Chapter 2
Encounters With Limits

Mathematics is not an arid land in the scientific universe. It is simultaneously the queen, maid and daughter of the observational sciences.

La mathématique ne constitue pas une terre aride dans l'univers scientifique. Elle est à la fois reine, servante et fille des sciences de l'observation.

Gustave Choquet (1915–2006)

Abstract The notion of limit is central in analysis. Thus the concept of convergence is presented in a general framework and then in the classes of topological spaces and metric spaces. Compactness, connectedness, completeness are studied in detail. Baire's Theorem is included as well as Ekeland's Variational Principle. The contraction theorem is proved and as an application an existence result for ordinary differential equations is presented.

The use of limits is the central theme of analysis. A detailed account of the history of limits and infinitesimals may be found in [28] in the context of nonstandard analysis. The rules in $\mathbb{R}$, $\mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$, $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$, $\mathbb{C}$, $\mathbb{R}^d$ serve as models for more general spaces. There are several ways to reach such generalizations. A first approach consists in selecting some rules for convergence; we evoke it briefly. A second approach uses some real-valued functions such as metrics, norms or seminorms in order to reduce convergence in an abstract set to convergence in $\mathbb{R}$. A third approach consists in introducing topologies, a set theoretical approach that is particularly versatile; it focusses attention on closed sets, i.e. sets that are stable when taking limits. However, it is usually formulated in terms of open sets, i.e. sets whose complements are closed.

The present chapter is devoted to a first glimpse of such approaches. In later chapters a more detailed study will be undertaken when special structures are at hand.

As an illustration let us consider the intuitive fact that if we write the letter Q with a segment that is shorter and shorter this letter "converges" to the letter O: Since the addresses we write on envelopes are read by machines and since these machines have a limited resolution, there is some risk that our letter will go to a wrong address if we are lazier and lazier in drawing the little slanted segment (Fig. 2.1).

J.-P. Penot, *Analysis*, Universitext, DOI 10.1007/978-3-319-32411-1_2

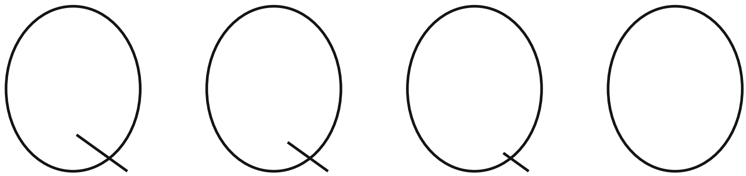

Fig. 2.1 Deformed letters

2.1 Convergences

The first approach to convergences appeared after the first quarter of the twentieth century. It is not often used nowadays, even if in some cases it is simpler than a topological approach (this is the case for pointwise convergence and for the second example given below). Since the selected rules are prevalent in practice, it is worth stating them in a formal definition. Let us first recall that a *sequence* in a set X is a map s from $\mathbb{N}$ to X, hence an element of $X^{\mathbb{N}}$. Setting $x_n := s(n)$ for $n \in \mathbb{N}$ we often write $(x_n)_{n\in\mathbb{N}}$, $(x_n)_{n\geq 0}$ or just (x_n) instead of s. A *subsequence* of $s := (x_n)$ is a sequence $s' := (x'_n)$ of X such that there exists a strictly increasing map $k : \mathbb{N} \to \mathbb{N}$ satisfying $s' = s \circ k$, i.e. $x'_n = x_{k(n)}$. Thus one obtains s' by deleting some terms of s and by reindexing the remaining terms.

Definition 2.1 A *space with limits* is a set X such that a relation denoted by $\to$ is defined between the sets $X^{\mathbb{N}}$ and X and read as (x_n) *converges* to x or x is a limit of (x_n), the relation $\to$ being required to satisfy the following properties:

(L1) any constant sequence with value x converges to x;
(L2) if $(x_n) \to x$, then any subsequence (x'_n) of (x_n) converges to x;
(L3) if $x \in X$ and $(x_n) \in X^{\mathbb{N}}$ are such that any subsequence (x'_n) of (x_n) has a subsequence (x''_n) converging to x, then $(x_n) \to x$.

For some needs, it is useful to add a uniqueness condition:

(U) if $(x_n) \to x$ and $(x_n) \to x'$ then $x = x'$.

Clearly, the convergences in $\mathbb{R}$, $\mathbb{R}_\infty$, $\overline{\mathbb{R}} := \mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^d$ satisfy these four conditions.

Example If X is the set of real-valued functions on a set S, or, in other terms, if $X = \mathbb{R}^S$, then *pointwise convergence* on X is defined by $(x_n) \to x$ if and only if for all $s \in S$ one has $(x_n(s))_n \to x(s)$ as $n \to +\infty$. Here the element $x := (x_s)_{s\in S}$ of $\mathbb{R}^S$ is identified with the function $f : S \to \mathbb{R}$ defined by $f(s) := x_s$ and is also denoted by $(x(s))_{s\in S}$.

Example *Uniform convergence* on the set X of real-valued functions on S is defined in the following way: $(f_n) \overset{u}{\to} f$ if and only if for all $\varepsilon > 0$ one can find some $n(\varepsilon) \in \mathbb{N}$ such that $|f_n(s) - f(s)| \leq \varepsilon$ for all $n \geq n(\varepsilon)$ and all $s \in S$.

Example Let X be the set of continuous functions on $\mathbb{R}^d$ that vanish outside a bounded subset. Declare that a sequence $(f_n) \to f$ if there exists a bounded subset B of $\mathbb{R}^d$ such that the functions f_n and f are null on $\mathbb{R}^d \backslash B$ and if the net of restrictions to B uniformly converges to the restriction of f: $(f_n \mid B) \xrightarrow{u} f \mid B$. Variants of such a convergence are used in the theory of distributions.

Sometimes sequences are inadequate and one must replace them with generalized sequences, also called nets (see Exercise 5). A *net* in a set X is a map s from a directed set $(I, \leq)$ into X. Setting $x_i := s(i)$, one also denotes it by $(x_i)_{i\in I}$. A *subnet* is a net $s' := (x'_j)_{j\in J}$ such that there exists a filtering map $h : J \to I$ such that $x'_j = x_{h(j)}$ for all $j \in J$. Note that, in contrast to what occurs for subsequences, one takes for J a directed set that may differ from I. It is often of the form $J := I \times K$, where K is another directed set, or a subset of $I \times K$, h being the first projection or a restriction of it. In some simple cases one can take for J a cofinal subset of I and for h the injection of J into I, but that it not always possible. The axioms we select are the analogues of those of the preceding definition.

Definition 2.2 A *convergence space* is a set X such that for every directed set I there is a relation denoted by $\to$ between the set X^I of nets of X indexed by I and X itself in such a way that the following conditions are satisfied:

(C1) for every $x \in X$ the constant net with value x converges to x;
(C2) if $(x_i)_{i\in I} \to x$ and if $(x'_j)_{j\in J}$ is a subnet of $(x_i)_{i\in I}$, then $(x'_j)_{j\in J} \to x$;
(C3) if $x \in X$ and $(x_i)_{i\in I}$ is a net in X such that for every subnet $(x'_j)_{j\in J}$ of $(x_i)_{i\in I}$ there exists a subnet $(x''_k)_{k\in K}$ of $(x'_j)_{j\in J}$ that converges to x, then $(x_i)_{i\in I} \to x$.

The preceding examples can be adapted to convergence spaces.

A map $f : X \to Y$ between two convergent spaces is said to be *continuous* at $x \in X$ if for any net $(x_i)_{i\in I}$ converging to x the net $(f(x_i))_{i\in I}$ converges to $f(x)$. It is continuous on some subset A of X if for all $x \in A$ it is continuous at x.

If $(W, \to)$ is a convergence space and if X is a subset of W, the induced convergence on X is defined by $(x_i)_{i\in I} \to_X x$ if $(x_i)_{\in I} \to x$ in W. It is easy to verify that the conditions (C1)–(C3) are satisfied and that a map $f : V \to X$ from a convergence space $(V, \to_V)$ into $(X, \to_X)$ is continuous at $\overline{v} \in V$ if and only if f is continuous at $\overline{v}$ when considered as a map from V into W.

If $(X_a)_{a\in A}$ is a family of convergence spaces, the product $X := \prod_{a\in A} X_a$ is made a convergence space by requiring that a net $(x_i)_{i\in I} \to x$ if for all $a \in A$ the a-component $(p_a(x_i)) \to p_a(x)$. Then, a map $f : W \to X$ from a convergence space W to X is continuous at $\overline{w} \in W$ if and only if for all $a \in A$ the a-component $f_a := p_a \circ f$ is continuous at $\overline{w}$.

The *lower limit* (resp. *upper limit*) of a net $(r_i)_{i\in I}$ of real numbers is defined by

$$\liminf_{i\in I} r_i := \sup_{h\in I} \inf_{i\in I,\, i\geq h} r_i \qquad (\text{resp. } \limsup_{i\in I} r_i := \inf_{h\in I} \sup_{i\in I,\, i\geq h} r_i)$$

These substitutes for the limit always exist in $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$.

Whereas it is useful to evoke the use of limits, the two concepts of limit spaces and of convergence spaces are not of great use in the present state of analysis. So, in the next sections we turn to other means of dealing with limits.

Exercises

1. Given a net $(r_i)_{i\in I}$ of real numbers, show that there exist subnets $(s_j)_{j\in J}$ and $(t_k)_{k\in K}$ such that $\limsup_{i\in I} r_i = \lim_{j\in J} s_j$ and $\liminf_{i\in I} r_i = \lim_{k\in K} t_k$.

2. Given a net $(r_i)_{i\in I}$ of real numbers, show that for any subnet $(q_h)_{h\in H}$ of $(r_i)_{i\in I}$ that converges one has $\liminf_{i\in I} r_i \leq \lim_{h\in H} q_h \leq \limsup_{i\in I} r_i$.

3. Define pointwise convergence of nets for functions from a set S into a convergence space $(X, \to)$ and verify conditions (C1)–(C3). Do the same for uniform convergence of functions with values in $\mathbb{R}^d$.

4. Let X be the set of real-valued continuous functions on $\mathbb{R}^d$ that vanish outside a bounded subset. Declare that a net $(f_i) \to f$ if there exist a bounded subset B of $\mathbb{R}^d$ and some $\bar{i} \in I$ such that for all $i \geq \bar{i}$ in I the functions f_i and f are null on $\mathbb{R}^d \backslash B$ and if the net of restrictions to B uniformly converges to the restriction of f: $(f_i \mid B) \overset{u}{\to} f \mid B$. Verify conditions (C1)–(C3).

5. (**Sequences do not suffice**). Let S be an infinite uncountable set and let X be the set of real-valued functions on S equipped with pointwise convergence. Let Y be the subset of X formed by those $f \in X$ that are null off a finite subset. Show that Y is *dense* in X in the sense that every $f \in X$ is the limit of a net $(f_i)_{i\in I}$ of Y. Verify that if $f \in X$ is the limit of a sequence $(f_n)_{n\in\mathbb{N}}$ of Y, then the set $S_f := \{s \in S : f(s) \neq 0\}$ is countable. Deduce from this the fact that the *sequential closure* of Y, i.e. the set of limits of sequences in Y, is different from its closure X.

6. Using the notion of topology displayed in the next section, define a convergence $\to$ on a topological space $(X, \mathcal{O})$ by setting $(x_i)_{i\in I} \to x$ if for any $O \in \mathcal{O}$ there exists some $h \in I$ such that $x_i \in O$ for all $i \in I$ satisfying $i \geq h$.

7. Conversely, associate to any convergence space $(X, \to)$ a topology $\mathcal{O}$ by taking for $\mathcal{O}$ the set of subsets O of X such that for any $x \in O$ and any net $(x_i)_{i\in I} \to x$ one has $x_i \in O$ whenever $i \geq h$ for some $h \in I$. Verify the assumptions (O1), (O2) of the next definition. Prove that the convergence $\to$ is stronger than (i.e. implies) the convergence $\to_{\mathcal{O}}$ associated with the topology $\mathcal{O}$. Show that for any map $f : X \to Y$ with values in a topological space $(Y, \mathcal{O}_Y)$ the map f is continuous from $(X, \mathcal{O})$ into $(Y, \mathcal{O}_Y)$ if and only if f is continuous for the convergence $\to$ on X and the convergence associated with the topology of Y.

8*. Find an additional condition on a convergence in order that it is the convergence associated with a topology. [See [174].]

2.2 Topologies

The success of topology is due to two features: first, convergences are defined through an intuitive notion of neighborhoods for each point; second, the formalism and the rules of set theory can be used efficiently.

2.2.1 General Facts About Topologies

A topology on a set X is obtained by selecting a family of subsets called the family of closed subsets having some stability property. Equivalently, one usually introduces topologies by considering the family of complements of closed sets. These sets are called open sets.

Definition 2.3 A *topology* on a set X is the data comprising a family $\mathcal{O}$ of so-called *open* subsets that satisfies the following two requirements:

(O1) the union of any subfamily of $\mathcal{O}$ belongs to $\mathcal{O}$;
(O2) the intersection of any finite subfamily of $\mathcal{O}$ belongs to $\mathcal{O}$.

By convention, we admit that these two conditions include the requirements that X and the empty set $\varnothing$ belong to $\mathcal{O}$. A topological space $(X, \mathcal{O})$ is also denoted by X if the choice of the topology $\mathcal{O}$ is unambiguous. A subset F of X is declared to be *closed* if $X \backslash F$ belongs to $\mathcal{O}$.

Exercise Give conditions characterizing the family of closed (resp. open) subsets of a topological space in terms of nets for the convergence associated to $\mathcal{O}$ defined in Exercise 6 of the preceding section and reminded a few lines below.

A subset V of a topological space $(X, \mathcal{O})$ is a *neighborhood* of some $\overline{x} \in X$ if there exists some $U \in \mathcal{O}$ such that $\overline{x} \in U \subset V$. For $x \in X$ we denote by $\mathcal{N}(x)$ the family of neighborhoods of x. The topology $\mathcal{O}$ is determined by $(\mathcal{N}(x))_{x\in X}$, as shown by the exercises: O is open iff for all $x \in O$ one has $O \in \mathcal{N}(x)$.

To any topology on X one can associate a convergence $\to$ by setting:

$$((x_i)_{i\in I} \to x) \iff (\forall\ V \in \mathcal{N}(x)\ \exists\ i_V \in I\ :\ i \in I,\ i \geq i_V \Rightarrow x_i \in V)\,.$$

When a limit is unique one also writes $x = \lim_{i\in I} x_i$.

Exercise Verify that the conditions (C1), (C2), (C3) are satisfied. Moreover, a subset C of X is closed if and only if for any net $(x_i)_{i\in I}$ of C and any $x \in X$ satisfying $(x_i)_{i\in I} \to x$ one has $x \in C$.

Definition 2.4 A map $f : (X, \mathcal{O}) \to (X', \mathcal{O}')$ between two topological spaces is said to be *continuous* at $\overline{x} \in X$ if for any $V' \in \mathcal{N}(f(\overline{x}))$ there exists some $V \in \mathcal{N}(\overline{x})$ such that $f(V) \subset V'$. The map f is said to be continuous on some subset A of X if it is continuous at all $\overline{x} \in A$.

The definition of continuity of f at $\overline{x}$ is natural: in order that $f(x)$ be close enough to $f(\overline{x})$ it suffices to take x close enough to $\overline{x}$. However, one has to remember that in this condition, the neighborhood V' of $f(\overline{x})$ should be prescribed first. Continuity of f at $\overline{x}$ can be expressed by requiring that for any $V' \in \mathcal{N}(f(\overline{x}))$ one has $f^{-1}(V') \in \mathcal{N}(\overline{x})$.

Exercise Show that $f : X \to X'$ is continuous (on X) if and only if for all $O' \in \mathcal{O}'$ its inverse image $f^{-1}(O') := \{x \in X : f(x) \in O'\}$ belongs to $\mathcal{O}$.

The composition of two continuous maps is clearly continuous.

A bijection that is continuous and whose inverse is continuous is called a *homeomorphism*. Topology is the study of properties that are preserved under homeomorphisms.

A topology $\mathcal{O}'$ on X is said to be *weaker* than a topology $\mathcal{O}$ if the *identity map* $I_X : (X, \mathcal{O}) \to (X, \mathcal{O}')$ is continuous, i.e. if any member of $\mathcal{O}'$ is in $\mathcal{O}$, i.e. if $\mathcal{O}' \subset \mathcal{O}$. One also says that $\mathcal{O}$ is *finer* or stronger than $\mathcal{O}'$.

Example On a set X there is a topology that is weaker than any other one, the *rough topology*: its family of open sets is $\mathcal{O}_R := \{\varnothing, X\}$. There is a topology that is finer than any other one, the *discrete topology*, for which any subset is open: $\mathcal{O}_D := \mathcal{P}(X)$.

Given a family $\mathcal{G}$ of subsets of a set X, there is a topology $\mathcal{O}$ on X that is the weakest among those containing $\mathcal{G}$. It is obtained as the intersection of the family of topologies $\mathcal{O}_i$ satisfying $\mathcal{G} \subset \mathcal{O}_i$. Then one says that $\mathcal{G}$ *generates* $\mathcal{O}$. If $\mathcal{B} \subset \mathcal{O}$ is such that any element of $\mathcal{O}$ is a union of elements of $\mathcal{B}$, one says that $\mathcal{B}$ is a *base* of $\mathcal{O}$. It is easy to verify that when $\mathcal{G}$ generates $\mathcal{O}$, the family $\mathcal{B}$ of finite intersections of elements of $\mathcal{G}$ is a base of $\mathcal{O}$. A family $\mathcal{U}$ of subsets of X is a *base of neighborhoods* of $\overline{x}$ if $\mathcal{U}$ is contained in the family $\mathcal{N}(\overline{x})$ of neighborhoods of $\overline{x}$ and if for any $V \in \mathcal{N}(\overline{x})$ there exists some $U \in \mathcal{U}$ such that $U \subset V$. Given $\mathcal{B} \subset \mathcal{O}$, we see that $\mathcal{B}$ is a base of $\mathcal{O}$ if and only if for all $x \in X$, $\mathcal{B}(x) := \{U \in \mathcal{B} : x \in U\}$ is a base of neighborhoods of x.

The notion of continuity can be localized by using neighborhood bases. A map $f : (X, \mathcal{O}) \to (X', \mathcal{O}')$ is continuous at $\overline{x} \in X$ if for any neighborhood W in some neighborhood base of $f(\overline{x})$ in $(X', \mathcal{O}')$ there exists some $V \in \mathcal{N}(\overline{x})$ such that $f(V) \subset W$.

Given a set X, a family $(X_a, \mathcal{O}_a)_{a \in A}$ of topological spaces, and a family $(g_a)_{a \in A}$ of maps $g_a : X \to X_a$, among all the topologies on X for which all g_a are continuous, there is one that is weaker than any other. It is the topology $\mathcal{O}_X$ generated by the sets $g_a^{-1}(G_a)$ for $a \in A$ and $G_a \in \mathcal{O}_a$. It is easy to verify that a map $f : W \to X$ from a topological space $(W, \mathcal{O}_W)$ into X is continuous with respect to $\mathcal{O}_X$ if and only if $g_a \circ f : W \to X_a$ is continuous for all $a \in A$. When X is the product $\Pi_{a \in A} X_a$ and g_a is the canonical projection, the topology $\mathcal{O}_X$ is called the *product topology*. When X is a subset of a topological space $(Y, \mathcal{O}_Y)$ and one considers for the family $(g_a)_{a \in A}$ the sole canonical injection $j : X \to Y$ one says that $\mathcal{O}_X$ is the *induced topology*. Then $O \subset X$ belongs to $\mathcal{O}_X$ if and only if there exists some $G \in \mathcal{O}_Y$ such that $O = G \cap X$. It is easy to verify that the associated convergence to $\mathcal{O}_X$ is the induced convergence.

Besides the notion of limit associated with a topology $\mathcal{O}$ on X, one disposes of a weaker notion. One says that a net $(x_i)_{i\in I}$ of X has a *cluster point* $\overline{x} \in X$ if for any $V \in \mathcal{N}(\overline{x})$ and any $i \in I$ one can find some $j \in I$ such that $j \geq i$ and $x_j \in V$.

One can show that $\overline{x} \in X$ is a cluster point of $(x_i)_{i\in I}$ if and only if there exists a subnet of $(x_i)_{i\in I}$ that converges to $\overline{x}$. The if condition is immediate. For the necessary condition one can take $J := \{(i, V) \in I\times\mathcal{N}(\overline{x}) : x_i \in V\}$, a cofinal subset of $I\times\mathcal{N}(\overline{x})$ for the product order, and define $h : J \to I$ by $h(i, V) := i$.

A topology $\mathcal{O}$ on a set X is said to be *Hausdorff* if for every pair (x, x') of distinct points of X one can find neighborhoods $V \in \mathcal{N}(x)$, $V' \in \mathcal{N}(x')$ that are disjoint. This property is equivalent to uniqueness of limits of nets.

Proposition 2.1 *A topology $\mathcal{O}$ on a set X is* Hausdorff *if and only if all nets in X have at most one limit. In fact, if $\mathcal{O}$ is Hausdorff and if a net $(x_i)_{i\in I}$ of X has a limit $\overline{x}$, it cannot have a different cluster point.*

Proof Suppose $\mathcal{O}$ is Hausdorff and a net $(x_i)_{i\in I}$ of X has a limit $\overline{x}$ and a cluster point $\overline{y} \neq \overline{x}$. Let $V \in \mathcal{N}(\overline{x})$, $W \in \mathcal{N}(\overline{y})$ be such that $V \cap W = \varnothing$. By definition, there exists an $i_V \in I$ such that $x_i \in V$ for all $i \geq i_V$. Thus one cannot find some $j \geq i_V$ such that $x_j \in W$, contradicting the assumption that $\overline{y}$ is a cluster point of $(x_i)_{i\in I}$.

Now suppose $\mathcal{O}$ is not Hausdorff: there exists a pair $(\overline{x}, \overline{y})$ of distinct points of X such that for any $V \in \mathcal{N}(\overline{x})$, $W \in \mathcal{N}(\overline{y})$ one has $V\cap W \neq \varnothing$. Denoting $\mathcal{N}(\overline{x})\times\mathcal{N}(\overline{y})$ by I and giving to I the order opposite to inclusion, for $i := (V, W) \in I$ one can pick $x_i \in V \cap W$. Then the net $(x_i)_{i\in I}$ converges to $\overline{x}$ and to $\overline{y}$. □

Corollary 2.1 *In a Hausdorff topological space $(X, \mathcal{O})$ finite subsets, in particular singletons, are closed.*

Proof It suffices to show that a singleton $S := \{\overline{x}\}$ is closed. If $\overline{y} \in X\setminus S$ one can find $V \in \mathcal{N}(\overline{x})$ and $W \in \mathcal{N}(\overline{y})$ that are disjoint. Then W is contained in $X\setminus S$. This proves that $X\setminus S$ is open and S is closed. □

A topology $\mathcal{O}$ on X is uniquely determined by its associated convergence for nets: a subset C of X is closed if and only if it contains the limits of its convergent nets. In general $\mathcal{O}$ is not determined by the convergence of sequences. The following proposition shows that nets may be convenient. It also shows that continuity in topological spaces coincides with continuity for the induced convergence spaces.

Proposition 2.2 *A map $f : X \to Y$ between two topological spaces is continuous at $\overline{x} \in X$ if and only if for any net $(x_i)_{i\in I}$ of X converging to $\overline{x}$, the net $(f(x_i))_{i\in I}$ converges to $f(\overline{x})$.*

When $\mathcal{N}(\overline{x})$ has a countable base sequences can be used instead of nets.

Proof Necessity is immediate. Let us show sufficiency. Suppose f is not continuous at $\overline{x}$. Then, there exists a $V \in \mathcal{N}(f(\overline{x}))$ such that for all $U \in \mathcal{N}(\overline{x})$ there exists some $x_U \in U$ with $f(x_U) \notin V$. Then, the net $(x_U)_{U\in\mathcal{N}(\overline{x})} \to \overline{x}$ but $(f(x_U))_{U\in\mathcal{N}(\overline{x})}$ does not converge to $f(\overline{x})$. □

The *closure* $\mathrm{cl}(S)$ of a subset S of a topological space $(X, \mathcal{O})$ is the intersection of the family of all closed subsets of X containing S. It is clearly the smallest closed subset of $(X, \mathcal{O})$ containing S. The *interior* $\mathrm{int}(S)$ of S is the union of all the open subsets of $(X, \mathcal{O})$ contained in S. It is the largest open subset of X contained in S. Thus $\mathrm{int}(S) = X \backslash \mathrm{cl}(X \backslash S)$. The *boundary* or frontier of S is $\mathrm{bdry}(S) := \mathrm{cl}(S) \backslash \mathrm{int}(S)$.

Exercise Prove that the closure $\mathrm{cl}(S)$ of a subset S of a topological space $(X, \mathcal{O})$ is the set of limits of nets of S that converge in X.

Proposition 2.3 *The closure* $\mathrm{cl}(S)$ *of a subset S of a topological space* $(X, \mathcal{O})$ *is the set of points* $x \in X$ *such that for any neighborhood V of x one has* $S \cap V \neq \varnothing$.

Proof If $\overline{x} \in X \backslash \mathrm{cl}(S)$ there exists some closed set C containing S such that $\overline{x} \in X \backslash C$. Then $V := X \backslash C$ is an open neighborhood of $\overline{x}$ and $S \cap V = \varnothing$. Conversely, if for some $V \in \mathcal{N}(\overline{x})$ one has $S \cap V = \varnothing$, taking some open subset U satisfying $\overline{x} \in U \subset V$ we see that $C := X \backslash U$ is a closed subset containing S (since $U \subset V$ and $S \cap V = \varnothing$) and $\overline{x} \notin C$, hence $\overline{x} \notin \mathrm{cl}(S)$. □

Corollary 2.2 *A point* $\overline{x}$ *of a topological space* $(X, \mathcal{O})$ *is a cluster point of a net* $(x_i)_{i \in I}$ *of X if* $\overline{x} \in \cap_{i \in I} C_i$*, where* $C_i := \mathrm{cl}(\{x_j : j \in I,\ j \geq i\})$.

Proof This follows from the fact that $\overline{x}$ is a cluster point of $(x_i)_{i \in I}$ if and only if for all $i \in I$ and all $V \in \mathcal{N}(\overline{x})$ one has $V \cap \{x_j : j \in I,\ j \geq i\} \neq \varnothing$. □

A subset D of $(X, \mathcal{O})$ is said to be *dense* in a subset E of X if $D \subset E$ and if E is contained in the closure of D. A topological space is said to be *separable* if it contains a countable dense subset.

Example Given a directed set $(I, \leq)$, let $I_\infty := I \cup \{\infty\}$, where ∞ is an additional element satisfying $i \leq \infty$ for all $i \in I$. Then one can endow I_∞ with the topology $\mathcal{O}$ defined by $G \in \mathcal{O}$ if either G is contained in I, else if there exists some $h \in I$ such that $i \in G$ for all $i \in I_\infty$ such that $i \geq h$. Thus I is dense in I_∞. Given a topological space $(X, \mathcal{O})$, $x \in X$ and a net $(x_i)_{i \in I}$ of X, one easily checks that $(x_i)_{i \in I} \to x$ if and only if the map $f : I_\infty \to X$ given by $f(i) := x_i, f(\infty) := x$ is continuous at ∞.

Definition 2.5 Given two topological spaces $(W, \mathcal{O})$, $(X', \mathcal{O}')$, a subset X of W and $w \in \mathrm{cl}(X)$, one says that $f : X \to X'$ has a limit $\overline{x}'$ as $x \to_X w$ (i.e. $x \to w$ with $x \in X$), or that f converges to $\overline{x}'$ as $x \to_X w$ and one writes $\overline{x}' = \lim_{x \to_X w} f(x)$, if for any $V' \in \mathcal{N}(\overline{x}')$ there exists some $V \in \mathcal{N}(w)$ such that $f(V \cap X) \subset V'$.

Thus f converges to $\overline{x}'$ as $x \to_X w$ iff for any net $(x_i)_{i \in I}$ of X satisfying $(x_i)_{i \in I} \to w$ in W one has $(f(x_i))_{i \in I} \to \overline{x}'$ in X'. If $X = W$, one just writes $\overline{x}' = \lim_{x \to w} f(x)$. Thus, f is continuous at w if and only if f has the limit $f(w)$ as $x \to w$. We invite the reader to verify that the notion $(x_n) \to 0_+$ in $\mathbb{R}$ (i.e. $(x_n) \to 0$ and $x_n > 0$ for all $n \in \mathbb{N}$) corresponds to the case when $W := \mathbb{R}$, $X := \mathbb{P}$, the set of positive real numbers.

The preceding definition is a special case of a more general concept. Given another map $g : X \to Y$ with values in another topological space $(Y, \mathcal{G})$ and some $\overline{y} \in Y$, one says that f has a limit $\overline{x}'$ as $g(x) \to \overline{y}$, or that f converges to $\overline{x}'$ as $g(x) \to \overline{y}$ and one writes $\overline{x}' = \lim_{g(x) \to \overline{y}} f(x)$, if for any $V' \in \mathcal{N}(\overline{x}')$ there exists a $W \in \mathcal{N}(\overline{y})$

such that $f(x) \in V'$ for all $x \in g^{-1}(W)$. Taking for g the canonical injection of X into $(Y, \mathcal{G}) := (W, \mathcal{O})$, one recovers the preceding notion of limit.

The next result is often used for uniqueness purposes.

Proposition 2.4 *Let $(W, \mathcal{O})$, $(X', \mathcal{O}')$ be two topological spaces, let X be a dense subset of W and let f, $g : W \to X'$ be two continuous maps. If the restrictions of f and g to X coincide, then f and g coincide.*

Proof The set $Z := \{w \in W : f(w) = g(w)\}$ is a closed subset of W containing X. Since X is dense in W, we have $Z = W$ since $W = \mathrm{cl}(X) \subset Z$. □

Exercises

1. Let $f : \mathbb{R} \to \mathbb{R}$ be a map such that $f(r + s) = f(r) + f(s)$ for all $r, s \in \mathbb{R}$. Show that $f(q) = qf(1)$ for all $q \in \mathbb{Q}$. Prove that f is linear over $\mathbb{R}$ when moreover f is continuous or monotone.
2. Write the alphabet with capital letters and decide which letters are mutually homeomorphic.
3. Let $(X, \mathcal{O})$ be a Hausdorff topological space and let $f : X \to X$ be a continuous map. Show that the set $F := \{x \in X : f(x) = x\}$ is closed in X.
4. Let $(X, \mathcal{O}_X)$, $(Y, \mathcal{O}_X)$ be topological spaces and let $f : X \to Y$ be a continuous map. Show that the set $G := \{(x, y) \in X \times Y : y = f(x)\}$ is homeomrphic to X.
5. Show that a topological space $(X, \mathcal{O})$ is Hausdorff if and only if the diagonal $\Delta_X := \{(x, x') \in X^2 : x = x'\}$ is closed in X^2.
6. Let X and Y be topological spaces, let $f : X \to Y$ and let $(X_i)_{i \in I}$ be a covering of X, i.e. a family of subsets of X whose union is X. Suppose the restriction f_i of f to X_i is continuous. Show that if every X_i is open, then f is continuous.
7. With the notation of the preceding exercise, suppose that I is finite and that every X_i is closed. Show that f is continuous if every f_i is continuous. Give an example showing that the assumption that I is finite cannot be dropped. Give an example showing that the assumption that every X_i is closed cannot be dropped.
8. Let X and Y be topological spaces, let $s \in X$, and let $f : X \times Y \to \mathbb{R}$ be separately continuous (i.e. f is continuous in each of its two variables). For a subset T of Y let
$$V_s(T) := \{x \in X : f(x, s) \leq f(x, t)\ \forall t \in T\}$$
with $V_s(\varnothing) := X$ be the Voronoi cell associated with s and T as in Exercise 17 of Sect. 1.1. Show that $V_s(T)$ is closed and that $V_s(\mathrm{cl}(T)) = V_s(T)$ for each $T \in \mathcal{P}(Y)$.

2.2.2 Connectedness

It is sometimes useful to know that a space is not made of several pieces: this can be used to globalize some properties or for uniqueness results. For example, if X is an open subset of $\mathbb{R}$ and if $f : X \to \mathbb{R}$ is a differentiable function whose derivative is 0 we cannot conclude that f is constant because X can be the union of disjoint open intervals. We must give a precise definition.

Definition 2.6 A topological space $(X, \mathcal{O})$ is said to be *connected* if $\varnothing$ and X are the only subsets of X that are both open and closed.

The space $(X, \mathcal{O})$ is *arcwise connected* if any two points x_0, x_1 of X can be joined by a continuous arc, i.e. if for any x_0, x_1 there is a continuous map $c : [0, 1] \to X$ such that $c(0) = x_0, c(1) = x_1$.

Connectedness is a more important notion than arcwise connectedness, but the latter is more intuitive and the proofs of several propositions below for arcwise connectedness are very easy. The reader is invited to verify this assertion. The definition of connectedness can be rephrased by saying that X is connected if any partition of X into two open subsets is improper, i.e. if one of the subsets is empty and the other one is the whole space, a *partition* of X being a covering by disjoint subsets. The reader must be aware that there are topological spaces that are not connected; in such a space, showing that a subset is not closed does not prove that it is open (a frequent mistake). For instance, if X is the union of two open disjoint intervals A, B of $\mathbb{R}$, A and B are also closed in X since their complements are open.

Example A bounded, closed interval $X := [a, b]$ of $\mathbb{R}$ is connected (with respect to the induced topology). To prove this, let us consider a nonempty subset C of X that is both open and closed and let us show that $C = X$. We may suppose $a \in C$ (otherwise we consider $C' := X \backslash C$). Let $s := \sup T$ with $T := \{t \in X : [a, t] \subset C\}$. Since C is open in X we have $s > a$. Moreover, $s \in T$ since there exists a sequence (t_n) of $T \subset C$ converging to s so that $[a, s[= \cup_n [a, t_n[\subset C$ and since C is closed we have $[a, s] \subset C$. Let us show that assuming $s < b$ leads to a contradiction. Since C is open in X and $s \in C$, we can find $\varepsilon > 0$ such that $s + \varepsilon \leq b$ and $[s, s + \varepsilon] \subset C$. Then since $s \in T$ we get $[a, s + \varepsilon] = [a, s] \cup [s, s + \varepsilon] \subset C$; this means that $s + \varepsilon \in T$, contradicting the definition of s.

It follows from Proposition 2.7 below that any interval of $\mathbb{R}$ is connected. $\square$

The use of connectedness for existence results is illustrated by the following properties. The second one is often called the Intermediate Value Theorem.

Proposition 2.5 (Customs Lemma) *Let C be a connected subset of a topological space $(X, \mathcal{O})$ and let S be a subset of X. If $C \cap \mathrm{int}(S)$ and $C \cap (X \backslash \mathrm{cl}(S))$ are nonempty, then C contains some point of the boundary of S (Fig. 2.2).*

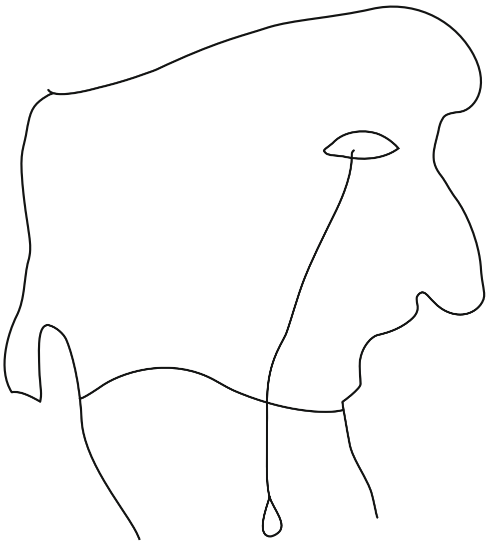

Fig. 2.2 The customs lemma

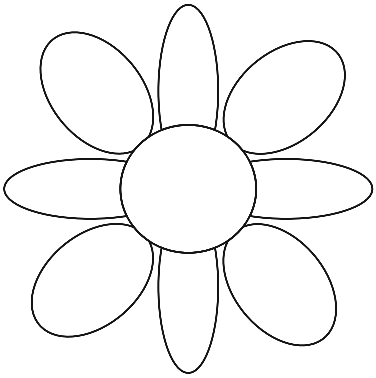

Fig. 2.3 The Daisy property

Proof If, on the contrary, C does not meet the boundary of S then C is the union of the two disjoint sets $C \cap \text{int}(S)$ and $C \cap (X \backslash \text{cl}(S))$, contradicting the connectedness of C. $\square$

Proposition 2.6 (Bolzano) *Let $(X, \mathcal{O})$ be a connected topological space and let $f : X \to \mathbb{R}$ be a continuous map. Let $a < b < c$ in $\mathbb{R}$ be such that $a \in f(X)$, $c \in f(X)$. Then there exists some $x \in X$ such that $f(x) = b$.*

Proof If $f^{-1}(b) = \varnothing$, the sets $f^{-1}(]-\infty, b[)$ and $f^{-1}(]b, +\infty[)$ form a partition of X into two open subsets, an impossibility if X is connected. $\square$

The following property can be used as a convenient criterion (Fig. 2.3).

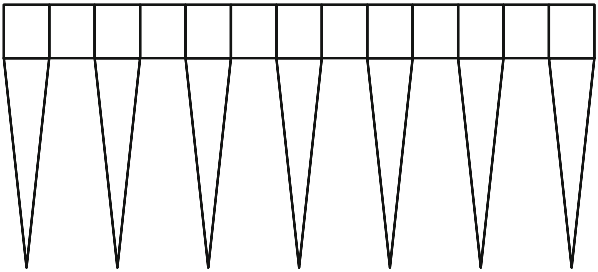

Fig. 2.4 The Comb or Rake property

Proposition 2.7 (The Daisy Property) *Let $(X, \mathcal{O})$ be a topological space that is the union of a family $(X_i)_{i\in I}$ of connected subspaces. If $\cap_{i\in I} X_i$ is nonempty, then X is connected.*

Proof Let $\overline{x} \in \cap_{i\in I} X_i$. Let C be a subset of X that is closed and open. Changing C into $X\backslash C$ we may suppose $\overline{x} \in C$. Then $C \cap X_i$ is open and closed in X_i and contains $\overline{x}$. Since X_i is connected, we have $C \cap X_i = X_i$. Thus $C \supseteq X_i$ for all $i \in I$, hence $C = X$. □

A slight refinement can be given (Fig. 2.4).

Corollary 2.3 (The Comb or Rake Property) *Let $(X, \mathcal{O})$ be a topological space that is the union of a family $(X_i)_{i\in I}$ of connected subspaces. If there is some nonempty connected subspace Y of X such that $X_i \cap Y \neq \varnothing$ for all $i \in I$, X is connected.*

Proof Set $Y_i := X_i \cup Y$. Then, by the preceding proposition, Y_i is connected. Since $X = \cup_{i\in I} Y_i$ and $\cap_{i\in I} Y_i$ contains $Y \neq \varnothing$, then X is connected. □

Proposition 2.8 *A product of two connected spaces is connected.*

Proof Let $Z := X \times Y$, the spaces X, Y being connected and nonempty. Let $x_0 \in X$. Then $Y_0 := \{x_0\} \times Y$ is homeomorphic to Y, hence is connected and for all $y \in Y$ the subspace $X_y := X \times \{y\}$ is connected and meets Y_0 as $(x_0, y) \in X_y \cap Y_0$. Since $Z = \cup_{y\in Y} X_y$, Z is connected by the rake property (2.3).

Now let us consider the general case of an arbitrary product $X := \prod_{i\in I} X_i$ of connected spaces. Given two nonempty open subsets, A, B of X satisfying $A \cup B = X$, the construction of the product topology ensures that there exist a finite subset J of I and open subsets A_J, B_J of $X_J := \prod_{j\in J} X_j$ such that $A = A_J \times X_{I\backslash J}$ and $B = B_J \times X_{I\backslash J}$. Then $A_J \cup B_J = X_J$. Since X_J is connected in view of the first part of the proof and of an induction, we have $A_J \cap B_J \neq \varnothing$. Thus $A \cap B \neq \varnothing$ and X is connected. □

The preceding proof used the obvious fact that when two topological spaces are homeomorphic, both are connected when one of them is connected. This fact is also a consequence in a more general property that is obviously valid for arcwise connectedness.

Proposition 2.9 *Let $(X, \mathcal{O}_X)$, $(Y, \mathcal{O}_Y)$ be two topological spaces and let $f : X \to Y$ be continuous. If X is connected, then $f(X)$ is connected with respect to the induced topology.*

Proof There is no loss of generality in assuming $f(X) = Y$. Then, if $G \in \mathcal{O}_Y$ is nonempty, open and closed, then so is $f^{-1}(G)$. Since X is connected we have $f^{-1}(G) = X$, hence $G = f(f^{-1}(G)) = f(X) = Y$. □

Corollary 2.4 *An arcwise connected space is connected.*

Proof Let $(X, \mathcal{O})$ be an arcwise connected space and let $x_0 \in X$. By definition, for all $x \in X$ there exists a continuous map $f_x : [0, 1] \to X$ such that $f_x(0) = x_0$ and $f_x(1) = x_1$. Since $C_x := f_x([0, 1])$ is connected and since $X = \cup_{x \in X} C_x$ with $x_0 \in C_x$ for all x, X is connected. □

Given a topological space $(X, \mathcal{O})$ and $x \in X$, Proposition 2.7 implies that the union $C(x)$ of all connected subsets of X containing x is connected. It is clearly the largest connected subset containing x. Moreover, X can be split into a partition of connected subsets called the *connected components* of X by taking those $C(x)$ that are disjoint (note that if $C(x) \cap C(x') \neq \varnothing$ then $C(x) = C(x')$). It follows from Exercise 1 below that the connected components of X are closed subsets.

A topological space $(X, \mathcal{O})$ is said to be *locally connected* if every point of X has a base of neighborhoods formed by connected sets. Clearly $\mathbb{R}$ is locally connected, but $\mathbb{Q}$ and $\mathbb{R} \backslash \mathbb{Q}$ are not locally connected. It is easy to show that a space is locally connected if and only if the connected components of any open subset are open.

Exercises

1. Let A and B be two subsets of a topological space $(X, \mathcal{O})$ such that $A \subset B \subset \mathrm{cl}(A)$. Show that B is connected whenever A is connected. Deduce from this that $\mathrm{cl}(A)$ is connected and that any connected component of X is closed.
2. Let $(X, \mathcal{O})$ be a topological space such that X is the union of a sequence $(X_n)_n$ of connected subsets satisfying $X_n \cap X_{n+1} \neq \varnothing$ for all $n \in \mathbb{N}$. Show that X is connected.
3. Show that the connected subsets of $\mathbb{R}$ are the intervals.
4. Prove that any open subset of $\mathbb{R}$ is the union of a finite or countable family of disjoint open intervals.
5. Verify that a topological space $(X, \mathcal{O})$ is connected if and only if any continuous map $f : X \to \mathbb{Z}$ is constant.
6. Let A and B be two nonempty closed subsets of a topological space. Show that if $A \cap B$ and $A \cup B$ are connected, then A and B are connected. Show by an example of two subsets of $\mathbb{R}$ that the assumption that A and B are closed cannot be omitted.
7. Let $G := \{(r, \sin(1/r)) : r \in]0, 1]\}$ and let X be its closure in $\mathbb{R}^2$. Prove that X is connected but that X is not arcwise connected and not locally connected.

2.2.3 Lower Semicontinuity

In order to deal with minimization problems, one may use a one-sided weakening of continuity when a continuity assumption is not realistic. A precise definition is as follows.

Definition 2.7 A function $f : X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ on a topological space X is said to be *lower semicontinuous* (l.s.c.) at some $\overline{x} \in X$ if for every real number $r < f(\overline{x})$ there exists some member V of the family $\mathcal{N}(\overline{x})$ of neighborhoods of $\overline{x}$ such that $r < f(v)$ for all $v \in V$. A function f is upper semicontinuous (u.s.c.) at $\overline{x}$ whenever $-f$ is l.s.c. at $\overline{x}$.

The function f is said to be lower semicontinuous (l.s.c.) on some subset S of X if f is lower semicontinuous at each point of S.

We observe f is automatically l.s.c. at $\overline{x}$ when $f(\overline{x}) = -\infty$; when $f(\overline{x}) = +\infty$ the lower semicontinuity of f means that the values of f can be as large as required provided one remains in a small enough neighborhood of $\overline{x}$. When $f(\overline{x})$ is finite the definition amounts to assigning to any $\varepsilon > 0$ a neighborhood V_ε of $\overline{x}$ such that $f(v) > f(\overline{x}) - \varepsilon$ for each $v \in V_\varepsilon$. Thus, lower semicontinuity allows sudden upward changes of the value of f but excludes sudden downward changes. Obviously, f is continuous at $\overline{x}$ iff it is both l.s.c. and u.s.c. at $\overline{x}$.

Example The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = 1$ for $x < 0$, $f(x) = 0$ for $x \in \mathbb{R}_+$ is l.s.c. but not continuous at 0.

Example The *indicator function* ι_A of a subset A of X, defined by $\iota_A(x) = 0$ for $x \in A$, $\iota_A(x) = +\infty$ for $x \in X \backslash A$ is l.s.c. if and only if, A is closed, as is easily seen. Such a function is of great use in optimization theory and nonsmooth analysis.

Example The *characteristic function* 1_A of a subset A of X, defined by $1_A(x) = 1$ for $x \in A$, $1_A(x) = 0$ for $x \in X \backslash A$ is l.s.c. if and only if, A is open. Such a function is of primary importance in integration theory.

Example *The Length Function* Given a metric space (M, d), let $X := C(T, M)$ be the space of continuous maps (curves or arcs of M) from $T := [0, 1]$ to M. Given a subdivision $\sigma := \{t_0 = 0 < t_1 < \ldots < t_n = 1\}$ of T, let us set for $x \in X$

$$\ell_\sigma(x) := \sum_{i=1}^{n} d(x(t_{i-1}), x(t_i)),$$

and let $\ell(x)$ be the supremum of $\ell_\sigma(x)$ as σ varies in the set of finite subdivisions of T. The properties devised below yield that ℓ is l.s.c. when X is endowed with the topology of uniform convergence (and even when X is endowed with the topology of pointwise convergence). However ℓ is not continuous: one can increase ℓ by following a nearby curve which makes many small changes (a fact any dog knows, when tied with a lease). Details are given in Exercise 3 below.

The following characterizations are global ones.

Proposition 2.10 *For a function $f : X \to \overline{\mathbb{R}}$ the following assertions are equivalent:*

(a) f is l.s.c.;
(b) the epigraph $E := \{(x,r) \in X \times \mathbb{R} : r \geq f(x)\}$ of f is closed;
(c) for each $s \in \mathbb{R}$ the sublevel set $S(s) := \{x \in X : f(x) \leq s\}$ is closed.

Proof (a)⇒(b) It suffices to prove that $(X \times \mathbb{R}) \setminus E$ is open when f is l.s.c. Given $(\overline{x}, \overline{r}) \in (X \times \mathbb{R}) \setminus E$, i.e. such that $\overline{r} < f(\overline{x})$, for any $r \in]\overline{r}, f(\overline{x})[$ one can find a neighborhood V of $\overline{x}$ such that $r < f(v)$ for all $v \in V$. Then $V \times]-\infty, r[$ is a neighborhood of $(\overline{x}, \overline{r})$ in $X \times \mathbb{R}$ which does not meet E. Hence $(X \times \mathbb{R}) \setminus E$ is open.

(b)⇒(c) It suffices to observe that for each $s \in \mathbb{R}$ one has $S(s) \times \{s\} = E \cap (X \times \{s\})$.

(c)⇒(a) Given $\overline{x} \in X$ and $r \in \mathbb{R}$ such that $r < f(\overline{x})$ one has $\overline{x} \in X \setminus S(r)$ which is open and for all $v \in V := X \setminus S(r)$ one has $r < f(v)$. □

The notion of lower semicontinuity is intimately tied to the concept of lower limit (denoted by liminf), which is a one-sided concept of limit which can be used even when there is no limit. In the following definition we assume X is a subspace of a larger space W, $w \in \mathrm{cl}(X)$ (a situation which will be encountered later, for instance when $X = \mathbb{P} :=]0, \infty[$, $W = \mathbb{R}$ and $w = 0$) and we denote by $\mathcal{N}(w)$ the family of neighborhoods of the point w in W.

Definition 2.8 Given a topological space W, a subspace X and a point w in the closure of X, the lower limit of a function $f : X \to \overline{\mathbb{R}}$ at w is the extended real number

$$\liminf_{x \to_X w} f(x) := \sup_{V \in \mathcal{N}(w)} \inf_{v \in V \cap X} f(v).$$

Setting $m_V := \inf f(V \cap X)$, the supremum over $V \in \mathcal{N}(w)$ of the family $(m_V)_V$ can also be considered as the limit of the net $(m_V)_V$; this explains the terminology. One can show that $\sup_V m_V$ is also the least cluster point of $f(x)$ as $x \to w$ in X. When W is metrizable, one can replace the family $\mathcal{N}(w)$ by the family of balls centered at w, so that $\liminf_{x \to w} f(x) = \sup_{r>0} m_r$, with $m_r := \inf f(B(w,r) \cap X)$, is the limit of a sequence.

Exercise Deduce from the preceding that the lower limit of a function $f : X \to \overline{\mathbb{R}}$ at w is the least of the limits of the converging nets $(f(x_i))_{i \in I}$ where $(x_i)_{i \in I}$ is a net in X converging to w.

Lower semicontinuity can be characterized using the notion of lower limit (here $W = X$).

Lemma 2.1 *A function $f : X \to \overline{\mathbb{R}}$ on a topological space X is l.s.c. at some $w \in X$ iff one has $f(w) \leq \liminf_{x \to w} f(x)$.*

Proof Clearly, when f is l.s.c. at w, one has $f(w) \leq \liminf_{x \to w} f(x)$. Conversely, when this inequality holds, for any $r < f(w)$, by the definition of the supremum over $\mathcal{N}(w)$, one can find $V \in \mathcal{N}(w)$ such that $r < \inf_{v \in V} f(v)$, so that f is l.s.c. at w. □

One can also use nets for such a characterization.

Lemma 2.2 *A function $f : X \to \overline{\mathbb{R}}$ on a topological space X is l.s.c. at some $\overline{x} \in X$ iff for any net $(x_i)_{i \in I}$ in X converging to $\overline{x}$ one has $f(\overline{x}) \leq \liminf_{i \in I} f(x_i)$. When $\overline{x}$ has a countable base of neighborhoods, one can replace nets by sequences in that characterization.*

Proof The condition is necessary: if a net $(x_i)_{i \in I}$ in X converges to $\overline{x}$, for any $r < f(\overline{x})$ there exists some $V \in \mathcal{N}(\overline{x})$ such that $f(v) > r$ for all $v \in V$, and there exists some $h \in I$ such that $x_i \in V$ for $i \geq h$, so that $\inf_{i \geq h} f(x_i) \geq r$, hence $\liminf_{i \in I} f(x_i) \geq \inf_{i \geq h} f(x_i) \geq r$.

Conversely, suppose f is not l.s.c. at $\overline{x}$ and let $(V_i)_{i \in I}$ be a base of neighborhoods of $\overline{x}$: there exists some $r < f(\overline{x})$ such that for any $i \in I$ there exists some $x_i \in V_i$ such that $f(x_i) < r$. Ordering I by $j \geq i$ if $V_j \subset V_i$, we get a net $(x_i)_{i \in I}$ which converges to $\overline{x}$ and is such that $\liminf_{i \in I} f(x_i) \leq r$.

The second assertion follows from the fact that when $\overline{x}$ has a countable base of neighborhoods, one can take a decreasing sequence of neighborhoods for a base. □

Let us give some useful calculus rules (with the convention $0r = 0$ for all $r \in \overline{\mathbb{R}}$).

Exercise For any $\alpha \in \mathbb{R}_+$ and $f : X \to \overline{\mathbb{R}}$ one has $\liminf_{x \to \overline{x}} \alpha f(x) = \alpha \liminf_{x \to \overline{x}} f(x)$.

Exercise If $f, g : X \to \overline{\mathbb{R}}$ are such that $\{\liminf_{x \to \overline{x}} f(x), \liminf_{x \to \overline{x}} g(x)\} \neq \{-\infty, +\infty\}$, then

$$\liminf_{x \to \overline{x}} (f + g)(x) \geq \liminf_{x \to \overline{x}} f(x) + \liminf_{x \to \overline{x}} g(x).$$

The family of lower semicontinuous functions enjoys stability properties.

Proposition 2.11 *If $(f_i)_{i \in I}$ is a family of functions which are l.s.c. at $\overline{x}$, then the function $f := \sup_{i \in I} f_i$ is l.s.c. at $\overline{x}$.*

For any $\alpha \in \mathbb{R}_+$ and f, g which are l.s.c. at $\overline{x}$, the functions $\inf(f, g)$ and αf are l.s.c. at $\overline{x}$; the same is true for $f + g$ provided $\{f(\overline{x}), g(\overline{x})\} \neq \{-\infty, +\infty\}$. If moreover f and g are nonnegative, then fg is l.s.c. at $\overline{x}$.

If $f : X \to \overline{\mathbb{R}}$ is finite and l.s.c. at $\overline{x}$ and if $g : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ is nondecreasing and l.s.c. at $f(\overline{x})$, then $g \circ f$ is l.s.c. at $\overline{x}$.

One may observe that in the first assertion one cannot replace lower semicontinuity by continuity, as shown by the above example of arc length.

Proof Let $r \in \mathbb{R}$ be such that $r < f(\overline{x})$. There exists some $j \in I$ such that $r < f_j(\overline{x})$, hence one can find some $V \in \mathcal{N}(\overline{x})$ such that $r < f_j(v) \leq f(v)$ for all $v \in V$. The proofs of the other assertions are also straightforward or follow from the preceding lemma. □

Proposition 2.12 *For any function $f : X \to \overline{\mathbb{R}}$ on a topological space X, the family of l.s.c. functions majorized by f has a greatest element $\bar{f}$ called the lower semicontinuous hull of f (in short, the l.s.c. hull of f). Its epigraph is the intersection with $X \times \mathbb{R}$ of the closure* $\mathrm{cl}(\mathrm{epi} f)$ *of the epigraph of f in $X \times \overline{\mathbb{R}}$. The function $\bar{f}$ is given by*

$$\bar{f}(x) = \liminf_{u \to x} f(u) = \sup_{U \in \mathcal{N}(x)} \inf_{u \in U} f(u).$$

Proof The first assertion is a direct consequence in Proposition 2.11. The second assertion easily stems from the fact that setting $g(x) = \min\{r : (x, r) \in \mathrm{cl}(\mathrm{epi} f)\}$ one defines a lower semicontinuous function that is the greatest lower semicontinuous function majorized by f. The proof of the explicit expression of $\bar{f}$ is left as an exercise. □

The treatment of the following example requires the knowledge of some material in integration theory to be found in Sects. 7.4 and 8.5. Its importance justifies its presence here.

Example Let E be a Banach space, let $(S, \mathcal{S}, \mu)$ be a measure space and let $L : E \times S \to \mathbb{R}$ be a measurable function such that for a null set N in S the function $L_s : e \mapsto L(e, s)$ is lower semicontinuous whenever $s \in S \backslash N$. Suppose that for some $p \in [1, \infty[$ and some $a \in \mathbb{R}$, $b \in \mathcal{L}_1(S)$ one has

$$L(e, s) \geq b(s) - a \|e\|^p \qquad \forall (e, s) \in E \times (S \backslash N).$$

Then the function $j : L_p(S, E) \to \overline{\mathbb{R}}$ given by $j(x) := \int_S L(x(s), s) d\mu(s)$ is lower semicontinuous.

We first prove this assertion in the case $a = 0$, $b = 0$. Let (x_n) be a sequence in $L_p(S, E)$ converging to some $x \in L_p(S, E)$. Taking a subsequence if necessary we may suppose $(j(x_n))$ converges to $\liminf_n j(x_n)$. Taking a further subsequence we may assume that $(x_n(s)) \to x(s)$ a.e. in S. Our lower semicontinuity assumption on L ensures that

$$\liminf_n L(x_n(s), s) \geq L(x(s), s) \qquad \text{a.e.},$$

so that, by Fatou's lemma, we obtain

$$\liminf_n \int_S L(x_n(s), s) d\mu(s) \geq \int_S \liminf_n L(x_n(s), s) d\mu(s) \geq \int_S L(x(s), s) d\mu(s),$$

or $\liminf_n j(x_n) \geq j(x)$, the required lower semicontinuity property.

In the general case we set $M(e, s) := L(e, s) - b(s) + a \|e\|^p \in \mathbb{R}_+$ and $m(x) := \int_S M(x(s), s) d\mu(s)$ for $x \in L_p(S, E)$. Given $x \in L_p(S, E)$ and a sequence $(x_n) \to x$ in

$L_p(S,E)$, we have

$$\liminf_n j(x_n) = \liminf_n (m(x_n) - a \int_S \|x_n(s)\|^p \, d\mu(s) + \int_S b(s) d\mu(s)$$
$$\geq m(x) - a \|x\|_p^p + \int_S b(s) d\mu(s) = j(x).$$

That shows that j is lower semicontinuous. □

Exercises

1. Using the relation $E = \bigcap_{i \in I} E_i$, where E_i is the epigraph of a function f_i and E is the epigraph of $f := \sup_{i \in I} f_i$, show that f is l.s.c. on X when each f_i is l.s.c. on X. Use a similar argument with sublevel sets.

2. Suppose X is a metric space. Show that f is l.s.c. at $\overline{x}$ iff for any sequence (x_n) converging to $\overline{x}$ one has $f(\overline{x}) \leq \liminf_n f(x_n)$.

3. Let (M,d) be a metric space and let $X := C(T,M)$, where $T := [0,1]$. Given some $x \in X$ and some element s of the set S of nondecreasing sequences $s := (s_n)_{n \geq 0}$ satisfying $s_0 = 0$, $s_n = 1$ for n large, let

$$\ell_s(x) := \sum_{n \geq 0} d(x(s_n), x(s_{n+1}))$$

(observe that the preceding sum contains only a finite number of non-zero terms). Define the length of a curve $x \in X$ by $\ell(x) := \sup_{s \in S} \ell_s(x)$. Show that $\ell_s : X \to \mathbb{R}$ is continuous when X is endowed with the metric of uniform convergence (and even when X is provided with the topology of pointwise convergence). Conclude that the length ℓ is a l.s.c. function on X.

Show that ℓ is not continuous by taking $M := \mathbb{R}^2$, $\overline{x}$ given by $\overline{x}(t) := (t,0)$ and by showing that there is some $x_n \in X$ such that $d(x_n, \overline{x}) \to 0$ and $\ell(x_n) \geq \sqrt{2}$ [Hint: for $n > 0$ define $x_n(t) = t - \frac{2k}{2n}$ for $t \in [\frac{2k}{2n}, \frac{2k+1}{2n}]$, $k \leq n$ and $x_n(t) = -t + \frac{2k+2}{2n}$ for $t \in [\frac{2k+1}{2n}, \frac{2k+2}{2n}]$, $k \leq n-1$].

4. Show that the infimum of an infinite family of l.s.c. functions is not necessarily l.s.c. [Hint: observe that any function f on X is the infimum of the family $(f_a)_{a \in X}$ of functions given by $f_a(x) = f(a)$ if $x = a$, $+\infty$ else].

5. Let $f : X \to \mathbb{R}_\infty$ be a l.s.c. function on a topological space X and let A be a nonempty subset of X. Show that $\inf f(A) = \inf f(\mathrm{cl} A)$. Can one replace inf by sup?

6. Show that the supremum of a family of continuous functions is not necessarily continuous.

7. (**Ritz's method**) Let $f : X \to \mathbb{R}$ be an upper semicontinuous function on a topological space X and let $(X_n)_{n \geq 0}$ be a sequence in subspaces such that for all $x \in X$ there exists a sequence $(x_n) \to x$ satisfying $x_n \in X_n$ for all n. Let $m := \inf f(X)$, $m_n := \inf f(X_n)$. Show that $m = \lim_n m_n$.

2.2.4 Compactness

The existence of limits being a frequent aim, the following definition is of interest.

Definition 2.9 A topological space $(X, \mathcal{O})$ is said to be *compact* if it is Hausdorff and if any net in X has a convergent subnet.

Equivalently, since a cluster point of net is the limit of some subnet, a topological space $(X, \mathcal{O})$ is compact if every net in X has a cluster point. Moreover, if a net $(x_i)_{i\in I}$ of a compact space $(X, \mathcal{O})$ has at most one cluster point $\overline{x}$, then it converges to $\overline{x}$: if this were not the case, one could find an open neighborhood V of $\overline{x}$ and a cofinal subset J of I such that $x_j \notin V$ for all $j \in J$ and then $(x_j)_{j\in J}$ would have a convergent subnet whose limit would be in $X\backslash V$, contradicting the uniqueness of the cluster point of $(x_i)_{i\in I}$.

The property of the definition can be characterized in different ways. The usual one deals with open coverings of X, i.e. families $(O_i)_{i\in I}$ of open subsets whose union is X. Another one deals with families $(C_i)_{i\in I}$ satisfying the *finite intersection property*, i.e. such that for any finite subset J of I one has $\cap_{j\in J} C_j \neq \varnothing$.

Theorem 2.1 *A Hausdorff topological space $(X, \mathcal{O})$ is compact if and only if every open covering $(O_i)_{i\in I}$ of X has a finite covering, if and only if any family $(C_i)_{i\in I}$ of closed subsets with the finite intersection property has a nonempty intersection.*

Proof Setting $O_i := X\backslash C_i$ (and conversely $C_i := X\backslash O_i$) one sees that the last two properties are equivalent, since for all $J \subset I$ the family $(O_j)_{j\in J}$ is a covering of X if and only if $\cap_{j\in J} C_j$ is empty. Let us assume the last property is satisfied and let $(x_i)_{i\in I}$ be a net in X. Setting $C_i := \mathrm{cl}\{x_j : j \geq i\}$, we see that for every finite subset J of I and for any $k \in I$ such that $k \geq j$ for all $j \in J$ (such a k exists since I is directed) one has $\cap_{j\in J} C_j \supset C_k \neq \varnothing$, so that $\cap_{i\in I} C_i$ is nonempty. Since by Proposition 2.3 $\cap_{i\in I} C_i$ is the set of cluster points of $(x_i)_{i\in I}$, we get a cluster point of $(x_i)_{i\in I}$.

Now let $(C_i)_{i\in I}$ be a family of closed subsets satisfying the finite intersection property. Let $\mathcal{J}$ be the family of finite subsets of I and for $J \in \mathcal{J}$ let $x_J \in C_J := \cap_{j\in J} C_j$. Since $\mathcal{J}$ is directed with respect to inclusion, we get a net $(x_J)_{J\in\mathcal{J}}$ in X. Let us show that if $\cap_{i\in I} C_i = \varnothing$ the net $(x_J)_{J\in\mathcal{J}}$ cannot have a cluster point. Suppose $\overline{x}$ is a cluster point of $(x_J)_{J\in\mathcal{J}}$. Since $\overline{x} \notin \cap_{i\in I} C_i$ there exists some $k \in I$ such that $\overline{x} \in V := X\backslash C_k$. Then, for $J \in \mathcal{J}$ satisfying $J \supset K := \{k\}$ we cannot have $x_J \in V$ since $x_J \in C_J \subset C_k$. Thus $\overline{x}$ cannot be a cluster point of $(x_J)_{J\in\mathcal{J}}$. □

Corollary 2.5 *Let X be a subset of a Hausdorff topological space $(W, \mathcal{O}_W)$ endowed with the induced topology $\mathcal{O}_X := \{O \cap X : O \in \mathcal{O}_W\}$. Then $(X, \mathcal{O}_X)$ is compact if and only if every covering $(O_i)_{i\in I}$ of X by members of $\mathcal{O}_W$ has a finite subcovering.*

Here a family $(W_i)_{i\in I}$ of subsets of W is called a *covering* of X if $X \subset \cup_{i\in I} W_i$ and a subcovering is a subfamily $(W_j)_{j\in J}$ of $(W_i)_{i\in I}$ that is still a covering of X.

Proof If $(X, \mathcal{O}_X)$ is compact, for every covering $(O_{i\in I})$ of X by open subsets of W, the family $(O_i \cap X)_{i\in I}$ being a covering of X for the induced topology $\mathcal{O}_X$, one can find a finite subset J of I such that $(O_j \cap X)_{j\in J}$ is a covering of X. Then $(O_j)_{j\in J}$ is a finite subcovering of X.

Conversely, suppose every covering $(O_i)_{i\in I}$ of X by members of $\mathcal{O}_W$ has a finite subcovering $(O_j)_{j\in J}$. Then, if $(G_i)_{i\in I}$ is an open covering of $(X, \mathcal{O}_X)$, picking $O_i \in \mathcal{O}_W$ such that $G_i = O_i \cap X$, we can find a finite subset J of I such that $X \subset \cup_{j\in J} O_j$ and then $(G_j)_{j\in J}$ is a finite subcovering of $(G_i)_{i\in I}$. Thus $(X, \mathcal{O}_X)$ is compact. □

Example A discrete space, i.e. a set endowed with the discrete topology, is compact if and only if it is finite.

Example If (x_n) is a convergent sequence in a Hausdorff topological space $(W, \mathcal{O})$ and if $\overline{x} := \lim_n x_n$, then the set $X := \{x_n : n \in \mathbb{N}\} \cup \{\overline{x}\}$ is compact with respect to the induced topology. In fact, given a covering $(O_i)_{i\in I}$ of X by open subsets of W we can find some $k \in I$ such that $\overline{x} \in O_k$. Since $(x_n) \to \overline{x}$ there exists some $m \in \mathbb{N}$ such that $x_n \in O_k$ for $n > m$. Then, taking for all $j \in \mathbb{N}, j \leq m$ some $i(j) \in I$ such that $x_j \in O_{i(j)}$, we obtain a finite subcover of X by taking the family $\{O_{i(j)} : 0 \leq j \leq m\} \cup \{O_k\}$. □

Another important example is given by the following theorem.

Theorem 2.2 (Heine-Borel-Lebesgue) *Every closed bounded interval of* $\mathbb{R}$ *is compact.*

Proof Let $X := [a, b]$ with $a \leq b$ in $\mathbb{R}$ and let $(O_i)_{i\in I}$ be an open covering of X by open subsets of $\mathbb{R}$. Let A be the set of $x \in X$ such that $[a, x]$ is covered by a finite number of members of $(O_i)_{\in I}$. Then A is nonempty (since $a \in A$), hence it has a least upper bound $c \leq b$. Let $h \in I$ be such that $c \in O_h$. Suppose $c < b$. Given $\varepsilon > 0$ such that $c + \varepsilon \leq b$ and $[c - \varepsilon, c + \varepsilon] \subset O_h$, one can find some $x \in A$ such that $c - \varepsilon < x \leq c$. By definition of A one can find a finite subset J of I such that $[a, x] \subset \cup_{j\in J} O_j$. Then $c \in A$ and $[a, c + \varepsilon] \subset \cup_{k\in K} O_k$ for $K := J \cup \{h\}$, so that $c + \varepsilon \in A$, a contradiction. Thus $c = b$ and $b \in A : [a, b]$ can be covered by a finite subfamily of $(O_i)_{i\in I}$. □

Corollary 2.6 *The space* $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ *is compact.*

Proof This follows from the fact that there exists a homeomorphism h from $\overline{\mathbb{R}}$ onto $[-1, 1]$, for instance $h : r \mapsto r/(1 + |r|)$ for $r \in \mathbb{R}$, $h(-\infty) := -1$, $h(+\infty) := 1$. □

Let us give some permanence properties.

Proposition 2.13 *Let X be a closed subset of a compact topological space $(W, \mathcal{O}_W)$. Then, denoting by $\mathcal{O}_X := \{O \cap X : O \in \mathcal{O}_W\}$ the induced topology, $(X, \mathcal{O}_X)$ is compact.*

Proof If $(x_i)_{i\in I}$ is a net in X, it has a subnet $(x_j)_{j\in J}$ that converges in W. But since X is closed in W, the limit $\overline{x}$ of $(x_j)_{j\in J}$ belongs to X and $(x_j)_{j\in J}$ converges in $(X, \mathcal{O}_X)$. □

Proposition 2.14 *Let X be a subset of a Hausdorff topological space $(W, \mathcal{O}_W)$. If X is compact with respect to the induced topology $\mathcal{O}_X$, then X is closed in $(W, \mathcal{O}_W)$.*

Proof Let $\overline{w}$ be an element of the closure of X. Then, there exists a net $(x_i)_{i\in i}$ of X that converges to $\overline{w}$. But since $(X, \mathcal{O}_X)$ is compact, $(x_i)_{i\in I}$ has a cluster point $\overline{x} \in X$. Then Proposition 2.1 ensures that $\overline{x} = \overline{w}$, so that $\overline{w} \in X$. □

Corollary 2.7 *The compact subsets of $\mathbb{R}$ (with respect to the induced topology) are the closed bounded subsets of $\mathbb{R}$.*

Proof Since $\mathbb{R}$ is Hausdorff, the preceding proposition shows that a compact subset X of $\mathbb{R}$ is closed. It is bounded since otherwise we could find a sequence (x_n) in X satisfying $|x_n| > n$; such a sequence cannot have a cluster point.

Conversely, let X be a closed bounded subset of $\mathbb{R}$. Then there exists a closed bounded interval $W := [a, b]$ containing X. Since W is compact and X is closed in W, X is compact by Proposition 2.13. □

Exercise Show that in a Hausdorff topological space the union of a finite family of compact subsets is compact and the intersection of a family of compact subsets is compact.

Theorem 2.3 *Let $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ be two Hausdorff topological spaces, and let $f : X \to Y$ be a continuous map. If X is compact, then $Z := f(X)$ is compact.*

Proof Let $(z_i)_{i\in I}$ be a net in Z. If $x_i \in X$ is such that $z_i = f(x_i)$, then $(x_i)_{i\in I}$ has a convergent subnet $(x_j)_{j\in J}$. Then $(z_j)_{j\in J}$ is a subnet of $(z_i)_{i\in I}$ that converges to $f(\lim_j x_j)$. □

Corollary 2.8 *Let $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ be two Hausdorff topological spaces, and let $f : X \to Y$ be a continuous injective map. If X is compact, then f is a homeomorphism from X onto $f(X)$.*

Proof If C is a closed subset of X, C is compact and $f(C)$ is compact too, so that $f(C)$ is closed in $f(X)$. □

The proof of the following theorem requires Zorn's Lemma when the product has an infinite number of factors.

Theorem 2.4 (Tykhonov) *The product of a family of compact topological spaces is compact.*

Proof We admit the result when the product has an infinite number of factors. The case of a finite number of factors is reduced to the case of two factors by an induction. Let $Z := X \times Y$ be the product of two compact spaces and let $(z_i)_{i\in I} := ((x_i, y_i))_{i\in I}$ be a net in Z. A subnet $(x_j)_{j\in J}$ of $(x_i)_{i\in I}$ converges. In turn, a subnet $(y_k)_{k\in K}$ of $(y_j)_{j\in J}$ converges. Then $(z_k)_{k\in K}$ is a convergent subnet of $(z_i)_{i\in I}$. □

Corollary 2.9 *The compact subsets of $\mathbb{R}^d$ (with respect to the induced topology) are the closed bounded subsets of $\mathbb{R}^d$.*

Here a subset of $\mathbb{R}^d$ is said to be *bounded* if its projections are bounded.

Proof If S is a compact subset of $\mathbb{R}^d$, it is closed and its projections are compact, hence bounded. Conversely, if S is a closed bounded subsets of $\mathbb{R}^d$ then S is contained in a product of closed bounded intervals. Thus S is a closed subset of a compact space, hence S is compact. □

A subset X of a topological space $(W, \mathcal{O}_W)$ is said to be *relatively compact* if its closure is compact. Thus, any subset of a relatively compact subset is relatively compact.

Since fixed point results enable us to solve equations, the interest of the notion of compactness is illustrated by the following theorem. It has been given many proofs; an elementary one can be found in [197, 222] and in the appendix. Recall that a subset C of a vector space is said to be *convex* if for all $x_0, x_1 \in C$ and $t \in [0, 1]$ one has $(1-t)x_0 + tx_1 \in C$.

Theorem 2.5 (Brouwer) *Let X be a compact convex subset of $\mathbb{R}^d$ (or of a finite dimensional normed vector space) and let $f : X \to X$ be a continuous map. Then there exists some $\overline{x} \in X$ such that $f(\overline{x}) = \overline{x}$.*

Corollary 2.10 (Hairy Ball Theorem) *Let X be a (finite dimensional) Euclidean space with scalar product $\langle \cdot \mid \cdot \rangle$, unit closed ball B_X, unit sphere S_X, let $r > 0$, and let $g : rB_X \to X$ be continuous and pointing inside rB_X, i.e. such that*

$$\langle g(x) \mid x \rangle \leq 0 \qquad \forall x \in rS_X.$$

Then there exists some $\overline{z} \in rB_X$ such that $g(\overline{z}) = 0$.

Proof Suppose on the contrary that $g(x) \neq 0$ for all $x \in rB_X$. Then $h : rB_X \to X$ given by

$$h(x) := \frac{r}{\|g(x)\|} g(x) \qquad x \in rB_X$$

is continuous, takes its values in $rS_X \subset rB_X$, hence has a fixed point $\overline{x} \in rB_X$ by Brouwer's Theorem. Then we get the contradiction

$$r^2 = \|h(\overline{x})\|^2 = \langle h(\overline{x}) \mid \overline{x} \rangle = \frac{r}{\|g(\overline{x})\|} \langle g(\overline{x}) \mid \overline{x} \rangle \leq 0.$$

This contradiction proves that g has a zero in rB_X. □

Corollary 2.11 *Let X be a (finite dimensional) Euclidean space, let $b, r \in \mathbb{R}_+$, and let $f : X \to X$ be continuous and such that $\langle f(x) \mid x \rangle \geq b \|x\|$ for all $x \in rS_X$. Then for all $y \in bB_X$ the equation $f(x) = y$ has a solution $x \in rB_X$.*

If $\langle f(x) \mid x\rangle \geq c(\|x\|)\,\|x\|$ *for some function* $c(\cdot)$ *such that* $c(r) \to \infty$ *as* $r \to \infty$, *then* $f(X) = X$. *Moreover, there exists a function* $k : \mathbb{R}_+ \to \mathbb{R}_+$ *such that for all* x, $y \in X$ *satisfying* $f(x) = y$ *one has* $\|x\| \leq k(\|y\|)$.

Proof Given $y \in bB_X$ let us set $g(\cdot) := y - f(\cdot)$, so that g is continuous and for $x \in rS_X$

$$\langle g(x) \mid x\rangle = \langle y \mid x\rangle - \langle f(x) \mid x\rangle \leq b\,\|x\| - b\,\|x\| \leq 0.$$

Thus, there exists some $x \in rB_X$ such that $g(x) = 0$ or $f(x) = y$.

Assuming that $\langle f(x) \mid x\rangle \geq c(\|x\|)\,\|x\|$ for $c(\cdot)$ satisfying $\lim_{r\to\infty} c(r) = \infty$, setting

$$k(s) := \sup\{r \in \mathbb{R}_+ : c(r) \leq s\} \qquad s \in \mathbb{R}_+$$

we see that k takes finite values and that whenever x is such that $f(x) = y$ we have $c(\|x\|).\,\|x\| \leq \langle f(x) \mid x\rangle = \langle y \mid x\rangle \leq \|y\|\,.\,\|x\|$, hence $\|x\| \leq k(\|y\|)$. □

Some topological spaces of interest have a local behavior involving compactness.

Definition 2.10 A topological space $(X, \mathcal{O}_X)$ is said to be locally compact if it is Hausdorff and if each point of X has a compact neighborhood.

A compact space is obviously locally compact. A discrete space (i.e. a space in which each subset is open) is locally compact, but it is not compact if it is infinite. The space $\mathbb{R}$ with its usual topology is locally compact but not compact. An open subset of a compact space or locally compact space is locally compact, as the following proposition shows.

Proposition 2.15 *In a compact space, or more generally in a locally compact space, each point has a base of neighborhoods formed by compact sets.*

Proof Let $(X, \mathcal{O}_X)$ be a compact space, let $\overline{x} \in X$ and let $U \in \mathcal{N}(\overline{x})$. We want to show that there exists some closed $V \in \mathcal{N}(\overline{x})$ contained in U. Without loss of generality we may assume U is open. If no such V exists, since $\mathcal{N}(\overline{x})$ is stable under finite intersections, the family $\{V \backslash U : V \in \mathcal{N}(\overline{x}),\ \mathrm{cl}(V) = V\}$ of closed subsets of $X \backslash U$ has the finite intersection property. Since $X \backslash U$ is compact by Proposition 2.13, one can find some $\overline{y} \in X \backslash U$ belonging to any closed neighborhood V of $\overline{x}$. Then we have $\overline{y} \neq \overline{x}$ and since X is Hausdorff we can find some $V \in \mathcal{N}(\overline{x})$, $W \in \mathcal{N}(\overline{y})$ with $V \cap W = \varnothing$. Since we may assume W is open, replacing V with its closure we may suppose V is closed. Then $\overline{y} \in V$. Since $\overline{y} \in W$ and $V \cap W = \varnothing$ we get a contradiction.

Now suppose X is locally compact. Let $\overline{x} \in X$ and let $U \in \mathcal{N}(\overline{x})$. By assumption there is some $W \in \mathcal{N}(\overline{x})$ that is compact. The preceding yields some compact neighborhood V of $\overline{x}$ in W contained in $U \cap W$. Then V is a neighborhood of $\overline{x}$ in X and is contained in U. □

Proposition 2.16 *Every open or closed subset X of a locally compact space $(W, \mathcal{O}_W)$ is locally compact.*

Proof If X is open, for every $\overline{x} \in X$ one has $X \in \mathcal{N}(\overline{x})$, so that there is some compact $V \in \mathcal{N}(\overline{x})$ contained in X and V is a neighborhood of $\overline{x}$ with respect to the induced topology on X.

If X is closed in W and if $\overline{x} \in X$, taking a neighborhood V of $\overline{x}$ in W that is compact, we see that $W \cap X$ is compact and is a neighborhood of $\overline{x}$ in X with respect to the induced topology. Thus X is locally compact. □

Proposition 2.17 *The product of a finite family of locally compact spaces is locally compact. In particular, $\mathbb{R}^d$ is locally compact.*

Proof It suffices to prove that the product $Z := X \times Y$ of two locally compact spaces is locally compact. Given $\overline{z} := (\overline{x}, \overline{y}) \in Z$ we pick $U \in \mathcal{N}(\overline{x})$, $V \in \mathcal{N}(\overline{y})$ that are compact. Then $U \times V$ is a compact neighborhood of $\overline{z}$. □

Given a topological space $(X, \mathcal{O})$ there are different ways of embedding it into a compact space. When $(X, \mathcal{O})$ is locally compact the simplest approach consists in adding a point $\overline{w}$ to X and declaring that a subset O of $W := X \cup \{\overline{w}\}$ is open if either it belongs to $\mathcal{O}$ or if $\overline{w} \in O$ and $X \backslash O$ is compact. It is easy to see that the resulting topological space is compact. It is called the *Alexandroff compactification* of X or one point compactification of X.

Exercise Show that the Alexandroff compactification of $\mathbb{R}^d$ is homeomorphic to the unit sphere $\mathbb{S}^d$ of $\mathbb{R}^{d+1}$. [Hint: use the stereographic projection $p : \mathbb{S}^d \setminus \{n\} \to \mathbb{R}^d$, where n is the "north pole" $n := (0, \ldots, 0, 1)$ of $\mathbb{S}^d$ defined as follows: for $x \in \mathbb{S}^d \setminus \{n\}$, $p(x)$ is the point of $\mathbb{R}^d \times \{0\}$ that belongs to the half-line $x + \mathbb{R}_+(x - n)$.]

The following theorem is the main existence result in optimization theory.

Theorem 2.6 (Weierstrass) *Let $f : X \to \overline{\mathbb{R}}$ be a lower semicontinuous function on a (nonempty) compact topological space X. Then the set $M := \{w \in X : f(w) \leq f(x)\ \forall x \in X\}$ of minimizers of f is nonempty.*

Proof We may suppose $m := \inf f(X) < \infty$ for otherwise f is constant with value ∞. Setting $S_f(r) := \{x \in X : f(x) \leq r\}$ for $r \in \mathbb{R}$, the family $\{S_f(r) : r > m\}$ is formed of nonempty closed subsets and any finite subfamily has a nonempty intersection: $\cap_{1 \leq i \leq k} S_f(r_i) = S_f(r_j)$ where $r_j := \min_{1 \leq i \leq k} r_i$. Therefore $M = \cap_{r > m} S_f(r)$ is nonempty. □

Given a topological space X, one may try to weaken its topology in order to enlarge the family of compact subsets. Then, a continuous function on X may not remain continuous. There are interesting cases, for instance making use of convexity assumptions, for which the function still remains lower semicontinuous, so that the preceding generalization of the classical existence of a minimizer under a continuity assumption is of interest.

Let us give a criterion for the lower semicontinuity of a function obtained by minimization.

Proposition 2.18 *Let W, X be topological spaces, let $\overline{w} \in W$ and let $f : W \times X \to \overline{\mathbb{R}}$ be a function which is lower semicontinuous at $(\overline{w}, \overline{x})$ for every $\overline{x} \in X$. If the following compactness assumption is satisfied, then the performance function $p : W \to \overline{\mathbb{R}}$ given by $p(w) := \inf_{x \in X} f(w, x)$ is lower semicontinuous at $\overline{w}$:*

(C) for any net $(w_i)_{i \in I} \to \overline{w}$ there exist a subnet $(w_j)_{j \in J}$, a convergent net $(x_j)_{j \in J}$ in X and $(\varepsilon_j)_{j \in J} \to 0_+$ such that $f(w_j, x_j) \le p(w_j) + \varepsilon_j$ for all $j \in J$.

Proof Given a net $(w_i)_{i \in I} \to \overline{w}$ such that $(p(w_i))_{i \in I}$ converges, let $(w_j)_{j \in J}$ be a subnet of $(w_i)_{i \in I}$, and let $(x_j)_{j \in J}$ and $(\varepsilon_j) \to 0_+$ be as in (C). Then, if $\overline{x}$ is the limit of $(x_j)_{j \in J}$, one has

$$p(\overline{w}) \le f(\overline{w}, \overline{x}) \le \liminf_{j \in J} f(w_j, x_j) \le \liminf_{j \in J} \big(p(w_j) + \varepsilon_j\big) = \lim_{i \in I} p(w_i).$$

Since $\liminf_{w \to \overline{w}} p(w)$ is the limit of $(p(w_i))_{i \in I}$ for some net $(w_i)_{i \in I} \to \overline{w}$ such that $(p(w_i))_{i \in I}$ converges, these inequalities show that $p(\overline{w}) \le \liminf_{w \to \overline{w}} p(w)$. □

Corollary 2.12 *Let W and X be topological spaces, X being compact, and let $f : W \times X \to \mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous. Then the performance function p defined as above is lower semicontinuous.*

Proof We give a proof in the case when f is just lower semicontinuous at each point of $\{\overline{w}\} \times X$; when f is lower semicontinuous on $W \times X$, a simpler proof can be given using the Weierstrass' Theorem. Condition (C) is clearly satisfied when X is compact since for any net $(w_i)_{i \in I} \to \overline{w}$ and for any sequence $(\alpha_n) \to 0_+$ one can take $H := I \times \mathbb{N}$, $w_h := w_i$, $\varepsilon_h := \alpha_n$ for $h := (i, n)$ and pick $x_h \in X$ satisfying $f(w_h, x_h) \le p(w_h) + \varepsilon_h$, and take a subnet $(x_j)_{j \in J}$ of $(x_h)_{h \in H}$ which converges in X.

□

Exercises

1. For every net $(r_i)_{i \in I}$ in the compact space $\overline{\mathbb{R}}$ show that $\liminf_{i \in I} r_i$ is the least cluster point of $(r_i)_{i \in I}$ and $\limsup_{i \in I} r_i$ is the greatest cluster point of $(r_i)_{i \in I}$.
2. Let X and Y be topological spaces and let Z be a closed subset of $X \times Y$. Show that if Y is compact, then the projection $p_X(Z)$ of Z on X is closed. Give an example showing that one cannot drop the assumption that Y is compact.
3. Let X and Y be topological spaces, Y being Hausdorff and let $f : X \to Y$ be a map. Show that if f is continuous then the graph $G := \{(x, f(x)) : x \in X\}$ of f is closed in $X \times Y$. Give an example with $X = Y = \mathbb{R}$ showing that the converse is not true. Prove that if Y is compact, then the converse is true.
4. Let $(K_n)_n$ be a decreasing sequence in nonempty compact subsets of a topological space X. Show that $K := \cap_n K_n$ is nonempty and that for any open subset G of X containing K there exists some $m \in \mathbb{N}$ such that $K_m \subset G$. Give a generalization to a filtered family $(K_i)_{i \in I}$ of nonempty compact subsets.

5. Let X be a Hausdorff topological space and let A and B be two disjoint compact subsets of X. Show that there exist open subsets U, V of X such that $A \subset U$, $B \subset V$ and $U \cap V = \varnothing$. [Hint: start with the case when B is a singleton.]

6. Let $X := [0, 1] \times [0, 1]$ be endowed with the lexicographic order $\preceq$. Let $\mathcal{O}$ be the topology generated by the open intervals $T_{w,z} := \{x \in X : w \prec x \prec z\}$ of X and the intervals $\{x \in X : x \prec z\}$, $\{x \in X : w \prec x\}$, where $x \prec x'$ means that $x \preceq x'$ and $x \neq x'$. Show that X is compact with respect to this topology.

7. Show that for any compact subset K of a locally compact space X and any neighborhood U of K there exists some compact neighborhood V of K contained in U.

8. Show that the intersection of two locally compact subsets of a topological space is locally compact.

9. Prove that the union of two locally compact subsets of a topological space is not always locally compact. [Hint: in $\mathbb{R}$ take $A := \{a\}$ where a is the limit of a sequence (a_n) of $\mathbb{R}\backslash\{a\}$ and $B := \mathbb{R}\backslash(\{a_n : n \in \mathbb{N}\} \cup \{a\})$ or in $\mathbb{R}^2$ take $A := \mathbb{R} \times \mathbb{P}$, $B := \{(0, 0)\}$.]

10. Show by an example that the image of a locally compact space X under a continuous map $f : X \to Y$ is not necessarily locally compact. [Hint: take a bijection f from $\mathbb{N}$ onto $\mathbb{Q}$.]

11*. Prove the odd **Hairy Ball Theorem**: for d odd there is no continuous vector field on the unit sphere $\mathbb{S}^{d-1} := S_{\mathbb{R}^d}$ tangent to $S_{\mathbb{R}^d}$ [Hint: use the appendix.]

12. (**Beals**) Let B be the closed unit ball of the space $X := c_0$ of sequences $x := (x_n)_{n\geq 0}$ with limit 0 endowed with the supremum norm. Prove that the map $f : B \to B$ given by $f(x) = (1, x_0, x_1, \ldots)$ for $x := (x_0, x_1, \ldots) \in B$ is nonexpansive, i.e. does not increase distances, and has no fixed point.

13. (**Stone**) A Boolean ring is a ring A such that $a^2 = a$ for all $a \in A$. A topological space $(X, \mathcal{O})$ is said to be *totally disconnected* if the family $\mathcal{B}$ of subsets of X which are simultaneously open and closed forms a base of its topology. Show that $\mathcal{B}$ forms a Boolean ring with unit when the product is $\cap$ and the addition is Δ as in Definition 1.2. Conversely, show that any Boolean ring with unit is isomorphic to the ring of subsets of a compact topological space which are simultaneously open and closed. [See [106, p. 41].]

2.3 Metric Spaces

A usual means of studying convergence on a set X is to reduce the question to the case of convergence in $\mathbb{R}$. That can be done if one disposes of functions from $X \times X$ to $\mathbb{R}$ that allow such a transfert. Metrics (also called distances) and semimetrics are such means.

2.3.1 General Facts About Metric Spaces

Definition 2.11 A *semimetric* on a set X is a function $d : X \times X \to \mathbb{R}_+ := [0, +\infty[$ satisfying the properties:

(SM1) for all $x \in X$ one has $d(x, x) = 0$;
(SM2) for all $x, x' \in X$ one has $d(x, x') = d(x', x)$;
(SM3) for all $x, x', x'' \in X$, one has $d(x, x'') \leq d(x, x') + d(x', x'')$.

A *metric* is a semimetric d such that for all $x, x' \in X$, $d(x, x') = 0 \Rightarrow x = x'$.
A *pseudo-metric* on X is a function $d : X \times X \to \overline{\mathbb{R}}_+ := [0, +\infty]$ satisfying (SM1)–(SM3).

The relation in (SM3) is called the *triangle inequality*.

A *metric space* (resp. *semimetric space*) is a pair (X, d) formed by a set X and a metric (resp. semimetric) d on X. When there is no ambiguity on d, we just write X. A *uniform space* is a pair $(X, (d_a)_{a \in A})$, where $(d_a)_{a \in A}$ is a family of pseudo-metrics on X. In a metric space (X, d) (resp. a uniform space $(X, (d_a)_{a \in A})$) one can introduce a convergence by setting

$$(x_i)_{i \in I} \to x \quad \Longleftrightarrow \quad (d(x_i, x))_{i \in I} \to 0$$

$$(\text{resp. } (x_i)_{i \in I} \to x \quad \Longleftrightarrow \quad \forall a \in A \quad (d_a(x_i, x))_{i \in I} \to 0).$$

Exercise Verify the axioms (C1), (C2), (C3) of convergence spaces (Definition 2.2)

In fact, a semimetric d induces a topology $\mathcal{O}$ on X defined by: $G \in \mathcal{O}$ iff for all $x \in G$ there exists some $r > 0$ such that the *open ball*

$$B(x, r) := \{x' \in X : d(x, x') < r\}$$

is contained in G. Thus $\mathcal{O}$ is the topology generated by the family of open balls. This family is a base of $\mathcal{O}$ and for all $\overline{x} \in X$, the family of open balls centered at $\overline{x}$ is a base of neighborhoods of $\overline{x}$. In the sequel, the *closed ball* with center x and radius $r \in \mathbb{R}_+$ is the set

$$B[x, r] := \{x' \in X : d(x, x') \leq r\}.$$

The family of closed balls centered at x with positive radius is also a base of neighborhoods of x. A topology can also be associated with a uniform space $(X, (d_a)_{a \in A})$: it is the topology generated by the balls $B_a(x, r) := d_a(x, \cdot)^{-1}([0, r[)$ for $a \in A$, $x \in X$, $r > 0$. It is easy to show that the convergence on (X, d) or $(X, (d_a)_{a \in A})$ described above is the convergence associated with the topology we just defined. Moreover, when d is a metric, the associated topology is Hausdorff: given $x, x' \in X$ such that $x \neq x'$, for $r \in]0, d(x, x')/2[$ the balls $B(x, r)$ and $B(x', r)$ are disjoint in view of the triangle inequality. Hence convergent nets or sequences have a unique limit by Proposition 2.1. The existence of a metric d on X implies a noticeable property of accumulation points. We propose it in the next exercise. A point a is called an *accumulation point* of a subset S of a topological space X if every neighborhood V of a contains some point $x \in S$, $x \neq a$.

Exercise Show that every neighborhood V of an accumulation point a of a subset S of a metric space (X, d) contains an infinite family of points of S. [Hint: by induction define a sequence (x_n) of $S \backslash \{a\}$ such that $d(x_{n+1}, a) < d(x_n, a)$.]

Given a subset S of a metric space (X, d) the *diameter* of S is $\mathrm{diam}(S) := \sup\{d(x, y) : x, y \in S\}$ and the *distance to* S is the function $d_S : X \to \mathbb{R}_+$ given by

$$d_S(x) := \inf\{d(x, y) : y \in S\}, \qquad x \in X.$$

This notion is often convenient. If A, B are two subsets of X, their *gap* $g(A, B)$ is

$$\mathrm{gap}(A, B) := \inf\{d(a, b) : a \in A,\ b \in B\}.$$

This number is sometimes called the distance between A and B but this terminology is improper as $(A, B) \mapsto \mathrm{gap}(A, B)$ is not a metric on the set $\mathcal{P}(X)$ of subsets of X or on the set of closed subsets of X.

In metric spaces continuity can be expressed with the help of ε's and δ's: a map $f : (X, d) \to (X', d')$ is continuous at $\overline{x} \in X$ if and only if for all $\varepsilon > 0$ one can find some $\delta > 0$ such that $d'(f(x), f(\overline{x})) < \varepsilon$ for all $x \in X$ satisfying $d(x, \overline{x}) < \delta$. In metric spaces, one can avoid nets and just use sequences in a convenient way:

Proposition 2.19 *For a map $f : (X, d) \to (X', d')$ between two metric spaces and $\overline{x} \in X$ the following assertions are equivalent:*

(a) f is continuous at $\overline{x} \in X$;
(b) for every sequence $(x_n) \to \overline{x}$ one has $(f(x_n)) \to f(\overline{x})$;
(c) for every sequence $(x_n) \to \overline{x}$ one can find a subsequence $(x_{k(n)})$ such that $(f(x_{k(n)})) \to f(\overline{x})$.

Proof (a)⇒(b) Sequences being particular nets, the implication stems from Proposition 2.2. (b)⇒(c) being obvious, it remains to show that (c)⇒(a).

(c)⇒(a) If f is not continuous at $\overline{x}$ there exists some $\varepsilon > 0$ such that for all $\delta > 0$ one has $f(B(\overline{x}, \delta)) \not\subseteq B(f(\overline{x}), \varepsilon)$. Taking a sequence $(\delta_n) \to 0_+$ we can find some $x_n \in B(\overline{x}, \delta_n)$ such that $f(x_n) \notin B(f(\overline{x}), \varepsilon)$. Then we have $(x_n) \to \overline{x}$ but for any subsequence $(x_{k(n)})$ the sequence $(f(x_{k(n)}))$ does not converge to $f(\overline{x})$. □

Proposition 2.20 *In a metric space (X, d) the closure $\mathrm{cl}(S)$ of a subset S is the set of limits of sequences in S.*

Proof If x is the limit of a sequence (x_n) of S, then clearly $x \in \mathrm{cl}(S)$. Conversely, if $x \in \mathrm{cl}(S)$, for any $n \geq 1$ the set $S \cap B(x, 1/n)$ is nonempty. Picking $x_n \in S \cap B(x, 1/n)$ we get a sequence (x_n) converging to x. □

Proposition 2.21 *In a metric space (X, d) any cluster point of a sequence in X is the limit of a subsequence.*

Proof Let $\overline{x}$ be a cluster point of a sequence (x_n) of X. Given a sequence $(r_n) \to 0_+$ an induction on n gives a sequence $(k(n))$ of $\mathbb{N}$ such that $k(n+1) > k(n)$ and $x_{k(n)} \in B(\overline{x}, r_n)$ for all n. Then $(x_{k(n)})_n$ is a subsequence in (x_n) that converges to $\overline{x}$. □

A map $f : (X, d) \to (X', d')$ is said to be *uniformly continuous* if for all $\varepsilon > 0$ one can find some $\delta > 0$ such that $d'(f(w), f(x)) < \varepsilon$ for all $w, x \in X$ satisfying $d(w, x) < \delta$. Such a map is clearly continuous at each $\overline{x} \in X$ and one sees that δ does not depend on $\overline{x}$. If there exists a *modulus* μ, i.e. a function $\mu : \mathbb{R}_+ \to \overline{\mathbb{R}}_+ := [0, +\infty]$ satisfying $\mu(t) \to 0$ as $t \to 0$, such that $d'(f(w), f(x)) \leq \mu(d(w, x))$ for all $w, x \in X$, then f is uniformly continuous. The converse is true and we invite the reader to produce a modulus satisfying the preceding property (and even the least such modulus μ_f called the modulus of uniform continuity of f). The case when μ is linear deserves some attention, but there are other cases of interest, for example the case $\mu(t) = ct^\alpha$ with $c \in \mathbb{R}_+, \alpha > 0$ (then f is said to be *Hölderian*).

A map $f : (X, d) \to (X', d')$ is said to be *Lipschitzian* if there exists some $c \in \mathbb{R}_+$ such that $d'(f(x_1), f(x_2)) \leq cd(x_1, x_2)$ for all $x_1, x_2 \in X$. The constant c is called a Lipschitz constant (or rate, or rank). The least such constant is called the (exact) *Lipschitz rate* of f. If this rate is 1, f is said to be *nonexpansive*. If this rate is less than 1 one says that f is contractive or a *contraction*. If f is a bijection and if for all $x_1, x_2 \in X$ one has $d'(f(x_1), f(x_2)) = d(x_1, x_2)$ one says that f is an *isometry*; then f^{-1} is also an isometry. For $x \in X$ the *Lipschitz rate* of f *at* x is the infimum of the Lipschitz rates of the restrictions of f to the neighborhoods of x (and $+\infty$ if there is no neighborhood of x on which f is Lipschitzian). If for all $x \in X$ there is a neighborhood V of x such that the restriction $f \mid V$ is Lipschitzian, f is said to be *locally Lipschitzian*.

If $(X, (d_a)_{a \in A})$ and $(Y, (d_b)_{b \in B})$ are two uniform spaces, a map $f : X \to Y$ is said to be *uniformly continuous* if for all $b \in B$ and all $\varepsilon \in \mathbb{P}$ one can find a finite subset $A(b, \varepsilon)$ of A and $\delta > 0$ such that $d_b(f(x), f(x')) \leq \varepsilon$ whenever $x, x' \in X$ satisfy $d_a(x, x') \leq \delta$ for all $a \in A(b, \varepsilon)$.

Exercise Show that the composition of two uniformly continuous maps is uniformly continuous.

Different equivalence properties can be defined on the set of metrics on a set X. Two metrics d, d' are said to be *topologically equivalent* if the topologies they define coincide. They are said to be *uniformly equivalent* (resp. *metrically equivalent*) if the identity map from (X, d) into (X, d') and its inverse are uniformly continuous (resp. Lipschitzian).

Proposition 2.22 *A metric space (X, d) is separable if and only if its topology has a countable base.*

Proof If the topology of (X, d) has a countable base $\mathcal{B} := \{B_n : n \in \mathbb{N}\}$, picking an arbitrary point $a_n \in B_n$ for all n we get a dense subset $A := \{a_n : n \in \mathbb{N}\}$ since for all $x \in X$ and any open subset U containing x there exists some $n \in \mathbb{N}$ such that $B_n \subset U$.

Conversely, suppose X contains a countable dense subset $A := \{a_n : n \in \mathbb{N}\}$. Then we claim that the family $\mathcal{B} := \{B(a_n, q) : n \in \mathbb{N},\ q \in \mathbb{Q},\ q > 0\}$ is a base of the topology of (X, d). In fact, given $\overline{x} \in X$ and $r > 0$ we can find $q \in \mathbb{Q}$ such that $0 < q < r/2$ and if $n \in \mathbb{N}$ is such that $a_n \in B(\overline{x}, q)$, for all $x \in B(a_n, q)$ we have $x \in B(\overline{x}, r)$ by the triangle inequality. $\square$

Example The set $\mathbb{R}$ of real numbers endowed with its usual metric has a countable base since the set $\mathbb{Q}$ of rational numbers is dense in $\mathbb{R}$.

Corollary 2.13 *A subspace of a separable metric space is separable.*

Proof This follows from the proposition and from the fact that if $\mathcal{B}$ is a countable base of (X, d), then for a subspace Y, the family $\mathcal{B}_Y := \{B \cap Y : B \in \mathcal{B}\}$ is a base for the induced topology on Y. $\square$

On the product $Z := X \times Y$ of two metric spaces (X, d_X), (Y, d_Y), a metric d is called a *product metric* if the canonical projections $p_X : Z \to X$, $p_Y : Z \to Y$ and the insertions $j_b : x \mapsto (x, b)$, $j_a : y \mapsto (a, y)$ are nonexpansive.

Exercise Show that a metric d on $Z := X \times Y$ is a product metric if and only if for all (u, v), $(x, y) \in X \times Y$ one has

$$\max(d_X(u, x), d_Y(v, y)) \leq d((u, v), (x, y)) \leq d_X(u, x) + d_Y(v, y).$$

The left (resp. right) side defines a convenient metric usually denoted by d_∞ (resp. d_1). Describe its balls.

Whereas the product X of a family of metric spaces (X_s, d_s) ($s \in S$, an arbitrary set) cannot be provided with a (sensible, i.e. inducing the product topology) metric in general, we have seen that a product of topological spaces $(X_s, \mathcal{O}_s)$ ($s \in S$) can always be endowed with a topology $\mathcal{O}$ that makes the projections $p_s : X \to X_s$ continuous and that is as weak as possible, namely it is the topology generated by the sets $p_s^{-1}(O_s)$ for $s \in S$, $O_s \in \mathcal{O}_s$. Its associated convergence is componentwise convergence. When $(X_s, \mathcal{O}_s) := (Y, \mathcal{O}_Y)$ for all $s \in S$, identifying the product X with the set Y^S of maps from S to Y, the convergence associated with the product topology $\mathcal{O}$ on X coincides with *pointwise convergence*: $(f_i)_{i \in I} \to f$ in Y^S if, for all $s \in S$, $(f_i(s))_{i \in I} \to f(s)$. When $\mathcal{O}_Y$ is the topology associated with a metric d_Y on Y, a stronger convergence can be defined on Y^S: it is the so-called *uniform convergence* for which $(f_i)_{i \in I} \to f$ iff $(d_\infty(f_i, f)) := (\sup_{s \in S} d_Y(f_i(s), f(s))) \to 0$. This convergence is adapted to bounded functions, but it can be considered for any set of maps from a set S into a metric space Y. It enjoys better preservation properties, such as the following one.

Theorem 2.7 *Let X be a topological space and let (Y, d) be a metric space. Let $(f_i)_{i \in I}$ be a net (or a sequence) of continuous functions from X into Y that converges uniformly to some map $f : X \to Y$. Then f is continuous.*

Proof Let $\overline{x} \in X$ and let $\varepsilon > 0$ be given. We can find $k \in I$ such that for $i \geq k$ we have $\sup_{x \in X} d(f_i(x), f(x)) \leq \varepsilon/3$. Since f_k is continuous there exists a neighborhood V of $\overline{x}$ such that $d(f_k(x), f_k(\overline{x})) \leq \varepsilon/3$ for all $x \in V$. Then, for $x \in V$ we have

$$d(f(x), f(\overline{x})) \leq d(f(x), f_k(x)) + d(f_k(x), f_k(\overline{x})) + d(f_k(\overline{x}), f(\overline{x})) \leq \varepsilon.$$

This proves that f is continuous at $\overline{x}$. □

Some constructions can be obtained by using metrics.

Lemma 2.3 (Urysohn's Lemma) *Let A and B be two disjoint nonempty closed subsets of a metric space (X, d). Then there exists inuous function $h : X \to [-1, 1]$ such that $h(x) = -1$ for all $x \in A$ and $h(x) = 1$ for all $x \in B$.*

If C is a closed subset of X and if $f : C \to [-1, 1]$ is a continuous function, there exists a continuous function $g : X \to [-1/3, 1/3]$ such that $|f(x) - g(x)| \leq 2/3$ for all $x \in C$.

Proof Since A and B are disjoint, for all $x \in X$ we have $g(x) := d(x, A) + d(x, B) > 0$. Setting $h(x) := (d(x, A) - d(x, B))/g(x)$ we obtain the required function.

Let us set $A := \{x \in C : f(x) \leq -1/3\}$ and $B := \{x \in C : f(x) \geq 1/3\}$. If A is empty we take $g := 1/3$; if B is empty we take $g := -1/3$. If A and B are both nonempty we take $g := (1/3)h$ where h is as in the first assertion. We obtain the relation $|f(x) - g(x)| \leq 2/3$ for all $x \in C$ by considering the three cases $x \in A$, $x \in B$, $x \in C \backslash (A \cup B)$. □

Theorem 2.8 (Tietze-Urysohn) *Let C be a closed subset of a metric space (X, d) and let $f : C \to \mathbb{R}$ a bounded continuous function. Then there exists a continuous function $g : X \to \mathbb{R}$ such that $g \mid_C = f$, $\inf g(X) = \inf f(C)$, and $\sup g(X) = \sup f(C)$.*

Proof The result is obvious when f is constant. Changing f into $af + b$, where a and b are appropriate real numbers, we may assume $\inf f(C) = -1$ and $\sup f(C) = 1$.

The second assertion of the lemma yields a continuous function g_0 with values in $[-1/3, 1/3]$ such that $|f(x) - g_0(x)| \leq 2/3$ for all $x \in C$.

Let us suppose that inductively we have defined for $n \in \mathbb{N}_k := \{1, \ldots, k\}$ a continuous function g_n such that

$$|g_n| \leq 1 - (\frac{2}{3})^{n+1} \qquad |g_n \mid_C - f| \leq (\frac{2}{3})^{n+1} \tag{2.1}$$

Applying the second assertion of the lemma to the function $(\frac{3}{2})^{k+1}(g_k \mid_C - f)$, we get a continuous function h_{k+1} with values in $[-2^{k+1}/3^{k+2}, 2^{k+1}/3^{k+2}]$ such that for all $x \in C$

$$|f(x) - g_k(x) - h_{k+1}(x)| \leq (\frac{2}{3})^{k+2}. \tag{2.2}$$

Setting $g_{k+1} := g_k + h_{k+1}$ we obtain the two inequalities of relation (2.1) for $n = k+1$. Since for $x \in X$ we have $|g_{k+1}(x) - g_k(x)| \leq 2^{k+1}/3^{k+2}$ for all $k \in \mathbb{N}$, the sequence (g_k) converges uniformly on X to a function g that is continuous by Theorem 2.7. Moreover, for $x \in C$, passing to the limit in relation (2.2) we get $g(x) = f(x)$. □

Metrics can be used in order to tackle optimization problems with constraints: a natural idea consists in introducing some penalty terms, replacing the objective f by a penalized objective. If one has to minimize f on an admissible subset A of a metric space (X,d) one may consider the minimization of

$$f_s := f + sd_A(\cdot)$$

on the whole space X, expecting that for a large penalization rate s the effect will be similar. Here $d_A(x) := d(x,A) := \inf\{d(x,w) : w \in A\}$ for $x \in X$. In general one has to replace s by an infinite sequence $(s_n) \to \infty$, so that one has to solve a sequence of unconstrained problems. In some favorable cases a single penalized problem suffices as in the simple situation presented in the following result.

Proposition 2.23 (Exact Penalization) *Let X be a metric space and let $f : X \to \mathbb{R}$ be a Lipschitzian function with rate r. Then, for any $s \geq r$ and any nonempty subset A of X*

$$\inf_{x\in A} f(x) = \inf_{x\in X} \left(f(x) + sd_A(x)\right). \tag{2.3}$$

Moreover, $\overline{x} \in A$ is a minimizer of f on A if and only if, $\overline{x}$ is a minimizer of $f_s := f + sd_A$ on X. If A is closed and if $s > r$, any minimizer z of f_s belongs to A.

A local version can be given by replacing X with a neighborhood of $\overline{x}$.

Proof Since $f_s = f$ on A, we have $m := \inf f(A) \geq \inf f_s(X)$. If we had strict inequality we could find $x \in X$ such that $f_s(x) < m$. Then we would have $sd_A(x) < m - f(x)$, so that we could pick $x' \in A$ such that $sd(x,x') < m - f(x)$. Since f is Lipschitzian with rate $r \leq s$ we would get $f(x') \leq f(x) + sd(x,x') < m$, a contradiction. The second assertion follows from (2.3).

Suppose now that A is closed and that for some $s > r$ a minimizer z of f_s does not belong to A. Then $d_A(z)$ is positive, so that, by the relations $\inf f(A) = \inf f_s(X) = f(z) + sd_A(z)$,

$$rd_A(z) < sd_A(z) = \inf f(A) - f(z)$$

one can find $a \in A$ such that $rd(a,z) < \inf f(A) - f(z)$, contradicting the relations $\inf f(A) = \inf f_r(X) \leq f(z) + rd(a,z)$. □

When the admissible set A is defined by equalities or inequalities, it is sensible to take these relations into account. In the case when the admissible set A is defined as $A := g^{-1}(C)$, where $g : X \to W$ is a map with values in another metric space (W,d') and C is a closed subset of W such that for some $c \in \mathbb{P}$ one has $d_A(x) \leq cd'(g(x),C)$

for all $x \in X$, one can replace f_s with $f + scd'(g(x), C)$. Since d_C is often easier to compute than the distance to the implicitly defined set $A := g^{-1}(C)$, as is the case when $W := \mathbb{R}^m$, $C := \mathbb{R}^m_-$, such a penalized problem is often more tractable.

Exercises

1. A function $h : \mathbb{R}_+ \to \mathbb{R}_+$ is said to be *subadditive* if it satisfies $h(r+s) \leq h(r) + h(s)$ for all $r, s \in \mathbb{R}_+$. Let H be the set of subadditive functions $h : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $h(0) = 0$ and $h(r) > 0$ for all $r > 0$. Verify that for $h, k \in H$ one has $h + k \in H$, $h \vee k \in H$, $h \circ k \in H$ and that H is stable under pointwise limits and suprema. Prove that if $h : \mathbb{R}_+ \to \mathbb{R}_+$ is concave, increasing and such that $h(0) = 0$, then $h \in H$. Show that for every metric d on a set X, the function $h \circ d$ is a metric on X. Using the functions $r \mapsto r/(1+r)$ and $r \mapsto \min(r, 1)$ show that any metric d on X is uniformly equivalent to a bounded metric.
2. Give examples of disjoint closed subsets A, B of a metric space (X, d) such that $\mathrm{gap}(A, B) = 0$. Show that the triangle inequality $\mathrm{gap}(A, C) \leq \mathrm{gap}(A, B) + \mathrm{gap}(B, C)$ is not valid.
3. (**Hausdorff-Pompeiu metric**) Let (X, d) be a metric space and let $\mathcal{B}_0$ be the set of nonempty bounded subsets of X. For $A, B \in \mathcal{B}_0$ let $\mathrm{e}(A, B)$ be the *excess* of A over B defined by $\mathrm{e}(A, B) := \sup\{d(a, B) : a \in A\}$. Verify that $\mathrm{e}(A, B) = 0$ if and only if $A \subset \mathrm{cl}(B)$. Verify that $d_H : (A, B) \mapsto \max(\mathrm{e}(A, B), \mathrm{e}(B, A))$ is a semimetric on $\mathcal{B}_0$ and a metric on the set $\mathcal{F}_b$ of nonempty closed bounded subsets of X.
4. Let (X, d) a metric space and let $(Y, \mathcal{O}_Y)$ be a topological space. Show that a map $f : X \to Y$ is continuous at $\overline{x} \in X$ if and only if for any sequence $(x_n) \to \overline{x}$ one has $(f(x_n)) \to f(\overline{x})$.
5. Prove that if $f : \mathbb{R}^d \to \mathbb{R}$ is uniformly continuous, then there exist $a, b \in \mathbb{R}_+$ such that $f(x) \leq ad(x, 0) + b$ for all $x \in \mathbb{R}^d$.
6. Let S be a subset of a metric space (X, d). For $r > 0$ let $B(S, r) := \{x \in X : d(x, S) < r\}$. Verify that $B(S, r)$ is the union over $x \in S$ of the balls $B(x, r)$ and that $\cap_{r>0} B(S, r) = \mathrm{cl}(S)$. Examine whether similar conclusions hold for $B[S, r] := \{x \in X : d(x, S) \leq r\}$.
7. Let (M, d) be a metric space in which closed balls are compact. Suppose X is arcwise connected. Show that any pair of points x_0, x_1 in X can be joined by a curve with least length (a so-called *geodesic*) [see [75]]. Identify such a curve when M is the unit sphere $\mathbb{S}^{d-1}$ of $\mathbb{R}^d$ and when M is a circular cylinder in $\mathbb{R}^3$. Such curves have prompted the development of differential geometry.
8. Given a metric space (M, d), let X be the set of subsets of M. For a net $(S_i)_{i \in I}$ in X define

$$\liminf_{i \in I} S_i := \{x \in M : \lim_{i \in I} d(x, S_i) = 0\},$$

$$\limsup_{i \in I} S_i := \{x \in M : \liminf_{i \in I} d(x, S_i) = 0\}$$

and write $(S_i)_{i\in I} \to S$ if $\liminf_{i\in I} S_i = \limsup_{i\in I} S_i = S$. Verify the axioms of convergence spaces.

9. Devise another proof of the Tietze-Urysohn extension theorem that gives an explicit expression for the extension g of f in the case $\inf f(C) = 1$, $\sup f(C) = 2$:

$$g(x) := \frac{1}{d(x, C)} \inf_{w\in C} f(w)d(w, x) \qquad x \in X\backslash C$$

and $g(x) = f(x)$ for $x \in C$. Show that g is well defined on X, satisfies $\inf g(X) = \inf f(C)$, $\sup g(X) = \sup f(C)$, and is continuous. [Hint: note that $x \mapsto \inf_{w\in C} f(w)d(w, x)$ is Lipschitzian with rate 2 and that for $x \in X\backslash C$ and $D(x) := C \cap B(x, 2d(x, C))$ one has $g(x) = (1/d(x, C)) \inf_{w\in D(x)} f(w)d(w, x)$.]

10*. Prove that any continuous function f on a metric space (X, d) can be uniformly approximated by a locally Lipschitz function [see [133].]

11. Verify the following counterexample showing that a continuous function f on a metric space (X, d) cannot always be uniformly approximated by a Lipschitz function: take $X := \mathbb{R}$ with its usual distance and f continuous such that $f(n) = 0$ for $n \in \mathbb{N}$, $f(n + r_n) = 1$ for $n \in \mathbb{N}$, where (r_n) is a sequence in $]0, 1[$) with limit 0. This counterexample and the preceding reference have been communicated to the author by G. Beer.

12. Let (X, d) be a metric space, let $s \in X$ and for a subset T of X let

$$V_s(T) := \{x \in X : d(x, s) \leq d(x, t)\ \forall t \in T\}$$

be the Voronoi cell associated with s and T (with $V_s(\varnothing) := X$ and $f := d$). Verify that $s \in \mathrm{int} V_s(T)$ if and only if $s \in X\backslash \mathrm{cl}(T)$.

13. Let (X, d) be a metric space such that for all $w, x \in X$ and $r > 0$ such that $r < d(w, x)$ one has $d(x, B[w, r]) < d(x, w)$. Using the notation of the preceding exercise, show that when $s \in X\backslash \mathrm{cl}(T)$ one has $V_s(T) = V_s(\mathrm{bdry}(T))$.

When $s \in \mathrm{cl}(T)$ show that the relation $V_s(T) = V_s(\mathrm{bdry}(T))$ may or may not hold. [Hint: for $X := \mathbb{R}^2$, $s = (0, 0)$, $T = \mathbb{R}^2_+$ one has $V_s(T) = V_s(\mathrm{bdry}(T)) = \mathbb{R}^2_-$ whereas for $T := \mathbb{R} \times \mathbb{R}_+$ one has $V_s(T) = \{0\} \times \mathbb{R}_-$, $V_s(\mathrm{bdry}(T)) = \{0\} \times \mathbb{R}$.]

Show that the assumption on (X, d) is satisfied whenever for any $w, x \in X$ with $w \neq x$ there exists a connected set S containing w and x such that $d(w, z) + d(z, x) = d(w, x)$ for all $z \in S$.

14. (**McShane**, **Whitney**, 1934) Let W be a nonempty subset of a metric space (X, d) and let $f : W \to \mathbb{R}$ be a Lipschitz function with rate r. Show that the functions $f^\flat, f^\# : X \to \mathbb{R}$ defined by

$$f^\flat(x) := \inf_{w\in W} (f(w) - rd(w, x)), \qquad f^\#(x) := \sup_{w\in W} (f(w) + rd(w, x))$$

are Lipschitzian with rate r and extend f. Prove that any Lipschitzian extension g of f with rate r satisfies $f^\flat \leq g \leq f^\#$.

2.3.2 Complete Metric Spaces

The structure of metric space is richer than the structure of topological space. In particular, one disposes of the notion of a Cauchy sequence: a sequence (x_n) of (X, d) is called a *Cauchy sequence* if $(d(x_n, x_p)) \to 0$ as $n, p \to +\infty$. A metric space is said to be *complete* if its Cauchy sequences are convergent. The interest of such a notion is the fact that one can assert the convergence of such a sequence without knowing its limit. The following result is worth noting; its proof is immediate.

Proposition 2.24 *If a Cauchy sequence (x_n) in a metric space (X, d) has a converging subsequence, then (x_n) is converging.*

It is often convenient to replace Cauchy sequences with more special sequences. We call a sequence (x_n) in a metric space (X, d) an *Abel sequence* if there exists some $c \in \mathbb{R}_+$ and some $r \in]0, 1[$ such that $d(x_n, x_{n+1}) \leq cr^n$ for all $n \in \mathbb{N}$. We observe that Abel sequences can be a substitute to Cauchy sequences in view of the following lemma, the easy proof of which is left to the reader.

Lemma 2.4 *Any Abel sequence in a metric space is a Cauchy sequence.*
Any Cauchy sequence has a subsequence that is an Abel sequence.

Corollary 2.14 *A metric space (X, d) is complete if and only if any Abel sequence in X is convergent.*

Let us give some permanence properties.

Proposition 2.25 *If (X, d_X) is a subspace of a metric space (W, d_W) with the induced metric and if (X, d_X) is complete, then X is closed in W.*

Conversely, any closed subset X of a complete metric space (W, d_W) is complete with respect to the induced metric.

Proof Let us show that any point $\overline{w}$ in the closure $\mathrm{cl}(X)$ of X belongs to X. Proposition 2.20 ensures that some sequence (x_n) of X converges to $\overline{w}$. Such a sequence is a Cauchy sequence, hence has a limit $\overline{x} \in X$. Uniqueness of limits in metric spaces implies that $\overline{w} = \overline{x} \in X$.

For the converse, let X be a closed subset of a complete metric space (W, d_W). Since a Cauchy sequence in X with respect to the induced metric d_X is a Cauchy sequence in (W, d_W), it converges in W, and in fact in X since X is closed. Thus X is complete with respect to d_X. □

Other permanence properties concern function spaces. It is extremely useful to consider metric spaces formed with functions or maps.

Proposition 2.26 *Let S be a set and let (M, d) be a metric space. On the set $B(S, M)$ of $f : S \to M$ that are bounded, i.e. such that $f(S)$ is bounded in (M, d), one can* *define a metric by setting for* $f, g \in B(S, M)$

$$d_\infty(f, g) := \sup_{s \in S} d(f(s), g(s)).$$

If (M, d) *is complete,* $(B(S, M), d_\infty)$ *is complete.*

Proof It is easy to see that d_∞ is a metric on $B(S, M)$. Let $(f_n)_n$ be a Cauchy sequence in $(B(S, M), d_\infty)$. Since for every $x \in S$ the evaluation map $f \mapsto f(x)$ is nonexpansive, the sequence $(f_n(x))_n$ is a Cauchy sequence in (M, d). When (M, d) is complete, this sequence converges. Let $f(x)$ be its limit. If $\varepsilon \mapsto k(\varepsilon)$ is such that $d_\infty(f_m, f_n) \le \varepsilon$ for $n \ge m \ge k(\varepsilon)$, passing to the limit as $n \to \infty$ we see that $d(f_m(x), f(x)) \le \varepsilon$ for $m \ge k(\varepsilon)$ and all $x \in S$. Equivalently we have $d_\infty(f_m, f) \le \varepsilon$ for $m \ge k(\varepsilon)$, so that $(f_n) \to f$ in $(B(S, M), d_\infty)$ since f is bounded as $\sup_{x,x' \in S} d(f(x), f(x')) \le \sup_{x,x' \in S} d(f_m(x), f_m(x')) + 2\varepsilon < +\infty$. □

When a sequence (f_n) converges to f for d_∞ one says that (f_n) *converges uniformly* to f. This property is stronger than *pointwise convergence*, which means that for all $x \in S$ one has $(f_n(x))_n \to f(x)$.

Proposition 2.27 *Let* X *be a topological space and let* (M, d) *be a metric space. The subspace* $C_b(X, M)$ *of bounded continuous functions from* X *to* M *is closed in* $B(X, M)$ *with respect to the metric* d_∞*. Thus, if* (M, d) *is complete,* $(C_b(X, M), d_\infty)$ *is complete.*

Proof We have to prove that the uniform limit f of a sequence (f_n) in $C_b(X, M)$ belongs to $C_b(X, M)$. We know from the preceding proof that f is bounded. The continuity of f is established in Theorem 2.7. □

Completeness can be used with great success for extension results.

Theorem 2.9 *Let* (W, d_W) *and* (Y, d_Y) *be complete metric spaces and let* $f : X \to Y$ *be uniformly continuous, where* X *is a dense subset of* W*. Then* f *can be extended uniquely to a uniformly continuous map* $\overline{f} : W \to Y$*.*

Proof Uniqueness of the extension stems from Proposition 2.4. Let us prove the existence of a uniformly continuous extension. Let $m : \mathbb{R}_+ \to \overline{\mathbb{R}}_+$ be a modulus of uniform continuity of f in the sense that $d_Y(f(x), f(x')) \le m(d_W(x, x'))$ for all $x, x' \in X$. Given $w \in W$, let (x_n) be a sequence in X with limit w. Since (x_n) is a Cauchy sequence in X, $(f(x_n))$ is a Cauchy sequence in (Y, d_Y). Since (Y, d_Y) is complete, $(f(x_n))$ has a limit $y \in Y$. This limit does not depend on the choice of the sequence (x_n) since given two such sequences (x_n), (x'_n) one has $d_Y(f(x_n), f(x'_n)) \le m(d_W(x_n, x'_n))$, so that, passing to the limit, we get $(f(x'_n)) \to y$. Denoting this limit by $\overline{f}(w)$ we define a map $\overline{f} : W \to Y$. For all $t > 1$ this map satisfies $d_Y(\overline{f}(w), \overline{f}(w')) \le m(td_W(w, w'))$ since for any sequences $(x_n) \to w$, $(x'_n) \to w'$ in X one has $d_W(x_n, x'_n) \le td_W(w, w')$ for n large enough, so that $d_Y(\overline{f}(w), \overline{f}(w')) = \lim_n d_Y(f(x_n), f(x'_n)) \le m(td_W(w, w'))$. Let us observe that when m is continuous (and not just continuous at 0), m is also a modulus of uniform continuity of the map $\overline{f}$. □

The notion of a complete metric space enables us to present a powerful fixed point theorem. Here we deal with *contraction maps* , i.e. Lipschitzian maps whose Lipschitz rate is less than 1 and the process is called the method of *successive approximations*.

Theorem 2.10 (Contraction Theorem, Picard, Banach) *Let (X, d) be a complete metric space and let $f : X \to X$ be a contraction. Then f has a fixed point $\overline{x}$. Moreover, $\overline{x}$ is unique and for every $x_0 \in X$ one has $(f^{(n)}(x_0)) \to \overline{x}$ for $f^{(n)} := f \circ f^{(n-1)}$.*

Proof Let $c \in [0, 1[$ be the Lipschitz rate of f. Uniqueness of the fixed point stems from the contraction property: if $\overline{x}$ and $\overline{y}$ satisfy $f(\overline{x}) = \overline{x}, f(\overline{y}) = \overline{y}$ we have $d(\overline{x}, \overline{y}) = d(f(\overline{x}), f(\overline{y})) \leq cd(\overline{x}, \overline{y})$ and as $c \in [0, 1[$ this is possible only if $d(\overline{x}, \overline{y}) = 0$, i.e. $\overline{x} = \overline{y}$.

Given $x_0 \in X$ let us define inductively $x_{n+1} = f(x_n)$ for $n \in \mathbb{N}$. Since for $n \geq 1$ we have $x_n = f(x_{n-1})$ and $x_{n+1} = f(x_n)$, the contraction property yields

$$d(x_{n+1}, x_n) \leq cd(x_n, x_{n-1}).$$

By induction, this relation entails that $d(x_{n+1}, x_n) \leq c^n d(x_1, x_0)$ for all $n \in \mathbb{N}$. The sequence (x_n) is thus an Abel sequence (hence a Cauchy sequence), hence has a limit $\overline{x} \in X$. Passing to the limit in the relation $d(f(x_n), x_n) \leq c^n d(x_1, x_0)$ we get $d(f(\overline{x}), \overline{x}) = 0$, hence $f(\overline{x}) = \overline{x}$. □

Let us observe that the convergence of (x_n) to $\overline{x}$ is rapid enough since

$$d(x_{n+p}, x_n) \leq \sum_{k=n+1}^{n+p} d(x_k, x_{k-1}) \leq d(x_1, x_0) \sum_{k=n}^{\infty} c^k = \frac{d(x_1, x_0)}{1-c} c^n \tag{2.4}$$

and, passing to the limit as $p \to +\infty$, $d(\overline{x}, x_n) \leq (1-c)^{-1} c^n d(f(x_0), x_0)$.

Corollary 2.15 *Let W be a topological space, let (X, d) be a complete metric space and let $f : W \times X \to X$ be a continuous map such that for some $c \in [0, 1[$ and all $w \in W$ the partial map $f_w : x \mapsto f(w, x)$ is Lipschitzian with rate c. Then there exists a unique continuous map $g : W \to X$ such that $g(w) = f(w, g(w))$ for all $w \in W$.*

Proof For $w \in W$, let $g(w)$ be the unique fixed point of f_w. We have to show that g is continuous. Let $z \in W$ and given $\varepsilon > 0$ let $V \in \mathcal{N}(z)$ be such that $d(f(w, g(z)), f(z, g(z))) \leq \varepsilon$ for all $w \in V$. The triangle inequality yields for $w \in V$

$$\begin{aligned} d(g(w), g(z)) &= d(f_w(g(w)), f_z(g(z))) \\ &\leq d(f_w(g(w)), f_w(g(z))) + d(f_w(g(z)), f_z(g(z))) \\ &\leq cd(g(w), g(z)) + \varepsilon, \end{aligned}$$

so that $d(g(w), g(z)) \leq (1-c)^{-1}\varepsilon$. This shows that g is continuous at z. □

A remarkable property of complete metric spaces is the Baire property.

Theorem 2.11 (Baire) *Let* (X, d) *be a complete metric space.*

If (G_n) *is a sequence in open dense subsets of* X*, then* $\cap_n G_n$ *is dense.*

If X *is the union of a countable family* (F_n) *of closed subsets, then one of them has a nonempty interior.*

Proof Let (G_n) be a sequence of open dense subsets of X. Let us show that $G := \cap_n G_n$ is dense. Let (s_n) be a sequence of positive numbers with limit 0. Given a nonempty open subset U of X, the set $G_n \cap U$ is nonempty and open; in particular, $G_0 \cap U$ contains some closed ball $B[x_0, r_0]$ with $r_0 \in]0, s_0]$. Assume by induction that we have constructed open balls $B(x_k, r_k)$ with $r_k \leq s_k$, $B[x_k, r_k] \subset B(x_{k-1}, r_{k-1}) \cap G_k$ for $k = 1, \ldots, n$. Since $B(x_n, r_n)$ meets G_{n+1}, we can find a closed ball $B[x_{n+1}, r_{n+1}] \subset G_{n+1} \cap B(x_n, r_n)$ with $r_{n+1} \leq s_{n+1}$. The sequence (x_n) obtained in this way is a Cauchy sequence (since $d(x_{n+p}, x_n) \leq s_n$ for all n, p). Its limit $\overline{x}$ belongs to $B[x_m, r_m] \subset G_m$ for all $m \in \mathbb{N}$ and in particular $\overline{x} \in B[x_0, r_0] \subset U$ and $\overline{x} \in G$, so that $G \cap U$ is nonempty: G is dense.

Now suppose $X = \cup_n F_n$, where each F_n is closed. Then $G_n := X \backslash F_n$ is open and if F_n has an empty interior then G_n is dense. If this happens for all $n \in \mathbb{N}$, then $\cap_n G_n$ is dense, an impossibility since $\cap_n G_n = \varnothing$. Thus, at least one F_n has a nonempty interior. □

The preceding result can be expressed in terms of genericity. A subset G of some topological space T is *generic* if it contains the intersection of a countable family of open subsets of T (a so-called $\mathcal{G}_\delta$ set, the notation being a reminder of the German term "Gebiete", while the notation F stems from the French "fermé" for closed) that are dense in T; other terminologies are that G is *residual* or that the complement of G is *meager* or a *set of first category*. It is convenient to say that a property involving a point is generic if it holds on a generic subset. The main feature of this notion is that the intersection of a finite (or countable) family of generic subsets is still generic, a property that does not hold for dense subsets (consider the set of rational numbers and the set of irrational numbers in $\mathbb{R}$). The (equivalent) properties of Theorem 2.11 can be phrased as follows: in a complete metric space any generic subset is dense. A topological space satisfying this property is called a Baire space. Locally compact topological spaces are also Baire spaces.

Complete metric spaces can be characterized by an approximate minimization principle that is extremely useful. Here, given $\varepsilon > 0$, we say that a point $\overline{x}$ of a set X is an *ε-minimizer* of a function $f : X \to \mathbb{R}_\infty$ if $f(\overline{x}) \leq \inf f(X) + \varepsilon$.

Theorem 2.12 (Ekeland) *Let* (X, d) *be a complete metric space and let* $f : X \to \mathbb{R}_\infty$ *be a bounded below lower semicontinuous function taking at least one finite value. Given* $\varepsilon > 0$*, an* ε*-minimizer* $\overline{x}$ *of* f*, and given* $c, r > 0$ *satisfying* $cr \geq \varepsilon$*, one can find* $u \in B[\overline{x}, r]$ *such that* $f(u) + cd(u, \overline{x}) \leq f(\overline{x})$ *and*

$$f(u) < f(x) + cd(u, x) \quad \textit{for all } x \in X \backslash \{u\}. \tag{2.5}$$

Thus, not too far from $\overline{x}$, we can find a strict mimimizer u of the modified function $f_{c,u} := f + cd(u, \cdot)$. Replacing the metric d with the metrics $d' := d/(1 + d)$ or $d'' := \min(d, 1)$ and c with the general term of a sequence $(c_n) \to 0_+$, we can ensure that the approximate function $f_{c,u}$ is as close to f as required with respect to the uniform metric. However, there is a trade off between the accuracies of the two approximating elements u, $f_{c,u}$: one cannot expect to get arbitrarily accurate approximations of f and of $\overline{x}$ at the same time.

Proof We associate to f and c an order on X defined by $w \preceq x$ if $f(w) + cd(w, x) \leq f(x)$. Let

$$B(x) := \{w \in X : f(w) + cd(w, x) \leq f(x)\} \qquad x \in X$$

be the set of elements below x. We have $x \in B(x)$ for all $x \in X$ and the relations $y \in B(x)$, $x \in B(y)$ imply $d(x, y) = 0$ or $x = y$. Let us verify that the relation B satisfies the transitivity property $B(y) \subset B(x)$ for all $x \in X$, $y \in B(x)$. We may assume $x \in \operatorname{dom} f := f^{-1}(\mathbb{R})$, so that $f(y) < \infty$. Then, for all $z \in B(y)$ we also have $f(z) < \infty$ and $cd(y, z) \leq f(y) - f(z)$. Since $y \in B(x)$, we also have $cd(x, y) \leq f(x) - f(y)$. Adding the respective sides of these two inequalities, and using the triangle inequality, we get $cd(x, z) \leq f(x) - f(z)$, or $z \in B(x)$. Thus B defines an order; we shall construct a minimal element.

Given $\overline{x} \in \operatorname{dom} f$, we can define inductively a sequence starting from $x_0 := \overline{x}$ by picking $x_{n+1} \in B(x_n)$ satisfying

$$f(x_{n+1}) \leq \frac{1}{2} f(x_n) + \frac{1}{2} \inf f(B(x_n)). \tag{2.6}$$

Such a choice is possible: it suffices to use the definition of an infimum when $\inf f(B(x_n)) < f(x_n)$ and to take $x_{n+1} = x_n$ when $\inf f(B(x_n)) = f(x_n)$. Since $x_n \in B(x_n)$, (2.6) ensures that the sequence $(f(x_n))$ is nonincreasing, hence is convergent as f is bounded below. Let $\ell := \lim_n f(x_n)$.

Since $x_{n+1} \in B(x_n)$ we have $cd(x_n, x_{n+1}) \leq f(x_n) - f(x_{n+1})$ and by induction

$$cd(x_n, x_{n+p}) \leq f(x_n) - f(x_{n+p}) \tag{2.7}$$

for all $n, p \geq 0$. Thus (x_n) is a Cauchy sequence, hence has a limit we denote by u.

Because f is lower semicontinuous, for each $n \in \mathbb{N}$ the set $B(x_n)$ is closed. Since relation (2.7) says that $x_{n+p} \in B(x_n)$ for all $p \geq 0$, we get $u \in B(x_n)$. In particular, taking $n = 0$ and remembering that $x_0 = \overline{x}$, we get

$$f(u) + cd(\overline{x}, u) \leq f(\overline{x}).$$

Moreover, by the transitivity property of relation B, for all $v \in B(u)$ and all $n \in \mathbb{N}$, we have $v \in B(x_n)$. Thus $\inf f(B(x_n)) + cd(x_n, v) \leq f(v) + cd(x_n, v) \leq f(x_n)$ and relation (2.6) yields

$$cd(x_n, v) \leq f(x_n) - \inf f(B(x_n)) \leq 2\,(f(x_n) - f(x_{n+1})) \to 0,$$

hence $d(v,u) = 0$. It follows that $B(u) = \{u\}$. This relation means that (2.5) is satisfied.

If $\overline{x}$ is such that $f(\overline{x}) \leq \inf f(X) + \varepsilon$, and $\varepsilon \leq cr$, we have

$$\inf f(X) + cd(u,\overline{x}) \leq f(u) + cd(u,\overline{x}) \leq f(\overline{x}) \leq \inf f(X) + cr,$$

so that $d(u,\overline{x}) \leq r$. □

Given a metric space (X,d) one may look for a complete metric space $(\widehat{X},\hat{d})$ that is close enough to (X,d). A precise answer can be given, even for semimetric spaces.

Theorem 2.13 *Given a semimetric space (X,d) there is a complete semimetric space $(\widehat{X},\hat{d})$ and an isometry j from X onto a dense subset of $(\widehat{X},\hat{d})$. Moreover, if (Y,d_Y) is a complete metric space and if $f : X \to Y$ is a uniformly continuous map, there is a unique uniformly continuous map $\hat{f} : \widehat{X} \to Y$ such that $f = \hat{f} \circ j$.*

Proof This last property ensures that the pair $(\widehat{X},j)$ is unique up to an isometry. In fact, if $(\widehat{X'},j')$ is another pair, there is a uniformly continuous map $\widehat{j'} : \widehat{X} \to \widehat{X'}$ such that $j' = \widehat{j'} \circ j$. Similarly, there is a uniformly continuous map $\hat{j} : \widehat{X'} \to \widehat{X}$ such that $j = \hat{j} \circ j'$. Then $j = \hat{j} \circ (\widehat{j'} \circ j) = I_{\widehat{X}} \circ j$, so that the two maps $\hat{j} \circ \widehat{j'}$ and $I_{\widehat{X}}$ coincide by the uniqueness requirement (or the density of $j(X)$ in $\widehat{X}$). Similarly, one shows that $\widehat{j'} \circ \hat{j}$ coincides with the identity map $I_{\widehat{X'}}$ on $\widehat{X'}$.

Several specific constructions can be given for $(\widehat{X},j)$. One consists in taking for $\widehat{X}$ the set of equivalence classes of Cauchy sequences for the relation $(x_n) \sim (x'_n)$ if $\lim_n d(x_n,x'_n) = 0$. We invite the reader to complete the construction by defining $\hat{d}$ and taking for $j(x)$ the class of the constant sequence with value x.

When d is a metric that is bounded above on X^2 we can also take an embedding j into the space $C_b(X) := C_b(X,\mathbb{R})$ endowed with the metric d_∞ defined by $d_\infty(f,g) = \sup_{x\in X} |f(x) - g(x)|$. Given $x \in X$ we define $j(x)$ as the function $w \mapsto d(w,x)$ on X. Then, for x, $y \in X$, the triangle inequality yields $d_\infty(j(x),j(y)) \leq d(x,y)$. In fact this inequality is an equality since $|j(x)(y) - j(y)(y)| = d(x,y)$. Taking for $\widehat{X}$ the closure of $j(X)$ in $C_b(X)$ for $\hat{d} := d_\infty$ we get a complete space in which $j(X)$ is dense. The last assertion of the statement is a consequence in Theorem 2.4. When d is an arbitrary metric, turning the requirement that j be isometric into the requirement that j be nonexpansive, i.e. Lipschitzian with rate 1, we can define j by setting $j(x)(w) := \min(d(w,x),1)$. □

If $(d_a)_{a\in A}$ is a family of semimetrics on X, one can consider on X the topology generated by the balls $B_a(x,r) := \{w \in X : d_a(w,x) < r\}$. However, X has a structure richer than a topology. It is called a uniformity and $(X,(d_a)_{a\in A}$ is called a *uniform space*. A sequence (x_n) of X is called a *Cauchy sequence* if for all $a \in A$ one has $(d_a(x_n,x_p)) \to 0$ as $n,p \to +\infty$. A uniform space is said to be complete if any Cauchy sequence is convergent.

2.3.3 Application to Ordinary Differential Equations

We intend to give a simple existence theorem for ordinary differential equations. One of its proofs relies on a generalization of the Contraction Theorem.

Proposition 2.28 *Let (X, d) be a complete metric space and let $f : X \to X$ be a continuous map such that for some $k \in \mathbb{N}\backslash\{0\}$ the k times iterated map $f^{(k)} := f \circ \ldots \circ f$ is a contraction. Then f has a unique fixed point.*

Proof If $\overline{x}$ and $\overline{y}$ are fixed points of f, then they are fixed points of $f^{(k)}$, hence $\overline{x} = \overline{y}$. Let $\overline{x}$ be the fixed point of $f^{(k)}$. We know that for any $x_0 \in X$ we have $(f^{(kn)}(x_0))_n \to \overline{x}$. In particular, taking $x_0 := f(\overline{x})$ we get that $(f^{(kn+1)}(\overline{x}))_n \to \overline{x}$. On the other hand, since $f^{(kn)}(\overline{x}) = \overline{x}$, we have $f(f^{(kn)}(\overline{x})) = f(\overline{x})$. By uniqueness of limits, we get $f(\overline{x}) = \overline{x}$. □

For this existence result, we anticipate to some notions of derivation and integration to be found later. The reader may suppose $E := \mathbb{R}$ for the sake of simplicity.

Theorem 2.14 *Let T be a bounded interval of $\mathbb{R}$, let E be a Banach space and let $f : T \times E \to E$ be a continuous map such that for some $c \in \mathbb{R}_+$ one has $\|f(t, e) - f(t, e')\| \leq c \|e - e'\|$ for all $t \in T$, e, $e' \in E$. Then, given $t_0 \in T$, $e_0 \in E$ there exists a unique solution $x \in C(T, E)$ with a continuous derivative x' to the equation*

$$x'(t) = f(t, x(t)) \quad t \in T, \qquad x(t_0) = e_0. \tag{2.8}$$

Proof We admit that $x \in C(T, E)$ satisfies (2.8) if and only if x satisfies the integral equation

$$x(t) = e_0 + \int_{t_0}^{t} f(s, x(s))ds. \tag{2.9}$$

Let us endow the space $X := C_b(T, E)$ of bounded continuous maps from T to E with the norm $\|\cdot\|_\infty$ and let us consider the map $F : X \to X$ given by

$$F(x)(t) := e_0 + \int_{t_0}^{t} f(s, x(s))ds.$$

For $x, y \in C_b(T, E)$ we have

$$\begin{aligned}\|F(x) - F(y)\|_\infty &= \sup_{t\in T} \left\| \int_{t_0}^{t} (f(s, x(s)) - f(s, y(s)))ds \right\| \\ &\leq \sup_{t\in T} \left| \int_{t_0}^{t} c \|x(s) - y(s)\| \, ds \right| \leq \ell(T)c \|x - y\|_\infty\end{aligned}$$

where $\ell(T)$ is the length of T. Thus, for $\ell(T)c < 1$, F is a contraction and F has a fixed point that is a solution to (2.9). In the general case it can be shown by induction that for all $k \in \mathbb{N}\backslash\{0\}$, $x, y \in C_b(T, E)$ one has

$$\left\| F^{(k)}(x) - F^{(k)}(y) \right\|_\infty \leq \frac{c^k \ell(T)^k}{k!} \left\| x - y \right\|_\infty .$$

For k large enough one has $c^k \ell(T)^k / k! < 1$ and we can apply the preceding proposition to find a fixed point of F, hence a solution to (2.8). □

Exercises

1. Observe that the proof we gave of Corollary 2.15 only uses the continuity of the partial map $w \mapsto f(w, g(z))$. Prove that such an assumption is equivalent to the continuity of f at $(z, g(z))$ in view of the hypothesis that for all $w \in W$ the map f_w is Lipschitzian with rate c.

2. Let (X, d) be a compact metric space and let $f : X \to X$ be a map such that for $x, x' \in X$ with $x \neq x'$ one has $d(f(x), f(x')) < d(x, x')$. Show that f has a unique fixed point that can be found by the method of successive approximations.

3. Let (X, d) be a complete metric space, let $c > 1$, and let $f : X \to X$ be a continuous map such that for $x, x' \in X$ one has $d(f(x), f(x')) \geq cd(x, x')$. Show that f has a unique fixed point.

4. Let $T := [a, b]$, let $X := L_2(T)$, let $k \in L_2(T \times T)$ be such that $\int_{T^2} k^2(s, t)dsdt < 1$, and let $f : \mathbb{R} \times T^2 \to \mathbb{R}$ be such that for all $(r, r', s, t) \in \mathbb{R}^2 \times T^2$ one has $|f(r, s, t) - f(r', s, t)| \leq k(s, t)\,|r - r'|$ and such that for all $x \in X$ the function $t \mapsto \int_a^b f(x(s), s, t)ds$ belongs to X. Prove that for every $y \in X$ the following *integral equation* has a unique solution

$$x(t) = y(t) + \int_a^b f(x(s), s, t)ds.$$

5. Let X be an open subset of a complete metric space (W, d), with $X \neq W$. For x, $x' \in X$, let

$$d_X(x, x') := \left| 1/d(x, W\backslash X) - 1/d(x', W\backslash X) \right| .$$

Show that d_X is a metric on X whose associated topology is the induced topology. Prove that (X, d_X) is complete. Does this contradicts the fact that in general X is not closed in W?

6. Show that on $\mathbb{R}$ the function $d' : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $d'(w, x) = \left| w^3 - x^3 \right|$ is a metric topologically equivalent to the usual metric d given by $d(w, x) := |w - x|$, but not uniformly equivalent to d. Verify that the Cauchy sequences for d and d' are the same.

7. Let (X, d) be a complete metric space and let $f, g : X \to X$ be two contractions with rate $c \in]0, 1[$. Let $\overline{x}$ (resp. $\overline{y}$) be the fixed point of f (resp. g). Prove that $d(\overline{x}, \overline{y}) \leq (1-c)^{-1} d_\infty(f, g)$ where $d_\infty(f, g) := \sup_{x \in X} d(f(x), g(x))$. [Hint: In relation (2.4) take $n = 0, x_0 := \overline{y}$, pass to the limit on p and note that $d(f(\overline{y}), \overline{y}) \leq d_\infty(f, g)$.]

2.3.4 Compact Metric Spaces

Some additional properties of compact spaces can be obtained when they are *metrizable*, i.e. when their topologies can be associated with metrics. They are even valid for sequentially compact spaces, a topological space X being called *sequentially compact* if every sequence in X has a convergent subsequence. First we note a uniformity property related to coverings.

Lemma 2.5 (Lebesgue) *Let (W, d) be a metric space and let X be a subset of W that is sequentially compact. Then for any family $(U_i)_{i \in I}$ of open subsets of W whose union contains X there exists some $r > 0$ such that for all $x \in X$ the ball $B(x, r)$ is contained in some U_i.*

Proof If no such r exists, for all $n \in \mathbb{N}\backslash\{0\}$ one can find some $x_n \in X$ such that for all $i \in I$ the ball $B(x_n, 1/n)$ is not contained in U_i. Let $\overline{x} \in X$ be the limit of a subsequence $(x_{k(n)})_n$ of (x_n). Let $r_n := d(x_{k(n)}, \overline{x})$ and let $j \in I$ be such that $\overline{x} \in U_j$. Let $s > 0$ be such that $B(\overline{x}, s) \subset U_j$. For n large enough we have $r_n + 1/k(n) < s$ so that $B(x_{k(n)}, 1/k(n))$ is contained in $B(\overline{x}, s)$, hence in U_j. This contradicts the choice of the balls $B(x_n, 1/n)$. □

The second property we consider captures the idea that such a space can be approximated by a finite set.

Definition 2.12 A metric space (X, d) is said to be *precompact* if for any $\varepsilon > 0$ there exists a finite subset F_ε of X such that for all $x \in X$ one has $d(x, F_\varepsilon) < \varepsilon$.

This means that for all $\varepsilon > 0$ there is a covering of X by a finite number of balls of radius ε. Such a space is clearly bounded and separable. But we have more.

Theorem 2.15 *For a metric space (X, d) the following properties are equivalent:*

(a) every sequence (x_n) of X has a cluster point;
(b) X is sequentially compact;
(c) X is precompact and complete;
(d) every infinite subset S of X has an accumulation point;
(e) X is compact.

Proof The equivalence (a)⇔(b) stems from Proposition 2.21 and the implication (e)⇒(a) is obvious. Let us prove (b)⇒(c). If X is sequentially compact then X is complete since every Cauchy sequence in X has a convergent subsequence, hence is convergent. If X is not precompact there exists some $\varepsilon > 0$ such that X is not covered by a finite number of balls of radius ε. Thus, by induction we build a sequence (x_n) such that, for all $n \in \mathbb{N}$, x_{n+1} is not contained in $B(x_0, \varepsilon) \cup \ldots \cup B(x_n, \varepsilon)$. Such a sequence cannot have a convergent subsequence. Now let us show that (c)$\Rightarrow$(d). For every $n \in \mathbb{N}$ there exists a finite subset F_n of X such that the balls with center in F_n and radius 2^{-n} cover X, hence covers S. One of these balls contains an infinite number of points of S. Let us denote by $x_n \in F_n$ its center. Similarly there exists some $x_{n+1} \in F_{n+1}$ such that $B(x_{n+1}, 2^{-n-1})$ contains an infinite number of points of $S \cap B(x_n, 2^{-n})$. By the triangle inequality we have $d(x_{n+1}, x_n) < 2^{-n+1}$. Since X is complete, the sequence (x_n) built in this way converges to some $\overline{x} \in X$ since it is a Cauchy sequence (in fact an Abel sequence). Again, the triangle inequality shows that $\overline{x}$ is an accumulation point of S. The implication (d)$\Rightarrow$(a) is immediate: given a sequence (x_n) of X either $S := \{x_n : n \in \mathbb{N}\}$ is finite and then a subsequence of (x_n) is constant, hence convergent, else S is infinite and any accumulation point a of S is a cluster point of (x_n) since we know that any neighborhood V of a contains an infinite number of points of S. It remains to show that (a)$\Rightarrow$(e). Let $(U_i)_{i \in I}$ be an open covering of X. The preceding lemma yields some $r > 0$ such that for all $x \in X$ there exists some $i(x) \in I$ such that $B(x, r) \subset U_{i(x)}$. On the other hand, since (a)$\Rightarrow$(c), X is precompact and there exists a finite subset F of X such that $\{B(x, r) : x \in F\}$ is a covering of X. Then $\{U_{i(x)} : x \in F\}$ is a finite covering of X: X is compact. □

Corollary 2.16 *A compact metric space is separable.*

Proof Let X be a compact metric space. Since X is precompact, for any sequence $(r_n) \to 0_+$ one can find a finite subset F_n of X such that $\{B(x, r_n) : x \in F_n\}$ is a covering of X. Then $D := \cup_n F_n$ is a dense countable subset of X. □

Another important consequence of compactness and metrizability follows.

Theorem 2.16 *Let (W, d_W) be a metric space and let X be a relatively compact subset of W. If $f : W \to Y$ is a continuous map with values in another metric space (Y, d_Y), then f is uniformly continuous around X in the following sense: for every $\varepsilon > 0$ there exists some $\delta > 0$ such that for all $w \in W$, $x \in X$ satisfying $d_W(w, x) < \delta$ one has $d_Y(f(w), f(x)) < \varepsilon$.*

Of course, if $X = W$ the conclusion is just usual uniform continuity of f.

Proof Given $\varepsilon > 0$, since f is continuous, for all $z \in \mathrm{cl}(X)$ there exists some open neighborhood U_z of z such that $f(U_z) \subset B(f(z), \varepsilon/2)$. Lemma 2.5 yields some $r > 0$ such that for all x in the compact set $\mathrm{cl}(X)$ one has $B(x, r) \subset U_{z(x)}$ for some $z(x) \in \mathrm{cl}(X)$. Then, for $w \in B(x, r)$ one has $d_Y(f(w), f(x)) \le d_Y(f(w), f(z(x))) + d_Y(f(z(x)), f(x)) < \varepsilon$ since $w, x \in B(x, r) \subset U_{z(x)}$. Thus we can take $\delta = r$. □

Exercise Give a proof by contradiction using sequences and subsequences.

Exercise Give a proof using sequences and cluster points.

Exercise Show that the result is valid if X is a semi-metric space and Y is a uniform space.

The compactness assumption in Theorem 2.6 can be relaxed by using a notion of coercivity. Let us say that a function $f : X \to \overline{\mathbb{R}}$ on a metric space X is *coercive* if for all $r \in \mathbb{R}$ the sublevel set $S_f(r) := f^{-1}(]-\infty, r])$ is bounded, or, equivalently, if $f(x) \to \infty$ as $d(x, x_0) \to \infty$ (x_0 being an arbitrary point of X). This notion is essentially used in the case when X is a normed space and $f(x) \to \infty$ as $\|x\| \to \infty$. We will say that f is *compactly coercive* if for all $r \in \mathbb{R}$ the sublevel set $S_f(r)$ is compact. When f is lower semicontinuous and the closed balls of X are compact, both notions coincide. For such a function, the existence of minimizers is ensured.

Lemma 2.6 *Let $f : X \to \mathbb{R}_\infty$ be a compactly coercive function on a metric space X. Then f attains its minimum. In particular, when the closed balls of X are compact, any coercive lower semicontinuous function on X attains its minimum.*

Proof The result being obvious when f only takes the value $+\infty$, let us take $r \in \mathbb{R}$ such that $S_f(r) := f^{-1}(]-\infty, r])$ is nonempty. By assumption, $S_f(r)$ is compact. By the Weierstrass' Theorem, f attains its infimum on $S_f(r)$. Since $\inf f(X) = \inf f(S_f(r))$, any minimizer of the restriction $f \mid S_f(r)$ of f is also a minimizer of f. □

Proposition 2.29 *Let C be a nonempty closed convex subset of a Euclidean space X and let $a \in X$. Then there exists some $p \in C$ called the best approximation of a in C such that $\|p - a\| \leq \|x - a\|$ for all $x \in C$. Moreover, p is characterized by the inequality*

$$\forall x \in C \qquad \langle a - p \mid x - p \rangle \leq 0. \tag{2.10}$$

Proof The function $x \mapsto \|x - a\|$ is continuous and coercive and the closed balls of X are compact since they are dilations of smaller balls and since X is locally compact. Setting $f(x) = \|x - a\|$ if $x \in C$ and $f(x) := +\infty$ if $x \in X \backslash C$, we get a coercive lower semicontinuous function on X. It attains its infimum at some point p of C, so that $\|p - a\| \leq \|x - a\|$ for all $x \in C$. Given $x \in C$, for $t \in]0, 1]$ we have $x_t := p + t(x - p)$ in C by convexity, hence

$$\|p - a\|^2 \leq \|x_t - a\|^2 = \|p - a\|^2 + 2t\langle p - a \mid x - p \rangle + t^2 \|x - p\|^2 .$$

Simplifying both sides and dividing by t and then passing to the limit we get (2.10). □

Exercises

1. Let X be a closed subset of $\mathbb{R}^d$ and let $f : X \to \overline{\mathbb{R}}$ be lower semicontinuous and *pseudo-coercive* in the sense that there exists some $x_0 \in X$ such that $f(x_0) < \liminf_{\|x\| \to \infty,\, x \in X} f(x)$. Show that f attains its infimum.

2. Let X be a closed subset of $\mathbb{R}^d$ and let $f : X \to \overline{\mathbb{R}}$. Assume f is *finitely minimizable* in the sense that there exists an $r \in \mathbb{R}_+$ such that, for any $t > m := \inf f(X)$, there exists some $x \in X$ with $\|x\| \leq r$, $f(x) < t$. Show that any pseudo-coercive function is finitely minimizable and that any finitely minimizable lower semicontinuous function on X attains its infimum at some point of $X \cap B[0, r]$, where r is the radius of essential minimization, i.e. the infimum of the real numbers r for which the above definition is satisfied.

3. Prove Weierstrass' Theorem in the case when X is a compact metric space by using a *minimizing sequence* of f, i.e. a sequence (x_n) of X such that $(f(x_n)) \to \inf f(X)$.

4. Show that any l.s.c. function f on $[0, 1]$ (or on a separable metric space X) is the supremum of the family of continuous functions majorized by f.

5. Prove Corollary 2.12 by using open subsets.

6. Show that among all cylindrical barrels of a given area s there is one with greatest volume.

7*. (**d'Alembert-Gauss Theorem**) Prove that any polynomial P with complex coefficients has at least one root in $\mathbb{C}$. [Hint: verify that $|P(\cdot)|$ is coercive and show that if z_0 is such that $|P(z_0)| = \inf\{|P(z)| : z \in \mathbb{C}\}$, then $P(z_0) = 0$.]

8. Show that the following properties of a locally compact metric space (X, d) are equivalent:

(a) X is the union of an increasing sequence (X_n) of open relatively compact subsets of X such that $\mathrm{cl}(X_n) \subset X_{n+1}$ for all $n \in \mathbb{N}$;
(b) X is the union of a countable family of compact subsets;
(c) X is separable.

Additional Reading

[59, 159, 163, 169, 174, 176, 177, 183, 225, 255]

http://www.springer.com/978-3-319-32411-1

Analysis
From Concepts to Applications
Penot, J.-P.
2016, XXIII, 669 p. 26 illus.,
ISBN: 978-3-319-32411-1