin(\phi)\cos(\theta) & r\sin(\phi)\sin(\theta) & r\cos(\phi) \\ \cos(\phi)\cos(\theta) & \cos(\phi)\sin(\theta) & -\sin(\phi) \\ -\frac{\sin(\theta)}{\sin(\phi)} & \frac{\cos(\theta)}{\sin(\phi)} & 0 \end{pmatrix}. \end{aligned}$$

5.4 Differentiation under the integral sign

Consider a map $f : [a,b] \times [c,d] \to \mathbb{R}$, where a,b,c,d are real numbers, and suppose that the partial derivatives of f exist (almost) everywhere. The aim of this section is to give sufficient conditions on f in order to have that

$$\frac{d}{dx_2}\left(\int_a^b f(x_1,x_2)dx_1\right) = \int_a^b \frac{\partial f}{\partial x_2} f(x_1,x_2)dx_1. \tag{5.4.1}$$

Note that the meaning of the previous equality is that the map $\Phi : [c,d] \to \mathbb{R}$ given by

$$\Phi(x_2) = \int_a^b f(x_1, x_2) dx_1$$

is differentiable and its derivative is the right member of (5.4.1).

Remark 5.4.1. *This type of situation appears typically in physics. For instance, consider we have a time-dependent arrangement of charges distributed in a straight line, which we suppose to be one of the coordinate axes. In this case, we have a linear density $\lambda(x,t)$ of charges depending on time t and the position x of the axis. The total charge in the segment $I = [a,b]$ of the axis at a fixed time t will be thus*

$$Q_I(t) = \int_a^b \lambda(x,t) dx.$$

The rate of change of the charge inside of region I is precisely the derivative of Q_I with respect to time and it is part of the ***continuity equation****.*

However, (5.4.1) need not be true, as the following example, which typically appears in the theory of shock waves, shows.

Example 5.4.2. *Consider the map $f : \mathbb{R}^2 \to \mathbb{R}$ given by*

$$f(x,t) = \begin{cases} 1, & \text{if } x < t, \\ 0, & \text{else.} \end{cases}$$

Define $\Phi :]0,1[\to \mathbb{R}$ by

$$\Phi(t) = \int_0^1 f(x,t) dx.$$

Hence,

$$\Phi(t) = \int_0^1 f(x,t) dx = \int_0^t f(x,t) dx = t.$$

The derivative of Φ at every $t \in]0,1[$ exists and it is the constant function 1.

On the other side, we have that

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

for all $x \neq t$. As a consequence,

$$\int_0^1 \frac{\partial f}{\partial x} f(x,t) dx = \int_0^t \frac{\partial f}{\partial x} f(x,t) dx + \int_t^1 \frac{\partial f}{\partial x} f(x,t) dx = 0.$$

A possible complain against the previous example is that the partial derivative of f is not defined everywhere. The next (more involved) function shows that the derivative may well exist without having the equality in (5.4.1).

Example 5.4.3. *Consider the map $f : \mathbb{R}^2 \to \mathbb{R}$ given by*

$$f(x_1, x_2) = \begin{cases} \operatorname{sgn}(x_1) \frac{x_1^2 - x_2^2}{x_1^2}, & \text{if } 0 < |x_2| < |x_1|, \\ 0, & \text{else.} \end{cases}$$

Then

$$\Phi(x_1) = \int_{-1}^1 f(x_1, x_2) dx_2 = \frac{4x_1}{3},$$

for $-1 \leq x_1 \leq 1$. *The derivative at any point* $x_1 \in]-1,1[$ *is just* $4/3$. *In particular* $\Phi'(0) = 4/3$.

On the other hand,

$$\frac{\partial f}{\partial x_2} f(0, x_2) = 0,$$

for all $x_2 \in \mathbb{R}$*, so*

$$\int_{-1}^{1} \frac{\partial f}{\partial x_2} f(0, x_2) = 0,$$

which does not coincide with $\Phi'(0) = 4/3$.

We have however the following result.

Proposition 5.4.4. *Let* $f : [a,b] \times [c,d] \to \mathbb{R}$ *be a* C^1 *map. Then the map* $\Phi : [a,b] \to \mathbb{R}$ *given by*

$$\Phi(x_1) = \int_c^d f(x_1, x_2) dx_2$$

is differentiable and

$$\frac{d}{dx_1}\left(\int_c^d f(x_1, x_2) dx_2\right) = \int_c^d \frac{\partial f}{\partial x_1} f(x_1, x_2) dx_2. \tag{5.4.2}$$

Proof. Fix $x_1 \in [a,b]$. If $x_1 + h \in [a,b]$, and $h \neq 0$, then

$$\frac{(\Phi(x_1 + h) - \Phi(x_1))}{h} = \int_c^d \frac{f(x_1 + h, x_2) - f(x_1, x_2)}{h} dx_2.$$

For each $c \leq x_2 \leq d$, we can apply the mean value theorem to the map $t \mapsto f(t, x_2)$ to obtain $0 < \theta_{x_2} < 1$ such that

$$f(x_1 + h, x_2) - f(x_1, x_2) = h \frac{\partial f}{\partial x_1}(x_1 + \theta_{x_2} h, x_2).$$

Let $\epsilon > 0$ be given. By the uniform continuity of $\partial f / \partial x$ on $[a,b] \times [c,d]$, due to Proposition 4.0.7, there exists a $\delta > 0$ for which

$$\left| \frac{\partial f}{\partial x_1}(s', t') - \frac{\partial f}{\partial x_1}(s, t) \right| < \epsilon,$$

if $|s' - s| < \delta$ and $|t' - t| < \delta$. Hence, if $0 < |h| < \delta$ we have

$$\left| \frac{(\Phi(x_1 + h) - \Phi(x_1))}{h} - \int_c^d \frac{\partial f}{\partial x_1}(x_1, x_2) dx_2 \right|$$
$$\leq \int_c^d \left| \frac{\partial f}{\partial x_1}(x_1 + \theta_{x_2} h, x_2) - \frac{\partial f}{\partial x_1}(x_1, x_2) \right| dx_2 \leq \epsilon(d - c).$$

The proposition is thus proved. □

5.5 Higher differentiability

5.5.1 Basic definitions and properties

Definition 5.5.1. *Let* $U \subseteq \mathbb{R}^n$ *be an open subset. Consider a map* $f : U \to \mathbb{R}^m$. *We write* $f(\vec{x}) = (f_1(\vec{x}), \ldots, f_m(\vec{x}))$, *for all* $\vec{x} \in U$. *Suppose that all partial derivatives*

$$\left\{ \frac{\partial f_i}{\partial x_j}(\vec{x}) : i = 1, \ldots, m, j = 1, \ldots, n \right\}$$

*exist for all $\vec{x} \in U$. We say that f **is of class** C^2 if all the previous partial derivatives of the components of f are of class C^1. More generally, given any integer $p \geq 2$, we say that f is of class C^p if if all the partial derivatives of all the components of f are of class C^{p-1}. Finally, we say that f **is of class** C^∞ if it is C^p for all $p \in \mathbb{Z}_{>0}$.*

Remark 5.5.2. *Note that the results of Lemma 5.1.4 hold with C^p instead of C^1, for $p \in \mathbb{Z}_{>0} \cup \{\infty\}$.*

The following result is attributed to K. Schwarz or to A.-C. Clairaut, and it shows the usefulness of being C^2.

Proposition 5.5.3. *Let $U \subseteq \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}$ be a map of class C^2 on U. Then, given any pair of indices $i, j = 1, \dots, n$ we have that*

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

for all $\vec{a} \in U$.

Proof. Without loss of generality we will assume that $n = 2$, $i = 1$ and $j = 2$. By the integral version of the mean theorem for the function $x_2 \mapsto (\partial f/\partial x_1)(a_1, x_2)$ we see that

$$\frac{\partial f}{\partial x_1}(a_1, x_2) = \frac{\partial f}{\partial x_1}(a_1, a_2) + \int_{a_2}^{x_2} \frac{\partial^2 f}{\partial x_2 \partial x_1}(a_1, y)dy. \tag{5.5.1}$$

The same argument applies to map $x_2 \mapsto f(x_1, x_2)$, where we consider that x_1 is fixed but arbitrary. We obtain thus

$$f(x_1, x_2) = f(x_1, a_2) + \int_{a_2}^{x_2} \frac{\partial f}{\partial x_2}(x_1, y)dy.$$

We now fix x_2 and take the (partial) derivative with respect to x_1 of the last maps appearing in the last equality to get

$$\begin{aligned} \frac{\partial f}{\partial x_1}(x_1, x_2) &= \frac{\partial f}{\partial x_1}(x_1, a_2) + \frac{\partial}{\partial x_1}\left(\int_{a_2}^{x_2} \frac{\partial f}{\partial x_2}(x_1, y)dy\right) \\ &= \frac{\partial f}{\partial x_1}(x_1, a_2) + \int_{a_2}^{x_2} \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, y)dy, \end{aligned}$$

where we have used Proposition 5.4.4 in the last equality. By choosing $x_1 = a_1$ in the previous identity we find thus

$$\frac{\partial f}{\partial x_1}(a_1, x_2) = \frac{\partial f}{\partial x_1}(a_1, a_2) + \int_{a_2}^{x_2} \frac{\partial^2 f}{\partial x_1 \partial x_2}(a_1, y)dy. \tag{5.5.2}$$

By comparing (5.5.1) and (5.5.2), we get

$$\int_{a_2}^{x_2} \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, y)dy = \int_{a_2}^{x_2} \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, y)dy,$$

and a final derivative with respect to x_2 and taking $x_2 = a_2$ gives the statement. □

5.5.2 Geometric interlude: Tangent planes

Definition 5.5.4. *Let $U \subseteq \mathbb{R}^2$ be an open subset. A **smooth parameterized surface** in $\mathbb{R}^n$ is a C^∞ map $\vec{\sigma} : U \to \mathbb{R}^n$. Write $\vec{\sigma}(x_1, x_2) = (\sigma_1(x_1, x_2), \dots, \sigma_n(x_1, x_2))$, for all $\vec{x} \in U$. We further say that it is **regular** if for all $\vec{x} \in U$ the set*

$$\left\{\frac{\partial \vec{\sigma}}{\partial x_1}(\vec{x}), \frac{\partial \vec{\sigma}}{\partial x_2}(\vec{x})\right\} \subseteq \mathbb{R}^n$$

is linearly independent, where

$$\frac{\partial \vec{\sigma}}{\partial x_i}(\vec{x}) = \left(\frac{\partial \sigma_1}{\partial x_i}(\vec{x}), \dots, \frac{\partial \sigma_n}{\partial x_i}(\vec{x}) \right),$$

for $i = 1, 2$.

Remark 5.5.5. *Let* $U \subseteq \mathbb{R}^2$ *be an open subset and* $f : U \to \mathbb{R}$ *be a* C^∞ *map. Then* $\vec{\sigma} : U \to \mathbb{R}^3$ *given by sending* (x_1, x_2) *to* $(x_1, x_2, f(x_1, x_2))$ *is a regular smooth parameterized surface in* $\mathbb{R}^3$*. Most of the examples we will consider are of this form. Furthermore, the Inverse Function Theorem (see* [3]*, Thm. XVIII.3.1) tells us that, at least locally, they are all the possible examples (up to permutation of the components of* $\vec{\sigma}$*).*

Example 5.5.6. *Let* $U = \mathbb{R}^2$ *and consider* $f : \mathbb{R}^2 \to \mathbb{R}$ *given by sending* (x_1, x_2) *to* $x_1^2 - x_2^2$*. The previous remark gives us a regular smooth parameterized surface* $\vec{\sigma} : \mathbb{R}^2 \to \mathbb{R}^3$ *given by sending* (x_1, x_2) *to* $(x_1, x_2, x_1^2 - x_2^2)$*.*

Fact 5.5.7. *Let* $U \subseteq \mathbb{R}^2$ *be an open subset and* $\vec{\sigma} : U \to \mathbb{R}^n$ *regular smooth parameterized surface in* $\mathbb{R}^n$*. The* ***tangent plane*** *to* $\vec{\sigma}$ *at* $\vec{\sigma}(\vec{a})$*, where* $\vec{a} \in U$*, is given by the regular smooth parameterized surface* $\vec{\varsigma} : \mathbb{R}^2 \to \mathbb{R}^n$ *defined as*

$$\vec{\varsigma}(s, t) = \vec{\sigma}(\vec{a}) + s \frac{\partial \vec{\sigma}}{\partial x_1}(\vec{a}) + t \frac{\partial \vec{\sigma}}{\partial x_2}(\vec{a}).$$

If $n = 3$*, the equation of this plane (see Subsubsection 2.2.2.2) is*

$$\left\{ \vec{x} \in \mathbb{R}^3 : \left(\vec{x} - \vec{\sigma}(\vec{a}) \right) \cdot \vec{n} = 0 \right\},$$

where

$$\vec{n} = \frac{\partial \vec{\sigma}}{\partial x_1}(\vec{a}) \times \frac{\partial \vec{\sigma}}{\partial x_2}(\vec{a}).$$

5.5.3 Applications: Thermodynamics (2)

This subsection is a continuation of Section 3.2. The following result indicates the solution of the optimisation problem stated in Remark 3.2.2. The proof is based on the theory of Lagrange multipliers, that will be explained in the sequel.

Proposition 5.5.8 ([5], Thm. 1.2.(i))**.** *Assume that the state function* Δ *of the standard thermodynamic system is* C^1*. If the vector* $({}^0T, {}^0P, {}^0\vec{\mu}, {}^0U, {}^0S, {}^0V, {}^0\vec{n})$ *is an equilibrium state, then there is a real number* Ω*, called the* ***system size****, such that*

$$ {}^0S = \Omega \frac{\partial \Delta}{\partial T}({}^0T, {}^0P, {}^0\vec{\mu}), \quad {}^0V = -\Omega \frac{\partial \Delta}{\partial P}({}^0T, {}^0P, {}^0\vec{\mu}), \text{ and } {}^0n_j = \Omega \frac{\partial \Delta}{\partial \mu_j}({}^0T, {}^0P, {}^0\vec{\mu}), \tag{5.5.3}$$

for all $j = 1, \dots, J$*. Note the second identity implies that* $\Omega > 0$*.*

The last collection of J equations will be written more compactly as

$${}^0\vec{n} = \Omega \frac{\partial \Delta}{\partial \vec{\mu}}({}^0T, {}^0P, {}^0\vec{\mu}).$$

Definition 5.5.9. *Given* $\omega \in \mathbb{R}_{>0}$*, a standard thermodynamic system is said to be of* ***(constant) size*** ω *if its equilibrium states are elements of the form*

$$\left\{ (T, P, \vec{\mu}, U, S, V, \vec{n}) \in \mathcal{E}q_\Delta : S = \omega \frac{\partial \Delta}{\partial T}(T, P, \vec{\mu}), V = -\omega \frac{\partial \Delta}{\partial P}(T, P, \vec{\mu}), \vec{n} = \omega \frac{\partial \Delta}{\partial \vec{\mu}}(T, P, \vec{\mu}) \right\}.$$

The following result is a corollary of Proposition 5.5.3.

Proposition 5.5.10 ([5], Thm. 1.2.(ii))**.** *Assume now that the state function* Δ *of the standard thermodynamic system is* C^2 *and that the standard thermodynamic system is of constant size. Then, we have the following identities, called the* ***Maxwell relations****,*

$$\frac{\partial S}{\partial P} = -\frac{\partial V}{\partial T}, \quad \frac{\partial S}{\partial \mu_j} = \frac{\partial n_j}{\partial T}, \quad \frac{\partial V}{\partial \mu_j} = -\frac{\partial n_j}{\partial P}, \quad \textit{and} \quad \frac{\partial n_j}{\partial \mu_i} = \frac{\partial n_i}{\partial \mu_j},$$

for all $j = 1, \dots, J$*, where the partial derivatives are evaluated at the intensive variables of any equilibrium state.*

5.6 Taylor expansion of degree 2

Let $U \subseteq \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}$ be a map of class C^3 on U. Consider $\vec{a} \in U$ and a nonzero vector $\vec{h} \in \mathbb{R}^n$ such that $\vec{a} + t\vec{h} \in U$ for all $t \in]-1, 2[$. Define the map $g :]-1, 2[\to \mathbb{R}$ given by $g(t) = f(\vec{a} + t\vec{h})$. It is of class C^3, for it is a composition of two C^3 maps. The integral expression of the residue for the Taylor expansion (of degree 2) of the theory of functions on one real variable tells us that

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(0) + \frac{1}{2}\int_0^1 (1-s)^2 g'''(s)ds.$$

On the other hand, the chain rule gives

$$g'(0) = \nabla f(\vec{a}) \cdot \vec{h} \quad \text{and} \quad g''(0) = \vec{h}.H_f(\vec{a}).\vec{h}^t, \tag{5.6.1}$$

where

$$H_f(\vec{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\vec{a}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\vec{a}) \\ \vdots & \dots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\vec{a}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\vec{a}) \end{pmatrix}$$

is called the **Hessian matrix** of f at $\vec{a}$. We note that the product in the last equation of (5.6.1) is the standard multiplication of matrices. Notice also that the condition that f is C^2 implies that the Hessian matrix is symmetric.

Moreover, the chain rule also implies that

$$g'''(s) = \sum_{i,j,k=1}^{n} h_i h_j h_k \underbrace{\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\vec{a} + s\vec{h})}_{\star_{i,j,k}(s)}, \tag{5.6.2}$$

for $s \in]-1, 2[$. The fact that f is C^3 implies that exists a constant C such that $|\star_{i,j,k}(s)| \leq C$ for all $s \in [0, 1]$ and all $i, j, k = 1, \dots, n$. Hence,

$$\left| \frac{1}{2}\int_0^1 (1-s)^2 g'''(s)ds \right| \leq \frac{1}{2} \sum_{i,j,k=1}^{n} \int_0^1 |h_i||h_j||h_k| C ds \leq \frac{C}{2} \sum_{i,j,k=1}^{n} |h_i||h_j||h_k|.$$

As a consequence,

$$\left| \frac{1}{2\|\vec{h}\|^2}\int_0^1 (1-s)^2 g'''(s)ds \right| \leq \frac{C}{2} \sum_{i,j,k=1}^{n} \underbrace{\frac{|h_i|}{\|\vec{h}\|}}_{\leq 1} \underbrace{\frac{|h_j|}{\|\vec{h}\|}}_{\leq 1} |h_k| \leq \frac{Cn^3}{2}\|\vec{h}\|,$$

where we have used the obvious inequality $|h_i| \leq \|\vec{h}\|$ for all $i = 1, \ldots, n$ and for any vector $\vec{h} = (h_1, \ldots, h_n)$ of $\mathbb{R}^n$.

We have thus proved the following result.

Proposition 5.6.1. *Let $U \subseteq \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}$ be a map of class C^3 on U. Consider $\vec{a} \in U$ and a nonzero vector $\vec{h} \in \mathbb{R}^n$ such that $\vec{a} + t\vec{h} \in U$ for all $t \in]-1, 2[$. Then,*

$$f(\vec{a} + \vec{h}) = f(\vec{a}) + \vec{h} \cdot \nabla f(\vec{a}) + \frac{1}{2}\vec{h}.H_f(\vec{a}).\vec{h}^t + \phi(\|\vec{h}\|),$$

where $\phi :]-\varepsilon, \varepsilon[\to \mathbb{R}$ is a function satisfying the $\mathrm{o}(t^2)$ condition, i.e.

$$\lim_{t \to 0} \frac{\phi(t)}{t^2} = 0.$$

5.7 Some differential operators

5.7.1 Basic definitions and properties

In this subsection we shall consider some examples of differential operators of interest in physics (and mathematics). We have already considered the gradient $\nabla f(\vec{x})$ of a scalar function $f : U \to \mathbb{R}$ of class C^1, where U is an open subset of $\mathbb{R}^n$. If f is also C^2, define the **Laplacian** of f as the continuous scalar field $\Delta f : U \to \mathbb{R}$ given by

$$\Delta f(\vec{x}) = \frac{\partial^2 f}{\partial x_1^2}(\vec{x}) + \cdots + \frac{\partial^2 f}{\partial x_n^2}(\vec{x}).$$

Let $\vec{f} : U \to \mathbb{R}^n$ be a vector field of class C^1, where U is an open subset of $\mathbb{R}^n$. We write $\vec{f}(\vec{x}) = (f_1(\vec{x}), \ldots, f_n(\vec{x}))$, for all $\vec{x} \in U$. The **Laplacian** of $\vec{f}$ is the continuous vector field $\Delta\vec{f} : U \to \mathbb{R}^n$ given by

$$\Delta\vec{f}(\vec{x}) = \big(\Delta f_1(\vec{x}), \ldots, \Delta f_n(\vec{x})\big).$$

The **divergence** $\operatorname{div}\vec{f} : U \to \mathbb{R}$ of $\vec{f}$ (also noted $\nabla \cdot \vec{f}$) is the continuous scalar field defined as

$$\operatorname{div}\vec{f}(\vec{x}) = \frac{\partial f_1}{\partial x_1}(\vec{x}) + \cdots + \frac{\partial f_n}{\partial x_n}(\vec{x}).$$

Consider further that $n = 3$. The **curl** (or **rotor**) $\operatorname{rot}\vec{f} : U \to \mathbb{R}^3$ of F (also noted $\nabla \times \vec{f}$) is the continuous vector field defined as

$$\operatorname{rot}\vec{f}(\vec{x}) = \left(\frac{\partial f_2}{\partial x_2}(\vec{x}) - \frac{\partial f_2}{\partial x_3}(\vec{x}), \frac{\partial f_1}{\partial x_3}(\vec{x}) - \frac{\partial f_3}{\partial x_1}(\vec{x}), \frac{\partial f_2}{\partial x_1}(\vec{x}) - \frac{\partial f_1}{\partial x_2}(\vec{x})\right).$$

Let $\vec{f} : U \to \mathbb{R}^n$ be a vector field, where U is an open subset of $\mathbb{R}^n$, and let $h : U \to \mathbb{R}$ be a scalar field of class C^1. Define the continuous scalar field $(\vec{f} \cdot \nabla)h : U \to \mathbb{R}$ by

$$(\vec{f} \cdot \nabla)h(\vec{x}) = f_1(\vec{x})\frac{\partial h}{\partial x_1}(\vec{x}) + \cdots + f_n(\vec{x})\frac{\partial h}{\partial x_n}(\vec{x}),$$

for all $\vec{x} \in U$. Moreover, if $\vec{h} : U \to \mathbb{R}^m$ is a vector map of class C^1, define the continuous vector map $(\vec{f} \cdot \nabla)\vec{h} : U \to \mathbb{R}^m$ by

$$(\vec{f} \cdot \nabla)\vec{h}(\vec{x}) = \big((\vec{f} \cdot \nabla)h_1(\vec{x}), \ldots, (\vec{f} \cdot \nabla)h_m(\vec{x})\big),$$

for all $\vec{x} \in U$, where we have written as usual $\vec{h}(\vec{x}) = (h_1(\vec{x}), \ldots, h_m(\vec{x}))$. The scalar and vector fields defined in the last two displayed equations are called the **material derivatives**.

The following properties are immediate.

Lemma 5.7.1. *Let $\vec{f}, \vec{g} : U \to \mathbb{R}^n$ be two vector field of class C^1, where U is an open subset of $\mathbb{R}^n$, and let $h : U \to \mathbb{R}$ be a scalar field of class C^1. Then,*

(i) $\nabla \cdot (\vec{f} + h\vec{g})(\vec{x}) = \nabla \cdot \vec{f}(\vec{x}) + h(\vec{x})\nabla \cdot \vec{g}(\vec{x}) + \nabla h(\vec{x}) \cdot \vec{g}(\vec{x})$;

(ii) $\nabla \cdot (\nabla h)(\vec{x}) = \Delta h(\vec{x})$, *if h is also assumed to be C^2;*

(iii) $\nabla \times (\vec{f} + h\vec{g})(\vec{x}) = \nabla \times \vec{f}(\vec{x}) + h(\vec{x})\nabla \times \vec{g}(\vec{x}) + \nabla h(\vec{x}) \times \vec{g}(\vec{x})$;

(iv) $\nabla \cdot (\nabla \times f)(\vec{x}) = 0$, *if f is also assumed to be C^2;*

(v) $\nabla \times (\nabla h)(\vec{x}) = 0$, *if h is also assumed to be C^2;*

(vi) $\nabla \times (\nabla \times \vec{f})(\vec{x}) = \nabla(\nabla \cdot \vec{f})(\vec{x}) - \Delta \vec{f}(\vec{x})$, *if $\vec{f}$ is also assumed to be C^2;*

(vii) $\nabla \times (\vec{f} \times \vec{g})(\vec{x}) = \vec{f}(\vec{x})(\nabla \cdot \vec{g})(\vec{x}) - \vec{g}(\vec{x})(\nabla \cdot \vec{f})(\vec{x}) + (\vec{g}(\vec{x}) \cdot \nabla)\vec{f}(\vec{x}) - (\vec{f}(\vec{x}) \cdot \nabla)\vec{g}(\vec{x})$;

(viii) $\nabla \cdot (\vec{f}\vec{g})(\vec{x}) = (\nabla \times \vec{f})(\vec{x}) \cdot \vec{g}(\vec{x}) - (\nabla \times \vec{g})(\vec{x}) \cdot \vec{f}(\vec{x})$;

(ix) $\nabla(\vec{f} \cdot \vec{g})(\vec{x}) = \vec{f}(\vec{x}) \times (\nabla \times \vec{g})(\vec{x}) + \vec{g}(\vec{x}) \times (\nabla \times \vec{f})(\vec{x}) + (\vec{f}(\vec{x}) \cdot \nabla)\vec{g}(\vec{x}) + (\vec{g}(\vec{x}) \cdot \nabla)\vec{f}(\vec{x})$;

for all $\vec{x} \in U$, where we have assumed that $n = 3$ in items (iii)-(ix).

5.7.2 Applications: Electromagnetism

The **electric** and **magnetic fields** are the mathematical construction which allows to describe the forces between charged particles. The are determined by a collection of identities, called the **Maxwell equations**.

Suppose that we are given a distribution of charges determined by a density function $\rho : U \times \mathbb{R} \to \mathbb{R}$, where $U \subseteq \mathbb{R}^3$ is a bounded open subset, and we have a collection of electric currents described by the time-dependent vector field $\vec{J} : U \times \mathbb{R} \to \mathbb{R}^3$, called the **current density**. The last argument of both fields is called the **time** t, so the typical argument will be of the form $(\vec{x}, t)$, for $\vec{x} \in U$. We recall that the current $\vec{j}$ originated by a moving charge q is $\vec{j} = q\vec{v}$, where $\vec{v}$ is the velocity of the charge. The time-dependent scalar field ρ and the vector field $\vec{J}$ are just the infinitesimal contribution (per volume) of the total charge and the total current. They satisfy the so-called **continuity equation** given by

$$\frac{\partial \rho}{\partial t}(\vec{x}, t) + \nabla \cdot \vec{J}(\vec{x}, t) = 0,$$

which is essentially states that charges are preserved.

The electric and magnetic fields will be time-dependent vector fields of the form $\vec{E}, \vec{B} : U \times \mathbb{R} \to \mathbb{R}^3$ that satisfy the **Maxwell equations** (in Gaussian units):

$$\begin{aligned} \nabla \cdot \vec{E}(\vec{x}, t) &= 4\pi\rho(\vec{x}, t), & \nabla \times \vec{E}(\vec{x}, t) &= -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}(\vec{x}, t), \\ \nabla \cdot \vec{B}(\vec{x}, t) &= 0, & \nabla \times \vec{B}(\vec{x}, t) &= -\frac{1}{c}\Big(4\pi\vec{J}(\vec{x}, t) + \frac{\partial \vec{E}}{\partial t}(\vec{x}, t)\Big), \end{aligned}$$

for all $\vec{x} \in U$ and $t \in \mathbb{R}$, where c is the speed of the light. One should further impose "reasonable" boundary conditions in order to assure existence and uniqueness of the solution.

Chapter 6

Basic integration theory

6.1 A crash course on Riemann integration

6.1.1 Basic definitions and properties

We recall that a subset $R \subseteq \mathbb{R}^n$ is called a **rectangle** in $\mathbb{R}^n$ if it is of the form $[a_1, b_1] \times \cdots \times [a_n, b_n]$, where $a_i < b_i$ are real numbers for all $i = 1, \dots, n$. Its **volume** $\mathrm{vol}(R)$ is just $(b_1 - a_1) \dots (b_n - a_n)$. A **partition** of $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ is given a family $P = \{a_i = {}^0a_i \leq {}^1a_i \leq \cdots \leq {}^Na_i = b_i : i = 1, \dots, n\}$, where $N \in \mathbb{Z}_{>0}$. A partition determines a collection $\mathrm{Rec}(P)$ of N^n rectangles, each of which is of the form $[{}^{j_1}a_1, {}^{j_1+1}a_1] \times \cdots \times [{}^{j_n}a_n, {}^{j_n+1}a_n]$, for $j_1, \dots, j_n \in \{0, \dots, N-1\}$. Let R be a rectangle and $f : R \to \mathbb{R}$ be a bounded function. Given a partition P of R as before we define the **lower** and **upper Riemann sums** of f as

$$s_R(P, f) = \sum_{T \in \mathrm{Rec}(P)} \inf\{f(\vec{x}) : \vec{x} \in T\} \mathrm{vol}(T),$$
$$S_R(P, f) = \sum_{T \in \mathrm{Rec}(P)} \sup\{f(\vec{x}) : \vec{x} \in T\} \mathrm{vol}(T),$$

respectively. Define the **lower** and **upper Riemann integrals** of f as

$$s_R(f) = \sup\{s_R(P, f) : P \text{ is a partition of } R\}, \tag{6.1.1}$$

and

$$S_R(f) = \inf\{s_R(P, f) : P \text{ is a partition of } R\}, \tag{6.1.2}$$

respectively.

Definition 6.1.1. *We say that a bounded function $f : R \to \mathbb{R}$ defined on a rectangle R is* ***Riemann integrable*** *if the lower and upper Riemann integrals given in* (6.1.1) *and* (6.1.2) *coincide. In this case, the* ***integral of*** *f* ***over*** *R is defined as $s_R(f) = S_R(f)$ and it is denoted by*

$$\int_R f(\vec{x}) d\vec{x}.$$

We have the following simple result (for a proof, see [3], XX.§1.1).

Lemma 6.1.2. *Let $R \subseteq \mathbb{R}^n$ be a fixed rectangle. Then,*

(i) if $f, g : R \to \mathbb{R}$ are two Riemann integrable functions and $\lambda \in \mathbb{R}$, then the function $f + \lambda g : R \to \mathbb{R}$ is also Riemann integrable and

$$\int_R (f + \lambda g)(\vec{x}) d\vec{x} = \int_R f(\vec{x}) d\vec{x} + \lambda \int_R g(\vec{x}) d\vec{x};$$

(ii) *if $f : R \to \mathbb{R}$ is Riemann integrable and $f(\vec{x}) \geq 0$, for all $\vec{x} \in R$, then*

$$\int_R f(\vec{x}) d\vec{x} \geq 0.$$

A subset S of $\mathbb{R}^n$ is called **negligible** if, given $\epsilon > 0$, there exists a finite collection of rectangles $R_1, \ldots, R_m$ such that

$$S \subseteq R_1 \cup \cdots \cup R_m \text{ and } \mathrm{vol}(R_1) + \cdots + \mathrm{vol}(R_m) < \epsilon.$$

Definition 6.1.3. *A function $f : R \to \mathbb{R}$ defined on a rectangle R is called **admissible** if it is bounded and the set of points of R where f is not continuous is negligible.*

We have a very simple criterion giving sufficient conditions to assure that a set S is negligible (for a proof see [3], Prop. XX.§2.2).

Lemma 6.1.4. *Let $m < n$ be two positive integers, and let $B \subseteq \mathbb{R}^m$ be a bounded set. Assume given a map $f : U \to \mathbb{R}^n$ of class C^1, where $U \subseteq \mathbb{R}^m$ is an open set containing B. Then $f(B)$ is negligible.*

The main result we have is the following (for a proof see [3], Thm. XX.§1.3).

Theorem 6.1.5. *Let R be a rectangle and $f : R \to \mathbb{R}$ be an admissible function. Then it is Riemann integrable. Moreover, if $g : R \to \mathbb{R}$ is another function such that g is bounded and*

$$\left\{\vec{x} \in R : g(\vec{x}) \neq f(\vec{x})\right\}$$

is negligible, then g is also admissible and

$$\int_R f(\vec{x}) d\vec{x} = \int_R g(\vec{x}) d\vec{x}.$$

We say that a set $A \subseteq \mathbb{R}^n$ is **admissible** if it is bounded and its boundary ∂A[1] is a negligible set.

We recall that the **characteristic function** $\chi_A : \mathbb{R}^n \to \mathbb{R}$ of a set $A \subseteq \mathbb{R}^n$ is defined as

$$\chi_A(\vec{x}) = \begin{cases} 1, & \text{if } \vec{x} \in A, \\ 0, & \text{else.} \end{cases}$$

The following result is trivial.

Fact 6.1.6. *The characteristic function χ_A of an admissible set A is admissible.*

Let $f : R \to \mathbb{R}$ be an admissible function defined on a rectangle R, and let A be an admissible subset of R. Since the product of admissible functions is admissible, $f\chi_A$ is admissible, and we define the **Riemann integral of f over A** as

$$\int_R f(\vec{x}) \chi_A(\vec{x}) d\vec{x}.$$

We will denote it by

$$\int_A f(\vec{x}) d\vec{x}.$$

Given an admissible set A and a map $f : A \to \mathbb{R}$, we will say that f is **admissible** (resp., **Riemann integrable**) if there exists a rectangle R including A such that the **extension of f to R by zero** (*i.e.* the map $\bar{f} : R \to \mathbb{R}$ is given by $\bar{f}(\vec{x}) = f(\vec{x})$ if $\vec{x} \in A$, and by $\bar{f}(\vec{x}) = 0$ else) is admissible (resp., Riemann integrable) in the sense of Definition 6.1.3 (resp., Definition 6.1.1). It is easy to show that the previous definitions are independent of the rectangle R.

[1] The boundary of a set $A \subseteq \mathbb{R}^n$ is the set of points $\vec{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n$ satisfying that $([a_1 - \epsilon, a_1 + \epsilon] \times \cdots \times [a_n - \epsilon, a_n + \epsilon]) \cap A \neq \emptyset$ and $([a_1 - \epsilon, a_1 + \epsilon] \times \cdots \times [a_n - \epsilon, a_n + \epsilon]) \cap (\mathbb{R}^n \setminus A) \neq \emptyset$, for all $\epsilon > 0$.

6.1.2 Fubini's theorem

Let $R \subseteq \mathbb{R}^n$ and $S \subseteq \mathbb{R}^m$ be two rectangles and let $f : R \times S \to \mathbb{R}$ be an integrable function. Given $\vec{x} \in R$ define the map $f_{\vec{x}} : S \to \mathbb{R}$ by $f_{\vec{x}}(\vec{y}) = f(\vec{x}, \vec{y})$. Suppose that there is a negligible subset X of R such that $f_{\vec{x}}$ is Riemann integrable for $\vec{x} \in R \setminus X$. We can define in this case

$$\int_S f_{\vec{x}}(\vec{y}) d\vec{y} \tag{6.1.3}$$

for all $\vec{x} \in R \setminus X$. Denote by

$$\int_S f(\vec{x}, \vec{y}) d\vec{y} \tag{6.1.4}$$

the map sending $\vec{x} \in R \setminus X$ to the value (6.1.3), and the elements of X to zero. Then the following result says that the map (6.1.4) is admissible, so integrable by Theorem 6.1.5. Moreover, we have the following (for a proof see [3], Thm. XX.§3.1).

Theorem 6.1.7. *Let $R \subseteq \mathbb{R}^n$ and $S \subseteq \mathbb{R}^m$ be two rectangles and let $f : R \times S \to \mathbb{R}$ be an integrable function. Assume that there is a negligible subset X of R such that $f_{\vec{x}}$ is Riemann integrable for $\vec{x} \in R \setminus X$. Then the map*

$$\int_S f(\vec{x}, \vec{y}) d\vec{y}$$

is integrable and

$$\int_R \left(\int_S f(\vec{x}, \vec{y}) d\vec{y} \right) d\vec{x} = \int_{R \times S} f(\vec{x}, \vec{y}) d(\vec{x}, \vec{y}).$$

Fact 6.1.8. ***(Basic computation)*** *Let $g_1, g_2 : [a, b] \to \mathbb{R}$ be two continuous functions satisfying that $g_1(x) \le g_2(x)$, for all $x \in [a, b]$. Define the closed set*

$$A = \Big\{ (x_1, x_2) \in [a, b] \times \mathbb{R} : g_1(x_1) \le x_2 \le g_2(x_1) \Big\}.$$

*It is not difficult to show that A is an admissible set (prove it!). A planar set of this form is called of **type** 1. Since g_1 and g_2 are continuous on $[a, b]$, they are bounded so A is included in a rectangle $R = [a, b] \times [c, d]$ of $\mathbb{R}^2$. Let $f : A \to \mathbb{R}$ be any continuous function on A, and consider the extension $\bar{f}$ of f by zero to R. Then $\bar{f}$ (or f) is Riemann integrable and the previous theorem tells us that*

$$\int_A f(x_1, x_2 d(x_1, x_2) = \int_R \bar{f}(x_1, x_2 d(x_1, x_2) = \int_a^b \left(\int_{g_1(x_1)}^{g_2(x_1)} f(x_1, x_2) dx_2 \right) dx_1.$$

*The analogous result holds for a planar set A' of **type** 2, i.e. A' is defined as*

$$A' = \Big\{ (x_1, x_2) \in \mathbb{R} \times [c, d] : h_1(x_2) \le x_1 \le h_2(x_2) \Big\},$$

*where $h_1, h_2 : [c, d] \to \mathbb{R}$ are two continuous functions satisfying that $h_1(y) \le h_2(y)$, for all $y \in [c, d]$. Finally, a planar set A is said to be of **type** 3 if it is of type 1 and 2.*

Example 6.1.9. *Suppose D is the triangle*

$$\Big\{ (x_1, x_2) \in \mathbb{R}^2 : 0 \le x_2 \le x_1 \text{ and } 0 \le x_1 \le \pi \Big\}.$$

We want to compute the integral

$$I = \int_D \frac{\sin(x_1)}{x_1} d(x_1, x_2).$$

Theorem 6.1.7 tells us on the one hand that

$$I = \int_0^\pi \Big(\int_{x_2}^\pi \frac{\sin(x_1)}{x_1} dx_1 \Big) dx_2.$$

This expression is completely useless however if we do not know any primitive of the function $\sin(x)/x$*. However, the same theorem also states that*

$$I = \int_0^\pi \Big(\int_0^{x_1} \frac{\sin(x_1)}{x_1} dx_2 \Big) dx_1.$$

The last integral can be directly computed, since

$$\int_0^\pi \Big(\int_0^{x_1} \frac{\sin(x_1)}{x_1} dx_2 \Big) dx_1 = \int_0^\pi \Big[\frac{\sin(x_1)}{x_1} x_2 \Big]_0^{x_1} dx_1 = \int_0^\pi \sin(x_1) dx_1 = 2.$$

6.2 Change of variables formula

6.2.1 The statement

Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}^n$ be a vector field of class C^1. Write $f(\vec{x}) = (f_1(\vec{x}), \ldots, f_n(\vec{x}))$, for all $\vec{x} \in U$, and recall that the Jacobian $J_f(\vec{x})$ of f at $\vec{x} \in U$ is the $n \times n$ matrix

$$J_f(\vec{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}) & \ldots & \frac{\partial f_1}{\partial x_n}(\vec{x}) \\ \vdots & \ldots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\vec{x}) & \ldots & \frac{\partial f_n}{\partial x_n}(\vec{x}) \end{pmatrix}.$$

Define $\Delta_f : U \to \mathbb{R}$ to be the scalar field sending every $\vec{x} \in U$ to the determinant $\det(J_f(\vec{x}))$ of the Jacobian matrix. It is called the **Jacobian determinant**.

The following result is one of the main important achievements of Riemann integration theory. Even though we will give a rather constraint form of it, it can be enhanced by means of the basic theory showed in the previous two subsections to deal with most of the applications. For a nice and simple proof we refer the reader to [6].

Theorem 6.2.1. *Let* $U, V \subseteq \mathbb{R}^n$ *be two bounded open sets and* $f : U \to V$ *be a bijective vector map of class* C^1 *such that its inverse is also* C^1*. If* $g : V \to \mathbb{R}$ *is an admissible map, then* $(g \circ f).|\Delta_f| : U \to \mathbb{R}$ *is also admissible and*

$$\int_V g(\vec{y}) d\vec{y} = \int_U (g \circ f)(\vec{x}) |\Delta_f(\vec{x})| d\vec{x}.$$

6.2.2 Example: A simple integral of Gaussian type

Let $a > 0$ be a positive real number. We want to compute the integral

$$I = \int_D e^{-x_1^2 - x_2^2} d(x_1, x_2),$$

where $D = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 0 \text{ and } x_1^2 + x_2^2 \leq a^2\}$. The first note that the boundary of D is negligible, by Lemma 6.1.4. Then

$$I = \int_V e^{-x_1^2 - x_2^2} d(x_1, x_2),$$

where $V = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0 \text{ and } x_1^2 + x_2^2 < a^2\}$. Note that V is an open set. Consider now the restriction of the map $\mathscr{P}$ defined in (2.4.1) to the open set $U =]0, a[\times]0, \pi[$, and call it f. The image of U under f is precisely V. By Theorem 6.2.1, we see that

$$\int_V e^{-x_1^2 - x_2^2} dx dy = \int_0^a \int_0^\pi e^{-r^2} r dr d\theta,$$

where we have used that $\Delta_f(r, \theta) = r$, for all $(r, \theta) \in U$. This last identity easily follows from the expression of the Jacobian matrix of $\mathscr{P}$ given in Subsection 5.3.1. The last integral is easily computed and it gives

$$\int_0^a \int_0^\pi e^{-r^2} r dr d\theta = \frac{\pi}{2}(1 - e^{-a^2}),$$

6.3 Line integrals

6.3.1 Basic theory

Let U be an open subset of $\mathbb{R}^n$ and let $\vec{f} : U \to \mathbb{R}^n$ be a continuous vector field. We recall that a **smooth parameterized curve with endpoints** is a map $\vec{\alpha} : [a, b] \to U$ satisfying that there exists a smooth parameterized curved $\vec{\beta} :]a - \epsilon, b + \epsilon[\to U$ for some $\epsilon > 0$, such that $\vec{\beta}(t) = \vec{\alpha}(t)$ for all $t \in [a, b]$. The **line integral of $\vec{f}$ along $\vec{\alpha}$** as

$$\int_a^b \vec{f}(\vec{\alpha}(t)) \cdot \vec{\alpha}'(t) dt$$

and it is denoted by

$$\int_{\vec{\alpha}} \vec{f} \quad \text{or} \quad \int_{\vec{\alpha}} \vec{f} \cdot \vec{\alpha}.$$

Example 6.3.1. *Let $U = \mathbb{R}^2$ and $\vec{f}(x_1, x_2) = (x_1^3 x_2, x_2^2)$ and let $\vec{\alpha} : [0, 1] \to U$ be given by $t \mapsto (t, t^2)$. Then*

$$\int_{\vec{\alpha}} \vec{f} = \int_0^1 (t^5, t^4) \cdot (1, t) dt = \int_0^1 2t^5 dt = \frac{1}{3}.$$

We have the following result stating the independence of parameterization of the line integral. It is a direct consequence of the chain rule of functions of one variable.

Proposition 6.3.2. *Let U be an open subset of $\mathbb{R}^n$, $\vec{f} : U \to \mathbb{R}^n$ a continuous vector field, and $\vec{\alpha} : [a, b] \to U$ a smooth parameterized curve with endpoints. Let $g : [c, d] \to [a, b]$ be a map such that $g(c) = a$, $g(d) = b$ and that there exists a C^∞ function $\tilde{g} :]c - \epsilon, d + \epsilon[\to \mathbb{R}$ satisfying that $\tilde{g}(t) = g(t)$ for all $t \in [c, d]$. Then, $\vec{\alpha} \circ g$ is a smooth parameterized curve with endpoints and*

$$\int_{\vec{\alpha} \circ g} \vec{f} = \int_{\vec{\alpha}} \vec{f}.$$

Let $\vec{\alpha} : [a, b] \to \mathbb{R}^n$ be a smooth parameterized curve with endpoints. Define the *opposite* curve $\vec{\alpha}^- : [a, b] \to \mathbb{R}^n$ by $\vec{\alpha}^-(t) = \vec{\alpha}(a + b - t)$, for all $t \in [a, b]$. The following result is direct.

Fact 6.3.3. *Let U be an open subset of $\mathbb{R}^n$, $\vec{f} : U \to \mathbb{R}^n$ a continuous vector field, and $\vec{\alpha} : [a, b] \to U$ a smooth parameterized curve with endpoints. Then,*

$$\int_{\vec{\alpha}^-} \vec{f} = -\int_{\vec{\alpha}} \vec{f}.$$

We will need a slight generalization of smooth parameterized curve with endpoints. A **piecewise smooth parameterized curve with endpoints** is a finite tuple $\mathcal{C} = ({}^1\vec{\alpha}, \ldots, {}^N\vec{\alpha})$ such that ${}^i\vec{\alpha} : [a_i, b_i] \to \mathbb{R}^n$ is a smooth parameterized curve with endpoints, for all $i = 1, \ldots, N$, satisfying that ${}^i\vec{\alpha}(b_i) = {}^{i+1}\vec{\alpha}(a_{i+1})$ for all $i = 1, \ldots, N-1$. The points ${}^1\vec{\alpha}(a_1)$ and ${}^N\vec{\alpha}(b_N)$ are called the **initial** and **final endpoints** of $\mathcal{C}$, respectively. We say the piecewise smooth parameterized curve with endpoints $\mathcal{C}$ is **included in an open set** $U \subseteq \mathbb{R}^n$ if the image of each ${}^i\vec{\alpha}$ is included in U. For a continuous vector field $\vec{f} : U \to \mathbb{R}^n$ and a piecewise smooth parameterized curve with endpoints $\mathcal{C}$ that is included in U, we extend the definition of line integral by the obvious formula

$$\int_{\mathcal{C}} \vec{f} = \int_{{}^1\vec{\alpha}} \vec{f} + \cdots + \int_{{}^N\vec{\alpha}} \vec{f}.$$

6.3.2 Conservative vector fields

We say that a continuous vector field $\vec{f} : U \to \mathbb{R}^n$ is **conservative** if there exists a C^1 scalar field $\phi : U \to \mathbb{R}$ such that $\vec{f} = -\nabla\phi$. In this case, ϕ is called the **potential function** for $\vec{f}$. A piecewise smooth parameterized curve with endpoints given by $({}^1\vec{\alpha}, \ldots, {}^N\vec{\alpha})$ is **closed** if ${}^1\vec{\alpha}(a_1) = {}^N\vec{\alpha}(b_N)$.

The following result shows that the line integral of a conservative vector field is independent of the curve chosen as long as the endpoints do not change. For a proof see [3], Thm. XV.§4.2.

Theorem 6.3.4. *Let U be a path-connected*[2] *open subset of $\mathbb{R}^n$ and let $\vec{f} : U \to \mathbb{R}^n$ be a continuous vector field. Then, the following conditions are equivalent:*

(i) $\vec{f}$ is conservative;

(ii) Let $\mathcal{C}$ and $\mathcal{D}$ be two piecewise smooth parameterized curve with endpoints having the same initial and final endpoints and that are included in U. Then,

$$\int_{\mathcal{C}} \vec{f} = \int_{\mathcal{D}} \vec{f}.$$

(iii) Let $\mathcal{C}$ be any closed piecewise smooth parameterized curve with endpoints that is included in U. Then,

$$\int_{\mathcal{C}} \vec{f} = 0.$$

Furthermore, if ϕ is any potential function for $\vec{f}$,

$$\int_{\mathcal{C}} \vec{f} = \phi(\vec{a}) - \phi(\vec{b}),$$

for any piecewise smooth parameterized curve with endpoints that is included in U and such that the initial enpoint is $\vec{a}$ and the final one is $\vec{b}$.

Example 6.3.5. *Let $U = \mathbb{R}^2 \setminus \{(0,0)\}$ and let $\vec{f} : U \to \mathbb{R}^2$ be the vector field*

$$\vec{f}(x_1, x_2) = \left(-\frac{x_2}{x_1^2 + x_2^2}, \frac{x_1}{x_1^2 + x_2^2} \right),$$

[2] A subset $S \subseteq \mathbb{R}^n$ is called **path-connected** if given two points $\vec{a}$ and $\vec{b}$ in S, there is smooth parameterized curve with endpoints $\vec{\alpha} : [a, b] \to S$ such that $\vec{\alpha}(a) = \vec{a}$ and $\vec{\alpha}(b) = \vec{a}$.

for all $(x_1, x_2) \in U$*. Let* $\vec{\alpha} : [0,1] \to U$ *be the smooth parameterized curve with endpoints defined as*

$$\vec{\alpha}(t) = (\cos(2\pi t), \sin(2\pi t)),$$

for all $t \in [0,1]$*. Then* $\vec{\alpha}$ *is closed. The integral*

$$\int_{\vec{\alpha}} \vec{f} = \int_0^1 2\pi dt = 2\pi \neq 0$$

is nonzero, so $\vec{f}$ *is not conservative. However, consider any set* $R =]a,b[\times]c,d[$ *that does not contain the origin (so it is included in* U*). We claim that* $\vec{f}|_R$ *has a potential function. Let us suppose that* $c \geq 0$*. Then the function* $\psi : R \to \mathbb{R}$ *given by*

$$\psi(x_1, x_2) = -\arccos\left(\frac{x_1}{\sqrt{x_1^2 + x_2^2}}\right)$$

satisfies that $\vec{f} = -\nabla\psi$*. We leave the three remaining cases* $d < 0$*,* $a > 0$ *and* $b < 0$ *to the reader.*

6.4 The Green-Riemann theorem

6.4.1 The statement

We state without proof the following result. The reader is referred to [4], Thm. X.§1.1.

Theorem 6.4.1 (Green-Riemann)**.** *Let* U *be an open subset of* $\mathbb{R}^2$ *and let* $\vec{f} : U \to \mathbb{R}^2$ *be a* C^1 *vector field, written componentwise as* $\vec{f}(\vec{x}) = (f_1(\vec{x}), f_2(\vec{x}))$*, for all* $\vec{x} \in U$*. Suppose that* $A \subseteq \mathbb{R}^2$ *is a planar set of type* 3 *such that* $A \cup \partial A \subseteq U$ *and such that the boundary* ∂A *is the image of a closed piecewise smooth parameterized curve with endpoints that we denote by* $\mathcal{C}$*. We suppose further that* $\mathcal{C}$ *is counterclockwise oriented*[3]*. Then,*

$$\int_{\mathcal{C}} \vec{f} = \int_A \left(\frac{\partial f_2}{\partial x_1}(\vec{x}) - \frac{\partial f_1}{\partial x_2}(\vec{x})\right) d\vec{x}.$$

By obvious reasons it is usual to call the integrand of the right member of the previous displayed equations the **curl** (or **rotor**) of the vector field $\vec{f}$. It is denoted by $\operatorname{rot}\vec{f}$.

We will now give a consequence of the Green-Riemann theorem. Let U be an open subset of $\mathbb{R}^2$ and let $\vec{f} : U \to \mathbb{R}^2$ be a continuous vector field. If $\vec{\alpha} : [a,b] \to U$ is a smooth parameterized curve with endpoints, written componentwise as $\vec{\alpha}(t) = (\alpha_1(t), \alpha_2(t))$ for all $t \in [a,b]$, set

$$\vec{N}_{\vec{\alpha}} : [a,b] \to \mathbb{R}^2$$

by $\vec{N}_{\vec{\alpha}}(t) = (-\alpha_2'(t), \alpha_1'(t))$, for all $t \in [a,b]$. Note that $\vec{N}_{\vec{\alpha}}(t) \cdot \vec{\alpha}'(t) = 0$, for all $t \in [a,b]$. We define

$$\int_{\vec{\alpha}^{\perp}} \vec{f} = \int_a^b \vec{f}(\vec{\alpha}(t)) \cdot \vec{N}_{\vec{\alpha}}(t) dt.$$

[3]This is in general a very hard notion to define in precise terms. However, for planar sets of type 1 or 2, it is immediate.

If $\mathcal{C} = ({}^1\vec{\alpha}, \dots, {}^N\vec{\alpha})$ is a piecewise smooth parameterized curve with endpoints that is included in U, define

$$\int_{\mathcal{C}^\perp} \vec{f} = \int_{{}^1\vec{\alpha}^\perp} \vec{f} + \cdots + \int_{{}^N\vec{\alpha}^\perp} \vec{f}.$$

We leave the following result as an exercise to the reader (assuming that Theorem 6.4.1 holds).

Corollary 6.4.2. *Assume the same hypothesis as in the previous theorem. Then,*

$$\int_{\mathcal{C}^\perp} \vec{f} = \int_A \operatorname{div} \vec{f}(\vec{x}) d\vec{x}.$$

6.4.2 Applications

6.4.2.1 The curl

Let U be an open subset of $\mathbb{R}^2$ and let $\vec{f} : U \to \mathbb{R}^2$ be a vector field of class C^1. Let $D_r \subseteq U$ be a closed disc of radius $r > 0$ centered at $\vec{a} \in U$ and let C_r be the corresponding boundary circle, described by the smooth parameterized curve $\vec{\alpha}_r : [0,1] \to \mathbb{R}^2$

$$\vec{\alpha}_r(t) = \vec{a} + (r\cos(2\pi t), r\sin(2\pi t)),$$

for all $t \in [0,1]$. The area of D_r is clearly πr^2. Then D_r is of type 3 and $\vec{\alpha}$ is parametrized counterclockwise. Hence, the previous theorem implies that

$$\frac{1}{\pi r^2} \int_{\vec{\alpha}_r} \vec{f} = \frac{1}{\pi r^2} \int_{D_r} \operatorname{rot} \vec{f}(\vec{x}) d\vec{x}.$$

Let us write

$$\frac{1}{\pi r^2} \int_{\vec{\alpha}_r} \vec{f} - \operatorname{rot} \vec{f}(\vec{a}) = \frac{1}{\pi r^2} \int_{D_r} \operatorname{rot} \vec{f}(\vec{x}) d\vec{x} - \operatorname{rot} \vec{f}(\vec{a}) = \frac{1}{\pi r^2} \int_{D_r} \left(\operatorname{rot} \vec{f}(\vec{x}) - \operatorname{rot} \vec{f}(\vec{a}) \right) d\vec{x}.$$

Hence,

$$\left| \frac{1}{\pi r^2} \int_{\vec{\alpha}_r} \vec{f} - \operatorname{rot} \vec{f}(\vec{a}) \right| \le \frac{1}{\pi r^2} \int_{D_r} \underbrace{\left| \operatorname{rot} \vec{f}(\vec{x}) - \operatorname{rot} \vec{f}(\vec{a}) \right|}_{\diamond(\vec{x})} d\vec{x}.$$

Since $\vec{f}$ is C^1, $\operatorname{rot} \vec{f}$ is continuous. As a consequence, given any $\epsilon > 0$ there is $r > 0$ such that $\diamond(\vec{x}) \le \epsilon$ for all $\vec{x} \in D_r$. This implies the following result.

Fact 6.4.3. *Let U be an open subset of $\mathbb{R}^2$ and let $\vec{f} : U \to \mathbb{R}^2$ be a vector field of class C^1. Then, for all $\vec{a} \in U$ we have*

$$\operatorname{rot} \vec{f}(\vec{a}) = \lim_{r \to 0} \frac{1}{\pi r^2} \int_{\vec{\alpha}_r} \vec{f}.$$

Since the velocity of the parameterization of the boundary of D_r points at all possible directions, we may interpret the previous identity as stating that:

> ***the rotor*** $\operatorname{rot} \vec{f}(\vec{a})$ ***is the rate at which*** $\vec{f}$ ***rotates per unit of area at*** $\vec{a}$.

6.4.2.2 The divergence

We assume the same hypotheses as in the previous paragraph.

By the previous corollary we get that

$$\frac{1}{\pi r^2}\int_{\vec{\alpha}_r^{\perp}} \vec{f} = \frac{1}{\pi r^2}\int_{D_r} \operatorname{div}\vec{f}(\vec{x})d\vec{x}.$$

We see that

$$\frac{1}{\pi r^2}\int_{\vec{\alpha}_r^{\perp}} \vec{f} - \operatorname{div}\vec{f}(\vec{a}) = \frac{1}{\pi r^2}\int_{D_r} \operatorname{div}\vec{f}(\vec{x})d\vec{x} - \operatorname{div}\vec{f}(\vec{a}) = \frac{1}{\pi r^2}\int_{D_r}\Big(\operatorname{rot}\vec{f}(\vec{x}) - \operatorname{div}\vec{f}(\vec{a})\Big)d\vec{x}.$$

So,

$$\left|\frac{1}{\pi r^2}\int_{\vec{\alpha}_r^{\perp}} \vec{f} - \operatorname{div}\vec{f}(\vec{a})\right| \leq \frac{1}{\pi r^2}\int_{D_r}\underbrace{\left|\operatorname{div}\vec{f}(\vec{x}) - \operatorname{div}\vec{f}(\vec{a})\right|}_{\clubsuit(\vec{x})}d\vec{x}.$$

As $\vec{f}$ is C^1, $\operatorname{div}\vec{f}$ is continuous. Hence, given any $\epsilon > 0$ there is $r > 0$ such that $\clubsuit(\vec{x}) \leq \epsilon$ for all $\vec{x} \in D_r$. This implies the following result.

Fact 6.4.4. *Let U be an open subset of $\mathbb{R}^2$ and let $\vec{f} : U \to \mathbb{R}^2$ be a vector field of class C^1. Then, for all $\vec{a} \in U$ we have*

$$\operatorname{div}\vec{f}(\vec{a}) = \lim_{r\to 0}\frac{1}{\pi r^2}\int_{\vec{\alpha}_r^{\perp}} \vec{f}.$$

Since the vector field $N_{\vec{\alpha}}$ is the outward vector of the parameterization of the boundary of D_r, we can interpret the previous equality as follows:

the divergence $\operatorname{div}\vec{f}(\vec{a})$ ***is the rate of outward flow of*** $\vec{f}$ ***per unit of area at*** $\vec{a}$.

Chapter 7

Optimization of functions

7.1 Basic definitions and properties

We will now use the previous recalled tools in order to study the local maxima and minima of functions assuming some smoothness assumptions. In particular, in this section we will solve the Problem 1 announced at the beginning of this course.

Definition 7.1.1. *Let $U \subseteq \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}$ be a map of class C^1 on U. Consider $\vec{a} \in U$. We say that it is*

*(i) a **critical point** of f if $\nabla f(\vec{a}) = \vec{0}$;*

*(ii) a (resp., **local minimum**) **local maximum** of f if there exists an open subset $V \subseteq U$ such that $\vec{a} \in V$ and (resp., $f(\vec{a}) \leq f(\vec{x})$) $f(\vec{a}) \geq f(\vec{x})$ for all $\vec{x} \in V$;*

*(iii) a **strict** (resp., local minimum) local maximum if we further have that (resp., $f(\vec{a}) < f(\vec{x})$) $f(\vec{a}) > f(\vec{x})$ for all $\vec{x} \in V \setminus \{\vec{a}\}$.*

*A (resp., **strict**) **local extremum** of f is a (resp., strict) local maximum or local minimum.*

The first step towards determining local extrema of functions is the following, which gives a necessary condition.

Proposition 7.1.2. *Let $U \subseteq \mathbb{R}^n$ be an open subset, $f : U \to \mathbb{R}$ be a map of class C^1 on U, and $\vec{a} \in U$. If $\vec{a}$ is a local extremum of f, then it is a critical point.*

Proof. Suppose that $\vec{a} + t\vec{e}_i \in U$ for all $t \in]-\epsilon, \epsilon[$ and all $i = 1, \dots, n$. Given any $i = 1, \dots, n$, consider the function of one variable $g_i :]-\epsilon, \epsilon[\to \mathbb{R}$ given by $t \mapsto f(\vec{a} + t\vec{e}_i)$. It is clearly of class C^1. Since $\vec{a}$ is a local extremum of f, $t = 0$ is a local extremum of g_i, which implies that $g_i'(0) = 0$, by the theory of functions of one real variable. However, the chain rule tells us that

$$g_i'(0) = \frac{\partial f}{\partial x_i}(\vec{a}).$$

The proposition is thus proved. □

7.2 The main result

We want now to give a criterion with sufficient conditions on a critical point in order to recognize when it is an extremum or not. We suppose from now on that f

is C^3 and suppose that $\vec{a}$ is critical point of it in U. Proposition 5.6.1 tells us that

$$f(\vec{a}+\vec{h}) = f(\vec{a}) + \vec{h}\cdot \underbrace{\nabla f(\vec{a})}_{=\vec{0}} + \frac{1}{2}\vec{h}.H_f(\vec{a}).\vec{h}^t + \phi(\|\vec{h}\|),$$

for $\vec{h}$ of sufficiently small norm and for a locally defined function ϕ of type $o(t^2)$. The vanishing condition on the gradient is due to the fact that $\vec{a}$ is a critical point of f. From the previous displayed equation we see that locally, the behaviour of f is determined by

$$\frac{1}{2}\vec{h}.H_f(\vec{a}).\vec{h}^t.$$

We introduce the following definition.

Definition 7.2.1. *Let $U \subseteq \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}$ be a map of class C^2 on U. A critical point $\vec{a} \in U$ of f is called **nondegenerate** if $\det(H_f(\vec{a})) \neq 0$. Note that this implies that all the eigenvalues of $H_f(\vec{a})$ are different from zero.*

The following result gives sufficient conditions to decide when a critical point is an extremum and of which nature. The proof is a direct consequence of Proposition 5.6.1.

Theorem 7.2.2. *Let $U \subseteq \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}$ be a map of class C^3 on U. Suppose given a nondegenerate critical point of $\vec{a} \in U$. Then,*

(i) if all the eigenvalues of $H_f(\vec{a})$ are (strictly) positive, then $\vec{a}$ is a strict local minimum;

(ii) if all the eigenvalues of $H_f(\vec{a})$ are (strictly) negative, then $\vec{a}$ is a strict local maximum;

(iii) if there is one positive and one negative eigenvalue of $H_f(\vec{a})$, then $\vec{a}$ is a neither minimum nor maximum.

*In the last case we say that $\vec{a}$ is a **saddle point**.*

In two dimensions there is a computationally direct way of recognizing the nature of the critical points, the proof of which is straightforward.

Corollary 7.2.3. *Assume the same hypothesis as in the previous theorem and also that $n = 2$. Let us write*

$$H_f(\vec{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\vec{a}) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{a}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{a}) & \frac{\partial^2 f}{\partial x_2^2}(\vec{a}) \end{pmatrix} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}.$$

The nondegeneracy assumption is precisely the inequality $ad - b^2 \neq 0$. Then,

(i) if $ad - b^2 > 0$ and $a > 0$, then $\vec{a}$ is a strict local minimum;

(ii) if $ad - b^2 > 0$ and $a < 0$, then $\vec{a}$ is a strict local minimum;

(iii) if $ad - b^2 < 0$, then $\vec{a}$ is a saddle point.

Example 7.2.4. *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by*

$$f(x_1, x_2) = \log(1 + x_1^2 + x_2^2).$$

It is clearly C^3, as it is a composition of C^3 functions. The only critical point of f is the origin, since

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = \frac{2x_1}{1 + x_1^2 + x_2^2}, \quad \text{and} \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = \frac{2x_1}{1 + x_1^2 + x_2^2}.$$

The second partial derivatives are

$$\frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) = 2\frac{1 - x_1^2 + x_2^2}{(1 + x_1^2 + x_2^2)^2},$$
$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) = 2\frac{1 + (x_1 - x_2)^2}{(1 + x_1^2 + x_2^2)^2},$$
$$\frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) = 2\frac{1 - x_2^2 + x_1^2}{(1 + x_1^2 + x_2^2)^2}.$$

The Hessian matrix at $(0,0)$ *is thus*

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

The previous corollary says thus that $(0,0)$ *is a strict local minimum. Since* $1 + x_1^2 + x_2^2 \geq 1$ *for all* $(x_1, x_2) \in \mathbb{R}^2$ *and* $\log$ *is an increasing function, we see that* $f(x_1, x_2) \geq \log(1) = f(0,0) = 0$*, so* $(0,0)$ *is also a global minimum.*

References

[1] Michel Chasles, *Note sur les propriétés générales du système de deux corps semblables entr'eux*, Bulletin de Sciences Mathematiques, Astronomiques Physiques et Chimiques, Baron de Ferussac, Paris (1830). ↑6

[2] Bob Palais and Richard Palais, *Chasles' fixed point theorem for Euclidean motions*, J. Fixed Point Theory Appl. **12** (2012), no. 1-2, 27–34. ↑6

[3] Serge Lang, *Undergraduate analysis*, 2nd ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1997. ↑15, 16, 18, 20, 25, 29, 30, 31, 34

[4] ———, *Calculus of several variables*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1987. ↑35

[5] Arnold Neumaier, *Phenomenological thermodynamics in a nutshell*, available at `http://www.mat.univie.ac.at/~neum/ms/phenTherm.pdf`. ↑11, 25, 26

[6] J. Schwartz, *The formula for change in variables in a multiple integral*, Amer. Math. Monthly **61** (1954), 81–85. ↑32