

2 Nonlinear Analysis Fundamentals

In this chapter we study convex sets and functions, convex cones, the Bouligand tangent cone to a set at a point, and semicontinuous functions. We prove fundamental results, which will be the main investigation tools in subsequent chapters. The Farkas Lemma and its consequences will be decisive for establishing optimality conditions in Karush-Kuhn-Tucker form, while the Banach fixed point Principle will be used for discussing some optimization algorithms. We also study some important inequalities related to the study of convex functions.

2.1 Convex Sets and Cones

Definition 2.1.1. *One says that a nonempty set $D \subset \mathbb{R}^p$ is convex if for every $x, y \in D$,*

$$[x, y] = \{\alpha x + (1 - \alpha)y \mid \alpha \in [0, 1]\} \subset D.$$

In other words, D is convex if and only if together with two points a_1, a_2 , it contains the whole segment $[a_1, a_2]$. It is sufficient to take $\alpha \in (0, 1)$. By mathematical induction one can show that if D is convex, then for every $n \in \mathbb{N}^*$, $x_1, x_2, \dots, x_n \in D$, $\alpha_1, \alpha_2, \dots, \alpha_n \in [0, 1]$ with $\sum_{i=1}^n \alpha_i = 1$:

$$\sum_{i=1}^n \alpha_i x_i \in D.$$

A sum like the one above is called a convex combination of the elements (x_i) . In \mathbb{R} , the convex sets are intervals.

One of the main objects for our study is defined next.

Definition 2.1.2. *A nonempty subset $K \subset \mathbb{R}^p$ is a cone if the next relation holds:*

$$\forall y \in K, \forall \lambda \in \mathbb{R}_+ := [0, \infty) : \lambda y \in K.$$

According to the definition, every cone contains the origin.

Proposition 2.1.3. *A cone C is convex if and only if $C + C = C$.*

Proof Suppose first that C is a convex cone. As $0 \in C$, it is clear that $C \subset C + C$. Let $u \in C + C$. Then there exist $c_1, c_2 \in C$ such that $c_1 + c_2 = u$. But, using the properties of C and the obvious relation

$$u = 2 \left(2^{-1}c_1 + 2^{-1}c_2 \right),$$

one can deduce that $u \in C$. For the converse implication, suppose that C is a cone which satisfies $C + C = C$. Fix $\alpha \in (0, 1)$ and $c_1, c_2 \in C$. Then, from the cone property of C , one knows that $\alpha c_1, (1 - \alpha)c_2 \in C$, therefore $\alpha c_1 + (1 - \alpha)c_2 \in C + C = C$, which ends the proof. \square

Definition 2.1.4. Let $A \subset \mathbb{R}^p$ be a nonempty set and $x \in \mathbb{R}^p$. One defines the distance from x to A by the relation:

$$d(x, A) := \inf\{\|x - a\| \mid a \in A\}.$$

We also consider the function $d_A : \mathbb{R}^p \rightarrow \mathbb{R}$ given by

$$d_A(x) := d(x, A).$$

We now introduce some basic properties of the distance from a point to a (nonempty) set.

Theorem 2.1.5. Let $A \subset \mathbb{R}^p$, $A \neq \emptyset$. Then:

- (i) $d(x, A) = 0$ if and only if $x \in \text{cl } A$.
- (ii) The function d_A is 1-Lipschitz.
- (iii) If A is closed, then for every $x \in \mathbb{R}^p$, there exists $a_x \in A$ such that $d(x, A) = \|x - a_x\|$. If, moreover, A is convex, then a_x having the previous property is unique and it is characterized by the relations

$$\begin{cases} a_x \in A \\ \langle x - a_x, u - a_x \rangle \leq 0, \forall u \in A. \end{cases}$$

Proof (i) The following equivalences hold

$$\begin{aligned} d(x, A) = 0 &\Leftrightarrow \inf_{a \in A} \|x - a\| = 0 \\ &\Leftrightarrow \exists (a_n) \subset A \text{ with } \lim_{n \rightarrow \infty} \|x - a_n\| = 0 \Leftrightarrow x \in \overline{A}. \end{aligned}$$

(ii) For every $x, y \in \mathbb{R}^p$ and every $a \in A$, these relations hold:

$$d(x, A) \leq \|x - a\| \leq \|x - y\| + \|y - a\|.$$

As a is taken arbitrary from A , one deduces

$$d(x, A) \leq \|x - y\| + d(y, A),$$

i.e.,

$$d(x, A) - d(y, A) \leq \|x - y\|.$$

By reversing the roles of x and y , one has:

$$|d(x, A) - d(y, A)| \leq \|x - y\|,$$

which is the desired conclusion.

(iii) If $x \in A$, then $a_x := x$ is the unique element having the previously introduced property. Let $x \notin A$. As $d(x, A)$ is a real number, there exists $r > 0$ such that $A_1 := A \cap D(x, r) \neq \emptyset$. Since A_1 is a compact set and the function $g : A_1 \rightarrow \mathbb{R}$, $g(y) = d(x, y)$ is continuous, according to Weierstrass Theorem, g attains its minimum on A_1 , i.e., there exists $a_x \in A_1$ with $g(a_x) = \inf_{y \in A_1} g(y) = d(x, A_1)$. Now, one can check that $d(x, A_1) = d(x, A)$ and the first conclusion follows. Suppose that, moreover, A is convex. If $x \in A$, there is nothing to prove. Take $x \notin A$. Consider $a_1, a_2 \in A$ with $d(x, A) = \|x - a_1\| = \|x - a_2\|$. Using the parallelogram law we know:

$$\|(x - a_1) + (x - a_2)\|^2 + \|(x - a_1) - (x - a_2)\|^2 = 2\|x - a_1\|^2 + 2\|x - a_2\|^2,$$

i.e.,

$$\|2x - a_1 - a_2\|^2 + \|a_2 - a_1\|^2 = 4d^2(x, A),$$

and dividing by 4 one gets

$$\left\|x - \frac{a_1 + a_2}{2}\right\|^2 + 4^{-1}\|a_2 - a_1\|^2 = d^2(x, A).$$

Since A is convex, $2^{-1}(a_1 + a_2) \in A$, hence $\left\|x - \frac{a_1 + a_2}{2}\right\|^2 \geq d^2(x, A)$. This relation and the previous equality show that $\|a_2 - a_1\| = 0$, hence $a_1 = a_2$. The proof of uniqueness is now complete. Let us prove now that a_x verifies the relation $\langle x - a_x, u - a_x \rangle \leq 0$ for any $u \in A$. For this, take $u \in A$. Then for every $\alpha \in (0, 1]$, one has

$$v = \alpha u + (1 - \alpha)a_x \in A.$$

Hence,

$$\|x - a_x\| \leq \|x - \alpha u - (1 - \alpha)a_x\| = \|x - a_x - \alpha(u - a_x)\|,$$

and, consequently,

$$\|x - a_x\|^2 \leq \|x - a_x\|^2 - 2\alpha \langle x - a_x, u - a_x \rangle + \alpha^2 \|u - a_x\|^2.$$

After reducing terms and dividing by $\alpha > 0$, we can see that

$$0 \leq -2 \langle x - a_x, u - a_x \rangle + \alpha \|u - a_x\|^2.$$

If we let $\alpha \rightarrow 0$, the desired inequality follows. For the converse, if an element $a \in A$ satisfies $\langle x - a, u - a \rangle \leq 0$ for any $u \in A$, then for every $v \in A$ one has

$$\|x - a\|^2 - \|x - v\|^2 = 2 \langle x - a, v - a \rangle - \|a - v\|^2 \leq 0,$$

hence a coincides with a_x . The proof is now complete. □

In the case when A is closed, then for $x \in \mathbb{R}^p$ one denotes the projection set of x on A by

$$\text{pr}_A x := \{a \in A \mid d(x, A) = \|x - a\|\}.$$

If, moreover, A is convex, then, according to the above theorem, this set consists of only one element, which we still denote by $\text{pr}_A x$, and we call this the projection of x on A .

Let $S \subset \mathbb{R}^p$ be a nonempty set. The polar of S is the set

$$S^- := \{u \in \mathbb{R}^p \mid \langle u, x \rangle \leq 0, \forall x \in S\}.$$

It is easy to observe that S^- is a closed convex cone and that, in general, $S \subset (S^-)^-$. If we consider the reverse inclusion, the next result follows.

Theorem 2.1.6. *Let $C \subset \mathbb{R}^p$ be a closed convex cone. Then $C = (C^-)^-$.*

Proof Consider $z \in (C^-)^-$ and $\bar{z} = \text{pr}_C z$. We will prove that $z = \bar{z}$. From the last part of Theorem 2.1.5, for any $c \in C$, one has

$$\langle z - \bar{z}, c - \bar{z} \rangle \leq 0.$$

As $0 \in C$ and $2\bar{z} \in C$, we deduce

$$-\langle \bar{z}, z - \bar{z} \rangle \leq 0, \quad \langle \bar{z}, z - \bar{z} \rangle \leq 0$$

hence

$$\langle z - \bar{z}, c \rangle \leq 0$$

for every $c \in C$, i.e., $z - \bar{z} \in C^-$. As $z \in (C^-)^-$, one gets

$$\langle z, z - \bar{z} \rangle \leq 0.$$

But

$$\|z - \bar{z}\|^2 = \langle z - \bar{z}, z \rangle - \langle z - \bar{z}, \bar{z} \rangle \leq 0$$

which means that $z = \bar{z}$, so $z \in C$. This establishes the theorem. \square

Example 2.1.7. 1. Consider $S = \{(x, 0) \in \mathbb{R}^2 \mid x \geq 0\}$. One can observe that $S^- = \{(x, y) \in \mathbb{R}^2 \mid x \leq 0\}$. Obviously, $(S^-)^- = S^-$.

2. The polar of $\mathbb{R}_+^2 := \{(x, y) \in \mathbb{R}^2 \mid x, y \geq 0\}$ is $\mathbb{R}_-^2 := \{(x, y) \in \mathbb{R}^2 \mid x, y \leq 0\}$. The polar of the set $S = \{(x, 0) \in \mathbb{R}^2 \mid x \geq 0\} \cup \{(0, y) \in \mathbb{R}^2 \mid y \geq 0\}$ is also \mathbb{R}_-^2 . From this example one can see that, in general, $S_1^- = S_2^-$ does not imply $S_1 = S_2$.

The next result, which has an algebraic character, it was obtained by the Hungarian mathematician Julius Farkas in 1902.

Theorem 2.1.8 (Farkas' Lemma). *Let $n \in \mathbb{N}^*$, $(\varphi_i)_{i \in \overline{1, n}} \subset L(\mathbb{R}^p, \mathbb{R})$ and $\varphi \in L(\mathbb{R}^p, \mathbb{R})$. Then*

$$\forall x \in \mathbb{R}^p : [\varphi_1(x) \leq 0, \dots, \varphi_n(x) \leq 0] \Rightarrow \varphi(x) \leq 0 \quad (2.1.1)$$

if and only if there exists $(\alpha_i)_{i \in \overline{1, n}} \subset [0, \infty)$ such that $\varphi = \sum_{i=1}^n \alpha_i \varphi_i$.

Proof The converse implication is obvious. We prove the other one by induction for $n \geq 1$. Define the proposition $P(n)$ which says that for every $\varphi, \varphi_1, \dots, \varphi_n \in L(\mathbb{R}^p, \mathbb{R})$ satisfying (2.1.1), there exist $(\alpha_i)_{i \in \overline{1, n}} \subset [0, \infty)$ such that $\varphi = \sum_{i=1}^n \alpha_i \varphi_i$.

Let us prove that $P(1)$ is true. Indeed, let $\varphi, \varphi_1 \in L(\mathbb{R}^p, \mathbb{R})$ such that

$$\varphi_1(x) \leq 0 \Rightarrow \varphi(x) \leq 0.$$

If $\varphi = 0$ then, obviously, $\varphi = 0\varphi_1$. Suppose $\varphi \neq 0$. Then, by the assumption:

$$\varphi_1(x) = 0 \Leftrightarrow [\varphi_1(x) \leq 0, \varphi_1(-x) \leq 0] \Rightarrow [\varphi(x) \leq 0, \varphi(-x) \leq 0] \Leftrightarrow \varphi(x) = 0,$$

hence $\text{Ker } \varphi_1 \subset \text{Ker } \varphi$. Since $\varphi \neq 0$, one has $\varphi_1 \neq 0$, so there exists $x_1 \in \mathbb{R}^p$ with $\varphi_1(\bar{x}) = -1$. Also by the assumption, $\varphi(\bar{x}) \leq 0$. Take $x \in \mathbb{R}^p$ arbitrarily. Then it is easy to verify that

$$x + \varphi_1(x)\bar{x} \in \text{Ker } \varphi_1,$$

hence

$$x + \varphi_1(x)\bar{x} \in \text{Ker } \varphi,$$

i.e.,

$$\varphi(x + \varphi_1(x)\bar{x}) = 0,$$

which proves that

$$\varphi(x) = -\varphi(\bar{x})\varphi_1(x).$$

Notice that x was arbitrarily chosen, so the desired relation is proved for $\alpha_1 := -\varphi(\bar{x}) \geq 0$.

Suppose now that $P(n)$ is true for a fixed $n \geq 1$ and we will try to prove that $P(n+1)$ is true.

Take $\varphi, \varphi_1, \dots, \varphi_n, \varphi_{n+1} \in L(\mathbb{R}^p, \mathbb{R})$ such that

$$\forall x \in \mathbb{R}^p : [\varphi_1(x) \leq 0, \dots, \varphi_n(x) \leq 0, \varphi_{n+1}(x) \leq 0] \Rightarrow \varphi(x) \leq 0. \quad (2.1.2)$$

If

$$\forall x \in \mathbb{R}^p : [\varphi_1(x) \leq 0, \dots, \varphi_n(x) \leq 0] \Rightarrow \varphi(x) \leq 0, \quad (2.1.3)$$

then, from $P(n)$, there exist $(\alpha_i)_{i \in \overline{1, n}} \subset [0, \infty)$ such that $\varphi = \sum_{i=1}^n \alpha_i \varphi_i$; take $\alpha_{n+1} := 0$ and the conclusion follows.

Suppose relation (2.1.3) is not satisfied. Then there exists $\bar{x} \in \mathbb{R}^p$ such that $\varphi(\bar{x}) > 0$ and $\varphi_i(\bar{x}) \leq 0$ for any $i \in \overline{1, n}$. As (2.1.2) holds, $\varphi_{n+1}(\bar{x}) > 0$; we may suppose (by multiplying by the appropriate positive scalar) that $\varphi_{n+1}(\bar{x}) = 1$. But

$$\varphi_{n+1}(x - \varphi_{n+1}(x)\bar{x}) = 0, \quad \forall x \in \mathbb{R}^p,$$

and from (2.1.2) we deduce

$$\forall x \in \mathbb{R}^p : [\varphi_1(x - \varphi_{n+1}(x)\bar{x}) \leq 0, \dots, \varphi_n(x - \varphi_{n+1}(x)\bar{x}) \leq 0] \Rightarrow \varphi(x - \varphi_{n+1}(x)\bar{x}) \leq 0. \quad (2.1.4)$$

Take $\varphi'_i := \varphi_i - \varphi_i(\bar{x})\varphi_{n+1}$ for $i \in \overline{1, n}$ and $\varphi' := \varphi - \varphi(\bar{x})\varphi_{n+1}$, and then relation (2.1.4) becomes

$$\forall x \in X : [\varphi'_1(x) \leq 0, \dots, \varphi'_n(x) \leq 0] \Rightarrow \varphi'(x) \leq 0.$$

As $P(n)$ is true, there exist $(\alpha_i)_{i \in \overline{1, n}} \subset [0, \infty)$ such that $\varphi' = \sum_{i=1}^n \alpha_i \varphi'_i$. We deduce that

$$\varphi - \varphi(\bar{x})\varphi_{n+1} = \sum_{i=1}^n \alpha_i [\varphi_i - \varphi_i(\bar{x})\varphi_{n+1}],$$

hence $\varphi = \sum_{i=1}^{n+1} \alpha_i \varphi_i$, where $\alpha_{n+1} = \varphi(\bar{x}) - \sum_{i=1}^n \alpha_i \varphi_i(\bar{x}) \geq 0$ from the choice of \bar{x} and from the fact that $\alpha_i \geq 0$ for any $i \in \overline{1, n}$. The proof is now complete. \square

Throughout this book we shall use several different concepts of tangent vectors to a set at a point. We introduce now one of these concepts.

Definition 2.1.9. Let $M \subset \mathbb{R}^p$ be a nonempty set and $\bar{x} \in \text{cl } M$. One says that a vector $u \in \mathbb{R}^p$ is tangent in the sense of Bouligand to the set M at \bar{x} if there exist $(t_n) \subset (0, \infty)$, $t_n \rightarrow 0$ and $(u_n) \rightarrow u$ such that for any $n \in \mathbb{N}$, one has

$$\bar{x} + t_n u_n \in M.$$

It is sufficient that the above inclusion holds for every $n \in \mathbb{N}$ sufficiently large.

Theorem 2.1.10. The set, denoted by $T_B(M, \bar{x})$, which contains all the tangent vectors to the set M at \bar{x} is a closed cone, which we call the Bouligand tangent cone (or the contingent cone) to the set M at the point \bar{x} .

Proof Let us prove first that $0 \in T_B(M, \bar{x})$. If $\bar{x} \in M$, then the assertion trivially follows, because it is sufficient to take (u_n) constantly equal to 0. If $\bar{x} \notin M$, then there exists $(x_n)_{n \in \mathbb{N}} \subset M$ such that $x_n \rightarrow \bar{x}$, and we consider $t_n := \sqrt{\|x_n - \bar{x}\|}$ and $u_n := \left(\sqrt{\|x_n - \bar{x}\|}\right)^{-1} (x_n - \bar{x})$ for every $n \in \mathbb{N}$. Since $t_n \rightarrow 0$ and $u_n \rightarrow 0$, one obtains the conclusion.

Consider now $u \in T_B(M, \bar{x})$ and $\lambda > 0$. According to the definition, there exist $(t_n) \subset (0, \infty)$, $t_n \rightarrow 0$ and $(u_n) \rightarrow u$ such that for every $n \in \mathbb{N}$,

$$\bar{x} + t_n u_n \in M.$$

This is equivalent to

$$\bar{x} + \frac{t_n}{\lambda} (\lambda u_n) \in M.$$

As $\left(\frac{t_n}{\lambda}\right) \rightarrow 0$ and $(\lambda u_n) \rightarrow \lambda u$, one deduces that $\lambda u \in T_B(M, \bar{x})$, hence $T_B(M, \bar{x})$ is a cone. We prove that the closure of $T_B(M, \bar{x})$ is contained in $T_B(M, \bar{x})$. Take $(u_n) \subset$

$T_B(M, \bar{x})$ and $(u_n) \rightarrow u$. One must prove that $u \in T_B(M, \bar{x})$. For every $n \in \mathbb{N}$, there exist $(t_n^k)_k \subset (0, \infty)$, $t_n^k \xrightarrow{k \rightarrow \infty} 0$ and $(u_n^k) \xrightarrow{k \rightarrow \infty} u_n$ such that for every $k \in \mathbb{N}$,

$$\bar{x} + t_n^k u_n^k \in M.$$

By using a diagonalization procedure, for every $n \in \mathbb{N}^*$, there exists $k_n \in \mathbb{N}$ such that the next relations hold:

$$t_n^{k_n} < \frac{1}{n}$$

$$\left\| u_n^{k_n} - u_n \right\| \leq \frac{1}{n}.$$

It is easy to observe that the positive sequence $(t_n^{k_n})_n$ converges to 0, and using the inequality

$$\left\| u_n^{k_n} - u \right\| \leq \left\| u_n^{k_n} - u_n \right\| + \|u_n - u\|,$$

one can deduce that $(u_n^{k_n}) \rightarrow u$. Moreover, for every $n \in \mathbb{N}$,

$$\bar{x} + t_n^{k_n} u_n^{k_n} \in M,$$

hence $u \in T_B(M, \bar{x})$ and the proof is complete. □

Our first example is given next.

Example 2.1.11. 1. Consider the ball $M \subset \mathbb{R}^2$, $M := \{(x, y) \in \mathbb{R}^2 \mid (x - 1)^2 + y^2 \leq 1\}$. Then $T_B(M, (0, 0)) = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0\}$.

2. One can easily observe that if $C \subset \mathbb{R}^p$ is a closed cone, then $T_B(C, 0) = C$.

Proposition 2.1.12. If $\emptyset \neq M \subset \mathbb{R}^p$ and $\bar{x} \in \text{cl } M$, then $T_B(M, \bar{x}) = T_B(\text{cl } M, \bar{x})$. If $\bar{x} \in \text{int } M$, then $T_B(M, \bar{x}) = \mathbb{R}^p$.

Proof The inclusion $T_B(M, \bar{x}) \subset T_B(\text{cl } M, \bar{x})$ is obvious by the use of $M \subset \text{cl } M$. Take $u \in T_B(\text{cl } M, \bar{x})$. There exist $(t_n) \subset (0, \infty)$, $t_n \rightarrow 0$ and $(u_n) \rightarrow u$ such that for every $n \in \mathbb{N}$,

$$\bar{x} + t_n u_n \in \text{cl } M.$$

Using the sequence characterization of the closure, for any fixed n , there exist $(v_n^k)_k \subset M$ such that

$$v_n^k \xrightarrow{k} \bar{x} + t_n u_n.$$

As above, for any fixed n , there exists $k_n \in \mathbb{N}$ such that

$$\left\| v_n^{k_n} - (\bar{x} + t_n u_n) \right\| \leq t_n^2.$$

Then, one can write

$$\left\| \frac{v_n^{k_n} - \bar{x}}{t_n} - u \right\| \leq \left\| \frac{v_n^{k_n} - \bar{x}}{t_n} - u_n \right\| + \|u_n - u\| \leq t_n + \|u_n - u\|,$$

hence $u'_n := \frac{v_n^{k_n} - \bar{x}}{t_n} \xrightarrow{n \rightarrow \infty} u$. But

$$\bar{x} + t_n u'_n = v_n^{k_n} \in M,$$

hence $u \in T_B(M, \bar{x})$. The second part of the conclusion easily follows: if $\bar{x} \in \text{int } M$, then for every $u \in \mathbb{R}^p$ and every $(t_n) \subset (0, \infty)$, $t_n \rightarrow 0$, one has $\bar{x} + t_n u_n \in M$ for any n sufficiently large. This shows, in particular, that $u \in T_B(M, \bar{x})$, and the conclusion follows. \square

In general, the Bouligand tangent cone is not convex and the relation $T_B(M, \bar{x}) = \mathbb{R}^p$ can be satisfied, even if $\bar{x} \notin \text{int } M$.

Example 2.1.13. 1. Consider the set $M \subset \mathbb{R}^2$, $M = \{(x, y) \mid x \geq 0, y = 0\} \cup \{(x, y) \mid x = 0, y \geq 0\}$. Then $T_B(M, (0, 0)) = M$ is not a convex set.

2. Let set M represent the plane domain bounded by the curve (the cardioid) which has the parametric representation

$$\begin{cases} x = -2 \cos t + \cos 2t + 1 \\ y = 2 \sin t - \sin 2t \end{cases}, \quad t \in [0, 2\pi].$$

Then $T_B(M, (0, 0)) = \mathbb{R}^2$, but $(0, 0) \notin \text{int } M$.

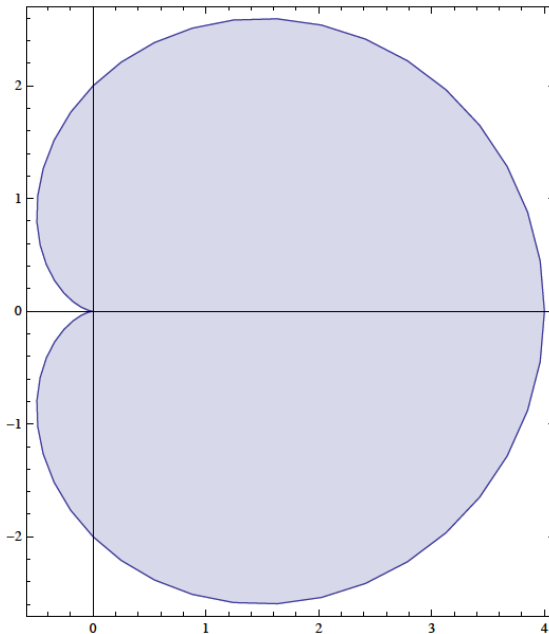


Figure 2.1: The cardioid.

Proposition 2.1.14. *Let $A_1, A_2 \subset \mathbb{R}^p$ be closed sets. Then the next relations hold:*

- (i) *if $\bar{x} \in A_1 \cap A_2$, then $T_B(A_1 \cup A_2, \bar{x}) = T_B(A_1, \bar{x}) \cup T_B(A_2, \bar{x})$;*
- (ii) *if $\bar{x} \in A_1 \cap A_2$, then $T_B(A_1 \cap A_2, \bar{x}) \subset T_B(A_1, \bar{x}) \cap T_B(A_2, \bar{x})$;*
- (iii) *if $\bar{x} \in \text{bd } A_1$, then $T_B(\text{bd } A_1, \bar{x}) = T_B(A_1, \bar{x}) \cap T_B(\mathbb{R}^p \setminus A_1, \bar{x})$.*

Proof The first two relations easily follow, as well as the inclusion $T_B(\text{bd } A_1, \bar{x}) \subset T_B(A_1, \bar{x}) \cap T_B(\mathbb{R}^p \setminus A_1, \bar{x})$ from (iii), which can be proved by using (ii) and Proposition 2.1.12. Let us prove now the other inclusion from (iii). Take $u \in T_B(A_1, \bar{x}) \cap T_B(\mathbb{R}^p \setminus A_1, \bar{x})$. According to the definition, there exist $(t_n), (t'_n) \subset (0, \infty)$, $t_n, t'_n \rightarrow 0$ and $(u_n), (u'_n) \rightarrow u$ such that for every $n \in \mathbb{N}$,

$$\bar{x} + t_n u_n \in A_1,$$

and

$$\bar{x} + t'_n u'_n \in \mathbb{R}^p \setminus A_1.$$

If an infinite number of terms from the first or the second relation are on the boundary of A_1 , there is nothing to prove. Suppose next, without loss of generality, that for every $n \in \mathbb{N}$, there exists $\lambda_n \in (0, 1)$ such that

$$\lambda_n (\bar{x} + t_n u_n) + (1 - \lambda_n)(\bar{x} + t'_n u'_n) \in \text{bd } A_1.$$

Consider the sequences

$$(t''_n) := (\lambda_n t_n + (1 - \lambda_n)t'_n) \subset (0, \infty)$$

$$(u''_n) := \frac{t_n \lambda_n}{t''_n} u_n + \frac{t'_n (1 - \lambda_n)}{t''_n} u'_n.$$

It is clear that $(t''_n) \rightarrow 0$. On the other hand,

$$\|u''_n - u\| \leq \|u_n - u\| + \|u'_n - u\|,$$

hence $(u''_n) \rightarrow u$. Since

$$\bar{x} + t''_n u''_n \in \text{bd } A_1,$$

one gets the desired conclusion. □

Denote by $N_B(M, \bar{x})$ the polar of $T_B(M, \bar{x})$ (i.e., $N_B(M, \bar{x}) := T_B(M, \bar{x})^\circ$), and we call this set the Bouligand normal cone to M at \bar{x} .

If the set M is convex, then the Bouligand tangent and normal cones have a special form.

Proposition 2.1.15. *Let $\emptyset \neq M \subset \mathbb{R}^p$ be a convex set and $\bar{x} \in M$. Then*

$$T_B(M, \bar{x}) = \text{cl } \mathbb{R}_+(M - \bar{x}),$$

and

$$N_B(M, \bar{x}) = \{u \in \mathbb{R}^p \mid \langle u, c - \bar{x} \rangle \leq 0, \forall c \in M\}.$$

Proof Take $c \in M$ and $d := c - \bar{x}$. Consider $(t_k)_k \rightarrow 0$. Then

$$\bar{x} + t_k d = (1 - t_k)\bar{x} + t_k c \in M,$$

hence $M - \bar{x} \subset T_B(M, \bar{x})$. Since $T_B(M, \bar{x})$ is a closed cone, one gets that $\text{cl } \mathbb{R}_+(M - \bar{x}) \subset T_B(M, \bar{x})$. Take $u \in T_B(M, \bar{x})$. Then there exist $(t_k) \subset (0, \infty)$, $t_k \rightarrow 0$ and $(u_k) \rightarrow u$ such that for every $k \in \mathbb{N}$,

$$x_k := \bar{x} + t_k u_k \in M.$$

Hence $u = \lim_k \frac{x_k - \bar{x}}{t_k}$. But $\left(\frac{x_k - \bar{x}}{t_k}\right)_k \subset \mathbb{R}_+(M - \bar{x})$. One can deduce that $T_B(M, \bar{x}) \subset \text{cl } \mathbb{R}_+(M - \bar{x})$. Recall that, by definition,

$$N_B(M, \bar{x}) = T_B(M, \bar{x})^\circ = \{u \in \mathbb{R}^p \mid \langle u, v \rangle \leq 0, \forall v \in T_B(M, \bar{x})\}.$$

Now, taking into account the particular form of $T_B(M, \bar{x})$, the conclusion follows. \square

For a reason we make clear later on, in the case of convex sets, we do not use the subscript B in the notation of these cones.

Example 2.1.16. We want to compute the tangent and normal cones, at different points, to the set $M \subset \mathbb{R}^p$,

$$M = \left\{ x = (x_1, x_2, \dots, x_p) \in \mathbb{R}^p \mid x_i \geq 0, \forall i \in \overline{1, p}, \sum_{i=1}^p x_i = 1 \right\},$$

which is called the unit simplex. This set is convex and closed. According to the previous result, for every $\bar{x} \in M$,

$$\begin{aligned} T(M, \bar{x}) &= \text{cl } \mathbb{R}_+(M - \bar{x}) \\ &= \text{cl } \{u \in \mathbb{R}^p \mid \exists \alpha \geq 0, x \in M, u = \alpha(x - \bar{x})\}. \end{aligned}$$

Take u from the right-hand set. It is clear that, on one hand, $\sum_{i=1}^p u_i = 0$, and, on the other hand, if $\bar{x}_i = 0$, then $u_i \geq 0$. Denote by $I(\bar{x}) := \{i \in \overline{1, p} \mid \bar{x}_i = 0\}$. It follows that

$$T(M, \bar{x}) \subset \left\{ u \in \mathbb{R}^p \mid \sum_{i=1}^p u_i = 0 \text{ and } u_i \geq 0, \forall i \in I(\bar{x}) \right\}.$$

Let us now prove the reverse inclusion. It is easy to verify that the right-hand set is closed. Take u from this set. If $u = 0$, then, obviously, $u \in T(M, \bar{x})$. If $u \neq 0$, then we must prove that there exists $\alpha > 0$ such that $\bar{x} + \alpha u \in M$. On one hand, it is clear that $\sum_{i=1}^p (\bar{x}_i + \alpha u_i) = 1$ is satisfied for any α . If there is no i with $u_i < 0$, then it is also easy to observe that $\bar{x}_i + \alpha u_i \geq 0$, for any $i \in \overline{1, p}$, hence $u \in T(M, \bar{x})$. Suppose that the set J of indices for which $u_j < 0$ is nonempty. Then $J \subset \overline{1, p} \setminus I(\bar{x})$, hence $\bar{x}_j > 0$ for any $j \in J$. One can choose the positive α such that

$$\alpha < \min\{-u_j^{-1} \bar{x}_j \mid j \in J\},$$

and again one has $\bar{x}_i + \alpha u_i \geq 0$, for any $i \in \overline{1, p}$. Hence, $u \in T(M, \bar{x})$, and the double inclusion follows.

We prove next that

$$N(M, \bar{x}) = \{ (a, a, \dots, a) \in \mathbb{R}^p \mid a \in \mathbb{R} \} + \{ v \in \mathbb{R}^p \mid v_i \leq 0, \forall i \in I(\bar{x}), v_i = 0, i \notin I(\bar{x}) \}.$$

For this, consider the elements

$$a_0 = (1, 1, \dots, 1), a_1 = -(1, 0, \dots, 0), \dots, a_n = -(0, 0, \dots, 1)$$

and observe that $T(M, \bar{x})$ can be equivalently written as:

$$T(M, \bar{x}) = \{ u \in \mathbb{R}^p \mid \langle a_0, u \rangle \leq 0, \langle -a_0, u \rangle \leq 0, \langle a_i, u \rangle \leq 0, \forall i \in I(\bar{x}) \}.$$

The polar of this set is

$$N(M, \bar{x}) = \left\{ \alpha a_0 - \beta a_0 + \sum_{i \in I(\bar{x})} \alpha_i a_i \mid \alpha, \beta, \alpha_i \geq 0, \forall i \in I(\bar{x}) \right\}.$$

Indeed, the fact that the right-hand set is contained in the normal cone is obvious, and the reverse inclusion follows from Farkas' Lemma (Theorem 2.1.8). We now obtain the desired form of the normal cone.

At the end of this section, we discuss the concepts of convex hull and conic hull of a set. Let $A \subset \mathbb{R}^p$ be a nonempty set. The convex hull of A is the set

$$\text{conv } A = \left\{ \sum_{i=1}^n \alpha_i x_i \mid n \in \mathbb{N}^*, (\alpha_i)_{i \in \overline{1, n}} \subset [0, \infty), \sum_{i=1}^n \alpha_i = 1, (x_i)_{i \in \overline{1, n}} \subset A \right\}.$$

It is not difficult to see that $\text{conv } A$ is a convex set which contains A . One can easily verify that $\text{conv } A$ is the smallest set (in the sense of inclusion) with these properties (see Problem 7.26).

The conic hull of the set A is

$$\text{cone } A := [0, \infty)A := \{ \alpha x \mid \alpha \geq 0, x \in A \}.$$

In fact, $\text{cone } A$ is the smallest cone which contains A .

We give next two results concerning these sets. The first one refers to the structure of the set $\text{conv } A$ and it is called the Carathéodory Theorem, after the name of the Greek mathematician Constantin Carathéodory, who proved this result in 1911 for compact sets.

Theorem 2.1.17 (Carathéodory Theorem). *Let $A \subset \mathbb{R}^p$ be a nonempty set. Then*

$$\text{conv } A = \left\{ \sum_{i=1}^{p+1} \alpha_i x_i \mid (\alpha_i)_{i \in \overline{1, p+1}} \subset [0, \infty), \sum_{i=1}^{p+1} \alpha_i = 1, (x_i)_{i \in \overline{1, p+1}} \subset A \right\}.$$

Proof We must prove that every element from $\text{conv } A$ can be written as a combination of at most $p + 1$ elements from A . Consider $x \in \text{conv } A$. According to the definition of $\text{conv } A$, x can be written as a convex combination of elements from A . Suppose, by means of contradiction, that the minimal number of elements from A which can form a convex combination equal to x is $n > p + 1$. So there exist $x_1, x_2, \dots, x_n \in A$, $\alpha_1, \alpha_2, \dots, \alpha_n \in (0, 1)$ with $\sum_{i=1}^n \alpha_i = 1$ such that $\sum_{i=1}^n \alpha_i x_i = x$. Then the elements $(x_i - x_n)_{i=1, \dots, n-1}$ are linearly dependent (their number is greater than the dimension p of the space), so there exist $(\lambda_i)_{i=1, \dots, n-1}$, not all equal to 0, such that

$$\sum_{i=1}^{n-1} \lambda_i (x_i - x_n) = 0,$$

which means

$$\sum_{i=1}^{n-1} \lambda_i x_i - \left(\sum_{i=1}^{n-1} \lambda_i \right) x_n = 0.$$

By denoting $-\left(\sum_{i=1}^{n-1} \lambda_i \right) = \lambda_n$, one has $\sum_{i=1}^n \lambda_i = 0$ and $\sum_{i=1}^n \lambda_i x_i = 0$. Then for every $t \in \mathbb{R}$,

$$x = \sum_{i=1}^n \alpha_i x_i + t \sum_{i=1}^n \lambda_i x_i = \sum_{i=1}^n (\alpha_i + t \lambda_i) x_i$$

and

$$\sum_{i=1}^n (\alpha_i + t \lambda_i) = 1.$$

As $\sum_{i=1}^n \lambda_i = 0$ and there is at least one nonzero element, there exists at least one negative value among the numbers $(\lambda_i)_{i \in \overline{1, n}}$. Denote $\bar{t} := \min\{-\alpha_i \lambda_i^{-1} \mid \lambda_i < 0\}$. Then all the values $(\alpha_i + \bar{t} \lambda_i)$ are in the interval $[0, \infty)$, and the corresponding value of the index which gives the minimum from above is zero, whence x is a convex combination of less than n elements from A , contradicting the minimality of n . Therefore, the assumption that we made was false, and the conclusion follows. \square

We now discuss the necessary conditions one needs in order that the conic hull of a set is closed. This does not happens automatically, as one can see from the example given by $A \subset \mathbb{R}^2$, $A := \{(x, y) \in \mathbb{R}^2 \mid (x - 1)^2 + y^2 = 1\}$, for which cone $A = \{(x, y) \in \mathbb{R}^2 \mid x > 0\} \cup \{(0, 0)\}$.

First, one defines for a nonempty set $A \subset X$ the asymptotic cone of A as

$$A^\infty = \{u \in X \mid \exists (t_n) \rightarrow 0, \exists (a_n) \subset A, t_n a_n \rightarrow u\}.$$

It is clear, by repeating the arguments from the case of the Bouligand tangent cone, that A^∞ is a closed cone. If A is bounded, then $A^\infty = \{0\}$, and the converse also holds (if A would contain an unbounded sequence (a_n) , then $\left(\frac{a_n}{\|a_n\|} \right)$ would also have a

subsequence which converges to a nonzero element, which must be from A^∞). Let us observe also that if A is a cone, then $A^\infty = T_B(A, 0) = \text{cl } A$.

The next result concerns decomposition.

Theorem 2.1.18. *Let $A \subset \mathbb{R}^p$ be a nonempty closed set.*

- (i) *If $0 \notin A$, then $\text{cl cone } A = \text{cone } A \cup A^\infty$.*
- (ii) *If $0 \in A$, then $\text{cl cone } A = \text{cone } A \cup A^\infty \cup T_B(A, 0)$.*

Proof (i) Suppose that $0 \notin A$. It is clear that $\text{cone } A \subset \text{cl cone } A$. From the definition of A^∞ , one also has that $A^\infty \subset \text{cl cone } A$. For the reverse inclusion, take $(u_n) \subset \text{cone } A$, $u_n \rightarrow u$. We must prove that $u \in \text{cone } A \cup A^\infty$. If $u = 0$, then the relation $u \in \text{cone } A$ is obvious. Suppose that $u \neq 0$, then for every $n \in \mathbb{N}$, there exist $t_n \geq 0$ and $a_n \in A$ such that $u_n = t_n a_n$. If (a_n) is unbounded, one can pass to a subsequence (a_{n_k}) , where $\|a_{n_k}\| \rightarrow \infty$. Hence $t_{n_k} \rightarrow 0$, and $u \in A^\infty$. Suppose (a_n) is bounded. Since $0 \notin A$ and A is closed, there exists $\gamma > 0$ such that $\|a_n\| \geq \gamma$ for every n . One can deduce that (t_n) is bounded, so it converges (on a subsequence (t_{n_k}) , eventually) to a number $t \geq 0$. If $t = 0$, then $u_{n_k} \rightarrow 0 = u$ (a situation which is excluded at this point of the proof). Accordingly, $t > 0$ and

$$\begin{aligned} \|a_{n_k} - t^{-1}u\| &= t^{-1} \|ta_{n_k} - u\| = t^{-1} \|t_{n_k}a_{n_k} - u + (t - t_{n_k})a_{n_k}\| \\ &\leq t^{-1} \|t_{n_k}a_{n_k} - u\| + |t - t_{n_k}| t^{-1} \|a_{n_k}\| \rightarrow 0. \end{aligned}$$

Hence $a_{n_k} \rightarrow t^{-1}u$ and since A is closed, $u \in \text{cone } A$. The proof of this part is complete.

(ii) The inclusion $\text{cone } A \cup A^\infty \cup T_B(A, 0) \subset \text{cl cone } A$ is obvious. Take $u \in \text{cl cone } A$. In the above it is possible as well that (a_n) converges to 0. Then $t_n \rightarrow \infty$, hence $u \in T_B(A, 0)$, which completes the proof. \square

We finish with the following characterization result:

Corollary 2.1.19. *Let $A \subset \mathbb{R}^p$ be a nonempty closed set.*

- (i) *If $0 \notin A$, then $\text{cone } A$ is closed if and only if $A^\infty \subset \text{cone } A$.*
- (ii) *If $0 \in A$, then $\text{cone } A$ is closed if and only if $A^\infty \cup T_B(A, 0) \subset \text{cone } A$.*

2.2 Convex Functions

2.2.1 General Results

In this section we present the special class of convex functions. These functions are defined on convex sets.

Definition 2.2.1. *Let $D \subset \mathbb{R}^p$ be a convex set. One says that a function $f : D \rightarrow \mathbb{R}$ is convex if*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in D, \quad \forall \lambda \in [0, 1]. \quad (2.2.1)$$

It is clear that in the above definition it is sufficient to take $\lambda \in (0, 1)$.

As said before, in \mathbb{R} the convex sets are exactly the intervals. In this framework, the convexity has the following geometric meaning: for every two points $x, y \in D$, $x < y$, the graph of the restriction of f to the $[x, y]$ interval lies below the line segment joining the points $(x, f(x))$ and $(y, f(y))$. This can be written as follows: for every $u \in [x, y]$,

$$f(u) \leq f(x) + \frac{f(y) - f(x)}{y - x}(u - x), \quad (2.2.2)$$

inequality which can be deduced from (2.2.1) by replacing λ with the value given by the relation $u = \lambda x + (1 - \lambda)y$. Therefore, (2.2.1) and (2.2.2) are equivalent (for functions defined on \mathbb{R}).

Definition 2.2.2. Let $D \subset \mathbb{R}^p$ be a convex set. One says that a function $f : D \rightarrow \mathbb{R}$ is concave if $-f$ is convex.

All of the properties of concave functions can be easily deduced from the similar properties of the convex functions, so in what follows we will consider only the later case.

We first deduce some general properties of the convex functions.

Proposition 2.2.3. Let $D \subset \mathbb{R}^p$ be a convex set and $f : D \rightarrow \mathbb{R}$. The following relations are equivalent:

- (i) f is convex;
- (ii) the epigraph of f ,

$$\text{epi } f := \{(x, t) \in D \times \mathbb{R} \mid f(x) \leq t\},$$

is a convex subset of $\mathbb{R}^p \times \mathbb{R}$;

- (iii) for any $x, y \in D$, define

$$I_{x,y} := \{t \in \mathbb{R} \mid tx + (1 - t)y \in D\};$$

then the function $\varphi_{x,y} : I_{x,y} \rightarrow \mathbb{R}$, $\varphi_{x,y}(t) = f(tx + (1 - t)y)$ is convex.

Proof We prove first the implication from (i) to (ii). Take $\lambda \in (0, 1)$ and $(x, t), (y, s) \in \text{epi } f$. By the convexity of D , one knows that $\lambda x + (1 - \lambda)y \in D$, and using the convexity of f , one can say:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \leq \lambda t + (1 - \lambda)s,$$

i.e., $(\lambda x + (1 - \lambda)y, \lambda t + (1 - \lambda)s) \in \text{epi } f$. Therefore, $\text{epi } f$ is a convex set.

We prove now the converse implication. Take $x, y \in D$ and $\lambda \in [0, 1]$. Then $(x, f(x)), (y, f(y)) \in \text{epi } f$ and by assumption, $\lambda(x, f(x)) + (1 - \lambda)(y, f(y)) \in \text{epi } f$, hence

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y),$$

which shows that f is a convex function.

We prove now the equivalence between (i) and (iii). Observe first that $I_{x,y}$ is an interval which contains $[0, 1]$. Suppose that f is convex and take $u, v \in I_{x,y}, \lambda \in [0, 1]$. One knows that:

$$\begin{aligned} \varphi_{x,y}(\lambda u + (1 - \lambda)v) &= f([\lambda u + (1 - \lambda)v]x + [1 - \lambda u - (1 - \lambda)v]y) \\ &= f(\lambda(ux + (1 - u)y) + (1 - \lambda)(vx + (1 - v)y)) \\ &\leq \lambda f(ux + (1 - u)y) + (1 - \lambda)f(vx + (1 - v)y) \\ &= \lambda \varphi_{x,y}(u) + (1 - \lambda)\varphi_{x,y}(v). \end{aligned}$$

For the converse implication, take $x, y \in D$ and $t \in [0, 1]$. Then $\varphi_{x,y}$ is convex, hence for every $\lambda \in [0, 1], u, v \in I_{x,y}$

$$\begin{aligned} \varphi_{x,y}(\lambda u + (1 - \lambda)v) &\leq \lambda \varphi_{x,y}(u) + (1 - \lambda)\varphi_{x,y}(v) \\ &= \lambda f(ux + (1 - u)y) + (1 - \lambda)f(vx + (1 - v)y). \end{aligned}$$

By taking $u = 1, v = 0, \lambda = t$ we deduce that

$$\varphi_{x,y}(t) \leq tf(x) + (1 - t)f(y),$$

hence f is convex. □

Theorem 2.2.4. *Let $D \subset \mathbb{R}^p$ be a convex set and $f : D \rightarrow \mathbb{R}$ be a convex function. Then f is continuous at every interior point of D .*

Proof Take $\bar{x} \in \text{int } D$. A translation permits us to consider the case $\bar{x} = 0$. We prove first that f is bounded on a neighborhood of 0. If we denote by $(e_i)_{i \in \overline{1,p}}$ the canonical base of \mathbb{R}^p , then there exists an $a > 0$ such that ae_i and $-ae_i$ are in D for any $i \in \overline{1,p}$. Under these conditions, the set

$$V := \left\{ x \in \mathbb{R}^p \mid x = \sum_{i=1}^p x_i e_i, |x_i| < \frac{a}{p}, \forall i \in \overline{1,p} \right\}$$

is a neighborhood of 0 contained in D . For $x \in V$, there exist $(x_i)_{i \in \overline{1,p}}$ with $|x_i| < \frac{a}{p}$, $i \in \overline{1,p}$ and $x = \sum_{i=1}^p x_i e_i$. Suppose first that $x_i \neq 0$ for any $i \in \overline{1,p}$. One has

$$\begin{aligned} f(x) &= f\left(\sum_{i=1}^p x_i e_i\right) = f\left(\sum_{i=1}^p \frac{|x_i|}{a} a \frac{x_i}{|x_i|} e_i + \left(1 - \sum_{i=1}^p \frac{|x_i|}{a}\right) 0\right) \\ &\leq \sum_{i=1}^p \frac{|x_i|}{a} f\left(a \frac{x_i}{|x_i|} e_i\right) + \left(1 - \sum_{i=1}^p \frac{|x_i|}{a}\right) f(0) \\ &\leq \max\{f(ae_i), f(-ae_i) \mid i \in \overline{1,p}\} + |f(0)|. \end{aligned}$$

We can now observe that if there are indices i for which $x_i = 0$, then these can be excluded from the above calculations, and the estimation holds.

Since the right-hand part is a constant (which we denote by M), the proof is finished. Take $\varepsilon \in (0, 1)$ and U a symmetric neighborhood of 0 such that $\varepsilon^{-1}U \subset V$. Then, for any $x \in U$,

$$f(x) = f\left(\varepsilon(\varepsilon^{-1}x) + (1 - \varepsilon)0\right) \leq \varepsilon f(\varepsilon^{-1}x) + (1 - \varepsilon)f(0) \leq \varepsilon M + (1 - \varepsilon)f(0),$$

i.e.,

$$f(x) - f(0) \leq \varepsilon M - \varepsilon f(0).$$

From the fact that U is symmetric, one deduces that for every $x \in U$

$$f(-x) \leq \varepsilon M + (1 - \varepsilon)f(0).$$

Moreover,

$$f(0) = f\left(\frac{1}{2}x + \frac{1}{2}(-x)\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(-x) \leq \frac{1}{2}f(x) + \frac{1}{2}(\varepsilon M + (1 - \varepsilon)f(0)),$$

hence

$$f(0) - f(x) \leq \varepsilon M - \varepsilon f(0).$$

This relation can be combined with the similar one from above, and gives

$$|f(x) - f(0)| \leq \varepsilon M - \varepsilon f(0).$$

This inequality proves the continuity of f at 0. □

We want to emphasize now some characterizations of differentiable convex functions. Some preliminary results on convex functions defined on a real intervals are necessary.

Proposition 2.2.5. *Let $I \subset \mathbb{R}$ be an interval and $f : I \rightarrow \mathbb{R}$ be a function. The next relations are equivalent:*

(i) f is convex;

(ii) for every $x_1, x_2, x_3 \in I$ satisfying the relation $x_1 < x_2 < x_3$ one has

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2};$$

(iii) for every $a \in \text{int} I$, the function $g : I \setminus \{a\} \rightarrow \mathbb{R}$ given by

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

is increasing.

Proof We prove the (i) \Rightarrow (ii) implication. Take $\lambda = \frac{x_2 - x_1}{x_3 - x_1} \in (0, 1)$. Then the equality $x_2 = \lambda x_3 + (1 - \lambda)x_1$ holds and one must prove now that

$$\frac{f(x_2) - f(x_1)}{\lambda(x_3 - x_1)} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1} \leq \frac{f(x_3) - f(x_2)}{(1 - \lambda)(x_3 - x_1)}.$$

After some calculations, one can show that:

$$f(x_2) \leq \lambda f(x_3) + (1 - \lambda)f(x_1).$$

The proof of the implication (ii) \Rightarrow (i) follows the inverse path of the proof of (i) \Rightarrow (ii), hence (i) and (ii) are equivalent.

In order to prove (ii) \Rightarrow (iii), we fix $x_1, x_2 \in I \setminus \{a\}$ with $x_1 < x_2$ and we find three situations. If $x_1 < x_2 < a$, then we apply (ii) for the triplet (x_1, x_2, a) . If $x_1 < a < x_2$, then we apply (ii) for the triplet (x_1, a, x_2) . Finally, if $a < x_1 < x_2$, then we apply (ii) for the triplet (a, x_1, x_2) .

We prove now (iii) \Rightarrow (i). Take $x, y \in I$ with $x < y$ and $\lambda \in (0, 1)$. Then $x < \lambda x + (1 - \lambda)y < y$, and by applying (iii) with $a = \lambda x + (1 - \lambda)y$, one deduces

$$\frac{f(x) - f(\lambda x + (1 - \lambda)y)}{x - \lambda x - (1 - \lambda)y} \leq \frac{f(y) - f(\lambda x + (1 - \lambda)y)}{y - \lambda x - (1 - \lambda)y}.$$

After some calculations, the relation follows from the definition of convexity. The fact that this relation holds for any $x, y \in I$ with $x < y$ and for any $\lambda \in (0, 1)$ is sufficient to prove the desired assertion. The proof is complete. \square

Proposition 2.2.6. *Let $I \subset \mathbb{R}$ be an interval and $f : I \rightarrow \mathbb{R}$ be a convex function. Then f admits lateral derivatives in every interior point of I and for every $x, y \in \text{int } I$ with $x < y$, one has*

$$f'_-(x) \leq f'_+(x) \leq f'_-(y) \leq f'_+(y).$$

Proof Fix $a \in \text{int } I$. Since the function $g : I \setminus \{a\} \rightarrow \mathbb{R}$ given by

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

is increasing (see the previous result) one deduces that g admits finite lateral limits, which implies the existence of the lateral derivatives of f at a . Moreover, $f'_-(a) \leq f'_+(a)$. For $x, y \in \text{int } I$, $x < y$ and for any $u, v \in (x, y)$, $u \leq v$, by using again the argument given by the last conclusion of Proposition 2.2.5, one deduces that

$$\frac{f(u) - f(x)}{u - x} \leq \frac{f(v) - f(x)}{v - x} = \frac{f(x) - f(v)}{x - v} \leq \frac{f(y) - f(v)}{y - v} = \frac{f(v) - f(y)}{v - y}.$$

Passing to the limit for $u \rightarrow x$ and $v \rightarrow y$, one gets $f'_+(x) \leq f'_-(y)$. \square

Here we characterize differentiable convex functions of one variable.

Theorem 2.2.7. *Let I be an open interval and $f : I \rightarrow \mathbb{R}$ be a function.*

(i) *If f is differentiable on I , then f is convex if and only if f' is increasing on I .*

(ii) *If f is twice differentiable on I , then f is convex if and only if $f''(x) \geq 0$ for every $x \in I$.*

Proof In the case of real functions of one variable, the equivalence between the monotonicity of f' and the sign of f'' is sufficient to prove that f is convex if and only if f' is increasing on I . If f is convex, the monotonicity of the derivative follows from Proposition 2.2.6. Conversely, suppose that f' is increasing and we prove that f is convex. Take $a, b \in I$. Define $g : [a, b] \rightarrow \mathbb{R}$ given by

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

Obviously, $g(a) = g(b) = 0$, and

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

The function f satisfies the conditions of Lagrange Theorem on $[a, b]$, hence there exists $c \in (a, b)$ such that

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

Consequently, $g'(x) = f'(x) - f'(c)$. From the monotonicity of f' , we deduce that g is decreasing on (a, c) and increasing on (c, b) , and since $g(a) = g(b) = 0$, we know that g is negative on the whole interval $[a, b]$. Take $x \in (a, b)$. Then there exists $\lambda \in (0, 1)$ such that

$$x = \lambda a + (1 - \lambda)b.$$

By replacing x in the expression of g and taking into account that $g(x) \leq 0$, we deduce that

$$f(\lambda a + (1 - \lambda)b) - f(a) - (1 - \lambda)(b - a) \frac{f(b) - f(a)}{b - a} \leq 0,$$

relation which reduces to the definition of the convexity. \square

Example 2.2.8. Based on the above result, one deduces the convexity of the following functions: $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = ax + b$, with $a, b \in \mathbb{R}$; $f : (0, \infty) \rightarrow \mathbb{R}$, $f(x) = -\ln x$; $f : (0, \infty) \rightarrow \mathbb{R}$, $f(x) = x \ln x$; $f : (0, \infty) \rightarrow \mathbb{R}$, $f(x) = x^a$, $a \geq 1$; $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = e^x$; $f : (-1, 1) \rightarrow \mathbb{R}$, $f(x) = -\sqrt{1 - x^2}$; $f : (0, \pi) \rightarrow \mathbb{R}$, $f(x) = \sin^{-1} x$.

Another example is given by the next result.

Proposition 2.2.9. Let $D \subset \mathbb{R}^p$ be a nonempty convex set. Then the function $d_D : \mathbb{R}^p \rightarrow \mathbb{R}$ given by $d_D(x) = d(x, D)$ is convex.

Proof Take $x, y \in \mathbb{R}^p$ and $\alpha \in [0, 1]$. For any $\varepsilon > 0$, there exist $d_{x,\varepsilon}, d_{y,\varepsilon} \in D$ such that

$$\begin{aligned} \|d_{x,\varepsilon} - x\| &< d_D(x) + \varepsilon \\ \|d_{y,\varepsilon} - y\| &< d_D(y) + \varepsilon. \end{aligned}$$

Using the convexity of D , one knows:

$$\begin{aligned} d_D(\alpha x + (1 - \alpha)y) &\leq \|\alpha x + (1 - \alpha)y - (\alpha d_{x,\varepsilon} + (1 - \alpha)d_{y,\varepsilon})\| \\ &\leq \alpha \|d_{x,\varepsilon} - x\| + (1 - \alpha) \|d_{y,\varepsilon} - y\| \\ &< \alpha d_D(x) + (1 - \alpha)d_D(y) + \varepsilon. \end{aligned}$$

As ε is arbitrarily chosen, we may pass to the limit for $\varepsilon \rightarrow 0$ and the conclusion follows. \square

We now characterize differentiable convex functions in the general case.

Theorem 2.2.10. *Let $D \subset \mathbb{R}^p$ be an open convex set and $f : D \rightarrow \mathbb{R}$ be a function.*

(i) *If f is differentiable on D , then f is convex if and only if for any $x, y \in D$,*

$$f(y) \geq f(x) + \nabla f(x)(y - x).$$

(ii) *If f is twice differentiable on D , then f is convex if and only if for every $x \in D$ and $y \in \mathbb{R}^p$, one has*

$$\nabla^2 f(x)(y, y) \geq 0.$$

Proof (i) Consider first the case when $p = 1$. Fix $x \in D$ with $y \neq x$ and take $\lambda \in (0, 1]$. Since f is convex, we get

$$\begin{aligned} f(x + \lambda(y - x)) &= f((1 - \lambda)x + \lambda y) \\ &\leq (1 - \lambda)f(x) + \lambda f(y) = f(x) + \lambda(f(y) - f(x)). \end{aligned}$$

Consequently,

$$\frac{f(x + \lambda(y - x)) - f(x)}{\lambda(y - x)} \cdot (y - x) \leq f(y) - f(x).$$

Passing to the limit for $\lambda \rightarrow 0$, one gets $f'(x)(y - x) \leq f(y) - f(x)$.

We pass now to the general case. For $x \in D$, consider the function $\varphi_{y,x}$ from Proposition 2.2.3, which we already know that is convex. Moreover, $\varphi_{y,x}$ is differentiable on the open interval $I_{y,x}$, and

$$\varphi'_{y,x}(t) = \nabla f(ty + (1 - t)x)(y - x).$$

According to the preceding step,

$$\varphi_{y,x}(1) \geq \varphi_{y,x}(0) + \varphi'_{y,x}(0)$$

which means that

$$f(y) \geq f(x) + \nabla f(x)(y - x).$$

Conversely, fix $x, y \in D$ and $\lambda \in [0, 1]$. Therefore, by assumption,

$$f(x) \geq f(\lambda x + (1 - \lambda)y) + (1 - \lambda)\nabla f(\lambda x + (1 - \lambda)y)(x - y)$$

and

$$f(y) \geq f(\lambda x + (1 - \lambda)y) + \lambda\nabla f(\lambda x + (1 - \lambda)y)(y - x).$$

Multiplying the first inequality by λ , the second one by $(1 - \lambda)$, and summing up the new inequalities, one obtains

$$\lambda f(x) + (1 - \lambda)f(y) \geq f(\lambda x + (1 - \lambda)y),$$

which proves that f is convex.

(ii) The case $p = 1$ is proved in Theorem 2.2.7. Now, in order to pass to the general case, take $x \in D$, $y \in \mathbb{R}^p$. Suppose f is convex. Since D is open, there exists an $\alpha > 0$ such that $u := x + \alpha y \in D$. According to the assumption, $\varphi_{u,x}$ is convex, and taking into account the case that we have already studied, $\varphi''_{u,x}(t) \geq 0$ for any $t \in I_{u,x}$. For $t = 0$, one deduces that

$$0 \leq \varphi''_{u,x}(0) = \nabla^2 f(x)(u - x, u - x),$$

and the conclusion follows. Conversely, for $x, y \in D$ and $t \in I_{x,y}$, $\varphi''_{x,y}(t) \geq 0$. From the case $p = 1$, we get that $\varphi_{x,y}$ is convex, hence f is convex. The proof is now complete. \square

From these results, one may observe that some properties of the convex functions have a global character, an aspect which will persist in subsequent sections.

At the end of this subsection, we will discuss a property which is stronger than convexity.

Definition 2.2.11. Let $D \subset \mathbb{R}^p$ be a convex set. One says that a function $f : D \rightarrow \mathbb{R}$ is strictly convex if

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in D, \quad x \neq y, \quad \forall \lambda \in (0, 1).$$

Definition 2.2.12. Let $D \subset \mathbb{R}^p$ be a convex set. One says that a function $f : D \rightarrow \mathbb{R}$ is strictly concave if $-f$ is strictly convex.

Every strictly convex function is convex, but the converse is false. To see this, consider a convex function which is constant on an interval. Again, the properties of the strictly concave functions easily follow from the corresponding ones of the strictly convex functions.

By the use of very similar arguments as in the proofs of preceding results, one can deduce the next characterizations.

Theorem 2.2.13. Let I be an open interval and $f : I \rightarrow \mathbb{R}$ be a differentiable function. The next assertions are equivalent:

- (i) f is strictly convex;
- (ii) $f(x) > f(a) + f'(a)(x - a)$, for any $x, a \in I, x \neq a$;
- (iii) f' is strictly increasing.

If, moreover, f is twice differentiable (on I), then one more equivalence holds :

(iv) $f''(x) \geq 0$ for any $t \in I$ and $\{x \in I \mid f''(x) = 0\}$ does not contain any proper interval.

Example 2.2.14. Using this result, one gets the strict convexity of the following functions: $f : (0, \infty) \rightarrow \mathbb{R}, f(x) = -\ln x$; $f : (0, \infty) \rightarrow \mathbb{R}, f(x) = x \ln x$; $f : (0, \infty) \rightarrow \mathbb{R}, f(x) = x^a, a > 1$; $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = e^x$; $f : (0, \infty) \rightarrow \mathbb{R}, f(x) = (1 + x^p)^{\frac{1}{p}}, p > 1$.

Theorem 2.2.15. Let $D \subset \mathbb{R}^p$ an open convex set and $f : D \rightarrow \mathbb{R}$ be a differentiable function. The next assertions are equivalent:

- (i) f is strictly convex;
- (ii) $f(x) > f(a) + \nabla f(a)(x - a)$, for any $x, a \in D, x \neq a$.

If, moreover, f is twice differentiable (on D), then the preceding two items are implied by the relation:

- (iii) $\nabla^2 f(x)(y, y) > 0$ for any $x \in D$ and $y \in \mathbb{R}^p \setminus \{0\}$.

2.2.2 Convex Functions of One Variable

In this subsection we will focus on some properties and applications of convex functions defined on real intervals, even if some of the results hold in more general situations.

The class of convex functions is stable under several algebraic operations, an aspect which makes it very useful. Here are some of these operations:

Proposition 2.2.16. Let $I, J \subset \mathbb{R}$ be intervals.

(i) Let $n \in \mathbb{N}^*$ and $f_1, f_2, \dots, f_n : I \rightarrow \mathbb{R}$ be convex functions, and $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$. Then $\sum_{i=1}^n \lambda_i f_i$ is convex. If at least one of the functions is strictly convex, and the corresponding scalar is not zero, then $\sum_{i=1}^n \lambda_i f_i$ is strictly convex.

(ii) Let $f : I \rightarrow J$ be (strictly) convex, $g : J \rightarrow \mathbb{R}$ be convex and (strictly) increasing. Then $g \circ f$ is (strictly) convex.

(iii) Let $f : I \rightarrow J$ be strictly decreasing, (strictly) convex, and surjective. Then f^{-1} is (strictly) convex.

The next result emphasizes some monotonicity properties of convex functions of one variable.

Theorem 2.2.17. *Let I be a nondegenerate interval (i.e., not a singleton set) and $f : I \rightarrow \mathbb{R}$ be convex. Then either f is monotone on $\text{int } I$, or there exists $\bar{x} \in \text{int } I$ such that f is decreasing on $I \cap (-\infty, \bar{x}]$, and increasing on $I \cap [\bar{x}, \infty)$.*

Proof Because of relation (2.2.2), it is sufficient to restrict our attention to the case when I is open, i.e., $I = \text{int } I$. Suppose that f is not monotone on I . Then there exist $a, b, c \in I$, $a < b < c$ such that $f(a) > f(b) < f(c)$ or $f(a) < f(b) > f(c)$. The second situation cannot hold, since in that case, using (2.2.2), one would have

$$f(b) \leq f(a) + \frac{f(c) - f(a)}{c - a}(b - a) = \frac{f(a)(c - b) + f(c)(b - a)}{c - a} < f(b).$$

Hence, $f(a) > f(b) < f(c)$. As f is continuous on $[a, c]$, its minimum on this interval must be attained at a point \bar{x} (from the Weierstrass Theorem). Take $x \in I \cap (-\infty, a)$. According to Proposition 2.2.5,

$$\frac{f(x) - f(\bar{x})}{x - \bar{x}} \leq \frac{f(a) - f(\bar{x})}{a - \bar{x}},$$

i.e.,

$$(\bar{x} - a)f(x) \geq (x - a)f(\bar{x}) + (\bar{x} - x)f(a) \geq (\bar{x} - a)f(\bar{x}),$$

hence $f(\bar{x}) \leq f(x)$. Similarly, one can prove that $f(\bar{x}) \leq f(x)$ for $x \in I \cap (c, \infty)$. It follows that $f(\bar{x}) = \inf f(I)$. We next prove that f is decreasing on $I \cap (-\infty, \bar{x})$. Take $u, v \in I \cap (-\infty, \bar{x})$, $u < v$. On one hand,

$$\frac{f(v) - f(u)}{v - u} \leq \frac{f(\bar{x}) - f(u)}{\bar{x} - u},$$

and on the other hand,

$$\frac{f(\bar{x}) - f(u)}{\bar{x} - u} = \frac{f(u) - f(\bar{x})}{u - \bar{x}} \leq \frac{f(v) - f(\bar{x})}{v - \bar{x}} \leq 0,$$

hence $f(v) - f(u) \leq 0$, which is exactly what we wanted to prove. Similarly, one can show that f is increasing on $I \cap (\bar{x}, +\infty)$. The continuity of f on I finalizes the proof. \square

As we saw before, a convex function defined on an interval can have discontinuities only at the extremities of the interval. The previous theorem allows us to consider a different function in those eventual discontinuity points, without losing the convexity. In this way one obtains the next consequence.

Corollary 2.2.18. *Let $a, b \in \mathbb{R}$, $a < b$ and $f : [a, b] \rightarrow \mathbb{R}$ be a convex function. Then there exist $\lim_{x \rightarrow a^+} f(x)$ and $\lim_{x \rightarrow b^-} f(x)$, and the function*

$$\bar{f}(x) = \begin{cases} \lim_{x \rightarrow a^+} f(x), & x = a \\ f(x), & x \in (a, b) \\ \lim_{x \rightarrow b^-} f(x), & x = b \end{cases}$$

is convex and continuous on $[a, b]$.

From the proof of Theorem 2.2.17 one can also obtain the next result, which will be restated in a more general framework in the next chapter.

Corollary 2.2.19. *Let $I \subset \mathbb{R}$ be a nondegenerate interval. If $f : I \rightarrow \mathbb{R}$ is a convex and non-monotone function, then it has a global minimum on $\text{int } I$.*

The next proposition, sometimes called the Jensen inequality, follows by applying the definitions and the mathematical induction principle.

Proposition 2.2.20. *Let $D \subset \mathbb{R}^p$ be a convex set and $f : D \rightarrow \mathbb{R}$. If the function f is convex, then*

$$f(\lambda_1 x_1 + \dots + \lambda_m x_m) \leq \lambda_1 f(x_1) + \dots + \lambda_m f(x_m)$$

for any $m \in \mathbb{N}^*$, $x_1, \dots, x_m \in D$, $\lambda_1, \dots, \lambda_m \geq 0$, $\lambda_1 + \dots + \lambda_m = 1$. The inequality is strict if the function f is strictly convex, at least two of the points (x_k) are different and the corresponding scalars (λ_k) are strictly positive.

Actually, for the convexity of a continuous function it is sufficient that the inequality from the definition is satisfied for $\lambda = 2^{-1}$.

Theorem 2.2.21. *Let $I \subset \mathbb{R}$ be an interval and $f : I \rightarrow \mathbb{R}$ be a continuous function. The function f is convex if and only if*

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}, \quad \forall x, y \in I. \quad (2.2.3)$$

Proof The necessity of the condition (2.2.3) is obvious. Let us prove that it is also sufficient. Suppose, by contradiction, that f is not convex, which means (2.2.2) is not satisfied. Then there exist $x, y \in I$, $x < y$ and $u \in (x, y)$ such that

$$f(u) > f(x) + \frac{f(y) - f(x)}{y - x}(u - x). \quad (2.2.4)$$

Observe by the relation (2.2.4) that the cases $u = x$ and $u = y$ cannot hold. Consider then $g : [x, y] \rightarrow \mathbb{R}$,

$$g(t) = f(t) - f(x) - \frac{f(y) - f(x)}{y - x}(t - x),$$

which is continuous and satisfies the relations $g(x) = g(y) = 0$. By (2.2.4) and Weierstrass' Theorem, there exists $z \in (x, y)$ such that $g(z) = \sup_{t \in [x, y]} g(t) > 0$. Denote by

$$w := \inf\{z \in (x, y) \mid g(z) = \sup_{t \in [x, y]} g(t)\}.$$

By the continuity of g , it follows that $g(w) = \sup_{t \in [a, b]} g(t) > 0$, and hence $w \in (x, y)$. Consequently, there exists $h > 0$ such that $w + h, w - h \in (x, y)$. But $w = 2^{-1}(w + h) +$

$2^{-1}(w - h)$, and also $g(w) \geq g(w + h)$ and $g(w) > g(w - h)$. Accordingly,

$$\begin{aligned} \frac{g(w - h) + g(w + h)}{2} &< g(w) = f(w) - f(x) - \frac{f(y) - f(x)}{y - x}(w - x) \\ &\leq \frac{f(w + h) + f(w - h)}{2} - f(x) - \frac{f(y) - f(x)}{y - x} \left(\frac{w + h + w - h}{2} - x \right) \\ &= \frac{g(w - h) + g(w + h)}{2}, \end{aligned}$$

which is a contradiction. It follows that the assumption made is false, hence f is convex. \square

Corollary 2.2.22. *Let $I \subset \mathbb{R}$ be an interval and $f : I \rightarrow \mathbb{R}$ be a continuous function. The function f is convex if and only if for any $x \in I$ and $h > 0$ with $x + h, x - h \in I$, one has*

$$f(x + h) + f(x - h) - 2f(x) \geq 0.$$

In the case of strictly convex functions, the results are similar.

Proposition 2.2.23. *Let $I \subset \mathbb{R}$ be an interval and $f : I \rightarrow \mathbb{R}$ be a continuous function. The function f is strictly convex if and only if*

$$f\left(\frac{x + y}{2}\right) < \frac{f(x) + f(y)}{2}, \quad \forall x, y \in I, x \neq y.$$

Corollary 2.2.24. *Let $I \subset \mathbb{R}$ be an interval and $f : I \rightarrow \mathbb{R}$ be a continuous function. The function f is strictly convex if and only if for any $x \in I$ and $h > 0$ with $x + h, x - h \in I$, one has*

$$f(x + h) + f(x - h) - 2f(x) > 0.$$

For the case of triplets, a result was proved in 1965 by the Romanian mathematician Tiberiu Popoviciu.

Theorem 2.2.25. *Let $I \subset \mathbb{R}$ be an interval and $f : I \rightarrow \mathbb{R}$ be a continuous function. The function f is convex if and only if for any $x, y, z \in I$, one has*

$$\frac{2}{3} \left[f\left(\frac{x + y}{2}\right) + f\left(\frac{y + z}{2}\right) + f\left(\frac{z + x}{2}\right) \right] \leq f\left(\frac{x + y + z}{3}\right) + \frac{f(x) + f(y) + f(z)}{3}. \quad (2.2.5)$$

Proof We prove first the necessity of condition (2.2.5). As in the case of Theorem 2.2.21, continuity is not involved in this step. Without losing the generality, suppose that $x \leq y \leq z$. If $y \leq 3^{-1}(x + y + z)$, then

$$\frac{x + y + z}{3} \leq \frac{x + z}{2} \leq z \quad \text{and} \quad \frac{x + y + z}{3} \leq \frac{y + z}{2} \leq z,$$

hence there exist $s, t \in [0, 1]$ such that

$$\frac{x + z}{2} = s \frac{x + y + z}{3} + (1 - s)z$$

$$\frac{y+z}{2} = t \frac{x+y+z}{3} + (1-t)z.$$

By summation, one gets

$$(x+y-2z)(s+t-2^{-1}3) = 0.$$

If $x+y=2z$, then $x=y=z$ and (2.2.5) is obvious. If $s+t=2^{-1}3$, then by summing the inequalities

$$\begin{aligned} f\left(\frac{x+z}{2}\right) &\leq sf\left(\frac{x+y+z}{3}\right) + (1-s)f(z) \\ f\left(\frac{y+z}{2}\right) &\leq tf\left(\frac{x+y+z}{3}\right) + (1-t)f(z) \\ f\left(\frac{x+y}{2}\right) &\leq \frac{f(x)+f(y)}{2} \end{aligned}$$

and by multiplying with $3^{-1}2$, one gets (2.2.5). The case $y > 3^{-1}(x+y+z)$ is similar.

In order to prove the sufficiency, observe that for $y=z$ in (2.2.5), one gets

$$\frac{1}{4}f(x) + \frac{3}{4}f\left(\frac{x+2y}{3}\right) \geq f\left(\frac{x+y}{2}\right), \quad \forall x, y \in I.$$

From now on, one can follow the arguments from the proof of the sufficiency in Theorem 2.2.21. \square

If f is strictly convex, then the Popoviciu inequality (2.2.5) is strict, except the case when $x=y=z$.

2.2.3 Inequalities

We shall now formulate several inequalities which follow from the general results previously given. In many cases, these inequalities can be seen as examples of discrete optimization and have, in general, a wide applicability in different mathematical areas.

We begin with a refinement of the inequality from Theorem 2.2.21, widely known under the name of Hermite-Hadamard inequality.

Theorem 2.2.26 (Hermite-Hadamard). *Let $a, b \in \mathbb{R}$, $a < b$ and $f : [a, b] \rightarrow \mathbb{R}$ be a convex function. Then the next inequality holds:*

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a} \int_a^b f(x)dx \leq \frac{f(a)+f(b)}{2}.$$

The equality is obtained if and only if f is affine.

Proof Observe first that, because of the continuity of f on (a, b) , f is Riemann integrable on $[a, b]$. Moreover, from convexity, one has, for any $\lambda \in [0, 1]$,

$$f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b).$$

By integration with respect to λ , one gets

$$\int_0^1 f(\lambda a + (1 - \lambda)b) d\lambda \leq \frac{f(a) + f(b)}{2}.$$

On the other hand, for any λ ,

$$\begin{aligned} f\left(\frac{a+b}{2}\right) &= f\left(\frac{\lambda a + (1 - \lambda)b}{2} + \frac{(1 - \lambda)a + \lambda b}{2}\right) \\ &\leq \frac{1}{2}f(\lambda a + (1 - \lambda)b) + \frac{1}{2}f((1 - \lambda)a + \lambda b). \end{aligned}$$

By integrating again with respect to λ and changing the variable, it follows that

$$\begin{aligned} f\left(\frac{a+b}{2}\right) &\leq \frac{1}{2} \int_0^1 f(\lambda a + (1 - \lambda)b) d\lambda + \frac{1}{2} \int_0^1 f((1 - \lambda)a + \lambda b) d\lambda \\ &= \int_0^1 f(\lambda a + (1 - \lambda)b) d\lambda. \end{aligned}$$

Therefore,

$$f\left(\frac{a+b}{2}\right) \leq \int_0^1 f(\lambda a + (1 - \lambda)b) d\lambda \leq \frac{f(a) + f(b)}{2}.$$

Finally, if we change the variable to $\lambda a + (1 - \lambda)b = x$ in the integral, then we get the desired inequalities.

The second inequality can be deduced by the integration of (2.2.2) written for the interval $[a, b]$.

For the second conclusion, define $g : [a, b] \rightarrow \mathbb{R}$, where

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

In general, $f(x) \leq g(x)$ for any $x \in [a, b]$. If f is affine, then

$$f(x) = g(x), \quad \forall x \in [a, b],$$

and a direct calculation shows the equality in this case. Conversely, suppose that f is not affine. Then there exists $\bar{x} \in (a, b)$ with

$$f(\bar{x}) < g(\bar{x}).$$

Take $\alpha > 0$ such that

$$f(\bar{x}) < g(\bar{x}) - \alpha.$$

By using the continuity of the both parts in \bar{x} , there exists $\varepsilon > 0$ such that $(\bar{x} - \varepsilon, \bar{x} + \varepsilon) \subset (a, b)$ and

$$f(x) < g(x) - \alpha, \quad \forall x \in (\bar{x} - \varepsilon, \bar{x} + \varepsilon).$$

Hence

$$\begin{aligned} \int_a^b f(x)dx &= \int_a^{\bar{x}-\varepsilon} f(x)dx + \int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon} f(x)dx + \int_{\bar{x}+\varepsilon}^b f(x)dx \\ &\leq \int_a^{\bar{x}-\varepsilon} g(x)dx + \int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon} (g(x) - \alpha)dx + \int_{\bar{x}+\varepsilon}^b g(x)dx \\ &= \int_a^b g(x)dx - 2\alpha\varepsilon = (b - a)\frac{f(a) + f(b)}{2} - 2\alpha\varepsilon. \end{aligned}$$

Therefore,

$$\int_a^b f(x)dx < (b - a)\frac{f(a) + f(b)}{2},$$

which is a contradiction. □

By using the Jensen inequality (Theorem 2.2.20) for several functions, we can deduce some classical inequalities. Such an example is provided by the convex function $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = e^x$. Take $n \in \mathbb{N}^*$ and $x_1, x_2, \dots, x_n \in \mathbb{R}, \lambda_1, \lambda_2, \dots, \lambda_n > 0$ with $\sum_{k=1}^n \lambda_k = 1$. By applying Theorem 2.2.20, we deduce that

$$e^{\sum_{k=1}^n \lambda_k x_k} \leq \sum_{k=1}^n \lambda_k e^{x_k}, \tag{2.2.6}$$

with equality only in the case $x_1 = x_2 = \dots = x_n$. Take $a_1, a_2, \dots, a_n > 0$ and $x_k = \ln a_k$ for any $k \in \overline{1, n}$. Then, from (2.2.6), we deduce that

$$a_1^{\lambda_1} a_2^{\lambda_2} \dots a_n^{\lambda_n} \leq \sum_{k=1}^n \lambda_k a_k. \tag{2.2.7}$$

Again, equality holds if and only if $a_1 = a_2 = \dots = a_n$. This inequality is sometimes called the general means inequality, because if one takes $\lambda_1 = \lambda_2 = \dots = \lambda_n = n^{-1}$ in (2.2.7), one recovers the well-known inequality between the geometric and the arithmetic means:

$$(a_1 a_2 \dots a_n)^{\frac{1}{n}} \leq \frac{a_1 + a_2 + \dots + a_n}{n}.$$

By the transform $x_k \rightarrow x_k^{-1}$, instead of (2.2.7), one gets

$$a_1^{\lambda_1} a_2^{\lambda_2} \dots a_n^{\lambda_n} \geq \frac{1}{\sum_{k=1}^n \frac{\lambda_k}{a_k}},$$

and from here (again for $\lambda_1 = \lambda_2 = \dots = \lambda_n = n^{-1}$) the inequality between the harmonic and the geometric means:

$$\frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_n}} \leq (a_1 a_2 \dots a_n)^{\frac{1}{n}}.$$

Also from (2.2.7), for $n = 2$, $a_1 = u > 0$, $a_2 = v > 0$, $\lambda_1 = \frac{1}{p}$, $\lambda_2 = \frac{1}{q}$ with $p, q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, one deduces that

$$u^{\frac{1}{p}} v^{\frac{1}{q}} \leq \frac{u}{p} + \frac{v}{q}. \tag{2.2.8}$$

By taking $u = \varepsilon a^p$, $v = \frac{1}{\varepsilon} b^q$ with $a, b, \varepsilon > 0$, one has

$$ab \leq \varepsilon \frac{a^p}{p} + \frac{1}{\varepsilon} \frac{b^q}{q},$$

and for $\varepsilon = 1$,

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}, \quad \forall p, q > 1, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad a, b > 0. \tag{2.2.9}$$

The equality holds if and only if $a^p = b^q$. The relation (2.2.9) is called the Young inequality.

The Hölder inequality can be proved in several ways. We now deduce it as a consequence of (2.2.8). Take $p, q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, $x_1, x_2, \dots, x_n > 0$ and $y_1, y_2, \dots, y_n > 0$. One takes, for a natural k arbitrarily fixed, $u := \frac{x_k^p}{\sum_{k=1}^n x_k^p}$ and $v := \frac{y_k^q}{\sum_{k=1}^n y_k^q}$. From (2.2.8), one gets

$$\frac{x_k}{\left(\sum_{k=1}^n x_k^p\right)^{\frac{1}{p}}} \frac{y_k}{\left(\sum_{k=1}^n y_k^q\right)^{\frac{1}{q}}} \leq \frac{1}{p} \frac{x_k^p}{\sum_{k=1}^n x_k^p} + \frac{1}{q} \frac{y_k^q}{\sum_{k=1}^n y_k^q}.$$

We write this relation for $k \in \overline{1, n}$ and make the sum. We find the Hölder inequality:

$$\sum_{k=1}^n x_k y_k \leq \left(\sum_{k=1}^n x_k^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^n y_k^q\right)^{\frac{1}{q}}.$$

As one may observe, the equality holds if and only if the elements $(x_k^p)_{k \in \overline{1, n}}$ and $(y_k^q)_{k \in \overline{1, n}}$ are proportional.

From here one can deduce the Minkowski inequality. Take $p > 1$ and $x_1, x_2, \dots, x_n > 0$, $y_1, y_2, \dots, y_n > 0$. Then, by applying the previous inequality (and by taking $q = \frac{p}{p-1}$), we obtain

$$\sum_{k=1}^n (x_k + y_k)^p = \sum_{k=1}^n x_k (x_k + y_k)^{p-1} + \sum_{k=1}^n y_k (x_k + y_k)^{p-1}$$

$$\begin{aligned} &\leq \left(\sum_{k=1}^n x_k^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^n (x_k + y_k)^{(p-1)q} \right)^{\frac{1}{q}} + \left(\sum_{k=1}^n y_k^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^n (x_k + y_k)^{(p-1)q} \right)^{\frac{1}{q}} \\ &= \left(\left(\sum_{k=1}^n x_k^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^n y_k^p \right)^{\frac{1}{p}} \right) \left(\sum_{k=1}^n (x_k + y_k)^p \right)^{\frac{1}{q}}. \end{aligned}$$

From this, one gets

$$\left(\sum_{k=1}^n (x_k + y_k)^p \right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^n x_k^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^n y_k^p \right)^{\frac{1}{p}}.$$

Several inequalities can also be obtained based on the following observation, which follows from the Jensen inequality: if f is a convex function defined on $(0, \infty)$, and $x_1, x_2, \dots, x_n > 0, y_1, y_2, \dots, y_n > 0$, then

$$f \left(\frac{\sum_{k=1}^n x_k y_k}{\sum_{k=1}^n x_k} \right) \leq \frac{\sum_{k=1}^n x_k f(y_k)}{\sum_{k=1}^n x_k}.$$

Another example is provided by the inequality

$$(x_1 x_2 \dots x_n)^{\frac{x_1 + x_2 + \dots + x_n}{n}} \leq x_1^{x_1} x_2^{x_2} \dots x_n^{x_n}, \quad \forall x_1, x_2, \dots, x_n > 0,$$

which can be obtained by combining the means inequality and the Jensen inequality for the convex function $f : (0, \infty) \rightarrow \mathbb{R}, f(x) = x \ln x$.

Another result is obtained on the same basis as the previous ones, but is formulated in a slightly different way.

Theorem 2.2.27. Take $p > 1$ and $a, b > 0$. Then

- (i) $\inf_{t>0} \left[\frac{1}{p} t^{\frac{1}{p}-1} a + \left(1 - \frac{1}{p} \right) t^{\frac{1}{p}} b \right] = a^{\frac{1}{p}} b^{1-\frac{1}{p}}$.
- (ii) $\inf_{0<t<1} \left[t^{1-p} a^p + (1-t)^{1-p} b^p \right] = (a+b)^p$.

Proof (i) The function e^x is (strictly) convex on \mathbb{R} . Therefore, for every $t > 0$,

$$\begin{aligned} a^{\frac{1}{p}} b^{1-\frac{1}{p}} &= \left[t^{\frac{1}{p}-1} a \right]^{\frac{1}{p}} \left[t^{\frac{1}{p}} b \right]^{1-\frac{1}{p}} \\ &= e^{\frac{1}{p} \ln \left(t^{\frac{1}{p}-1} a \right) + \left(1 - \frac{1}{p} \right) \ln \left(t^{\frac{1}{p}} b \right)} \\ &\leq \frac{1}{p} e^{\ln \left(t^{\frac{1}{p}-1} a \right)} + \left(1 - \frac{1}{p} \right) e^{\ln \left(t^{\frac{1}{p}} b \right)} \\ &= \frac{1}{p} t^{\frac{1}{p}-1} a + \left(1 - \frac{1}{p} \right) t^{\frac{1}{p}} b. \end{aligned}$$

The equality is obtained when $t = \frac{a}{b}$.

(ii) For $p > 1$, the function $f : (0, \infty) \rightarrow \mathbb{R}$, $f(u) = u^p$ is strictly convex. Consequently,

$$\begin{aligned}(a+b)^p &= \left[t \frac{a}{t} + (1-t) \frac{b}{1-t} \right]^p \\ &\leq t \left(\frac{a}{t} \right)^p + (1-t) \left(\frac{b}{1-t} \right)^p = t^{1-p} a^p + (1-t)^{1-p} b^p\end{aligned}$$

for any $0 < t < 1$. The equality holds for $t = \frac{a}{a+b}$. \square

Besides these well-known inequalities, we present some refinements and generalizations. The next result is an improvement of Young inequality.

Proposition 2.2.28. *Take $p \in (1, 2]$ and $q \in \mathbb{R}$ such that $p^{-1} + q^{-1} = 1$. Then for every $x, y > 0$, one has*

$$\frac{1}{q} \left(x^{\frac{p}{2}} - y^{\frac{q}{2}} \right)^2 \leq \frac{x^p}{p} + \frac{x^q}{q} - xy \leq \frac{1}{p} \left(x^{\frac{p}{2}} - y^{\frac{q}{2}} \right)^2.$$

Proof For $p = 2$, one has equalities. Suppose $p \in (1, 2)$. Then one must have $q > 2$. Let us prove the first inequality, which reduces to

$$\frac{2-p}{p} x^p + \frac{2}{q} x^{\frac{p}{2}} y^{\frac{q}{2}} - xy \geq 0.$$

Consider the function $f : [0, \infty) \rightarrow \mathbb{R}$, $f(y) = \frac{2-p}{p} x^p + \frac{2}{q} x^{\frac{p}{2}} y^{\frac{q}{2}} - xy$. Its only critical point is $y = x^{p-1}$. Since $f''(y) = x^{\frac{p}{2}} y^{\frac{q}{2}-1} \left(\frac{q}{2} - 1 \right) > 0$ for any $y > 0$, one deduces that $y = x^{p-1}$ is a global minimum for f . From $f(x^{p-1}) = 0$, one gets the conclusion. The other inequality can be similarly proved. \square

We present next one of the most important and profound numerical inequalities, due to Hardy.

Theorem 2.2.29 (Hardy Inequality). *Take $p > 1$ and a sequence $(a_n)_{n \in \mathbb{N}^*} \subset [0, \infty)$. Then the next inequality holds:*

$$\sum_{n=1}^{\infty} \left(\frac{a_1 + a_2 + \dots + a_n}{n} \right)^p \leq \left(\frac{p}{p-1} \right)^p \sum_{n=1}^{\infty} a_n^p. \quad (2.2.10)$$

If the series $\sum_{n=1}^{\infty} a_n^p$ is convergent, the equality holds if and only if (a_n) is constantly equal to 0. Moreover, the constant of the right-hand side is sharp (cannot be made smaller).

Proof If all the terms a_n are equal to zero, the conclusion is obvious. Suppose that at least one of these terms is strictly positive. Moreover, suppose that the series $\sum_{n=1}^{\infty} a_n^p$

is convergent, because otherwise the result is again trivial. Denote, for every $n \in \mathbb{N}^*$, the partial sum of the sequence (a_n) by S_n , i.e., $S_n = a_1 + a_2 + \dots + a_n$. Also, denote

$$A_n := \frac{S_n}{n}, \quad \forall n \in \mathbb{N}^*.$$

For completeness, take $S_0 = A_0 = 0$. We will use the following elementary inequality:

$$(n + 1)xy^n \leq x^{n+1} + ny^{n+1}, \quad \forall x, y \geq 0, \quad n \in \mathbb{N}^*, \quad (2.2.11)$$

where equality holds if and only if $x = y$.

Then, for a fixed $n \in \mathbb{N}^*$, one has

$$\begin{aligned} A_n^p - \frac{p}{p-1} A_n^{p-1} a_n &= A_n^p - \frac{p}{p-1} (nA_n - (n-1)A_{n-1}) A_n^{p-1} \\ &= A_n^p \left(1 - \frac{np}{p-1} \right) + \frac{(n-1)p}{p-1} A_n^{p-1} A_{n-1} \\ &\leq A_n^p \left(1 - \frac{np}{p-1} \right) + \frac{(n-1)}{p-1} ((p-1)A_n^p + A_{n-1}^p) \\ &= \frac{1}{p-1} [(n-1)A_{n-1}^p - nA_n^p]. \end{aligned}$$

Therefore,

$$A_n^p - \frac{p}{p-1} A_n^{p-1} a_n \leq \frac{1}{p-1} [(n-1)A_{n-1}^p - nA_n^p], \quad \forall n \in \mathbb{N}^*. \quad (2.2.12)$$

Fix $N \in \mathbb{N}^*$. For $n = 1, 2, \dots, N$, we write the relations (2.2.12) and then compute their sum. We deduce then

$$\sum_{n=1}^N A_n^p - \frac{p}{p-1} \sum_{n=1}^N A_n^{p-1} a_n \leq -\frac{NA_N^p}{p-1} \leq 0,$$

from where

$$\sum_{n=1}^N A_n^p \leq \frac{p}{p-1} \sum_{n=1}^N A_n^{p-1} a_n.$$

For the right-hand side of this relation, we use the Hölder inequality to deduce

$$\sum_{n=1}^N A_n^p \leq \frac{p}{p-1} \left(\sum_{n=1}^N a_n^p \right)^{\frac{1}{p}} \left(\sum_{n=1}^N A_n^p \right)^{\frac{p-1}{p}}.$$

From here we get, after taking the power p ,

$$\sum_{n=1}^N A_n^p \leq \left(\frac{p}{p-1} \right)^p \sum_{n=1}^N a_n^p.$$

By making $N \rightarrow \infty$, one gets (2.2.10).

In order to have equality in the final relation, one must also have equality when inequality (2.2.11) is applied, which reduces to $A_n = A_{n-1}$, for any $n \in \mathbb{N}^*$. This proves that the sequence (a_n) is constantly equal to 0.

The same conclusion follows when one analyzes the case of equality in the Hölder inequality when it is applied at a particular point. Therefore, from this analysis, we must have that the sequences $(a_n^p)_{n \in \mathbb{N}^*}$ and $(A_n^p)_{n \in \mathbb{N}^*}$ are proportional. If a finite number of terms a_n are nonzero, this proportionality cannot hold, hence the inequality is strict. Otherwise, if all the terms a_n are strictly positive, then there must exist a $c \in \mathbb{R}$ such that $a_n = cA_n$ for any $n \in \mathbb{N}^*$. We deduce from this fact that $c = 1$ and, moreover, the sequence (a_n) is constant, but in this case the series $\sum_{n=1}^{\infty} a_n^p$ diverges. Therefore, if the series $\sum_{n=1}^{\infty} a_n^p$ converges, the equality holds if and only if (a_n) is constantly equal to 0.

In order to show that one cannot decrease the constant $\left(\frac{p}{p-1}\right)^p$ from the right-hand side, we use the following argument. Take $N \in \mathbb{N}^*$ and $a_n = n^{-\frac{1}{p}}$, for $1 \leq n \leq N$ and $a_n = 0$ for $n > N$. Then

$$\sum_{n=1}^{\infty} a_n^p = \sum_{n=1}^N \frac{1}{n},$$

and for $n \leq N$,

$$S_n = \sum_{i=1}^n i^{-\frac{1}{p}} > \int_1^n x^{-\frac{1}{p}} dx = \frac{p}{p-1} \left(n^{\frac{p-1}{p}} - 1 \right).$$

Therefore

$$\left(\frac{S_n}{n}\right)^p > \left(\frac{p}{p-1}\right)^p \left(\frac{n^{\frac{p-1}{p}} - 1}{n}\right)^p = \left(\frac{p}{p-1}\right)^p \left(n^{-\frac{1}{p}} - \frac{1}{n}\right)^p.$$

Then there exists $(\varepsilon_n) \subset [0, \infty)$ such that $\varepsilon_n \rightarrow 0$ and

$$\left(\frac{S_n}{n}\right)^p > \left(\frac{p}{p-1}\right)^p \left(\frac{1 - \varepsilon_n}{n}\right)^p,$$

an assertion which follows based on the fact that

$$\begin{aligned} \lim_{n \rightarrow \infty} n \left(n^{-\frac{1}{p}} - \frac{1}{n} \right)^p &= \lim_{n \rightarrow \infty} \frac{\left(\left(\frac{1}{n} \right)^{\frac{1}{p}} - \frac{1}{n} \right)^p}{\frac{1}{n}} = \lim_{x \rightarrow 0^+} \frac{\left(x^{\frac{1}{p}} - x \right)^p}{x} \\ &= \lim_{x \rightarrow 0^+} \frac{\left(x^{\frac{1}{p}} - x \right)^p}{\left(x^{\frac{1}{p}} \right)^p} = \lim_{x \rightarrow 0^+} \left(1 - x^{1-\frac{1}{p}} \right)^p = 1. \end{aligned}$$

Consequently,

$$\sum_{n=1}^{\infty} \left(\frac{S_n}{n}\right)^p > \sum_{n=1}^N \left(\frac{S_n}{n}\right)^p > \left(\frac{p}{p-1}\right)^p \sum_{n=1}^N \left(\frac{1 - \varepsilon_n}{n}\right)^p = \left(\frac{p}{p-1}\right)^p \sum_{n=1}^N (1 - \varepsilon_n) a_n^p.$$

Then there exists $v_N \xrightarrow{N \rightarrow \infty} 0$ such that

$$\sum_{n=1}^{\infty} \left(\frac{S_n}{n} \right)^p > (1 - v_N) \left(\frac{p}{p-1} \right)^p \sum_{n=1}^N a_n^p. \tag{2.2.13}$$

In order to justify this relation, we remark first that it reduces to the fact that there exists $v_N \xrightarrow{N \rightarrow \infty} 0$ such that

$$v_N \sum_{n=1}^N \frac{1}{n} > \sum_{n=1}^N \frac{\varepsilon_n}{n}.$$

On the other hand, this becomes clear by the use of the Stolz-Cesàro criterion (Proposition 1.1.24), one has

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N \frac{\varepsilon_n}{n}}{\sum_{n=1}^N \frac{1}{n}} = \lim_{N \rightarrow \infty} \frac{\frac{\varepsilon_{N+1}}{N+1}}{\frac{1}{N+1}} = 0.$$

Therefore, the relation (2.2.13) is true and, by taking into account the last part of the proof, we deduce that the value of the constant $\left(\frac{p}{p-1} \right)^p$ cannot be smaller. The proof is now complete. \square

We now present the Carleman inequality.

Theorem 2.2.30 (Carleman Inequality). *Let $(a_n)_{n \in \mathbb{N}^*} \subset [0, \infty)$ be a sequence. Then the next inequality holds:*

$$\sum_{n=1}^{\infty} (a_1 a_2 \dots a_n)^{\frac{1}{n}} \leq e \sum_{n=1}^{\infty} a_n.$$

Proof Again, without losing the generality, we may suppose that all the terms are strictly positive. We use the Hardy inequality, where a_n is replaced by $a_n^{\frac{1}{p}}$. Then we can write

$$\sum_{n=1}^{\infty} \left(\frac{a_1^{\frac{1}{p}} + a_2^{\frac{1}{p}} + \dots + a_n^{\frac{1}{p}}}{n} \right)^p \leq \left(\frac{p}{p-1} \right)^p \sum_{n=1}^{\infty} a_n.$$

But

$$\left(\frac{a_1^{\frac{1}{p}} + a_2^{\frac{1}{p}} + \dots + a_n^{\frac{1}{p}}}{n} \right)^p = e^{\frac{\ln(a_1^{\frac{1}{p}} + a_2^{\frac{1}{p}} + \dots + a_n^{\frac{1}{p}}) - \ln n}{\frac{1}{p}}} \xrightarrow{p \rightarrow \infty} (a_1 a_2 \dots a_n)^{\frac{1}{n}}.$$

For passing to the limit, we have utilized the fact that

$$\lim_{x \rightarrow 0} \frac{\ln(a_1^x + a_2^x + \dots + a_n^x) - \ln n}{x} = \frac{\ln(a_1 a_2 \dots a_n)}{n},$$

which easily follows based on L'Hôpital Rule. Since, on the other hand, $\left(\frac{p}{p-1} \right)^p \xrightarrow{p \rightarrow \infty} e$ (increasingly), we get the desired inequality. \square

We end this section by presenting the Kantorovici inequality.

Theorem 2.2.31 (Kantorovici Inequality). *Let A be a symmetric and positive definite square matrix of dimension p . Then for every $x \in \mathbb{R}^p$, one has*

$$\|x\|^4 \leq \langle (Ax^t)^t, x \rangle \cdot \langle (A^{-1}x^t)^t, x \rangle \leq \frac{1}{4} \left(\sqrt{\frac{\lambda_1}{\lambda_p}} + \sqrt{\frac{\lambda_p}{\lambda_1}} \right)^2 \|x\|^4,$$

where λ_1 and λ_p are respectively the greatest and the smaller eigenvalue of A .

Proof It is sufficient to prove the inequality for any x of unit norm. As we observed in the first chapter, the eigenvalues of A are strictly positive reals and, without loss of generality, one may decreasingly arrange them: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$. We denote the diagonal matrix having the eigenvalues on its main diagonal (in the mentioned order) by $D := \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$. We also know that there exists an orthogonal matrix B such that $A = B^tDB$. Then $A^{-1} = (B^tDB)^{-1} = B^tD^{-1}B$, and $D^{-1} = \text{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_p^{-1})$. Hence

$$\langle (Ax^t)^t, x \rangle = \langle (B^tDBx^t)^t, x \rangle = \langle (DBx^t)^t, (Bx^t)^t \rangle$$

and

$$\langle (A^{-1}x^t)^t, x \rangle = \langle (B^tD^{-1}Bx^t)^t, x \rangle = \langle (D^{-1}Bx^t)^t, (Bx^t)^t \rangle.$$

On the other hand, the mapping $x \mapsto (Bx^t)^t$ is a bijection from the unit sphere of \mathbb{R}^p into itself, hence in order to get the desired conclusion it is sufficient to prove that for every $u \in \mathbb{R}^p$ with $\|u\| = 1$, one has

$$1 \leq \langle (Du^t)^t, u \rangle \cdot \langle (D^{-1}u^t)^t, u \rangle \leq \frac{1}{4} \left(\sqrt{\frac{\lambda_1}{\lambda_p}} + \sqrt{\frac{\lambda_p}{\lambda_1}} \right)^2.$$

If $\lambda_1 = \lambda_p$, then one has the equality. Suppose that $\lambda_p < \lambda_1$. One has

$$\langle (Du^t)^t, u \rangle = \sum_{i=1}^p u_i^2 \lambda_i; \quad \langle (D^{-1}u^t)^t, u \rangle = \sum_{i=1}^p u_i^2 \frac{1}{\lambda_i}.$$

Then the first inequality becomes

$$1 \leq \left(\sum_{i=1}^p u_i^2 \frac{1}{\lambda_i} \right) \left(\sum_{i=1}^p u_i^2 \lambda_i \right),$$

i.e.,

$$\frac{1}{\left(\sum_{i=1}^p u_i^2 \lambda_i \right)} \leq \sum_{i=1}^p u_i^2 \frac{1}{\lambda_i}.$$

Since $\sum_{i=1}^p u_i^2 = 1$, the previous inequality follows from Jensen inequality applied to the convex function $(0, \infty) \ni x \mapsto \frac{1}{x}$.

For the second inequality, observe that for any $i \in \overline{1, p}$,

$$\frac{1}{\lambda_i} \leq \frac{1}{\lambda_1} + \frac{1}{\lambda_p} - \frac{\lambda_i}{\lambda_1 \lambda_p},$$

because $\lambda_p \leq \lambda_i \leq \lambda_1$. Hence

$$\sum_{i=1}^p u_i^2 \frac{1}{\lambda_i} \leq \frac{1}{\lambda_1} + \frac{1}{\lambda_p} - \frac{\sum_{i=1}^p u_i^2 \lambda_i}{\lambda_1 \lambda_p}.$$

One gets

$$\begin{aligned} \left(\sum_{i=1}^p u_i^2 \frac{1}{\lambda_i} \right) \left(\sum_{i=1}^p u_i^2 \lambda_i \right) &\leq \left(\sum_{i=1}^p u_i^2 \lambda_i \right) \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_p} - \frac{\sum_{i=1}^p u_i^2 \lambda_i}{\lambda_1 \lambda_p} \right) \\ &= \frac{\left(\sum_{i=1}^p u_i^2 \lambda_i \right) (\lambda_1 + \lambda_p - \sum_{i=1}^p u_i^2 \lambda_i)}{\lambda_1 \lambda_p}. \end{aligned}$$

The second-degree polynomial

$$\lambda \mapsto \frac{\lambda(\lambda_1 + \lambda_p - \lambda)}{\lambda_1 \lambda_p}$$

attains its maximum for $\lambda = \frac{\lambda_1 + \lambda_p}{2}$, hence

$$\left(\sum_{i=1}^p u_i^2 \frac{1}{\lambda_i} \right) \left(\sum_{i=1}^p u_i^2 \lambda_i \right) \leq \left(\frac{\lambda_1 + \lambda_p}{2} \right)^2 \frac{1}{\lambda_1 \lambda_p} = \frac{1}{4} \left(\sqrt{\frac{\lambda_1}{\lambda_p}} + \sqrt{\frac{\lambda_p}{\lambda_1}} \right)^2.$$

The proof is now complete. □

2.3 Banach Fixed Point Principle

This section is dedicated to the Banach fixed point theorem (also known as Contraction Principle, or Banach Principle), which is one of the fundamental results in nonlinear analysis.

Take $f : \mathbb{R}^p \rightarrow \mathbb{R}^p$. A point $x \in \mathbb{R}^p$ for which $f(x) = x$ is called a fixed point of f . By a fixed point result we understand a result concerning the existence of the fixed points for a given function.

Banach fixed point theorem is a result which provides, under certain conditions, both the existence and the uniqueness of the fixed point. This theorem also creates the basis for getting several other remarkable mathematical results, such as, to name a few, the implicit function theorem, or theorems about the existence and the uniqueness of solution for differential equations or systems of differential equations.

2.3.1 Contractions and Fixed Points

We begin with a definition:

Definition 2.3.1. Let $A \subset \mathbb{R}^p$. A mapping $f : A \rightarrow \mathbb{R}^q$ is a contraction on A if there exists a real constant $\lambda \in (0, 1)$ such that $\|f(x) - f(y)\| \leq \lambda \|x - y\|$ for any $x, y \in A$.

Remark that λ does not depend on x and y , and by applying the function f to a pair of points from A , the distance between them shrinks (contracts). The contraction notion is a particular case of the concept of Lipschitz function (Definition 1.2.25), hence, in particular, every contraction is a (uniformly) continuous function (Proposition 1.2.26).

The next assertion, often used in order to prove that a function is Lipschitz, holds: if $A \subset \mathbb{R}^p$ is an open convex set, $f : A \rightarrow \mathbb{R}^q$ is of class C^1 on A , and there exists $M > 0$ such that $\|\nabla f(x)\| \leq M$ for any $x \in A$, then f is Lipschitz on A . The proof of this fact relies on Taylor formula: for every two points $x, y \in A$ we can apply Theorem 1.3.4 to the function f on $[x, y]$, hence we deduce the existence of a point $c_{x,y} \in (x, y) \subset A$ such that

$$\|f(x) - f(y)\| = \|\nabla f(c_{x,y})(x - y)\| \leq M \|x - y\|.$$

Accordingly, f is Lipschitz on A . We now formulate a particular case of this observation.

Proposition 2.3.2. Let $A \subset \mathbb{R}^p$ be an open convex set and $f : A \rightarrow \mathbb{R}^q$ of class C^1 on A , satisfying the condition that there exists $M \in (0, 1)$ such that $\|\nabla f(x)\| \leq M$ for any $x \in A$. Then f is a contraction on A .

Obviously, for $p = 1$ we may consider in the previous proposition also closed, or half-open intervals.

Example 2.3.3. The function $f : \mathbb{R}^p \rightarrow \mathbb{R}^p$ defined by $f(x) = \frac{1}{2}x$ is a contraction of \mathbb{R}^p into itself, because:

$$\|f(x) - f(y)\| = \frac{1}{2} \|x - y\|.$$

Example 2.3.4. Take $A = [0, \infty)$ and the function $f : A \rightarrow A$, given by $f(x) = \frac{1}{1+x^2}$. The function f is a contraction on A , and this fact can be proven using Proposition 2.3.2. The derivative of the function f is

$$f'(x) = \left(\frac{1}{1+x^2} \right)' = -\frac{2x}{(1+x^2)^2},$$

hence

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

Define the auxiliary function $g : [0, \infty) \rightarrow [0, \infty)$ given by $g(x) = \frac{2x}{(1+x^2)^2}$. Its derivative is

$$\begin{aligned} g'(x) &= \frac{2(1+x^2)^2 - 2x \cdot 2(1+x^2) \cdot 2x}{(1+x^2)^4} \\ &= \frac{2(1+x^2)(1+x^2-4x^2)}{(1+x^2)^4} = 2 \frac{1-3x^2}{(1+x^2)^3}. \end{aligned}$$

Remark that $x = \frac{1}{\sqrt{3}}$ is the positive maximum for the function g and $g\left(\frac{1}{\sqrt{3}}\right) = \frac{9}{8\sqrt{3}} < 1$. Hence $|f'(x)| \leq \frac{9}{8\sqrt{3}} < 1$ for every $x \in [0, \infty)$. It follows from Proposition 2.3.2 that f is a contraction.

We now formulate the announced principle. We mention here that the Banach fixed point theorem was proved by the Polish mathematician Stefan Banach in 1922 in the framework of complete normed vector spaces.

Theorem 2.3.5 (Banach Principle). *Let $A \subset \mathbb{R}^p$ be a closed nonempty set and $f : A \rightarrow A$ be a contraction. Then f admits a unique fixed point.*

Proof As usual in the case of existence and uniqueness results, we divide the proof into two steps. We prove first the existence, and then the uniqueness of the fixed point. According to the definition, there exists a real number $\lambda \in (0, 1)$ such that

$$\|f(x) - f(y)\| \leq \lambda \|x - y\|$$

for any $x, y \in A$. Consider $x_0 \in A$ as an arbitrary point and denote $x_1 = f(x_0)$, $x_2 = f(x_1)$, ..., $x_n = f(x_{n-1})$, an operation which can be made for every natural n . Observe that $x_2 = f(f(x_0)) = f^2(x_0)$ and, in general, $x_n = f^n(x_0)$ (we have denoted f^2 instead of $f \circ f$, and, in general, f^n instead of n times $f \circ f \circ \dots \circ f$). We prove that $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence. The next relations hold

$$\begin{aligned} \|x_2 - x_1\| &= \|f(x_1) - f(x_0)\| \leq \lambda \|x_1 - x_0\|, \\ \|x_3 - x_2\| &= \|f(x_2) - f(x_1)\| \leq \lambda \|x_2 - x_1\| \leq \lambda^2 \|x_1 - x_0\|. \end{aligned}$$

Using an inductive procedure, one gets the inequality

$$\|x_{n+1} - x_n\| \leq \lambda^n \|x_1 - x_0\|$$

for any $n \in \mathbb{N}^*$. For $m, n \in \mathbb{N}^*$ arbitrarily taken, one can successively write:

$$\begin{aligned} \|x_{m+n} - x_n\| &\leq \|x_{n+1} - x_n\| + \|x_{n+2} - x_{n+1}\| + \dots + \|x_{n+m} - x_{n+m-1}\| \\ &\leq \lambda^n \|x_1 - x_0\| + \lambda^{n+1} \|x_1 - x_0\| + \dots + \lambda^{n+m-1} \|x_1 - x_0\| \\ &= \|x_1 - x_0\| (\lambda^n + \lambda^{n+1} + \dots + \lambda^{n+m-1}) = \|x_1 - x_0\| \lambda^n \frac{1 - \lambda^m}{1 - \lambda} \\ &\leq \|x_1 - x_0\| \frac{\lambda^n}{1 - \lambda}. \end{aligned}$$

Hence,

$$\|x_{n+m} - x_n\| \leq \frac{\lambda^n}{1 - \lambda} \|x_1 - x_0\| \quad (2.3.1)$$

for any $n, m \in \mathbb{N}^*$. If $\|x_1 - x_0\| = 0$, it follows that $f(x_0) = x_0$, i.e., x_0 is a fixed point and then the existence of the fixed point is assured. If $\|x_1 - x_0\| \neq 0$, then, using the fact that $\lambda \in (0, 1)$, one deduces that

$$\lim_{n \rightarrow \infty} \frac{\lambda^n}{1 - \lambda} = 0,$$

hence

$$\lim_{n \rightarrow \infty} \frac{\lambda^n}{1 - \lambda} \|x_1 - x_0\| = 0.$$

By using the ε characterization of this convergence, it follows that for any $\varepsilon > 0$, there exists $n_\varepsilon \in \mathbb{N}^*$, such that for any $n \geq n_\varepsilon$, one has

$$\frac{\lambda^n}{1 - \lambda} \|x_1 - x_0\| < \varepsilon.$$

By combining this relation with (2.3.1), it follows that for any $\varepsilon > 0$, there exists $n_\varepsilon \in \mathbb{N}^*$, such that for every $n \geq n_\varepsilon$, and every $m \in \mathbb{N}$, one has

$$\|x_{n+m} - x_n\| < \varepsilon.$$

This proves that (x_n) is a Cauchy sequence, and because \mathbb{R}^p is a complete space, the sequence $(x_n)_{n \in \mathbb{N}}$ is convergent, so there exists $\bar{x} \in \mathbb{R}^p$ with $\lim_{n \rightarrow \infty} x_n = \bar{x}$. Since $(x_n) \subset A$ and A is closed, we deduce that $\bar{x} \in A$. Recall that the sequence is given by the relation

$$x_0 \in A, f(x_n) = x_{n+1}, \forall n \in \mathbb{N}. \quad (2.3.2)$$

The function f is continuous (since every contraction has this property, see Proposition 1.2.26). Hence, from the properties of the continuous functions, the limit of the sequence $(f(x_n))_n$ exists and equals $f(\bar{x})$. In the relation (2.3.2), we pass to the limit for $n \rightarrow \infty$ and we get

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1},$$

i.e.,

$$f(\bar{x}) = \bar{x},$$

hence \bar{x} is a fixed point. The existence is proved.

In order to prove the uniqueness, suppose there are two different fixed points x and y . Then

$$\|x - y\| = \|f(x) - f(y)\| \leq \lambda \|x - y\|.$$

Since $\|x - y\| > 0$, it follows that $1 \leq \lambda$, which is absurd. Therefore, there exists a unique fixed point for f . \square

The above result is very important and deserves an extended comment. The analysis of the statement, and also of the proof, provides us with some useful conclusions.

Firstly, observe (in the proof of the existence) how the fixed point was obtained: for every initial point $x_0 \in A$, the sequence given by the relation (2.3.2) converges to the unique fixed point \bar{x} of the mapping f . Coming back to the above inequalities, the next relations hold:

$$\begin{aligned} \|x_n - x_0\| &\leq \|x_1 - x_0\| + \|x_2 - x_1\| + \cdots + \|x_n - x_{n-1}\| \\ &\leq \|x_1 - x_0\| + \lambda \|x_1 - x_0\| + \cdots + \lambda^{n-1} \|x_1 - x_0\| \\ &= \|x_1 - x_0\| (1 + \lambda + \lambda^2 + \cdots + \lambda^{n-1}) = \|x_1 - x_0\| \frac{1 - \lambda^n}{1 - \lambda}, \end{aligned}$$

and, by passing to the limit for $n \rightarrow \infty$, we deduce that

$$\|x_0 - \bar{x}\| \leq \frac{1}{1 - \lambda} \|x_0 - f(x_0)\|.$$

In this way, we see that for every $x \in A$, one has

$$\|x - \bar{x}\| \leq \frac{1}{1 - \lambda} \|x - f(x)\|. \quad (2.3.3)$$

Moreover, one gets the following estimations:

$$\|x_n - \bar{x}\| \leq \|x_1 - x_0\| \frac{\lambda^n}{1 - \lambda} \quad (2.3.4)$$

for any $n \in \mathbb{N}^*$, which follows from (2.3.1) passing to the limit for $m \rightarrow \infty$.

The “a priori” estimation we get in this way is useful in determining the maximum number of steps of the iteration (2.3.2) which one needs, in order to obtain the desired precision in the estimation of the fixed point, by knowing the initial value x_0 and the value $x_1 = f(x_0)$, a fact which we will use in the study of some algorithms. More exactly, for getting an error smaller than $\varepsilon > 0$, one needs that

$$\|x_1 - x_0\| \frac{\lambda^n}{1 - \lambda} < \varepsilon,$$

which drives us to the conclusion that we need a number n of iterations greater than the value

$$\left\lceil \frac{\ln \varepsilon + \ln(1 - \lambda) - \ln(\|x_1 - x_0\|)}{\ln \lambda} \right\rceil. \quad (2.3.5)$$

For example, if we need ε to be of the form 10^{-m} , the size order of n is of the type

$$\frac{m}{|\ln \lambda|} + \text{constant},$$

and for additional decimal places of accuracy, one needs to supplement the number of iterations by $|\ln \lambda|^{-1}$. Therefore, the closer λ approaches to 0, this value is smaller, and we will need to iterate less in order to obtain the desired precision. Also, one observes from the relation (2.3.5) that a smaller value of $\|x_1 - x_0\|$ implies that the decrease of the number of iterations which are necessary for a prescribed precision. Therefore, in

case of some algorithms, it is preferable to start from points as close as possible to the initial point. We will later discuss examples on the computational effects described here in the Chapter 6.

One can also get the following estimate:

$$\|x_n - \bar{x}\| \leq \frac{\lambda}{1 - \lambda} \|x_n - x_{n-1}\|, \quad (2.3.6)$$

which is obtained by passing to the limit for $m \rightarrow \infty$ in the relation

$$\begin{aligned} \|x_{n+m} - x_n\| &\leq \|x_{n+1} - x_n\| + \|x_{n+2} - x_{n+1}\| + \cdots + \|x_{n+m} - x_{n+m-1}\| \\ &\leq \lambda \|x_n - x_{n-1}\| + \lambda^2 \|x_n - x_{n-1}\| + \cdots + \lambda^m \|x_n - x_{n-1}\| \\ &= (\lambda + \lambda^2 + \dots + \lambda^m) \|x_n - x_{n-1}\|. \end{aligned}$$

The speed of convergence of the sequence (x_n) is given by the approximation

$$\|x_{n+1} - \bar{x}\| \leq \lambda \|x_n - \bar{x}\|,$$

which is deduced from the relations

$$\|x_{n+1} - \bar{x}\| = \|f(x_n) - f(\bar{x})\| \leq \lambda \|x_n - \bar{x}\|.$$

More details regarding the speed of convergence will be given in Chapter 6.

Another inequality which follows immediately from the definition of the contraction is

$$\|x_n - \bar{x}\| \leq \lambda^n \|x_0 - \bar{x}\|, \quad \forall n \in \mathbb{N}.$$

Besides these observations regarding the convergence type of the iterations towards the fixed point, we make some other remarks concerning the assumptions of the Banach Principle.

Remark 2.3.6. *The Banach fixed point theorem shows not only the existence, but the uniqueness of the fixed point and, at the same time, shows us a method to approximate the fixed point \bar{x} , and allows us to emphasize an estimation of the error produced by considering this approximation. This approximation method of the solution by the terms of the sequence $x_n = f^n(x_0)$ is called the successive approximations method, or the Picard method, after the name of the French mathematician Charles Émile Picard, which initiated it in 1890.*

Remark 2.3.7. *The assumption $\lambda < 1$ is essential both for the existence, as for the uniqueness of the fixed point. One can see, for instance, that for the identity map $f(x) = x$, for any $x \in \mathbb{R}$, every point of \mathbb{R} is a fixed point, while the map $f(x) = x + 1$, for any $x \in \mathbb{R}$, does not have any fixed point. In both cases, $\lambda = 1$.*

Remark 2.3.8. *If A is not closed, one loses the completeness argument and the conclusion of Banach Principle does not hold. For example, the mapping $f : (0, 1] \rightarrow (0, 1]$ given by $f(x) = \frac{x}{2}$ does not have any fixed point, although is a contraction.*

Aiming to further illustrate the assumptions of the Banach Principle, we introduce a weaker notion than the one of contraction.

Definition 2.3.9. A function $f : \mathbb{R}^p \rightarrow \mathbb{R}^p$ is called a weak contraction if

$$\forall x, y \in \mathbb{R}^p, x \neq y, \|f(x) - f(y)\| < \|x - y\|.$$

Example 2.3.10. There exist weak contractions, defined on closed sets, without fixed points. We consider the next example. Take $f : \mathbb{R} \rightarrow \mathbb{R}$, defined by

$$f(x) = 1 + x - \frac{x}{1 + |x|}.$$

Obviously, $f(x) > x$ for any $x \in \mathbb{R}$, hence f does not have any fixed point. On the other hand, f is a weak contraction, which one can show by considering the relation

$$|f(x) - f(y)| = \left| (x - y) - \left(\frac{x}{1 + |x|} - \frac{y}{1 + |y|} \right) \right|$$

in several situations.

If $x, y \geq 0$, then:

$$\begin{aligned} |f(x) - f(y)| &= \left| x - y - \left(\frac{x}{1 + x} - \frac{y}{1 + y} \right) \right| = \left| x - y - \frac{x - y}{(1 + x)(1 + y)} \right| \\ &= \left| (x - y) \left(1 - \frac{1}{(1 + x)(1 + y)} \right) \right| < |x - y| \end{aligned}$$

because $x \neq y$ and

$$1 - \frac{1}{(1 + x)(1 + y)} \in (0, 1).$$

If $x, y < 0$, one has:

$$\begin{aligned} |f(x) - f(y)| &= \left| x - y - \left(\frac{x}{1 - x} - \frac{y}{1 - y} \right) \right| = \left| (x - y) - \frac{x - y}{(1 - x)(1 - y)} \right| \\ &= \left| (x - y) \left(1 - \frac{1}{(1 - x)(1 - y)} \right) \right| < |x - y| \end{aligned}$$

because $x \neq y$ and

$$1 - \frac{1}{(1 - x)(1 - y)} \in (0, 1).$$

If $x > 0$ and $y < 0$, one has:

$$\begin{aligned} |f(x) - f(y)| &< \left| x - y - \left(\frac{x}{1 + x} - \frac{y}{1 - y} \right) \right| = \left| (x - y) - \frac{(x - y) - 2xy}{(1 + x)(1 - y)} \right| \\ &= x - y - \frac{(x - y) - 2xy}{(1 + x)(1 - y)} < x - y = |x - y|, \end{aligned}$$

where the last inequality can be obtained by direct calculation.

Observe that the Picard iterations associated to f diverge to $+\infty$ for any initial data x_0 : if $x_0 \geq 0$, then (x_n) is strictly increasing, has positive values and cannot have a finite limit, while if $x_0 < 0$, then the terms of the sequence become positive from a certain rank, and we are again in the previous framework.

The weak contractions may yet have a fixed point, if at least a sequence of Picard iterations has a convergent subsequence. More precisely, the next assertion holds.

Proposition 2.3.11. *Let $A \subset \mathbb{R}^p$ be a closed set and $f : A \rightarrow A$ be a function such that*

$$\|f(x) - f(y)\| < \|x - y\|, \quad \forall x, y \in A, \quad x \neq y,$$

i.e., f is a weak contraction on A . If there exists $a \in A$ such that the sequence of Picard iterations, having the initial data a , and given by $x_1 = f(a)$, $x_{n+1} = f(x_n)$, $n \geq 1$, has a subsequence which is convergent to a point $\bar{x} \in A$, then \bar{x} is the unique fixed point of f on A .

Proof Suppose, by contradiction, that $f(\bar{x}) \neq \bar{x}$. Consider

$$D := \{(x, y) \in A \times A \mid x \neq y\}$$

and the mapping $g : D \rightarrow \mathbb{R}$ given by

$$g(x, y) := \frac{\|f(x) - f(y)\|}{\|x - y\|}.$$

From the assumptions made, we have that $g(x, y) < 1$. But $(\bar{x}, f(\bar{x})) \in D$, hence $g(\bar{x}, f(\bar{x})) < 1$. Let $(x_{n_k})_{k \in \mathbb{N}}$ be a subsequence of $(x_n)_{n \in \mathbb{N}}$ which converges to \bar{x} . Then

$$\|f(x_{n_k}) - f(\bar{x})\| < \|x_{n_k} - \bar{x}\|, \quad \forall k \in \mathbb{N},$$

from which one deduces that $f(x_{n_k}) \rightarrow f(\bar{x})$. It follows that

$$\|x_{n_k} - f(x_{n_k})\| \rightarrow \|\bar{x} - f(\bar{x})\| \quad \text{and} \quad g(x_{n_k}, f(x_{n_k})) \rightarrow g(\bar{x}, f(\bar{x})) < 1.$$

Take r such that $g(\bar{x}, f(\bar{x})) < r < 1$. Taking into account the continuity of f (in particular, f is Lipschitz), there exists a $k_0 \in \mathbb{N}$ such that for every $k \geq k_0$,

$$\frac{1}{3} \|\bar{x} - f(\bar{x})\| < \|x_{n_k} - f(x_{n_k})\|$$

and $g(x_{n_k}, f(x_{n_k})) < r$ or, equivalently,

$$\|f(x_{n_k}) - f(f(x_{n_k}))\| < r \|x_{n_k} - f(x_{n_k})\|.$$

Then for every $i > k \geq k_0$, one has:

$$\frac{1}{3} \|\bar{x} - f(\bar{x})\| < \|x_{n_i} - f(x_{n_i})\| = \|f(x_{n_{i-1}}) - f(x_{n_i})\|$$

$$\begin{aligned} &< \|x_{n_{i-1}} - x_{n_i}\| < \dots < \|f(x_{n_{i-1}}) - f(f(x_{n_{i-1}}))\| \\ &< r \|x_{n_{i-1}} - f(x_{n_{i-1}})\| < \dots < r^{i-k} \|x_{n_k} - f(x_{n_k})\|. \end{aligned}$$

We make $i \rightarrow \infty$ and we get $\|\bar{x} - f(\bar{x})\| = 0$, which contradicts $(\bar{x}, f(\bar{x})) \in D$. The uniqueness of the fixed point is straightforward. \square

We present next some generalizations and consequences of the Banach Principle.

A first generalization of the Banach Principle mainly says that it is sufficient for one of the iterations of the function to be a contraction to obtain the conclusions of the Banach Principle. We will see that such an assumption is weaker than the condition that f is a contraction. Remark first that if f is a contraction with constant $\lambda < 1$, then f^n is a contraction with constant $\lambda^n < 1$, hence, in the assumptions of Banach Principle, has a unique fixed point. As every fixed point of f is a fixed point for every iteration, one deduces that f and all its iterations have the same fixed point. The role of the next theorem is also to formulate a partial converse of this observation.

Theorem 2.3.12. *Take $f : \mathbb{R}^p \rightarrow \mathbb{R}^p$. If there exists $q \in \mathbb{N}^*$ such that f^q is a contraction, then f has a unique fixed point. Moreover, for every initial data $x_0 \in \mathbb{R}^p$, the sequence of the Picard iterations converges to the fixed point of f .*

Proof Let $q \in \mathbb{N}^*$ be such that f^q is a contraction. From the Banach Principle, f^q has a unique fixed point, which we denote by \bar{x} . Observe, based on the previous comment, that \bar{x} is the sole candidate to be a fixed point for f (any fixed point for f is a fixed point for any iteration). The next relations hold:

$$f(\bar{x}) = f(f^q(\bar{x})) = f^q(f(\bar{x})),$$

i.e., $f(\bar{x})$ is a fixed point for f^q . As f^q has a unique fixed point, one deduces that $f(\bar{x}) = \bar{x}$. Let us now prove the assertion concerning the convergence of the Picard iterations. We start from a fixed element $x_0 \in \mathbb{R}^p$ and we construct the associated Picard sequence $(f^n(x_0))_n$. We must prove that this sequence converges to \bar{x} . Take $r \in \overline{0, q-1}$. Then, the set of the terms of sequence $(f^n(x_0))_n$ is the union of the sets of terms of subsequences of the type $(f^{qk+r}(x_0))_k$. On the other hand, $(f^{qk+r}(x_0))_k$ can be seen as the sequence of Picard iterations associated to f^q , with the initial point $f^r(x_0)$, because:

$$f^{qk+r}(x_0) = (f^q)^k(f^r(x_0)).$$

Since f^q is a contraction, we know from the Banach Principle that all the Picard iterations of f^q converge to the fixed point \bar{x} . Hence all the q subsequences which partition the initial sequence have the same limit (i.e., \bar{x}), which shows that $\lim_{n \rightarrow \infty} f^n(x_0) = \bar{x}$ and the theorem is completely proved. \square

As in the case of Banach Principle, the result holds if $f : A \rightarrow A$, where A is a closed (hence, complete) subset of \mathbb{R}^p . Observe that in this theorem is not necessary

for f to have any special property (even continuity of f is not assumed). For instance, the function $f : [0, 1] \rightarrow [0, 1]$ given by

$$f(x) = \begin{cases} 0 & \text{if } x \in [0, \frac{1}{2}] \\ \frac{1}{2} & \text{if } x \in (\frac{1}{2}, 1] \end{cases}$$

is discontinuous, but $f^2(x) = 0$ for any $x \in [0, 1]$, hence is a contraction. Of course, the unique fixed point of f is $\bar{x} = 0$.

Let us give now another example which proves that the assumption that an iterate of f is a contraction is weaker than the condition that f is a contraction. Consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = e^{-x}$. This function is not a contraction on \mathbb{R} : for instance

$$|f(-2) - f(0)| = e^2 - 1 > |-2 - 0| = 2.$$

But $f^2(x) = e^{-e^{-x}}$ is a contraction. To prove this, we evaluate the absolute value of the derivative (in order to apply Proposition 2.3.2):

$$\left| (f^2)'(x) \right| = \left| e^{-e^{-x}} e^{-x} \right| = e^{-x-e^{-x}} \leq e^{-1} < 1, \quad \forall x \in \mathbb{R},$$

based on the observation that

$$1 \leq x + e^{-x}, \quad \forall x \in \mathbb{R}.$$

Therefore, according to Theorem 2.3.12, there exists a unique fixed point \bar{x} of the function f (hence $\bar{x} = e^{-\bar{x}}$) which can be approximated using the Picard iterations. A numerical calculation (made on a computer, see Chapter 6) suggests the approximate value $\bar{x} \simeq 0.567$. We will return in Chapter 6 to the possibility of approximating the fixed points in some concrete situations.

At the end of this section, we consider another two interesting consequences of the Banach fixed point principle. The first one refers at the continuous dependence of the fixed point with respect to a parameter.

Theorem 2.3.13. *Let $g : \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}^p$ be a continuous function. Suppose that there exists $\alpha \in (0, 1)$ such that*

$$\|g(x, t) - g(y, t)\| \leq \alpha \|x - y\|$$

for every $t \in \mathbb{R}^q$ and $x, y \in \mathbb{R}^p$. For fixed t in \mathbb{R}^q , denote by $\mu(t)$ the unique fixed point of the contraction $g(\cdot, t)$. Then the mapping $\mu : \mathbb{R}^q \rightarrow \mathbb{R}^p$ is continuous.

Proof Take $t_0 \in \mathbb{R}^q$ and $\varepsilon > 0$. The continuity of g in $(\mu(t_0), t_0)$ implies the existence of a number $\delta > 0$ such that for any fixed t , with $\|t - t_0\| < \delta$, one has

$$\|g(\mu(t_0), t) - g(\mu(t_0), t_0)\| < \varepsilon(1 - \alpha),$$

which is equivalent to

$$\|g(\mu(t_0), t) - \mu(t_0)\| < \varepsilon(1 - \alpha).$$

From the relation (2.3.3) and the above relations, we deduce that for any t with $\|t - t_0\| < \delta$, one has

$$\|\mu(t) - \mu(t_0)\| \leq \frac{\|\mu(t_0) - g(\mu(t_0), t)\|}{1 - \alpha} < \varepsilon.$$

This proves the theorem. \square

Theorem 2.3.14. *Let $f : \mathbb{R}^p \rightarrow \mathbb{R}^p$ be a contraction. Then the mapping $v : \mathbb{R}^p \rightarrow \mathbb{R}^p$ given by $v(x) = x + f(x)$ is a bicontinuous bijection (i.e., is a homeomorphism of \mathbb{R}^p).*

Proof It is clear that v is continuous. Moreover, f is injective because the relation $v(x) = v(y)$ and the contraction property of f imply $x = y$. Consider $g : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ given by

$$g(x, y) = y - f(x).$$

It is clear that g satisfies the property from the assumption of the previous theorem, hence the mapping $x \mapsto g(x, y)$ has a unique fixed point $\mu(y)$ for any $y \in \mathbb{R}^p$. Therefore,

$$\mu(y) = y - f(\mu(y)),$$

i.e.,

$$y = v(\mu(y)), \tag{2.3.7}$$

hence v is surjective. It remains to prove that v^{-1} is continuous. Relation (2.3.7) shows that the functions v and μ are inverse one to each other, and the continuity of μ (hence of v^{-1}) is assured by the previous theorem. Hence, v is a homeomorphism. \square

2.3.2 The Case of One Variable Functions

We now present some fixed point results for real functions of one variable. Geometrically, the fixed points of a function f in this context are those points x for which $(x, f(x))$ lays on the first bisector, i.e., the abscissae of points where the graph of f intersects the first bisector.

Obviously, the first result which we must mention is the particular case of the Banach Principle, which works for contractions which map a closed subset of \mathbb{R} into itself. The first result which is different from the Banach Principle is the Knaster fixed point theorem.

Theorem 2.3.15 (Knaster). *Let $a, b \in \mathbb{R}$, $a < b$ and $f : [a, b] \rightarrow [a, b]$ be an increasing function. Then f has at least one fixed point.*

Proof Define the set

$$A := \{x \in [a, b] \mid f(x) \geq x\}.$$

It is clear, on one hand, that A is nonempty (because $a \in A$) and, on the other hand, that A is bounded (being a subset of $[a, b]$). Hence, according to the completeness axiom, A admits a supremum in \mathbb{R} . Denote this number by \bar{x} . Therefore, $\bar{x} = \sup A$ and it is clear that $\bar{x} \in [a, b]$. As $\bar{x} \geq x$ for any $x \in A$, the monotony of f allows us to write the inequality $f(\bar{x}) \geq f(x) \geq x$ for any $x \in A$. Hence, $f(\bar{x})$ is a majorant for A , so $f(\bar{x}) \geq \bar{x}$. It also follows from monotony that $f(f(\bar{x})) \geq f(\bar{x})$ and hence $f(\bar{x}) \in A$, i.e., $f(\bar{x}) \leq \bar{x}$. Consequently, one has the equality $f(\bar{x}) = \bar{x}$ and \bar{x} is a fixed point. \square

Remark 2.3.16. *If in the previous result one takes f to be decreasing, then the conclusion does not hold. One may consider the following counterexample: $f : [0, 1] \rightarrow [0, 1]$ given by*

$$f(x) = \begin{cases} 1 - x & \text{if } x \in [0, \frac{1}{2}) \\ \frac{1}{2} - \frac{x}{2} & \text{if } x \in [\frac{1}{2}, 1]. \end{cases}$$

It is clear that f is decreasing on $[0, 1]$, but still does not admit any fixed point.

The next result is simple, but will be useful in many situations.

Theorem 2.3.17. *Let $a, b \in \mathbb{R}$, $a < b$ and $f : [a, b] \rightarrow [a, b]$ be a continuous function. Then there exists $\bar{x} \in [a, b]$ such that $f(\bar{x}) = \bar{x}$, i.e., f has at least one fixed point.*

Proof Define the function $g : [a, b] \rightarrow \mathbb{R}$, given by $g(x) = f(x) - x$. It is obvious that g is continuous it is the difference of continuous functions and, in particular, it has the Darboux property. Obviously, because $f(a), f(b) \in [a, b]$, one has the inequalities

$$\begin{aligned} g(a) &= f(a) - a \geq 0 \\ g(b) &= f(b) - b \leq 0, \end{aligned}$$

hence $g(a) \cdot g(b) \leq 0$, and by the use of the Darboux property, one deduces that there exists a point $\bar{x} \in [a, b]$ such that $g(\bar{x}) = 0$. Therefore, $f(\bar{x}) = \bar{x}$ and the proof is now complete. \square

Remark 2.3.18. *In this situation, the uniqueness of the fixed point is not ensured. The immediate example is the identity function of the $[a, b]$ interval. It is also essential that the interval is closed. For example, the function $f : [0, 1) \rightarrow [0, 1)$ given by $f(x) = \frac{x+1}{2}$ does not have any fixed point. It is equally essential that the interval is bounded: the function $f : [1, +\infty) \rightarrow [1, +\infty)$ given by $f(x) = x + x^{-1}$ does not have any fixed point. Finally, the result does not hold anymore in the case where the function is not defined on an interval. For instance, $f : [-2, -1] \cup [1, 2] \rightarrow [-2, -1] \cup [1, 2]$, $f(x) = -x$ does not have any fixed point.*

In some situations, we can say something about the structure of the set of fixed points.

Theorem 2.3.19. *Let $f : [0, 1] \rightarrow [0, 1]$ be a 1-Lipschitz function on $[0, 1]$. Then the set of the fixed points of f is a (possibly degenerate) interval.*

Proof Let

$$F := \{x \in [0, 1] \mid f(x) = x\}$$

be the set of fixed points of f . It is clear, from the continuity of f , that F is a compact set, so it admits a minimum and a maximum, denoted respectively by a and b . Obviously, $F \subset [a, b]$. If we fix now an arbitrary $\bar{x} \in [a, b]$, it is sufficient to prove that \bar{x} is a fixed point of f , i.e., $\bar{x} \in F$. Since a is a fixed point of f and $a \leq \bar{x}$, one gets

$$f(\bar{x}) - a \leq |f(\bar{x}) - a| = |f(\bar{x}) - f(a)| \leq \bar{x} - a,$$

so $f(\bar{x}) \leq \bar{x}$. By a similar argument in the case of b , one has:

$$b - f(\bar{x}) \leq |b - f(\bar{x})| = |f(b) - f(\bar{x})| \leq b - \bar{x},$$

which shows that $f(\bar{x}) \geq \bar{x}$. Consequently, $f(\bar{x}) = \bar{x}$ and the proof is complete. \square

We now emphasize a supplemental condition, which assures the uniqueness of the fixed point.

Theorem 2.3.20. *Let $a, b \in \mathbb{R}$, $a < b$ and $f : [a, b] \rightarrow [a, b]$ be a continuous function on $[a, b]$, which is differentiable on (a, b) and has the property that $f'(x) \neq 1$ for any $x \in (a, b)$. Then f has a unique fixed point.*

Proof Theorem 2.3.17 assures the existence of the fixed point. If, by contradiction, there are two fixed points of f , denoted by x_1 and x_2 , then the Lagrange Theorem applied to f on the interval $[x_1, x_2]$, implies the existence of a point $c \in (x_1, x_2) \subset (a, b)$ such that

$$f(x_1) - f(x_2) = f'(c)(x_1 - x_2).$$

But this relation implies $f'(c) = 1$, which contradicts the assumption. This proves the result. \square

There is an important difference between the proof of Banach fixed point theorem and the proof of Theorem 2.3.17 above, which is that we are not given any information concerning the position of the fixed point, or a method to approximate it. Recall that the successive approximations method from the Banach's proof works like this: we start with an arbitrary point x_0 (this time, from the interval $[a, b]$), and we construct by recurrence a sequence (the sequence of Picard approximations) from $[a, b]$ by the relation $x_{n+1} = f(x_n)$ for every $n \in \mathbb{N}$. All of the hard work in the proof of the Banach

Principle was to show that (x_n) converges. In case we know this, the continuity of f furnishes the conclusion that the limit of x_n , denoted by \bar{x} , is a fixed point for f . Indeed, we may write:

$$f(\bar{x}) = f(\lim_{n \rightarrow \infty} x_n) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = \bar{x}. \quad (2.3.8)$$

In order to prove that sequence of the successive approximations is convergent, one usually shows that this sequence is fundamental.

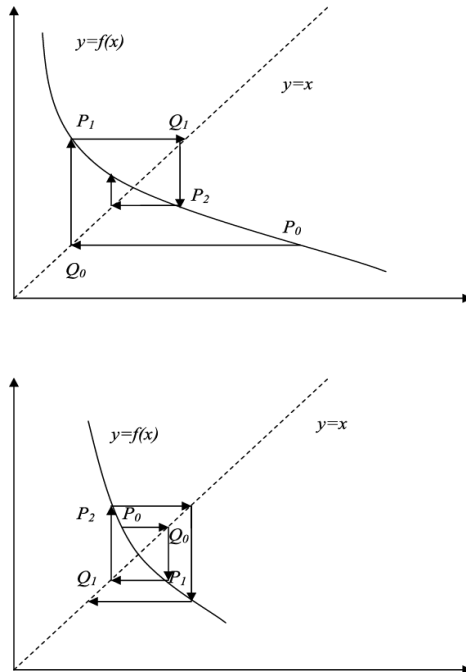


Figure 2.2: Behavior of the Picard iterations.

We can show that in certain conditions, the Picard iterations are convergent to a fixed point, outside the standard assumptions of the Banach Principle. This result is called the Picard convergence theorem, and its geometric meaning stems from the following discussion. Take $x_0 \in [a, b]$ arbitrary and consider the Picard iteration associated with the function f and the initial point x_0 , $x_{n+1} = f(x_n)$, for any $n \in \mathbb{N}$. We may describe this procedure of determining the Picard sequence geometrically. Consider first $P_0(x_0, f(x_0))$ on the curve of equation $y = f(x)$ (the graph of f). Take the point $Q_0(f(x_0), f(x_0))$ on the first bisector $y = x$, and we project it vertically on the curve $y = f(x)$ in order to find the point $P_1(x_1, f(x_1))$. We project the point P_1 horizontally in Q_1 , which lies on the first bisector, and continue this procedure by the vertical projection of Q_1 on $y = f(x)$ in the point $P_2(x_2, f(x_2))$. One can repeat this procedure

in such a way that a “spider net” is obtained, where the fixed point of f is “captured” (i.e., the intersection of the curves $y = f(x)$ and $y = x$). Actually, one may observe that this “capturing” happens if the slopes of the tangents to the curve $y = f(x)$ are smaller in absolute value than the slope of the line $y = x$ (which equals 1). If this condition concerning the slopes is not satisfied, then the points Q_0, Q_1, \dots move away from the fixed point (see Figure 2.2 for the two different situations).

Therefore, the Picard Theorem analytically transposes this geometric observation.

Theorem 2.3.21 (Picard Theorem). *Let $a, b \in \mathbb{R}$, $a < b$ and $f : [a, b] \rightarrow [a, b]$ be a continuous function on $[a, b]$ and differentiable on (a, b) , having the property that $|f'(x)| < 1$ for any $x \in (a, b)$. Then f has a unique fixed point, and the sequences of Picard iterations are convergent to the unique fixed point of f .*

Proof Theorem 2.3.20 assures the existence and the uniqueness of the fixed point. Denote it by \bar{x} . Take $x_0 \in [a, b]$ and let (x_n) be the Picard sequence which is generated starting at x_0 . If for $n \in \mathbb{N}$ we have $x_n = x_{n+1}$, then x_n is the fixed point, and the sequence becomes stationary. In this case the convergence is obvious. Suppose then that (x_n) is not stationary. From the Lagrange Theorem (applied to the function f on an interval having endpoints \bar{x} and x_n) one deduces the existence of an element $c_n \in (a, b)$ for which we can write the relation:

$$x_{n+1} - \bar{x} = f(x_n) - f(\bar{x}) = f'(c_n)(x_n - \bar{x}).$$

This inequality and the assumption on the derivative of f imply collectively that

$$|x_{n+1} - \bar{x}| < |x_n - \bar{x}|. \quad (2.3.9)$$

We prove now that (x_n) converges to \bar{x} . Since (x_n) is a bounded sequence (in $[a, b]$), it is sufficient to prove that every convergent subsequence has its limit equal to \bar{x} . Consider then a convergent subsequence $(x_{n_k})_k$ of (x_n) and denote by $x \in [a, b]$ its limit. Since (n_k) is strictly increasing, by applying inductively the inequality (2.3.9), one gets:

$$|x_{n_{k+1}} - \bar{x}| \leq |x_{n_{k+1}} - \bar{x}| < |x_{n_k} - \bar{x}|. \quad (2.3.10)$$

But $|x_{n_k} - \bar{x}| \xrightarrow{k \rightarrow \infty} |x - \bar{x}|$, and $|x_{n_{k+1}} - \bar{x}| \xrightarrow{k \rightarrow \infty} |x - \bar{x}|$, and also

$$|x_{n_{k+1}} - \bar{x}| = |f(x_{n_k}) - f(\bar{x})| \xrightarrow{k \rightarrow \infty} |f(x) - f(\bar{x})|.$$

Passing to the limit in (2.3.10), one gets then

$$|f(x) - f(\bar{x})| = |x - \bar{x}|.$$

If $x \neq \bar{x}$, by applying again the Lagrange Theorem, we have

$$|f(x) - f(\bar{x})| < |x - \bar{x}|,$$

which provides a contradiction. Therefore, $\bar{x} = x$, and the theorem is proved. \square

We want to emphasize now that if the condition $|f'(x)| < 1$ does not hold for any $x \in (a, b)$, then, even if f has a unique fixed point, the Picard sequence may not converge to this fixed point. For instance, consider the function $f : [-1, 1] \rightarrow [-1, 1]$ given by $f(x) = -x$. Obviously, f is differentiable on its domain, but the absolute value of its derivative equals 1. The sole fixed point of f is $\bar{x} = 0$. By construction the Picard iteration starting from $x_0 \neq 0$, the Picard sequence has the form $x_0, -x_0, x_0, -x_0, \dots$, hence is not convergent. Remark also that a function which satisfies the assumptions of the previous theorem is not necessarily a contraction, hence the Picard Theorem cannot be obtained from the Banach Principle. To illustrate this aspect, we have the following example. Take $f : [0, 1] \rightarrow [0, 1]$, $f(x) = \frac{x}{1+x}$. Obviously, $|f'(x)| \in (0, 1)$ for any $x \in (0, 1)$, but f is not a contraction because

$$\lim_{(x,y) \rightarrow (0,0), x \neq y} \frac{|f(x) - f(y)|}{|x - y|} = \lim_{(x,y) \rightarrow (0,0), x \neq y} \frac{1}{(x+1)(y+1)} = 1.$$

However, we could apply the Banach Principle for the function f if we restrict it to an interval on which it is a contraction. Observe that for $x \in [\frac{1}{2}, 1]$, $f(x) \in [\frac{1}{2}, \frac{2}{3}]$, so we can define $f : [\frac{1}{2}, 1] \rightarrow [\frac{1}{2}, 1]$ and remark that this restriction is a contraction, because $\sup_{x \in [\frac{1}{2}, 1]} |f'(x)| = \frac{4}{9} < 1$. Therefore, there exists a unique fixed point of f in $[\frac{1}{2}, 1]$, and since the initial function has a unique fixed point, the fixed point of the restriction coincides with it.

Also in the framework of this example, one can apply Theorem 2.3.12, because $f^2(x) = \frac{x+1}{x+2}$ is a contraction ($\sup_{x \in [0,1]} |(f^2)'(x)| = \frac{1}{4} < 1$).

In some cases, it may happen that we cannot restrict the function to a contraction, nor can we find an iteration which is a contraction, so one cannot apply any of the two methods from the preceding example. For example, take the function $f : [0, 1] \rightarrow [0, 1]$, $f(x) = \frac{x}{1+x}$. Again, $|f'(x)| \in (0, 1)$ for any $x \in (0, 1)$, and for every $n \in \mathbb{N}^*$ and every $x \in [0, 1]$, $f^n(x) = \frac{x}{1+nx}$. As above:

$$\lim_{(x,y) \rightarrow (0,0), x \neq y} \frac{|f^n(x) - f^n(y)|}{|x - y|} = \lim_{(x,y) \rightarrow (0,0), x \neq y} \frac{1}{(nx+1)(ny+1)} = 1$$

and hence f^n is not a contraction. Moreover, it is impossible to find any restriction of the function f which is a contraction from a set into itself, because the fixed point of f is $\bar{x} = 0$, so, such a restriction should be defined on an interval which contains 0, which is impossible in view of relation

$$\lim_{(x,y) \rightarrow (0,0), x \neq y} \frac{|f(x) - f(y)|}{|x - y|} = 1.$$

However, the function from the Picard Theorem is necessarily a weak contraction, and we will see in the subsequent sections how one can reobtain the Picard Theorem by using a more general result concerning weak contractions (Theorem 2.3.24 and Corollary 2.3.25).

In the assumptions of the Picard Theorem, the sequence of iterations is convergent to the fixed point of the function f for any possible choice of the initial term x_0 in the interval $[a, b]$. In order to obtain a quicker convergence, it is natural to try to start with this first term as close as we can to \bar{x} . Observe that, because of equality $\bar{x} = f(\bar{x})$, the point \bar{x} lies in the image of f , and since \bar{x} is a fixed point, it would be in the images of the functions f^n for every natural nonzero n . Therefore, \bar{x} will lie in the intersection of all the images of the above functions. Hence, if we could calculate $\bigcap_{n \in \mathbb{N}^*} \text{Im } f^n$, we could be close enough to the fixed point. In case this intersection consists of only one point, that point is the desired fixed point. For example, consider the function $f : [0, 1] \rightarrow [0, 1]$ given by $f(x) = \frac{x+1}{4}$. Inductively, one can show that for every $n \in \mathbb{N}^*$,

$$f^n(x) = \frac{3x + 4^n - 1}{3 \cdot 4^n},$$

and

$$\text{Im } f^n = \left[\frac{1}{3} \left(1 - \frac{1}{4^n}\right), \frac{1}{3} \left(1 + \frac{2}{4^n}\right) \right].$$

If we take the intersection of all these intervals, we get

$$\bigcap_{n \in \mathbb{N}^*} \text{Im } f^n = \left\{ \frac{1}{3} \right\},$$

hence $\bar{x} = \frac{1}{3}$ is the fixed point of the function f we are looking for.

The next result highlights the most important assumption of Picard Theorem, by emphasizing a partial converse.

Proposition 2.3.22. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function and $x_0 \in \mathbb{R}$. If the sequence of Picard iterations starting from x_0 converges to a number $l \in \mathbb{R}$, without being stationary, and f is differentiable at l , then $|f'(l)| \leq 1$.*

Proof Suppose by contradiction that $|f'(l)| > 1$. It is clear, from the continuity of f and from a previous comment, that l must be a fixed point of f . Since

$$\lim_{x \rightarrow l} \frac{f(x) - f(l)}{x - l} = f'(l)$$

we get

$$\lim_{x \rightarrow l} \left| \frac{f(x) - f(l)}{x - l} \right| = |f'(l)|.$$

Take

$$\varepsilon := \frac{|f'(l)| - 1}{2} > 0.$$

Then for this ε , there exists $\delta > 0$ such that for any $x \in (l - \delta, l + \delta) \setminus \{l\}$,

$$|f'(l) - \varepsilon < \left| \frac{f(x) - f(l)}{x - l} \right| < |f'(l) + \varepsilon$$

or

$$\frac{|f'(l) + 1}{2} < \left| \frac{f(x) - l}{x - l} \right|.$$

In particular, since $|f'(l)| > 1$,

$$|x - l| < |x - l| \frac{|f'(l) + 1}{2} < |f(x) - f(l)|$$

for every $x \in (l - \delta, l + \delta) \setminus \{l\}$. Since the sequence (x_n) of the Picard iterations starting from x_0 converges to l , without being stationary, there exists $n_\delta \in \mathbb{N}$ such that for any $n \geq n_\delta$,

$$x_n \in (l - \delta, l + \delta) \setminus \{l\}.$$

One gets from the above relations that

$$|x_n - l| < |f(x_n) - f(l)| = |x_{n+1} - l|$$

for every $n \geq n_\delta$. In particular, we get

$$|x_{n_\delta} - l| < |x_{n_\delta+1} - l| < |x_{n_\delta} - l|$$

for every $n > n_\delta + 1$. Passing to the limit in the last relation for $n \rightarrow \infty$, we get the contradiction:

$$|x_{n_\delta} - l| < |x_{n_\delta+1} - l| \leq 0.$$

Consequently, the assumption made is false, hence $|f'(l)| \leq 1$. \square

Remark 2.3.23. *In general, one cannot obtain the inequality on the derivative if the sequence is stationary. If one considers the function $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^3 + 2x$, the sequence of Picard iterations starting from 0 is convergent (even stationary) to 0, but $f'(0) = 2$.*

We now present a very interesting result due to Beardon, which says that in case of a weak contraction on \mathbb{R} (which, as we already saw in Example 2.3.10, may not have fixed points), the Picard iteration process gives the same limit point (possibly equal to $+\infty$ or $-\infty$), for any choice of the initial data.

Theorem 2.3.24 (Beardon's Theorem). *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function which is a weak contraction on \mathbb{R} , i.e., for any two distinct real numbers x and y , one has*

$$|f(x) - f(y)| < |x - y|.$$

Then there exists $\bar{x} \in \overline{\mathbb{R}}$ such that for any $x \in \mathbb{R}$, the Picard sequence generated by

$$x_n = f^n(x), \forall n \in \mathbb{N}^*$$

has the limit \bar{x} .

Proof Suppose firstly that f has a fixed point, $\bar{x} \in \mathbb{R}$. Obviously, this will be the unique fixed point of f . Without loss of generality, we may consider $\bar{x} = 0$, hence

$$|f(x)| < |x|$$

for any $x \neq 0$. Fix $x \in \mathbb{R}$. Then the sequence $(|f^n(x)|)_n$ is decreasing and hence is convergent to a number $\mu(x) \geq 0$. We will prove that $\mu(x) = 0$ for any initial data x . Suppose by contradiction that $\mu(x) > 0$. Then $f(\mu(x)) =: y_1$, and $f(-\mu(x)) =: y_2$, where $|y_1|, |y_2| < |\mu(x)|$. By the continuity of f , there are two neighborhoods of $\mu(x)$ and $-\mu(x)$, respectively, which are applied through f in the interval $I := (-\mu(x), \mu(x))$, which contains y_1 and y_2 . Then, for n sufficiently large, $f^n(x)$ also lies in I , which contradicts the inequality $|f^n(x)| \geq |\mu(x)|$. Therefore, for every $x \in \mathbb{R}$, one has the convergence $f^n(x) \rightarrow 0$.

Suppose now f does not have any fixed point. Then $f(x) > x$ or $f(x) < x$ for any real x . We will only prove the case $f(x) > x$, the other situation following analogously. It is clear that for any real x , the sequence $(f^n(x))$ is strictly increasing, so it has the limit in $(-\infty, +\infty]$. If the limit would be a real number l , then $f^n(x) \neq l$ for any natural nonzero n , and we get

$$|f^{n+1}(x) - f(l)| < |f^n(x) - l|$$

from where, by passing to the limit, we get that $f(l) = l$, i.e., l is a fixed point, which is a contradiction. Hence, $f^n(x) \rightarrow +\infty$. Obviously, in case $f(x) < x$, we get $f^n(x) \rightarrow -\infty$. The proof is now complete. \square

Remark also from the proof that if $\bar{x} \in \mathbb{R}$, then \bar{x} is necessarily the sole fixed point of f . Actually, if f has a fixed point (which is necessarily unique from the contraction condition), then all the iterations converge to this fixed point, and if f does not have a fixed point, then the Picard iterations converge to $+\infty$ or $-\infty$, according to $f(x) > x$ or $f(x) < x$, respectively.

Based on this result we can deduce the next corollary.

Corollary 2.3.25. *Let $a, b \in \mathbb{R}$, $a < b$ and $f : [a, b] \rightarrow [a, b]$ be a weak contraction. Then f has a unique fixed point, and for any initial data $x \in [a, b]$, the sequences of the Picard iterations are convergent to this fixed point.*

Proof The existence and the uniqueness of the fixed point are assured by the Theorem 2.3.17 and by the contraction condition. By repeating the arguments from the first part

of the proof of the previous theorem, one gets the conclusion concerning the Picard iterations. \square

The preceding corollary one can obtain also the Picard theorem of convergence (Theorem 2.3.21), because in the assumptions of that theorem, the function is a weak contraction.

2.4 Graves Theorem

We dedicate this section to an important result, known as the Graves Theorem, which gives sufficient conditions for a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ to be open (i.e., the image through f of an open set in \mathbb{R}^n is open in \mathbb{R}^m). If f is linear, this is an well-known and deep result, known in Functional Analysis as the Open Mapping Principle (and is applicable, as well as the Graves Theorem, in a much wider setting). We give this principle next and illustrate it through a very short and elementary proof (in our particular framework).

Theorem 2.4.1 (Open Mapping Principle). *Let $T : \mathbb{R}^p \rightarrow \mathbb{R}^m$ be a linear surjective map ($m, p \in \mathbb{N}^*$). Then T is open.*

Proof First of all, we recall that T is continuous. Let $\{e_1, e_2, \dots, e_m\}$ be the canonical base of \mathbb{R}^m . Taking into account the surjectivity of T , for any $i \in \overline{1, m}$, there exist $x_i \in \mathbb{R}^p$ with $T(x_i) = e_i$. Now take $\psi : \mathbb{R}^m \rightarrow \mathbb{R}^p$, $\psi(\alpha) = \psi(\alpha_1, \dots, \alpha_m) = \sum_{i=1}^m \alpha_i x_i$. Clearly, ψ is continuous and, moreover, $T \circ \psi : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is the identity of \mathbb{R}^m . In particular, ψ is injective. Now, consider $U \subset \mathbb{R}^p$ an open set. Since ψ is continuous, $\psi^{-1}(U)$ is open and since ψ is injective, $U = \psi(\psi^{-1}(U))$. But

$$T(U) = T(\psi(\psi^{-1}(U))) = (T \circ \psi)(\psi^{-1}(U)) = \psi^{-1}(U).$$

So, $T(U)$ is open. \square

Remark that, in fact, a stronger property than the usual openness can be deduced from the Open Mapping Principle: there exists $L > 0$ such that $B(Tx, Lr) \subset T(B(x, r))$ for every $x \in \mathbb{R}^p$ and $r > 0$.

To see this, observe that it is sufficient to prove the above relation for $x = 0$. In this case, since $T(B(0, 1))$ is an open set, there exists an $L > 0$ such that $B(0, L) \subset T(B(0, 1))$. The fact that $B(0, Lr) \subset T(B(0, r))$ for every $r > 0$ easily follows.

The property mentioned before is called linear openness, and the corresponding constant $L > 0$ is called the linear openness modulus.

We now present two results given in 1950 by the American mathematician Lawrence Murray Graves. For a general function $f : \mathbb{R}^p \rightarrow \mathbb{R}^m$, one says that f is linearly open at x_0 with modulus $L > 0$ if there is $\rho > 0$ such that $B(f(x_0), Lr) \subset f(B(x_0, r))$

for every $r \in (0, \rho)$. Of course, in case of linear operators, the linear openness at a single point is equivalent to the linear openness at every point.

Theorem 2.4.2 (Graves). *Let $T : \mathbb{R}^p \rightarrow \mathbb{R}^m$ be a linear surjective map, and denote its linear openness modulus by $L > 0$. Let $M \in (0, L)$, $r > 0$, $U \subset \mathbb{R}^p$ be an open set, $x_0 \in U$, and $f : U \rightarrow \mathbb{R}^m$ be a continuous function such that*

$$\|f(u) - f(v) - T(u - v)\| \leq M \|u - v\|$$

for any $u, v \in B(x_0, r)$. Then f is linearly open at x_0 with modulus $L - M > 0$.

Proof Take $y \in \mathbb{R}^m$ such that $\|y - f(x_0)\| < r(L - M)$. We construct inductively a sequence (ξ_n) as follows: take $\xi_0 := 0$ and, for $n \geq 1$, ξ_n satisfies

$$T(\xi_n - \xi_{n-1}) = y - f(x_0 + \xi_{n-1}) \quad (2.4.1)$$

and

$$\|\xi_n - \xi_{n-1}\| \leq L^{-1} \|y - f(x_0 + \xi_{n-1})\|. \quad (2.4.2)$$

We easily deduce

$$T(\xi_n - \xi_{n-1}) = T(\xi_{n-1} - \xi_{n-2}) - f(x_0 + \xi_{n-1}) + f(x_0 + \xi_{n-2}),$$

hence

$$\begin{aligned} \|\xi_n - \xi_{n-1}\| &\leq L^{-1} \|T(\xi_n - \xi_{n-1})\| \leq L^{-1} \|T(\xi_{n-1} - \xi_{n-2}) - f(x_0 + \xi_{n-1}) + f(x_0 + \xi_{n-2})\| \\ &\leq L^{-1} M \|\xi_{n-1} - \xi_{n-2}\|, \end{aligned} \quad (2.4.3)$$

if $\xi_n, \xi_{n-1} \in B(0, r)$. This is true for every n . Indeed, from (2.4.1) and (2.4.2), we obtain that

$$\|\xi_1\| \leq L^{-1} \|y - f(x_0)\| < r(1 - L^{-1}M) < r,$$

and since $\xi_0 = 0$, we obtain $\xi_1, \xi_0 \in B(0, r)$, hence (2.4.3) holds for $n = 2$. From this,

$$\|\xi_2\| = \|\xi_2 - \xi_1 + \xi_1\| \leq \|\xi_1\| (1 + L^{-1}M) < r,$$

hence (2.4.3) holds for $n = 3$. Suppose (2.4.3) holds until n , and deduce that

$$\|\xi_n\| \leq \|\xi_1\| (1 + L^{-1}M + \dots + (L^{-1}M)^{n-1}) < r, \quad (2.4.4)$$

which means that (2.4.3) holds for $n + 1$. Moreover, from (2.4.3) we know that the sequence (ξ_n) is Cauchy, and hence is convergent to an element $\xi \in \mathbb{R}^p$. From (2.4.1) and the continuity of f , $y = f(x_0 + \xi)$, and (2.4.4) shows that

$$\|\xi\| \leq \|\xi_1\| (1 - L^{-1}M)^{-1} \leq L^{-1} \|y - f(x_0)\| (1 - L^{-1}M)^{-1} < L^{-1}r(L - M)(1 - L^{-1}M)^{-1} = r.$$

But this means exactly that

$$B(f(x_0), (L - M)r) \subset f(B(x_0, r)).$$

Since this is true if one replaces r above with arbitrary $r' \in (0, r)$, we have the conclusion. \square

Corollary 2.4.3. *Let $T : \mathbb{R}^p \rightarrow \mathbb{R}^m$ be a linear surjective map, with linear openness modulus $L > 0$. If S is a linear map such that $\|S - T\| \leq L/2$, then S is surjective, hence linearly open, with the linear openness modulus $L/2$.*

Proof Apply the previous theorem for $f := S$ and $x_0 := 0$. \square

On the basis of the previous results, one deduces the celebrated Lyusternik-Graves Theorem.

Theorem 2.4.4 (Lyusternik-Graves). *Let $U \subset \mathbb{R}^p$ be an open set, $x_0 \in U$, and $f : U \rightarrow \mathbb{R}^m$ a Fréchet differentiable function, with ∇f continuous at x_0 , and $\nabla f(x_0)$ surjective. Then f is linearly open at x_0 .*

Proof Denote by L the openness modulus of $\nabla f(x_0)$, given by the Open Mapping Principle. Since ∇f is continuous at x_0 , there exists $\delta > 0$ such that for every $x \in B(x_0, \delta)$, one has

$$\|\nabla f(x) - \nabla f(x_0)\| \leq L/8 < L/2. \tag{2.4.5}$$

This means, on the basis of Corollary 2.4.3, that for every $x \in B(x_0, \delta)$, $\nabla f(x)$ is linearly open with modulus $L/2$. For such x , one can apply Lagrange Theorem to the function $y \mapsto f(y) + \nabla f(x)y$, to get that

$$\|f(u) - f(v) - \nabla f(x)(u - v)\| \leq \|u - v\| \sup_{t \in [x_0, x]} \|\nabla f(t) - \nabla f(x)\| \leq \frac{L}{4} \|u - v\|,$$

if $u, v \in B(x_0, \delta)$, where (2.4.5) was used in the last inequality. The conclusion now follows from Theorem 2.4.2, with $L/2$ instead of L , and $L/4$ instead of M . \square

The condition that $\nabla f(x_0)$ is surjective, as the iterative procedure from the proof of Theorem 2.4.2, was introduced in 1934 by the Russian mathematician Lazar Aronovich Lyusternik. For more extensions, developments and historical facts, see (Dontchev and Rockafellar, 2009) and (Klatte and Kummer, 2002).

2.5 Semicontinuous Functions

The aim of this section is to introduce and study some generalizations of the continuity concept, which will be useful in the subsequent discussion concerning the existence of the solutions of the optimization problems.

Observe that a function $f : \mathbb{R}^p \rightarrow \mathbb{R}$ is continuous at a if and only if the next two conditions simultaneously hold:

$$\forall \lambda \in \mathbb{R}, \lambda < f(a), \exists U \in \mathcal{V}(a), \forall x \in U, \lambda < f(x) \quad (2.5.1)$$

and

$$\forall \lambda \in \mathbb{R}, \lambda > f(a), \exists U \in \mathcal{V}(a), \forall x \in U, \lambda > f(x). \quad (2.5.2)$$

By taking each of these two conditions separately, we can define lower and upper semicontinuity. Therefore, the function $f : \mathbb{R}^p \rightarrow \mathbb{R}$ is lower semicontinuous at $a \in \mathbb{R}$ if the condition (2.5.1) is satisfied, and f is upper semicontinuous at a if the condition (2.5.2) is satisfied. Similarly, if the respective conditions hold in every point of \mathbb{R}^p , one says that f is lower semicontinuous, or respectively upper semicontinuous. Obviously, according to the definitions, a function is continuous at $a \in \mathbb{R}^p$ if and only if it is simultaneously upper and lower semicontinuous at a . It is easy to provide examples of functions which are semicontinuous without being continuous. For instance, the function $f : \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} 1, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

is lower semicontinuous (on \mathbb{R}), but is discontinuous at 0. Similarly, the function $f : \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} 0, & x \neq 0 \\ 1, & x = 0 \end{cases}$$

is upper semicontinuous (on \mathbb{R}), but is discontinuous at 0.

A more elaborate example is the Riemann function $f : [0, 1] \rightarrow \mathbb{R}$, given by

$$f(x) = \begin{cases} \frac{1}{n}, & x \in (0, 1], x = \frac{m}{n}, m, n \in \mathbb{N}^*, (m, n) = 1 \\ 0, & x \in [0, 1] \setminus \mathbb{Q} \text{ or } x = 0, \end{cases}$$

which is continuous on $[0, 1] \setminus \mathbb{Q} \cup \{0\}$ and discontinuous on $(0, 1] \cap \mathbb{Q}$. This function has its limit equal to 0 at every point. This shows that it is upper semicontinuous.

One can easily show that f is upper semicontinuous if and only if $-f$ is lower semicontinuous. As consequence, we will restrict our study to lower semicontinuous functions, since the results can be easily reformulated in the case of the upper semicontinuous functions.

Having a function $f : \mathbb{R}^p \rightarrow \mathbb{R}$, besides the epigraph of the function previously introduced (Proposition 2.2.3), that is

$$\text{epi } f = \{(x, t) \in \mathbb{R}^p \times \mathbb{R} \mid f(x) \leq t\},$$

we introduce now the level sets: if $v \in \mathbb{R}$,

$$N_v f := \{x \in \mathbb{R}^p \mid f(x) \leq v\} = f^{-1}((-\infty, v]).$$

Those functions which are globally lower semicontinuous, have the following characterization theorem.

Theorem 2.5.1. *Let $f : \mathbb{R}^p \rightarrow \mathbb{R}$. The next assertions are equivalent:*

- (i) f is lower semicontinuous (on \mathbb{R}^p);
- (ii) $N_\nu f$ is closed in \mathbb{R}^p for any $\nu \in \mathbb{R}$;
- (iii) $\text{epi } f$ is a closed set in $\mathbb{R}^p \times \mathbb{R}$;
- (iv) $\{x \in \mathbb{R}^p \mid f(x) > \beta\}$ is open in \mathbb{R}^p for any $\beta \in \mathbb{R}$.

Proof (i) \Rightarrow (ii) Take $\nu \in \mathbb{R}$. We prove that $N_\nu f$ has an open complement. Take $x \notin N_\nu f$, i.e., $f(x) > \nu$. Since f is lower semicontinuous at x , there exists U , a neighborhood of x , such that $f(y) > \nu$ for any $y \in U$. Therefore, $U \cap N_\nu f = \emptyset$, i.e., $U \subset \mathbb{R}^p \setminus N_\nu f$. It follows that the complement of $N_\nu f$ is open, i.e., $N_\nu f$ is closed.

(ii) \Rightarrow (iii) We prove that $(\mathbb{R}^p \times \mathbb{R}) \setminus \text{epi } f$ is open. Take $(x, t) \notin \text{epi } f$, which means $f(x) > t$. There exists ν such that $f(x) > \nu > t$. Then $x \notin N_\nu f$ and according to (ii), there exists U , a neighborhood of x , such that $U \cap N_\nu f = \emptyset$. It follows that $U \times (-\infty, \nu] \cap \text{epi } f = \emptyset$. Since $U \times (-\infty, \nu]$ is a neighborhood of (x, t) , we obtain the conclusion.

(iii) \Rightarrow (i) Take $x \in \mathbb{R}^p$ and $t \in \mathbb{R}$ such that $f(x) > t$. Then $(x, t) \notin \text{epi } f$ and hence there exists U , a neighborhood of x , and $\varepsilon > 0$, such that $U \times (t - \varepsilon, t + \varepsilon) \cap \text{epi } f = \emptyset$. Therefore, for every $y \in U$, $(y, t) \notin \text{epi } f$, i.e., $f(y) > t$. Accordingly, f is lower semicontinuous.

(ii) \Leftrightarrow (iv) The relation

$$\{x \in \mathbb{R}^p \mid f(x) > \beta\} = \mathbb{R}^p \setminus N_\beta f$$

proves the equivalence between (ii) and (iv). □

A function is upper semicontinuous if and only if for any $y \in \mathbb{R}$, the sets of the type $\{x \in \mathbb{R}^p \mid f(x) \geq y\}$ are closed, which is equivalent to the fact that the sets of the type $\{x \in \mathbb{R}^p \mid f(x) < y\}$ are open.

Let $f : A \subset \mathbb{R}^p \rightarrow \mathbb{R}$ be a function. Its lower and upper limits at $a \in \text{cl } A$ are given, respectively, by

$$\liminf_{x \rightarrow a} f(x) := \sup_{U \in \mathcal{V}(a)} \inf_{x \in U \cap A} f(x) \text{ and } \limsup_{x \rightarrow a} f(x) := \inf_{U \in \mathcal{V}(a)} \sup_{x \in U \cap A} f(x). \tag{2.5.3}$$

From their definitions, it is obvious that

$$\liminf_{x \rightarrow a} f(x) \leq \limsup_{x \rightarrow a} f(x) \text{ and } \limsup_{x \rightarrow a} f(x) = -\liminf_{x \rightarrow a} (-f)(x).$$

Moreover, for a sequence (x_n) , one defines its lower and upper limits as

$$\liminf_{n \rightarrow \infty} x_n := \sup_{n \in \mathbb{N}} \inf_{k \geq n} x_k \text{ and } \limsup_{n \rightarrow \infty} x_n := \inf_{n \in \mathbb{N}} \sup_{k \geq n} x_k. \tag{2.5.4}$$

In fact, the definitions (2.5.4) naturally follow from (2.5.3) for the function $f : \mathbb{N} \rightarrow \mathbb{R}$, $f(n) := x_n$, and $a := \infty \in \text{cl } \mathbb{N}$.

Theorem 2.5.2. *Let $f : \mathbb{R}^p \rightarrow \mathbb{R}$ and $x \in \mathbb{R}^p$. Then:*

- (i) *f is lower semicontinuous at x if and only if $f(x) = \liminf_{y \rightarrow x} f(y)$.*
- (ii) *If f is lower semicontinuous at x , then for every $(x_n) \rightarrow x$, one has $\liminf_{n \rightarrow \infty} f(x_n) \geq f(x)$.*

Proof (i) It is easy to observe that for every $U \in \mathcal{V}(x)$, $\inf_{y \in U} f(y) \leq f(x)$, hence $\liminf_{y \rightarrow x} f(y) \leq f(x)$, and this is true for an arbitrary function. Suppose f is lower semicontinuous at x and take $\lambda \in \mathbb{R}$, $\lambda < f(x)$. Then, according to the definition of the lower semicontinuity, there exists $U \in \mathcal{V}(x)$ such that $\lambda < f(y)$, for any $y \in U$. It means that $\lambda \leq \inf_{y \in U} f(y) \leq \liminf_{y \rightarrow x} f(y)$. Since λ was taken arbitrary and assumed to be smaller than $f(x)$, it follows that $f(x) \leq \liminf_{y \rightarrow x} f(y)$, and hence we have equality. Suppose now $f(x) = \liminf_{y \rightarrow x} f(y)$ and take $\lambda \in \mathbb{R}$, $\lambda < f(x)$. According to the definition of \liminf , we know that there is $U \in \mathcal{V}(x)$ such that $\lambda < \inf_{y \in U} f(y) < f(z)$, for every $z \in U$, hence f is lower semicontinuous at x .

(ii) Take arbitrary $U \in \mathcal{V}(x)$. Observe that since $(x_n) \rightarrow x$, there exists $n_U \in \mathbb{N}$ such that, for every $k \geq n_U$, one has $x_k \in U$. Accordingly, $\inf_{k \geq n_U} f(x_k) \geq \inf_{y \in U} f(y)$. It follows that

$$\sup_{n \in \mathbb{N}} \inf_{k \geq n} f(x_k) \geq \inf_{y \in U} f(y), \quad \forall U \in \mathcal{V}(x).$$

Passing to the supremum for $U \in \mathcal{V}(x)$, and taking into account (i), the conclusion follows. \square

For upper semicontinuous functions, similar results can be deduced.

The next result is a generalization of the Weierstrass Theorem.

Theorem 2.5.3. *Let $f : \mathbb{R}^p \rightarrow \mathbb{R}$ be a lower semicontinuous function and $K \subset \mathbb{R}^p$ be a compact set. Then f is lower bounded on K and it attains its minimum on K .*

Proof Take $(x_k) \subset K$ such that $f(x_k) \rightarrow \inf\{f(x) \mid x \in K\}$. From the compactness of K , the sequence (x_k) has a convergent subsequence to an element $\bar{x} \in K$. From the fact that the sets of the type $K \cap N_\nu f$ are closed for any $\nu \in \mathbb{R}$, $\nu > \inf\{f(x) \mid x \in K\}$, we deduce that \bar{x} lies in all these sets (because for every ν , the terms x_k are, from a certain rank, in $K \cap N_\nu f$). Then we deduce that $f(\bar{x}) \leq \nu$ for any $\nu > \inf\{f(x) \mid x \in K\}$. Consequently, $f(\bar{x}) \leq \inf\{f(x) \mid x \in K\}$. On one hand, this means that $\inf\{f(x) \mid x \in K\} \in \mathbb{R}$, hence f is lower bounded on K , and, on the other hand, that \bar{x} is the point we are looking for, which realizes the minimum of f on K . The proof is complete. \square

Obviously, for upper semicontinuous functions, we will have the other “half” of the Weierstrass Theorem.

Theorem 2.5.4. Let $f : \mathbb{R}^p \rightarrow \mathbb{R}$ be an upper semicontinuous function and $K \subset \mathbb{R}^p$ be a compact set. Then f is upper bounded on K and it attains its maximum on K .

We know that, in general, the (pointwise) supremum of a family of continuous functions is not a continuous function. For instance, if we define for every $n \in \mathbb{N}$, $f_n : [0, 1] \rightarrow \mathbb{R}$, $f_n(x) = -x^n$, then the pointwise supremum is the function

$$f(x) = \sup_{n \in \mathbb{N}} f_n(x) = \begin{cases} 0, & x \in [0, 1) \\ -1, & x = 1, \end{cases}$$

which is not continuous on $[0, 1]$. In turn, f is lower semicontinuous on $[0, 1]$, a fact which is not accidental, as the following result shows.

Theorem 2.5.5. Let I be a nonempty arbitrary family of indices, and $(f_i)_{i \in I}$ a family of lower semicontinuous functions from \mathbb{R}^p into \mathbb{R} . If for any $x \in \mathbb{R}^p$,

$$\sup\{f_i(x) \mid i \in I\} \in \mathbb{R},$$

then the function $f : \mathbb{R}^p \rightarrow \mathbb{R}$ given by $f(x) = \sup\{f_i(x) \mid i \in I\}$ is lower semicontinuous on \mathbb{R}^p .

Proof For any $\alpha \in \mathbb{R}$, one has:

$$\begin{aligned} f^{-1}((\alpha, \infty)) &= \{x \in \mathbb{R}^p \mid f(x) > \alpha\} \\ &= \bigcup_{i \in I} \{x \in \mathbb{R}^p \mid f_i(x) > \alpha\}. \end{aligned}$$

Since the functions f_i are lower semicontinuous, the sets $\{x \in \mathbb{R}^p \mid f_i(x) > \alpha\}$ are open according to Theorem 2.5.1. Since every union of open sets is an open set, we deduce that $f^{-1}((\alpha, \infty))$ is open, and by applying again Theorem 2.5.1, we get that f is lower semicontinuous on \mathbb{R}^p . \square