2 Preliminaries

In this chapter we introduce a number of fundamental structures in general spaces: topology, metric, norm, and scalar product. They are the natural generalizations of the corresponding concepts in the Euclidean space $\mathbb{R}^n$.

The most detailed structure is given by a scalar product in a $\mathbb{K}$-vector space, where here and throughout we take either $\mathbb{K} = \mathbb{R}$, i.e. $\mathbb{K}$ is the set of ***real numbers***, or $\mathbb{K} = \mathbb{C}$, i.e. $\mathbb{K}$ is the set of ***complex numbers***. For $\alpha \in \mathbb{K}$ we use the notation

$$|\alpha| := \sqrt{\alpha\overline{\alpha}} \quad \text{with} \quad \overline{\alpha} := \begin{cases} \mathrm{Re}\alpha - \mathrm{i}\ \mathrm{Im}\alpha & \text{for } \mathbb{K} = \mathbb{C}, \\ \alpha & \text{for } \mathbb{K} = \mathbb{R}, \end{cases}$$

and if $\alpha \in \mathbb{C}$ and for example

$$\alpha > 0, \quad \text{we implicitly assume that } \alpha \in \mathbb{R} \subset \mathbb{C}.$$

2.1 Scalar product. Let X be a $\mathbb{K}$-vector space. We call a map $(x_1, x_2) \mapsto (x_1\,,\,x_2)_X$ from $X \times X$ to $\mathbb{K}$ a ***sesquilinear form*** if for all $\alpha \in \mathbb{K}$ and for all $x, x_1, x_2, y, y_1, y_2 \in X$ one has

(S1) $(\alpha x\,,\,y)_X = \alpha\,(x\,,\,y)_X\,,$
$(x\,,\,\alpha y)_X = \overline{\alpha}\,(x\,,\,y)_X\,,$

(S2) $(x_1 + x_2\,,\,y)_X = (x_1\,,\,y)_X + (x_2\,,\,y)_X\,,$
$(x\,,\,y_1 + y_2)_X = (x\,,\,y_1)_X + (x\,,\,y_2)_X\,.$

This means that $(\bullet_1\,,\,\bullet_2)_X$ is ***linear*** in the first argument and ***conjugate linear*** in the second argument. Where no ambiguities arise, one can also write $(x_1\,,\,x_2)$ in place of $(x_1\,,\,x_2)_X$. The sesquilinear form is called ***symmetric*** (also called a ***Hermitian form***) if for all $x, y \in X$ one has

(S3) $(x\,,\,y)_X = \overline{(y\,,\,x)_X}$ (***Symmetry***).

A sesquilinear form is called ***positive semidefinite*** if for all $x \in X$

(S4') $(x\,,\,x)_X \geq 0$ (and then $(x\,,\,x)_X \in \mathbb{R}$) (***Positivity***)

and ***positive definite*** if for all $x \in X$

(S4) $(x\,,\,x)_X \geq 0$ and in addition: $(x\,,\,x)_X = 0 \iff x = 0\,.$

H.W. Alt, *Linear Functional Analysis*, Universitext,
DOI 10.1007/978-1-4471-7280-2_2

For Hermitian forms $(x\,,\,x)_X = \overline{(x\,,\,x)_X}$ is real-valued, and for positive semidefinite Hermitian forms it is always nonnegative, in which case we define

$$\|x\|_X := \sqrt{(x\,,\,x)_X}\,.$$

A positive definite Hermitian form is also called a ***scalar product*** or ***inner product***, and then the pair $(X, (\bullet_1\,,\,\bullet_2)_X)$ is called a ***pre-Hilbert space***. If this scalar product in the vector space X is fixed, then we also say that X is a ***pre-Hilbert space***.

The following lemma contains the fundamental properties of a scalar product.

2.2 Lemma. Let $(x_1, x_2) \mapsto (x_1\,,\,x_2)_X$ from $X \times X$ to $\mathbb{K}$ be a positive semidefinite Hermitian form and $\|x\| := \sqrt{(x\,,\,x)_X}$ for $x \in X$. Then it holds for all $x, y \in X$ and all $\alpha \in \mathbb{K}$ that

(1) $\|\alpha x\| = |\alpha| \cdot \|x\|$ (***Homogeneity***),
(2) $|(x\,,\,y)_X| \leq \|x\| \cdot \|y\|$ (***Cauchy-Schwarz inequality***),
(3) $\|x + y\| \leq \|x\| + \|y\|$ (***Triangle inequality***),
(4) $\|x + y\|^2 + \|x - y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big)$ (***Parallelogram law***).

Proof (1). $\|\alpha x\|^2 = (\alpha x\,,\,\alpha x)_X = \alpha\,(x\,,\,\alpha x)_X = \alpha\overline{\alpha}\,(x\,,\,x)_X = |\alpha|^2 \|x\|^2$. □

Proof (2). Let $x, y \in X$. It holds for $\alpha, \beta \in \mathbb{K} \setminus \{0\}$ (we want to set $\alpha = \|x\|$ and $\beta = \|y\|$) that

$$0 \leq \left\| \frac{x}{\alpha} - \frac{y}{\beta} \right\|^2 = \frac{\|x\|^2}{|\alpha|^2} + \frac{\|y\|^2}{|\beta|^2} - 2\mathrm{Re}\left(\frac{(x\,,\,y)_X}{\alpha\overline{\beta}}\right), \tag{2-1}$$

and hence for $\alpha > 0$ and $\beta > 0$, upon multiplying the inequality by $\alpha\beta > 0$, that

$$2\mathrm{Re}\,(x\,,\,y)_X \leq \frac{\beta}{\alpha}\|x\|^2 + \frac{\alpha}{\beta}\|y\|^2\,.$$

Setting $\alpha = \|x\| + \varepsilon$, $\beta = \|y\| + \varepsilon$ with $\varepsilon > 0$ yields that

$$\begin{aligned} 2\mathrm{Re}\,(x\,,\,y)_X &\leq (\|y\| + \varepsilon) \cdot \frac{\|x\|^2}{\|x\| + \varepsilon} + (\|x\| + \varepsilon) \cdot \frac{\|y\|^2}{\|y\| + \varepsilon} \\ &\leq (\|y\| + \varepsilon) \cdot \|x\| + (\|x\| + \varepsilon) \cdot \|y\|\,. \end{aligned}$$

As this holds for all $\varepsilon > 0$, it follows that

$$\mathrm{Re}\,(x\,,\,y)_X \leq \|x\| \cdot \|y\|\,.$$

On replacing x with $\overline{(x\,,\,y)_X}x$ we obtain

$$|(x\,,\,y)_X|^2 \leq |(x\,,\,y)_X| \cdot \|x\| \cdot \|y\|$$

and then cancelling in the case $(x\,,\,y)_X \neq 0$ gives the desired result. If $(x\,,\,y)_X = 0$, then the claim is trivial. □

Proof (3). On recalling (2) we have

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\,(x\,,\,y)_X \\ &\le \|x\|^2 + \|y\|^2 + 2\|x\|\cdot\|y\| = (\|x\|+\|y\|)^2\,.\end{aligned}$$

□

Proof (4). The first identity in the proof of (3) was

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\,(x\,,\,y)_X\ .$$

Replacing y by $-y$ yields, since $(x\,,\,-y)_X = -\,(x\,,\,y)_X$, that

$$\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\mathrm{Re}\,(x\,,\,y)_X\ . \tag{2-2}$$

Adding the two identities gives the result. □

2.3 Orthogonality. Let X be a pre-Hilbert space over $\mathbb{K}$ and for $x \in X$ let $\|x\|_X := \sqrt{(x\,,\,x)_X}$ as in 2.1.

(1) Let $x, y \in X$. If $(x\,,\,y)_X = 0$, we say that x and y are ***perpendicular***, or that they are ***orthogonal vectors***. Then

$$\|x-y\|_X^2 = \|x\|_X^2 + \|y\|_X^2 \qquad (\textbf{\textit{Pythagoras' theorem}}).$$

(2) If Y and Z are two subspaces (see 4.4(2)) of a vector space X, then the sum

$$Y+Z := \{y+z \in X\,;\ y \in Y \text{ and } z \in Z\}$$

is again a subspace. The sum is called a ***direct sum***, and we write $Y \oplus Z = Y + Z$, if $Y \cap Z = \{0\}$. If X is a pre-Hilbert space, then the subspaces are called ***orthogonal*** if $(y\,,\,z)_X = 0$ for all $y \in Y$ and $z \in Z$. Clearly it then holds that $Y \cap Z = \{0\}$ and we denote the subspace $Y \oplus Z$ also by $Y \perp Z$. The ***orthogonal complement*** of a subspace Y is defined by

$$Y^\perp := \{x \in X\,;\ (y\,,\,x)_X = 0 \text{ for all } y \in Y\}\ (\text{see also 9.17})\ .$$

It holds that $Y \cap Y^\perp = \{0\}$.

(3) For $x, y \in X \setminus \{0\}$ the Cauchy-Schwarz inequality 2.2(2) then reads

$$|\gamma| \le 1 \quad \text{with} \quad \gamma := \left(\frac{x}{\|x\|_X}\,,\,\frac{y}{\|y\|_X}\right)_X\,.$$

Here equality holds if and only if x and y are linearly dependent.

(4) If $\mathbb{K} = \mathbb{R}$, then in (3) there exists a unique

$$\theta \in [0,\pi] \quad \text{such that} \quad \gamma = \cos(\theta).$$

We call θ the ***angle*** between x and y. It follows from (3) that x and y are linearly dependent if and only if $\theta = 0$ or $\theta = \pi$, and they are orthogonal if and only if $\theta = \frac{\pi}{2}$.

Proof (1). The theorem of Pythagoras follows from (2-2). □

Proof (2). This essentially contains only definitions. □

Proof (3). If x and y are linearly dependent, it is obvious that $|\gamma| = 1$. If $|\gamma| = 1$, then on setting $\alpha = \|x\|_X$, $\beta = \overline{\gamma}\|y\|_X$, equation (2-1) becomes

$$0 \leq \left\| \frac{x}{\alpha} - \frac{y}{\beta} \right\|^2 = 2 - 2\mathrm{Re}\left(\frac{(x\,,\,y)_X}{\|x\|_X \cdot \gamma \|y\|_X} \right) = 0\,,$$

which implies

$$\frac{x}{\alpha} = \frac{y}{\beta}\,,$$

hence x and y are linearly dependent. □

Proof (4). By (3), the vectors x and y are linearly dependent if and only if $1 = |\gamma| = |\cos(\theta)|$, which means $\theta = 0$ or $\theta = \pi$. By (1), the vectors x and y are orthogonal if and only if $(x\,,\,y)_X = 0$, which means $\cos(\theta) = 0$, that is, $\theta = \frac{\pi}{2}$. □

The standard example is the n-dimensional ***Euclidean space*** $\mathbb{R}^n$. The ***Euclidean scalar product*** and the ***Euclidean norm*** (for clarity these will be denoted by special symbols) are defined by

$$x \bullet y := \sum_{i=1}^{n} x_i y_i \quad \text{and} \quad |x| := \sqrt{x \bullet x} = \Big(\sum_{i=1}^{n} x_i^2 \Big)^{\frac{1}{2}}$$

for $x = (x_i)_{i=1,\ldots,n} \in \mathbb{R}^n$, $y = (y_i)_{i=1,\ldots,n} \in \mathbb{R}^n$. In the complex space $\mathbb{C}^n$ we define correspondingly

$$z \bullet w := \sum_{i=1}^{n} z_i \overline{w_i} \in \mathbb{C} \quad \text{and} \quad |z| := \sqrt{z \bullet z} = \Big(\sum_{i=1}^{n} z_i \overline{z_i} \Big)^{\frac{1}{2}} \in \mathbb{R}$$

for $z = (z_i)_{i=1,\ldots,n} \in \mathbb{C}^n$, $w = (w_i)_{i=1,\ldots,n} \in \mathbb{C}^n$. The infinite-dimensional analogue of Euclidean space is the sequence space (see 2.23).

A fundamental step in the development of functional analysis was the introduction of norms $x \mapsto \|x\|_X$ that are not induced by a scalar product as in 2.1, but are instead only characterized by the homogeneity and the triangle inequality in 2.2.

2.4 Norm. Let X be a $\mathbb{K}$-vector space. The pair $(X, \|\cdot\|)$ is called a ***normed space*** if $\|\cdot\| : X \to \mathbb{R}$ satisfies the following conditions for $x, y \in X$ and $\alpha \in \mathbb{K}$:

(N1) $\|x\| \geq 0$ (***Positivity***),
and: $\|x\| = 0 \iff x = 0$,
(N2) $\|\alpha x\| = |\alpha| \cdot \|x\|$ (***Homogeneity***),
(N3) $\|x + y\| \leq \|x\| + \|y\|$ (***Triangle inequality***).

We then say that the map $\|\cdot\| : X \to \mathbb{R}$ is a ***norm*** on X. If a norm $\|\cdot\|_X : X \to \mathbb{R}$ is fixed on the vector space X, then we also call X a ***normed space***.

Note that the property $\big(x = 0 \Longrightarrow \|x\| = 0 \big)$ in (N1) follows independently from (N2) on setting $\alpha = 0$ there. We call $\|\cdot\|$ a ***seminorm*** if we take (N1) without the property $\big(\|x\| = 0 \Longrightarrow x = 0 \big)$. It then follows from (N2) and (N3) that the set $Z := \{z \in X ;\ \|z\| = 0\}$ is a subspace of X, and hence

$$x \sim y \quad :\Longleftrightarrow \quad x - y \in Z$$

defines an equivalence relation "$\sim$" on X. Now let $\widetilde{X}$ be the set X together with the equivalence relation

$$x = y \text{ in } \widetilde{X} \quad :\Longleftrightarrow \quad x \sim y \quad \Longleftrightarrow \quad x - y \in Z.$$

Then all the vector space properties carry over from X to $\widetilde{X}$, and $(\widetilde{X}, \|\cdot\|)$ is a normed space (see remark). A common notation for the ***factor space*** or ***quotient space*** $\widetilde{X}$ is X/Z.

Remark: Let X be an arbitrary set, with "$\sim$" an arbitrary equivalence relation on X, and then let $\widetilde{X}$ be the set X with this equivalence relation, that is,

$$x = y \text{ in } \widetilde{X} \quad :\Longleftrightarrow \quad x \sim y \text{ in } X.$$

A map $f : \widetilde{X} \to S$ to another set S is said to be ***well defined*** if

$$x = y \text{ in } \widetilde{X} \quad \Longrightarrow \quad f(x) = f(y) \text{ in } S. \tag{2-3}$$

Hence, when defining a map on $\widetilde{X}$, condition (2-3) always needs to be verified.

Similarly, given a map $f : X \to S$, then this also defines a map from $\widetilde{X}$ to S, if (2-3) is satisfied for f. Analogous results hold for maps defined on e.g. $X \times X$.

In the case of a seminorm as discussed above, it can be easily shown that this is satisfied for the maps $(x, y) \mapsto x + y$ from $X \times X$ to X and $(\alpha, x) \mapsto \alpha x$ from $\mathbb{K} \times X$ to X, as well as for the map $x \mapsto \|x\|$ from X to $\mathbb{R}$.

In Section 3 we will introduce the most important norms in spaces of continuous and integrable functions. These norms are derived from the following norms in $\mathbb{K}^n$.

2.5 Example. For $1 \le p \le \infty$ the ***p-norm*** on $\mathbb{K}^n$ is defined by

$$|x|_p := \begin{cases} \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} & \text{for } 1 \le p < \infty\,, \\ \max_{i=1,\ldots,n} |x_i| & \text{for } p = \infty\,, \end{cases}$$

where $x = (x_i)_{i=1,\ldots,n} \in \mathbb{K}^n$. For $p = 2$ the ***Euclidean norm*** of x is $|x|_2 = |x|$. Alternative notations for the ***maximum norm*** $|x|_\infty$ are $|x|_{\max}$ and $|x|_{\sup}$, while the ***sum norm*** $|x|_1$ is also denoted by $|x|_{\text{sum}}$.

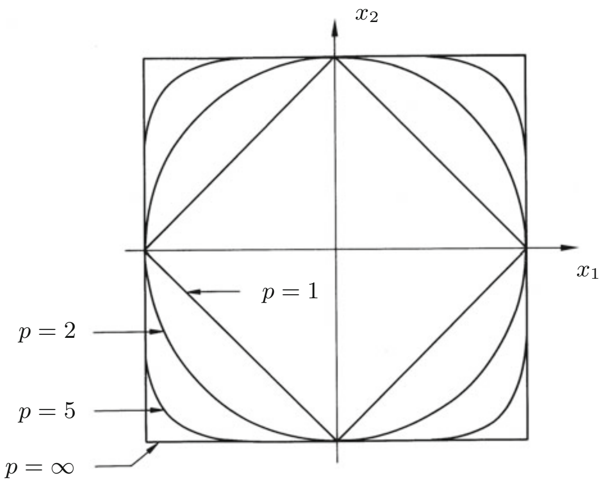

Fig. 2.1. *Unit spheres for p-norms in* $\mathbb{R}^2$

Proof. All of the norm axioms are easily verified, apart from the triangle inequality in the case $1 < p < \infty$ for $n \ge 2$. However, this follows from the ***Hölder inequality*** (proof to follow)

$$|x \bullet y| \le |x|_p \cdot |y|_{p'} \tag{2-4}$$

for $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, where p' is the ***dual exponent*** to p, i.e. it is defined by $\frac{1}{p} + \frac{1}{p'} = 1$.
Note: This inequality is a special case of the general Hölder inequality in 3.18 for the counting measure on $\{1, \ldots, n\}$. Here we give a different proof.

The inequality (2-4) can, for instance, be shown by induction, on employing the inequality for $n = 2$. To this end, let $x' := (x_1, \ldots, x_{n-1})$, $y' := (y_1, \ldots, y_{n-1})$ and observe that

$$\begin{aligned}|x \bullet y| &\le |x' \bullet y'| + |x_n| \cdot |y_n| \\ &\le |x'|_p \cdot |y'|_{p'} + |x_n| \cdot |y_n| \quad \text{(induction hypothesis)} \\ &\le \Big|(|x'|_p\,,\ |x_n|)\Big|_p \cdot \Big|(|y'|_{p'}\,,\ |y_n|)\Big|_{p'} \quad \text{(inequality for } n=2) \\ &= |x|_p \cdot |y|_{p'}\,.\end{aligned}$$

The inequality for $n = 2$ follows immediately from the elementary inequality

$$a_1 b_1 + a_2 b_2 \le (a_1^p + a_2^p)^{\frac{1}{p}} \cdot (b_1^{p'} + b_2^{p'})^{\frac{1}{p'}} \quad \text{for} \quad a_1, a_2, b_1, b_2 \ge 0\,. \tag{2-5}$$

This holds trivially if one of the numbers is equal to 0. Otherwise, dividing by $a_1 b_1$ and setting $\alpha := a_2^p a_1^{-p}$, $\beta := b_2^{p'} b_1^{-p'}$ yields the equivalent inequality

$$1 + \alpha^{\frac{1}{p}} \beta^{\frac{1}{p'}} \le (1+\alpha)^{\frac{1}{p}} \cdot (1+\beta)^{\frac{1}{p'}} \quad \text{for } \alpha, \beta > 0\,, \tag{2-6}$$

which we will prove now. For fixed $r := \alpha^{\frac{1}{p}} \cdot \beta^{\frac{1}{p'}}$ we have that

$$\alpha = \left(r \beta^{-\frac{1}{p'}}\right)^p = r^p \beta^{-\frac{p}{p'}} = r^p \beta^{1-p} =: \psi(\beta)\,,$$

since $\frac{p}{p'} = p - 1$. Then the inequality reads

$$1 + r \le \varphi(\beta) := (1 + \psi(\beta))^{\frac{1}{p}} \cdot (1+\beta)^{\frac{1}{p'}}\,,$$

and the right-hand side is minimal, if $\varphi'(\beta) = 0$. Now

$$\varphi'(\beta) = \varphi(\beta) \cdot \left(\frac{\psi'(\beta)}{p(1+\psi(\beta))} + \frac{1}{p'(1+\beta)}\right) = \frac{\varphi(\beta)}{p'\beta} \cdot \left(\frac{\beta}{1+\beta} - \frac{\psi(\beta)}{1+\psi(\beta)}\right)\,,$$

since $\psi'(\beta) = -\psi(\beta) \cdot \frac{p-1}{\beta}$. Hence $\varphi'(\beta) = 0$ means $\beta = \psi(\beta)$, and so $\beta = r$, $\alpha = r$. This proves (2-6), and therefore the Hölder inequality (2-5) follows.

On letting $z_i := |x_i + y_i|^{p-1}$, $z = (z_1, \dots, z_n)$, we have that

$$|x_i + y_i|^p \;\le\; |x_i|\; z_i + |y_i|\; z_i\,.$$

The Hölder inequality then implies, since $p' \cdot (p-1) = p$, that

$$\begin{aligned}|x+y|_p^p &\le (|x_i|)_{i=1,\dots,n} \bullet z + (|y_i|)_{i=1,\dots,n} \bullet z \\ &\le |x|_p \cdot |z|_{p'} + |y|_p \cdot |z|_{p'} = (|x|_p + |y|_p) \cdot |x+y|_p^{p-1}\,,\end{aligned}$$

which yields $|x + y|_p \le |x|_p + |y|_p$. □

We now interpret the norm $\|x\|$ of x as the distance of the point x from the origin 0 and replace $\|x\|$ with a value $d(x, 0)$, where $d : X \times X \to \mathbb{R}$ is a map for which only the triangle inequality has to hold. This notion of a distance can be defined in arbitrary sets.

2.6 Metric. A ***metric space*** is a pair (X,d), where X is a set and

$$d : X \times X \to \mathbb{R} \text{ for all } x,y,z \in X$$

has the following properties:

(M1) $d(x,y) \geq 0$ (***Positivity***),
and: $d(x,y) = 0 \iff x = y$,
(M2) $d(x,y) = d(y,x)$ (***Symmetry***),
(M3) $d(x,y) \leq d(x,z) + d(z,y)$ (***Triangle inequality***).

We then call $d(x,y)$ the ***distance*** between the points x and y. The map $d : X \times X \to \mathbb{R}$ is called a ***metric*** on X. If a metric $d_X : X \times X \to \mathbb{R}$ is fixed on the set X, then we also call X a ***metric space***. If (X,d) is a metric space and $A \subset X$, then (A,d) is also a metric space, with d restricted to $A \times A$.

Without the property $(\, d(x,y) = 0 \implies x = y\,)$ in (M1) we call d a ***semimetric***. Then the ***factor space*** of X with respect to d is given as follows: The properties of the semimetric imply that

$$x \sim y \quad :\iff \quad d(x,y) = 0$$

defines an equivalence relation "$\sim$" on X. Now let $\widetilde{X}$ be the set X equipped with the equivalence relation

$$x = y \text{ in } \widetilde{X} \quad :\iff \quad x \sim y \quad \iff \quad d(x,y) = 0.$$

Then (M3) implies that d is also well defined on $\widetilde{X} \times \widetilde{X}$, and that $(\widetilde{X}, d)$ is a metric space (see the remark in 2.4).

2.7 Fréchet metric. In vector spaces X, metrics d are often defined by

$$d(x,y) = \varrho(x-y) \quad \text{for } x,y \in X\,,$$

where $\varrho : X \to \mathbb{R}$ satisfies the following properties for all $x,y \in X$:

(F1) $\varrho(x) \geq 0$ (***Positivity***),
and: $\varrho(x) = 0 \iff x = 0$,
(F2) $\varrho(x) = \varrho(-x)$ (***Symmetry***),
(F3) $\varrho(x+y) \leq \varrho(x) + \varrho(y)$ (***Triangle inequality***).

A map $\varrho : X \to \mathbb{R}$ satisfying (F1)–(F3) is called a ***Fréchet metric***. Any norm $x \mapsto \|x\|$ on X is a Fréchet metric and hence defines the ***induced metric*** $d(x,y) := \|x-y\|$.

We begin with some elementary examples.

2.8 Examples of metrics.

(1) A bounded Fréchet metric on $\mathbb{K}^n$ that is not a norm is given by

$$\varrho(x) := \frac{|x|}{1+|x|} \quad \text{for } x \in \mathbb{K}^n\,.$$

(2) Let $-\infty$, $+\infty$ be two distinct elements that do not belong to $\mathbb{R}$. One can then define a metric on $\mathbb{R} \cup \{\pm\infty\}$ by

$$d(x, y) := |g(x) - g(y)| \quad \text{for } x, y \in \mathbb{R} \cup \{\pm\infty\},$$

where

$$g(x) := \begin{cases} -1 & \text{for } x = -\infty, \\ \dfrac{x}{1 + |x|} & \text{for } x \in \mathbb{R}, \\ +1 & \text{for } x = +\infty. \end{cases}$$

(3) Let ∞ be an element that does not belong to $\mathbb{R}^n$. One can then define a metric on $\mathbb{R}^n \cup \{\infty\}$ by

$$d(x, y) := |\tau_{\text{stereo}}(x) - \tau_{\text{stereo}}(y)|.$$

Here

$$\tau_{\text{stereo}} : \mathbb{R}^n \cup \{\infty\} \longrightarrow \left\{ y \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1} \; ; \; \left| y - (0, \tfrac{1}{2}) \right| = \tfrac{1}{2} \right\},$$

where the image is the ball $\mathrm{B}^{\mathbb{R}^{n+1}}_{\frac{1}{2}}\big((0, \frac{1}{2})\big)$ with respect to the Euclidean metric, is defined by

$$\tau_{\text{stereo}}(x) := \begin{cases} \dfrac{(x, |x|^2)}{1 + |x|^2} & \text{for } x \in \mathbb{R}^n, \\ (0, 1) & \text{for } x = \infty. \end{cases}$$

Remark: The inverse ${\tau_{\text{stereo}}}^{-1}$ is the ***stereographic projection***, i.e. $y = \tau_{\text{stereo}}(x)$ with $\left| y - (0, \frac{1}{2}) \right| = \frac{1}{2}$ and $y \neq (0, 1)$ is given by

$$(1 - a)(0, 1) + ay = (x, 0) \quad \text{for an } a \in \mathbb{R}.$$

Proof (1). The function $\varphi(s) := \frac{s}{1+s}$ for $s \geq 0$ satisfies

$$\varphi(s) \leq \varphi(\widetilde{s}) \quad \text{for } 0 \leq s \leq \widetilde{s},$$

$$\varphi(s_1 + s_2) = \frac{s_1}{1 + s_1 + s_2} + \frac{s_2}{1 + s_1 + s_2} \leq \varphi(s_1) + \varphi(s_2) \quad \text{for } s_1, s_2 \geq 0.$$

Apply the above for $s = |x + y| \leq |x| + |y| = \widetilde{s}$, $s_1 = |x|$, $s_2 = |y|$. □

Proof (2) *and* (3). Use that g, resp. τ_{stereo}, is injective and employ the triangle inequality in $\mathbb{R}$, resp. $\mathbb{R}^{n+1}$. □

With the help of the distance between two points we now define the distance between two sets. As a special case we obtain the definition of balls with respect to a given metric.

2.9 Balls and distance between sets. Let (X, d) be a metric space. For two sets $A_1, A_2 \subset X$ the ***distance*** between A_1 and A_2 is defined by

$$\operatorname{dist}(A_1, A_2) := \inf \{ d(x,y) \; ; \; x \in A_1, \; y \in A_2 \} ,$$

where $\inf \emptyset := \infty$ (so that $\operatorname{dist}(A, \emptyset) = \infty$). For $x, y \in X$ it holds that $d(x,y) = \operatorname{dist}(\{x\}, \{y\})$. For $x \in X$ the ***distance*** from x to $A \subset X$ is defined by

$$\operatorname{dist}(x, A) := \operatorname{dist}(\{x\}, A) = \inf \{ d(x,y) \; ; \; y \in A \} .$$

For $r > 0$ the ***r-neighbourhood of the set A*** is defined by

$$\mathrm{B}_r(A) := \{ x \in X \; ; \; \operatorname{dist}(x, A) < r \} ,$$

and $\mathrm{B}_r(x) := \mathrm{B}_r(\{x\})$ is called the ***ball around x with radius r*** or, alternatively, the ***r-neighbourhood of the point x***. We have

$$\mathrm{B}_r(x) = \{ y \in X \; ; \; d(y,x) < r \} .$$

The ***diameter*** of a subset $A \subset X$ is defined by

$$\operatorname{diam}(A) := \sup \{ d(x,y) \; ; \; x, y \in A \} ,$$

if $A \neq \emptyset$, and $\operatorname{diam}(\emptyset) := 0$ (or make the convention that $\sup \emptyset := 0$). A set $A \subset X$ is called ***bounded*** if $\operatorname{diam}(A) < \infty$.

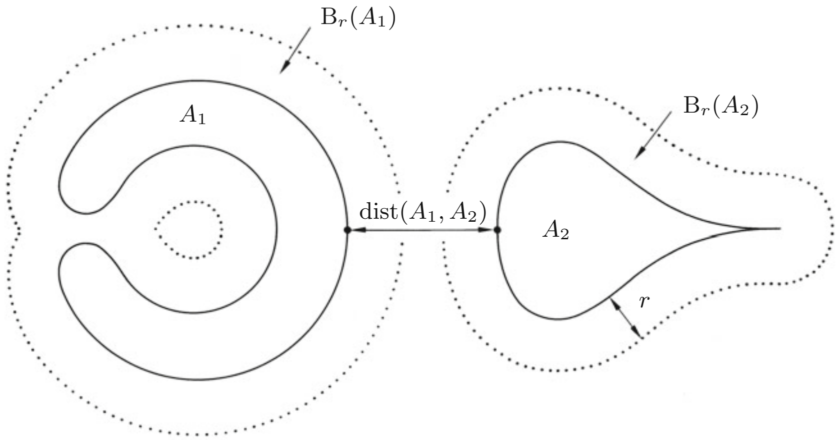

Fig. 2.2. *Metric definitions in* $\mathbb{R}^2$ *with respect to* $x \mapsto |x|_2$

The concept of a ball plays an important role in definitions and proofs for metric spaces. It can be used, for instance, to introduce the following notion of an "open subset" (see 2.10). In functional analysis the concept of open sets is applied to function spaces. Depending on the chosen distance the notion of open sets is different, therefore one obtains different results for the considered class of functions. In 3.2, 3.3 and 3.7 this applies to function spaces with respect to supremum norms, in 3.15 to function spaces with respect to integral norms, and in 3.13 to function spaces equipped with distances that are induced by a measure.

2.10 Open and closed sets. Let (X, d) be a metric space. For $A \subset X$ the ***interior*** of A (notation: $\mathrm{intr}_X(A)$ or $\mathrm{intr}(A)$ or $\mathring{A}$) is defined by

$$\mathrm{intr}(A) := \{\, x \in X \ ; \ \mathrm{B}_\varepsilon(x) \subset A \text{ for an } \varepsilon > 0 \,\} \subset A\,,$$

and the ***closure*** of A (or the ***closed hull***, notation: $\mathrm{clos}_X(A)$ or $\mathrm{clos}(A)$ or $\overline{A}$) is defined by

$$\mathrm{clos}(A) := \{\, x \in X \ ; \ \mathrm{B}_\varepsilon(x) \cap A \neq \emptyset \text{ for all } \varepsilon > 0 \,\} \supset A\,.$$

It holds that $x \in \mathrm{clos}(A)$ if and only if $\mathrm{dist}(x, A) = 0$. Using quantifiers, the above definitions can be written as

$$\begin{aligned} x \in \mathrm{intr}(A) \quad &\Longleftrightarrow \quad \exists\, \varepsilon > 0 \,:\, \mathrm{B}_\varepsilon(x) \setminus A = \emptyset\,, \\ x \in \mathrm{clos}(A) \quad &\Longleftrightarrow \quad \forall\, \varepsilon > 0 \,:\, \mathrm{B}_\varepsilon(x) \cap A \neq \emptyset\,, \end{aligned}$$

or

$$\begin{aligned} x \in \mathrm{intr}(A) \quad &\Longleftrightarrow \quad \exists\, \varepsilon > 0 \,:\, \forall\, y \in \mathrm{B}_\varepsilon(x) \,:\, y \in A\,, \\ x \in \mathrm{clos}(A) \quad &\Longleftrightarrow \quad \forall\, \varepsilon > 0 \,:\, \exists\, y \in \mathrm{B}_\varepsilon(x) \,:\, y \in A\,. \end{aligned}$$

A subset $A \subset X$ is called ***open*** if $\mathrm{intr}(A) = A$, and $A \subset X$ is called ***closed*** if $\mathrm{clos}(A) = A$. The complement of a closed set is open and the complement of an open set is closed. The ***boundary*** of $A \subset X$ (notation: $\mathrm{bdry}_X(A)$ or $\mathrm{bdry}(A)$ or ∂A) is defined by

$$\begin{aligned} \mathrm{bdry}(A) &:= \mathrm{clos}(A) \setminus \mathrm{intr}(A) \\ &= \mathrm{clos}(A) \cap \mathrm{clos}(X \setminus A) = \mathrm{bdry}(X \setminus A) \end{aligned}$$

and, being an intersection of closed sets, is a closed set. We have

$$X = \mathrm{intr}(A) \cup \mathrm{bdry}(A) \cup \mathrm{intr}(X \setminus A)\,,$$

where the union is disjoint.

We now consider on X only the class of open sets. This class is characterized by the fact that arbitrary unions of open sets and finite intersections of open sets are still open sets.

2.11 Topology. A ***topological space*** is a pair $(X, \mathcal{T})$, where X is a set and $\mathcal{T}$ is a system of subsets of X (the elements of $\mathcal{T}$ are called ***open sets***), with the following properties:

(T1) $\emptyset \in \mathcal{T}$, $X \in \mathcal{T}$,
(T2) $\mathcal{T}' \subset \mathcal{T} \implies \bigcup_{U \in \mathcal{T}'} U \in \mathcal{T}$,
(T3) $U_1, U_2 \in \mathcal{T} \implies U_1 \cap U_2 \in \mathcal{T}$.

A topological space is called a ***Hausdorff space*** if in addition the following ***separation axiom*** is satisfied:

(T4) For $x_1, x_2 \in X$ with $x_1 \neq x_2$ there exist $U_1, U_2 \in \mathcal{T}$ such that $x_1 \in U_1$, $x_2 \in U_2$, and $U_1 \cap U_2 = \emptyset$, the same with quantifiers: $\forall\, x_1, x_2 \in X,\ x_1 \neq x_2 : \exists\, U_1, U_2 \in \mathcal{T} : x_1 \in U_1,\ x_2 \in U_2,\ U_1 \cap U_2 = \emptyset$.

A subset $A \subset X$ is called ***closed*** with respect to $\mathcal{T}$ if $X \setminus A \in \mathcal{T}$, that is, with respect to $\mathcal{T}$, the complement of an open set is closed, and the complement of a closed set is open. We define for $A \subset X$ (note the remark in 2.12 below)

$$\mathrm{intr}_{(X,\mathcal{T})}(A) := \big\{ x \in X \,;\ U \subset A \text{ for some } U \in \mathcal{T} \text{ with } x \in U \big\} \subset A\,,$$
$$\mathrm{clos}_{(X,\mathcal{T})}(A) := \big\{ x \in X \,;\ U \cap A \neq \emptyset \text{ for all } U \in \mathcal{T} \text{ with } x \in U \big\} \supset A\,.$$

Alternative notations are $\mathrm{intr}\,(A) := \mathrm{intr}_{(X,\mathcal{T})}(A)$ or $\mathring{A} := \mathrm{intr}_{(X,\mathcal{T})}(A)$ and $\mathrm{clos}\,(A) := \mathrm{clos}_{(X,\mathcal{T})}(A)$ or $\overline{A} := \mathrm{clos}_{(X,\mathcal{T})}(A)$. It holds that

$$A = \mathrm{intr}_{(X,\mathcal{T})}(A) \iff A \in \mathcal{T}\,,$$
$$A = \mathrm{clos}_{(X,\mathcal{T})}(A) \iff X \setminus A \in \mathcal{T}\,.$$

If $A \subset X$, then $(A, \mathcal{T}_A)$ is a topological space with the ***relative topology***

$$\mathcal{T}_A := \{U \cap A\,;\ U \in \mathcal{T}\}\,.$$

The following is the standard construction of a topology and it shows that for a metric space the definitions of interior and closure in 2.11 (with respect to a topology) and in 2.10 (with respect to a metric) are the same.

2.12 Proposition. Let (X, d) be a metric space and, on recalling the definition of the interior of a set in 2.10 (we write $\mathrm{intr}_{(X,d)}(A)$ instead of $\mathrm{intr}_X(A)$), let

$$\mathcal{T} := \{A \subset X\,;\ \mathrm{intr}_{(X,d)}(A) = A\}\,.$$

Then $(X, \mathcal{T})$ is a topological space and, in particular, a Hausdorff space. We call $\mathcal{T}$ the topology ***induced*** by the metric d.
Remark: For all subsets $A \subset X$ it holds $\mathrm{intr}_{(X,d)}(A) = \mathrm{intr}_{(X,\mathcal{T})}(A)$ and $\mathrm{clos}_{(X,d)}(A) = \mathrm{clos}_{(X,\mathcal{T})}(A)$.

Proof of the proposition. In order to show axiom (T3), let $A_1, A_2 \in \mathcal{T}$ and $x \in A_1 \cap A_2$. Then $\mathrm{intr}_{(X,d)}(A_1) = A_1$ and $\mathrm{intr}_{(X,d)}(A_2) = A_2$ with the definition as in 2.10. Hence there exist $\varepsilon_1, \varepsilon_2 > 0$ such that $\mathrm{B}_{\varepsilon_1}(x) \subset A_1$ and $\mathrm{B}_{\varepsilon_2}(x) \subset A_2$. Setting $\varepsilon := \min(\varepsilon_1, \varepsilon_2) > 0$ yields $\mathrm{B}_\varepsilon(x) \subset A_1 \cap A_2$, and hence $A_1 \cap A_2 \in \mathcal{T}$. For the proof of (T4) let $x \neq y$. Then the triangle inequality yields that

$$\mathrm{B}_r(x) \cap \mathrm{B}_r(y) = \emptyset \text{ for } r := \tfrac{1}{2}d(x,y) > 0\,,$$

and $\mathrm{B}_r(x)\,, \mathrm{B}_r(y) \in \mathcal{T}$ (see E2.2(2)). □

2.13 Definition. Let $(X, \mathcal{T})$ be a topological space. A subset $A \subset X$ is called ***dense*** in X if $\operatorname{clos}(A) = X$, and X is called ***separable*** if X contains a countable dense subset. A subset $A \subset X$ is called separable if the relative topological space $(A, \mathcal{T}_A)$ is separable. Hence, if (X, d) is a metric space, a subset $A \subset X$ is separable if the metric space (A, d) is separable.

2.14 Comparison of topologies. Let $\mathcal{T}_1$, $\mathcal{T}_2$ be two topologies on a set X. We say that $\mathcal{T}_2$ is ***stronger*** (or ***finer***) than $\mathcal{T}_1$, or equivalently that $\mathcal{T}_1$ is ***weaker*** (or ***coarser***) than $\mathcal{T}_2$, if

$$\mathcal{T}_1 \subset \mathcal{T}_2\,.$$

Let d_1, d_2 be two metrics on X and $\mathcal{T}_1$, $\mathcal{T}_2$ the corresponding induced topologies (see 2.11). Then the metric d_2 is said to be ***stronger*** (***weaker***) than d_1 if $\mathcal{T}_2$ is stronger (weaker) than $\mathcal{T}_1$. The metrics d_1 and d_2 are called ***equivalent***, if $\mathcal{T}_1 = \mathcal{T}_2$. Similarly, a norm is said to be ***stronger*** (***weaker***) than another norm, and two norms are called ***equivalent*** if this holds for the induced metrics, respectively.

2.15 Comparison of norms. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a $\mathbb{K}$-vector space X. Then

(1) $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$ if and only if there exists a positive number C such that

$$\|x\|_1 \le C\|x\|_2 \quad \text{for all } x \in X\,.$$

(2) The two norms are equivalent if and only if there exist positive numbers c and C such that

$$c\|x\|_2 \le \|x\|_1 \le C\|x\|_2 \quad \text{for all } x \in X\,.$$

Proof (1). Let $\mathrm{B}^i_r(x)$ denote the balls and $\mathcal{T}_i$ the topologies with respect to the norms $\|\cdot\|_i$. Let $\mathcal{T}_1 \subset \mathcal{T}_2$. Since $\mathrm{B}^1_1(0) \in \mathcal{T}_1$ (see E2.2(2)), $\mathrm{B}^1_1(0)$ is open with respect to $\|\cdot\|_2$ and, in particular, 0 lies in the interior (with respect to $\|\cdot\|_2$) of $\mathrm{B}^1_1(0)$, i.e.

$$\mathrm{B}^2_\varepsilon(0) \subset \mathrm{B}^1_1(0) \quad \text{for some } \varepsilon > 0.$$

This means, for $x \in X$, $x \neq 0$, that

$$\left\| \frac{\varepsilon x}{2\|x\|_2} \right\|_2 = \frac{\varepsilon}{2} < \varepsilon\,, \qquad \text{therefore } \left\| \frac{\varepsilon x}{2\|x\|_2} \right\|_1 < 1\,,$$

$$\text{that is } \|x\|_1 \le \frac{2}{\varepsilon}\|x\|_2\,.$$

Conversely, if the inequality in assertion (1) holds, then

$$\mathrm{B}^2_r(x) \subset \mathrm{B}^1_{Cr}(x) \quad \text{for all } x \in X \text{ and } r > 0.$$

Let $A \in \mathcal{T}_1$. Then $A = \mathrm{intr}_{d_1}(A)$ with respect to $\mathcal{T}_1$, and for $x \in A$ there is an $\varepsilon > 0$ such that

$$\mathrm{B}^1_\varepsilon(x) \subset A\,, \quad \text{therefore} \quad \mathrm{B}^2_{\frac{\varepsilon}{C}}(x) \subset A\,.$$

This proves that $A \in \mathcal{T}_2$. □

Proof (2). Apply (1) twice. □

2.16 Examples.

(1) The p-norms on $\mathbb{K}^n$ defined in 2.5 are pairwise equivalent, since for $1 \le p < \infty$

$$|x|_\infty \le |x|_p \le n^{\frac{1}{p}} |x|_\infty\,.$$

(2) The Euclidean norm and the Fréchet metric in 2.8(1) induce the same topology on $\mathbb{K}^n$, since for $y \in \mathbb{K}^n$

$$|y| \le 2\varrho(y) \text{ if } \varrho(y) \le \tfrac{1}{2}\,, \quad \varrho(y) \le |y|\,.$$

Hence, $\mathrm{B}^{\mathrm{metric}}_{\frac{r}{2}}(x) \subset \mathrm{B}^{\mathrm{norm}}_r(x) \subset \mathrm{B}^{\mathrm{metric}}_r(x)$ for $0 < r \le 1$.

(3) For open sets $U \subset \mathbb{R} \cup \{\pm\infty\}$ with respect to the metric in 2.8(2) it holds that

$$\begin{aligned} x \in U \cap \mathbb{R} \quad &\Longleftrightarrow \quad]\,x-\varepsilon, x+\varepsilon\,[\, \subset U \quad \text{for an } \varepsilon > 0\,,\\ +\infty \in U \quad &\Longleftrightarrow \quad]\,\tfrac{1}{\varepsilon}, +\infty] \subset U \quad \text{for an } \varepsilon > 0\,,\\ -\infty \in U \quad &\Longleftrightarrow \quad [-\infty, -\tfrac{1}{\varepsilon}\,[\, \subset U \quad \text{for an } \varepsilon > 0\,. \end{aligned}$$

(4) For open sets $U \subset \mathbb{K}^n \cup \{\infty\}$ with respect to the metric in 2.8(3) it holds that

$$\begin{aligned} x \in U \cap \mathbb{K}^n \quad &\Longleftrightarrow \quad \{y \in \mathbb{K}^n\,;\ |y-x| < \varepsilon\} \subset U \quad \text{for an } \varepsilon > 0\,,\\ \infty \in U \quad &\Longleftrightarrow \quad \{y \in \mathbb{K}^n\,;\ |y| > \tfrac{1}{\varepsilon}\} \subset U \quad \text{for an } \varepsilon > 0\,. \end{aligned}$$

One of the most important concepts in analysis is the notion of a limit and the resulting concept of continuity. Given a mapping $f : X \to Y$ between Hausdorff spaces X and Y, then f is ***continuous*** at $x_0 \in X$ (see 2.17(4) below) if

$$f(x_0) = \lim_{x \to x_0} f(x) \quad \text{in } Y.$$

This is the well-known notion of continuity in the analysis of Euclidean spaces. We now generalize this concept as follows: Given sets S, X, Y and mappings $\varphi : S \to X$, $f : S \to Y$, we consider two points $x_0 \in X$, $y_0 \in Y$ and the question is whether the function values $f(s)$ are "close to" y_0 if $\varphi(s)$ is "close to" x_0. In metric spaces we can define the notion of closeness with the help of balls around x_0 and y_0, and similarly in topological spaces with the help of open sets that contain x_0 and y_0, respectively.

Usually we have that $S \subset X$ and $\varphi(s) = s$ for $s \in S$. But often this is not the case. A nontrivial example is given in A3.17 (there S is a system of sets with $f(E) := |\nu(E)|$ and $\varphi(E) := \mu(E)$ for $E \in S$, hence $X = \mathbb{R}$ and $Y = \mathbb{R}$).

2.17 Convergence and continuity. Let S be a set, $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ Hausdorff spaces, and

$$\varphi : S \to X, \quad x_0 \in X, \quad f : S \to Y, \quad y_0 \in Y.$$

We say that

$$\begin{aligned}&f(s) \textbf{\textit{ converges to }} y_0 \text{ in } Y \text{ (with respect to } \mathcal{T}_Y) \\ &\textbf{\textit{as }} \varphi(s) \textbf{\textit{ goes to }} x_0 \text{ in } X \text{ (with respect to } \mathcal{T}_X),\end{aligned}$$

and use the notation

$$f(s) \to y_0 \text{ in } Y \quad \text{as} \quad \varphi(s) \to x_0 \text{ in } X\,,$$

if the following holds for all $U_0 \subset X$, $V_0 \subset Y$:

$$\begin{array}{l} x_0 \in U_0 \in \mathcal{T}_X, \\ y_0 \in V_0 \in \mathcal{T}_Y \end{array} \quad \Longrightarrow \quad \begin{array}{l} \text{There exists a } U \in \mathcal{T}_X \text{ such that } x_0 \in U \subset U_0\,, \\ \varphi^{-1}(U) \neq \emptyset \text{ and } f\big(\varphi^{-1}(U)\big) \subset V_0\,. \end{array}$$

The conclusion states that for a $U \in \mathcal{T}_X$ with $x_0 \in U \subset U_0$ it holds that

$$s \in S,\ \varphi(s) \in U \quad \Longrightarrow \quad f(s) \in V_0\,,$$

and that $\varphi(s) \in U$ for at least one $s \in S$. We have (see E2.4):

(1) Given x_0, f, φ, there exists at most one such $y_0 \in Y$. Hence we write

$$y_0 = \lim_{\varphi(s) \to x_0} f(s)\,,$$

and call y_0 the ***limit*** of $f(s)$ as $\varphi(s)$ goes to x_0.

(2) $x_0 \in \operatorname{clos}(\varphi(S))$ and $y_0 \in \operatorname{clos}(f(S))$.

(3) The most important special case is: $S \subset X$ and $\varphi(s) = s$ for $s \in S$. Then, for points $x_0 \in \operatorname{clos}(S)$ and $y_0 \in Y$, the definition

$$f(x) \to y_0 \text{ in } Y \text{ as } x \to x_0 \text{ in } X\,, \text{ i.e. } y_0 = \lim_{x \to x_0} f(x)\,,$$

is equivalent to

$$V \in \mathcal{T}_Y\,,\ y_0 \in V \quad \Longrightarrow \quad \begin{array}{l} \text{There exists a } U \in \mathcal{T}_X \text{ such that } x_0 \in U \\ \text{and } f(U \cap S) \subset V\,, \end{array}$$

in words: For every open set V containing y_0 there exists an open set U containing x_0 such that $f(U \cap S)$ is contained in V.

(4) If in (3) in addition $x_0 \in S$, then it follows that $y_0 = f(x_0)$, i.e.

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

In this case f is called ***continuous at the point*** $\boldsymbol{x_0}$.
(5) If $S = X$, then $f : X \to Y$ is called a ***continuous map*** if f is continuous at all points $x_0 \in X$. This is equivalent to

$$V \in \mathcal{T}_Y \quad \Longrightarrow \quad f^{-1}(V) \in \mathcal{T}_X\,,$$

in words: The mapping f has the property that the inverse image of each open set in Y is open in X.

2.18 Convergence in metric spaces. Let (X, d_X) and (Y, d_Y) be metric spaces and let $A \subset X$ and $f : A \to Y$.

(1) Let $x_0 \in \operatorname{clos}(A)$ and $y_0 \in Y$. Then

$$f(x) \to y_0 \text{ in } Y \text{ as } x \to x_0 \text{ in } X$$

if and only if:

$$\begin{array}{l}\text{For all } \varepsilon > 0 \text{ there exists a } \delta > 0 \text{, such that} \\ x \in A,\ d_X(x, x_0) < \delta \Longrightarrow d_Y\big(f(x), y_0\big) < \varepsilon\,,\end{array}$$

i.e. if and only if

$$d_Y(f(x), y_0) \to 0 \text{ as } d_X(x, x_0) \to 0 \qquad (\text{in } \mathbb{R}).$$

Using quantifiers this property can be written as:

$$\forall\, \varepsilon > 0 \,:\, \exists\, \delta > 0 \,:\, \forall\, x \in A \,:\, d_X(x, x_0) < \delta \Longrightarrow d_Y\big(f(x), y_0\big) < \varepsilon\,.$$

(2) Let $X = \mathbb{K}^n \cup \{\infty\}$ (equipped with the metric in 2.8(3)) and let $A \subset \mathbb{K}^n$ be unbounded. Then $\infty \in \operatorname{clos}(A)$, and $x \to \infty$ in $\mathbb{K}^n \cup \{\infty\}$ means that $|x| \to +\infty$ in $\mathbb{R} \cup \{\pm\infty\}$ (equipped with the metric in 2.8(2)). Let $y \in Y$. Then

$$f(x) \to y \text{ in } Y \text{ as } |x| \to +\infty$$

if and only if:

$$\begin{array}{l}\text{For all } \varepsilon > 0 \text{ there exists a } \delta > 0 \text{, such that} \\ x \in A,\ |x| > \frac{1}{\delta} \Longrightarrow d_Y\big(f(x), y\big) < \varepsilon\,.\end{array}$$

Using quantifiers this property can be written as:

$$\forall\, \varepsilon > 0 \,:\, \exists\, \delta > 0 \,:\, \forall\, x \in A \,:\, |x| > \tfrac{1}{\delta} \Longrightarrow d_Y\big(f(x), y\big) < \varepsilon\,.$$

(3) Let $X = \mathbb{R} \cup \{\pm\infty\}$ (equipped with the metric in 2.8(2)) and let $A = \mathbb{N}$, i.e. $(y_j)_{j \in \mathbb{N}}$ with $y_j := f(j)$ is a sequence in Y. It then holds for $y \in Y$ that

$$y_j \to y \text{ in } Y \text{ as } j \to +\infty$$

if and only if:

$$\text{For all } \varepsilon > 0 \text{ there exists a } k \in \mathbb{N} \text{ such that}$$
$$j \in \mathbb{N},\ j > k \Longrightarrow d_Y(y_j, y) < \varepsilon\,.$$

Using quantifiers this property can be written as:

$$\forall\, \varepsilon > 0 \,:\, \exists\, k \in \mathbb{N} \,:\, \forall\, j \in \mathbb{N} \,:\, j > k \Longrightarrow d_Y(y_j, y) < \varepsilon\,.$$

(4) In metric spaces convergence is equivalent to ***sequential convergence***, that is, the convergence in (1) holds if and only if for all sequences $(x_j)_{j\in\mathbb{N}}$ in A:

$$x_j \to x_0 \text{ as } j \to \infty \quad \Longrightarrow \quad d_Y\big(f(x_j), y_0\big) \to 0 \text{ as } j \to \infty\,. \tag{2-7}$$

Proof (1). Use the fact that balls $\mathrm{B}_\varepsilon(y_0)$ belong to the topology $\mathcal{T}_Y$ induced by d_Y and that

$$y_0 \in V \in \mathcal{T}_Y \quad \Longrightarrow \quad \mathrm{B}_\varepsilon(y_0) \subset V \text{ for an } \varepsilon > 0\,.$$

Likewise in X, every $\mathrm{B}_\delta(x_0) \in \mathcal{T}_X$, and if $x_0 \in U \in \mathcal{T}_X$, then $\mathrm{B}_\delta(x_0) \subset U$ for some $\delta > 0$. □

Proof (2). Follows from (1), on noting that for $0 < \delta' < 1$ for the ball $\mathrm{B}_{\delta'}(\infty)$ with respect to the stereographic projection the following is true:

$$x \in \mathrm{B}_{\delta'}(\infty) \quad \Longleftrightarrow \quad |x| > \sqrt{\delta'^{-2} - 1} =: \delta^{-1}\,.$$

□

Proof (3). Similarly to (2), by choosing $\frac{1}{\delta} \le k < \frac{1}{\delta} + 1$. □

Proof (4). Assume that (1) holds and that $x_j \to x_0$ in X as $j \to \infty$. Then given $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_Y(f(x), y_0) < \varepsilon$ for $x \in A$ with $d_X(x, x_0) < \delta$. Then (3) yields the existence of a $k \in \mathbb{N}$ such that $d_X(x_j, x_0) < \delta$ for $j > k$. Consequently $d_Y(f(x_j), y_0) < \varepsilon$. This proves the claim in (4).

Conversely, assume that the convergence statement in (1) is not true. Then we have to negate the assertion

$$\forall\, \varepsilon > 0 \,:\, \exists\, \delta > 0 \,:\, \forall\, x \in A \,:\, d_X(x, x_0) < \delta \Longrightarrow d_Y\big(f(x), y_0\big) < \varepsilon\,.$$

The negation is:

$$\exists\, \varepsilon > 0 \,:\, \forall\, \delta > 0 \,:\, \exists\, x \in A \,:\, d_X(x, x_0) < \delta \text{ and } d_Y\big(f(x), y_0\big) \ge \varepsilon\,.$$

Consequently there exist an $\varepsilon > 0$ and, for $\delta_j := \frac{1}{j}$, $j \in \mathbb{N}$, an $x_j \in A$ such that

$$d_X(x_j, x_0) < \delta_j \text{ and } d_Y\big(f(x_j), y_0\big) \ge \varepsilon.$$

In particular, $x_j \to x_0$ in X as $j \to \infty$, but $d_Y(f(x_j), y_0) \ge \varepsilon$ for all $j \in \mathbb{N}$. This contradicts (2-7). □

2.19 Note. In 2.18(3) we identified sequences in Y with maps from $\mathbb{N}$ to Y. This can be generalized to arbitrary sets I and A. Here the notation $(a_i)_{i\in I}$, with $a_i \in A$ for $i \in I$, defines a map $i \mapsto a_i$ from I to A. The set of all of these maps is denoted by A^I and I is also called ***index set***,

$$A^I := \{(a_i)_{i\in I} \; ; \; \forall\, i \in I \, : \, a_i \in A\} .$$

In this book, I is usually a subset of $\mathbb{N}$. Examples are the sequence space $\mathbb{K}^{\mathbb{N}}$ in 2.23 and the set $X^{\mathbb{N}}$ in 2.24. In addition one can identify $\mathbb{K}^n$ with $\mathbb{K}^{\{1,\ldots,n\}}$. In general it is important to note that $(a_i)_{i\in I}$ is well distinguished from the subset $\{a_i \in A\,;\; i \in I\} \subset A$ (relevant in e.g. 9.3).

The analysis of limits in metric spaces is often based on inequalities, which we also call "estimates" or "bounds"; this is especially true in function spaces. Usually performing the limit is not trivial and consists of a "nested limit".

2.20 Note (Nested limits). We make the following remark on convergence proofs. By a ***nested limit*** for sequences defined on $\mathbb{N}$ we understand the following. Let $a_i \geq 0$, $b_{k,i} \geq 0$, $c_k \geq 0$ for $i, k \in \mathbb{N}$ with the property

$$a_i \leq \underbrace{b_{k,i}}_{\substack{\to 0 \text{ as } i \to \infty \\ \text{for a fixed } k}} + \underbrace{c_k}_{\to 0 \text{ as } k \to \infty} .$$

From this we deduce that $(a_i)_{i\in\mathbb{N}}$ is a ***null sequence***, i.e.

$$a_i \to 0 \quad \text{as } i \to \infty .$$

To see this, assume that the inequality $a_i \leq b_{k,i} + c_k$ holds for $i, k \in \mathbb{N}$, that $c_k \to 0$ as $k \to \infty$ and that for each $k \in \mathbb{N}$ we have that $b_{k,i} \to 0$ as $i \to \infty$. For an arbitrary $\varepsilon > 0$ we can then choose a $k_\varepsilon \in \mathbb{N}$ such that $c_{k_\varepsilon} < \varepsilon$. Moreover, for this k_ε there exists an i_ε such that $b_{k_\varepsilon,i} < \varepsilon$ for all $i > i_\varepsilon$. Hence we have that

$$a_i \leq b_{k_\varepsilon,i} + c_{k_\varepsilon} < 2\varepsilon \quad \text{for all } i > i_\varepsilon .$$

This proves that $(a_i)_{i\in\mathbb{N}}$ is a null sequence.

This book contains many such limit considerations. A first example you can find in the proof of 2.23(2). In these cases the detailed argumentation will either be omitted, or dramatically shortened to something like:

First choose k large, then choose i large.

Also nested limits with more than two indices are used.

One of the most important concepts in metric spaces is the

2.21 Completeness. Let (X,d) be a metric space.

(1) A sequence $(x_k)_{k\in\mathbb{N}}$ in X is called a ***Cauchy sequence*** if

$$d(x_k,x_l)\to 0 \qquad \text{as } (k,l)\to(\infty,\infty).$$

One usually writes $k,l\to\infty$ in place of $(k,l)\to(\infty,\infty)$.
Remark: Here convergence of $(k,l)\in\mathbb{N}^2\subset(\mathbb{R}\cup\{\pm\infty\})^2$ is understood with respect to the product metric $d_2(a,b):=d_1(a_1,b_1)+d_1(a_2,b_2)$ for $a=(a_1,a_2)$ and $b=(b_1,b_2)$ in $(\mathbb{R}\cup\{\pm\infty\})^2$, where d_1 is the metric on $\mathbb{R}\cup\{\pm\infty\}$ as defined in 2.8(2).

(2) If $(x_k)_{k\in\mathbb{N}}$ is a sequence in X, then a point $x\in X$ is called a ***cluster point*** of this sequence if there exists a ***subsequence*** $(x_{k_i})_{i\in\mathbb{N}}$ (i.e. a sequence $(k_i)_{i\in\mathbb{N}}$ in $\mathbb{N}$ with $k_i\to\infty$ as $i\to\infty$) such that $x=\lim_{i\to\infty}x_{k_i}$.
Remark: The set of all cluster points of a sequence $(x_k)_{k\in\mathbb{N}}$ in X is identical to the closed set

$$\bigcap_{m\in\mathbb{N}}\operatorname{clos}_X\left(\{x_k\in X\,;\;k\geq m\}\right)\,. \tag{2-8}$$

(3) The metric space (X,d) is called ***complete*** if every Cauchy sequence in X has a cluster point in X.
Remark: Because every Cauchy sequence can have at most one cluster point, this means that every Cauchy sequence in X has a limit in X.

2.22 Banach spaces and Hilbert spaces.

(1) A normed $\mathbb{K}$-vector space X is called a ***Banach space*** if it is complete with respect to the induced metric.

(2) A Banach space X is called a ***Banach algebra*** if it is an algebra satisfying

$$\|xy\|_X\leq\|x\|_X\cdot\|y\|_X \qquad \text{for all } x,y\in X. \tag{2-9}$$

Here X is an ***algebra*** if a product $(x,y)\mapsto xy\in X$ is defined on X which satisfies the associative law, the distributive law and $\alpha(xy)=(\alpha x)y=x(\alpha y)$ for all $\alpha\in\mathbb{K}$ and all $x,y\in X$. The algebra is called ***commutative*** if $xy=yx$ for all $x,y\in X$.

(3) A pre-Hilbert space that is complete with respect to the induced metric is called ***Hilbert space***.

The basic example of a complete space is the space of real numbers $\mathbb{R}$, where the axiom of completeness in $\mathbb{R}$ is precisely the additional axiom compared to the space of rational numbers $\mathbb{Q}$. From the completeness of $\mathbb{R}$ one can then deduce (see E2.6) that $\mathbb{R}^n$ and $\mathbb{C}^n$ are complete (with respect to any of the metrics introduced in 2.5 and 2.8). As the simplest infinite-dimensional example we now consider

2.23 Sequence spaces. We denote by $\mathbb{K}^{\mathbb{N}}$ the set of all sequences (defined on $\mathbb{N}$) with values in $\mathbb{K}$:

$$\mathbb{K}^{\mathbb{N}} := \left\{ x = (x_i)_{i\in\mathbb{N}} \ ; \ x_i \in \mathbb{K} \text{ for } i \in \mathbb{N} \right\}.$$

The canonical unit vectors in $\mathbb{K}^{\mathbb{N}}$ are given by

$$\mathbf{e}_i := (0,\ldots,0,\underset{\substack{\uparrow \\ i\text{-th component}}}{1},0,\ldots) \quad \text{for } i \in \mathbb{N}.$$

Then:

(1) The set $\mathbb{K}^{\mathbb{N}}$ becomes a metric space with the Fréchet metric

$$\varrho(x) := \sum_{i\in\mathbb{N}} 2^{-i} \frac{|x_i|}{1+|x_i|} \qquad \text{for } x = (x_i)_{i\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}.$$

(2) Let $x^k = \left(x_i^k\right)_{i\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ and $x = (x_i)_{i\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$. Then

$$\begin{aligned} &\varrho(x^k - x) \to 0 \text{ as } k \to \infty \\ \iff &\text{For every } i : (\ x_i^k \to x_i \text{ as } k \to \infty\ ). \end{aligned}$$

(3) The set $\mathbb{K}^{\mathbb{N}}$ equipped with this metric is complete.

(4) For $x = (x_i)_{i\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ we define

$$\begin{aligned} \|x\|_{\ell^p} &:= \Big(\sum_{i\in\mathbb{N}} |x_i|^p\Big)^{\frac{1}{p}} \in [0,\infty]\,, \qquad \text{if } 1 \le p < \infty, \\ \|x\|_{\ell^\infty} &:= \sup_{i\in\mathbb{N}} |x_i| \in [0,\infty]\,, \end{aligned}$$

and consider for $1 \le p \le \infty$ the set (for the case $0 < p < 1$ see E4.11)

$$\ell^p(\mathbb{K}) := \left\{ x \in \mathbb{K}^{\mathbb{N}} \ ; \ \|x\|_{\ell^p} < \infty \right\}.$$

Then the set $\ell^p(\mathbb{K})$ with the norm $x \mapsto \|x\|_{\ell^p}$ is a Banach space.

(5) If $p = 2$, then $\ell^2(\mathbb{K})$ becomes a Hilbert space with the scalar product

$$(x\,,\,y)_{\ell^2} := \sum_{i\in\mathbb{N}} x_i\,\overline{y_i} \qquad \text{for } x, y \in \ell^2(\mathbb{K}).$$

Proof (1). Let $\varrho_0(s) := \frac{|s|}{1+|s|}$ for $s \in \mathbb{K}$. Then

$$\varrho(x) = \sum_{i=1}^{\infty} 2^{-i}\varrho_0(x_i) \le \sum_{i=1}^{\infty} 2^{-i} = 1\,,$$

and hence $\varrho(x)$ is always finite. The triangle inequality for ϱ follows as in 2.8(1). □

Proof (2). Let $x^k, x \in \mathbb{K}^{\mathbb{N}}$ with $\varrho(x^k - x) \to 0$ as $k \to \infty$. Then $\varrho_0(x_i^k - x_i) \leq 2^i \varrho(x^k - x) \to 0$ for all i, and hence $\left|x_i^k - x_i\right| \to 0$ as $k \to \infty$. Conversely, assuming that $x_i^k \to x_i$ as $k \to \infty$ for all i yields that

$$\varrho(x^k - x) \leq \underbrace{\sum_{i=1}^{j} 2^{-i} \varrho_0(x_i^k - x_i)}_{\to 0 \text{ as } k \to \infty \text{ for any } j} + \underbrace{2^{-j}}_{\to 0 \text{ as } j \to \infty} .$$

Consequently, $\varrho(x^k - x) \to 0$ as $k \to \infty$. □

Proof (3). If $\left(x^k\right)_{k \in \mathbb{N}}$ is a Cauchy sequence in $\mathbb{K}^{\mathbb{N}}$, then, similarly to the above, it follows that $\left(x_i^k\right)_{k \in \mathbb{N}}$ is a Cauchy sequence in $\mathbb{K}$ for any i. Hence there exist the limit

$$x_i := \lim_{k \to \infty} x_i^k \quad \text{in } \mathbb{K} .$$

On setting $x := (x_i)_{i \in \mathbb{N}}$ it follows from (2) that $\varrho(x^k - x) \to 0$ as $k \to \infty$. □

Proof (4). This is a special case of the more general result in 3.16 for the counting measure on $\mathbb{N}$. Here we give a separate proof.

Let $x = (x_i)_{i \in \mathbb{N}}$ and $y = (y_i)_{i \in \mathbb{N}}$ be in $\ell^p(\mathbb{K})$ and for $n \in \mathbb{N}$ define $x^n := (x_1, \dots, x_n)$, $y^n := (y_1, \dots, y_n)$. It follows from 2.5 that

$$|x^n + y^n|_p \leq |x^n|_p + |y^n|_p \leq \|x\|_{\ell^p} + \|y\|_{\ell^p} < \infty .$$

Letting $n \to \infty$ this implies that $x + y \in \ell^p(\mathbb{K})$, with

$$\|x + y\|_{\ell^p} \leq \|x\|_{\ell^p} + \|y\|_{\ell^p} .$$

Hence $\ell^p(\mathbb{K})$ is a normed space. In order to show completeness, let $\left(x^k\right)_{k \in \mathbb{N}}$, with $x^k = \left(x_i^k\right)_{i \in \mathbb{N}} \in \ell^p(\mathbb{K})$, be a Cauchy sequence in $\ell^p(\mathbb{K})$. As $\left|x_i^k - x_i^l\right| \leq \left\|x^k - x^l\right\|_{\ell^p}$ we have that $\left(x_i^k\right)_{k \in \mathbb{N}}$ are Cauchy sequences in $\mathbb{K}$, and hence there exist $x_i := \lim_{k \to \infty} x_i^k \in \mathbb{K}$. This implies for $n \in \mathbb{N}$ in the case $p < \infty$ that as $l \to \infty$

$$\sum_{i=1}^{n} \left|x_i^k - x_i\right|^p \longleftarrow \sum_{i=1}^{n} \left|x_i^k - x_i^l\right|^p \leq \left\|x^k - x^l\right\|_{\ell^p}^p ,$$

and so

$$\left(\sum_{i=1}^{n} \left|x_i^k - x_i\right|^p\right)^{\frac{1}{p}} \leq \limsup_{l \to \infty} \left\|x^k - x^l\right\|_{\ell^p} =: \varepsilon_k < \infty$$

for all n. Hence $x^k - x \in \ell^p(\mathbb{K})$, and consequently $x \in \ell^p(\mathbb{K})$, and it holds that $\left\|x^k - x\right\|_{\ell^p} \leq \varepsilon_k \to 0$ as $k \to \infty$. In the case $p = \infty$ we can argue analogously. □

The set of real numbers $\mathbb{R}$ may be defined as the completion of the rational numbers $\mathbb{Q}$. This procedure can be generalized to arbitrary metric spaces.

2.24 Completion. Let (X,d) be a (not necessarily complete) metric space. Consider the set $X^{\mathbb{N}}$ of all sequences in X and define

$$\widetilde{X} := \left\{ \widetilde{x} = (x_j)_{j\in\mathbb{N}} \in X^{\mathbb{N}} \;;\; (x_j)_{j\in\mathbb{N}} \text{ is a Cauchy sequence in } X \right\}$$

with the equivalence relation

$$(x_j)_{j\in\mathbb{N}} = (y_j)_{j\in\mathbb{N}} \text{ in } \widetilde{X} \quad :\Longleftrightarrow \quad (d(x_j,y_j))_{j\in\mathbb{N}} \text{ is a null sequence.}$$

Then $(\widetilde{X},\widetilde{d})$ is a complete metric space, where $\widetilde{d}$ is defined by

$$\widetilde{d}\big((x_j)_{j\in\mathbb{N}}, (y_j)_{j\in\mathbb{N}}\big) := \lim_{j\to\infty} d(x_j,y_j)\,.$$

Moreover, the rule $J(x) := (x)_{j\in\mathbb{N}}$ defines an injective map $J : X \to \widetilde{X}$ which is ***isometric***, i.e.

$$\widetilde{d}\big(J(x),J(y)\big) = d(x,y) \quad \text{for all } x,y\in X\,.$$

For $(x_j)_{j\in\mathbb{N}} \in \widetilde{X}$ it holds that $\widetilde{d}\big((x_j)_{j\in\mathbb{N}}, J(x_i)\big) \to 0$ as $i\to\infty$, and so $J(X)$ is dense in $\widetilde{X}$.

Conclusion: The above shows that for any metric space (X,d) there exist a complete metric space $(\widetilde{X},\widetilde{d})$ and an injective isometric map $J : X \to \widetilde{X}$ such that $J(X)$ is dense in $\widetilde{X}$. It is then natural to identify elements $x\in X$ with $J(x)\in\widetilde{X}$.

Proof. For $\widetilde{x} = (x_i)_{i\in\mathbb{N}}$ and $\widetilde{y} = (y_i)_{i\in\mathbb{N}}$ in $\widetilde{X}$ we have

$$\begin{aligned}
&|d(x_j,y_j) - d(x_i,y_i)| \le |d(x_j,y_j) - d(x_i,y_j)| + |d(x_i,y_j) - d(x_i,y_i)| \\
&\le d(x_j,x_i) + d(y_j,y_i) \qquad \text{(triangle inequality)} \\
&\to 0 \quad \text{as } i,j\to\infty\,,
\end{aligned}$$

and hence there exists

$$\widetilde{d}(\widetilde{x},\widetilde{y}) := \lim_{i\to\infty} d(x_i,y_i)\,.$$

Similarly, it follows for $\widetilde{x}^1 = \widetilde{x}^2$ in $\widetilde{X}$ and $\widetilde{y}^1 = \widetilde{y}^2$ in $\widetilde{X}$ that

$$\big|d(x_i^2,y_i^2) - d(x_i^1,y_i^1)\big| \to 0 \quad \text{as } i\to\infty.$$

This shows that $\widetilde{d} : \widetilde{X}\times\widetilde{X} \to \mathbb{R}$ is well defined (see the remark in 2.4). Furthermore, it follows that $\widetilde{d}(\widetilde{x},\widetilde{y}) = 0$ if $\widetilde{x} = \widetilde{y}$ in $\widetilde{X}$, and the triangle inequality carries over from d to $\widetilde{d}$. Hence $\widetilde{d}$ is a metric on $\widetilde{X}$.

In order to show completeness, let $\big(x^k\big)_{k\in\mathbb{N}}$ be a Cauchy sequence in $\widetilde{X}$, where $x^k = \big(x_j^k\big)_{j\in\mathbb{N}}$ for $k\in\mathbb{N}$. Given $k\in\mathbb{N}$ choose j_k such that

$$d(x_i^k, x_j^k) \le \tfrac{1}{k} \quad \text{for } i, j \ge j_k\,.$$

Then

$$\begin{aligned} d(x_{j_k}^k, x_{j_l}^l) &\le d(x_{j_k}^k, x_j^k) + d(x_j^k, x_j^l) + d(x_j^l, x_{j_l}^l) \\ &\le \frac{1}{k} + d(x_j^k, x_j^l) + \frac{1}{l} \quad \text{for } j \ge j_k, j_l \\ &\to \frac{1}{k} + \widetilde{d}\big(x^k, x^l\big) + \frac{1}{l} \quad \text{as } j \to \infty \\ &\to 0 \quad \text{as } k, l \to \infty. \end{aligned} \tag{2-10}$$

Hence we have that $x^\infty := \big(x_{j_l}^l\big)_{l\in\mathbb{N}} \in \widetilde{X}$ and

$$\begin{aligned} \widetilde{d}(x^l, x^\infty) &\leftarrow d(x_k^l, x_k^\infty) \quad \text{as } k \to \infty \\ &\le d(x_k^l, x_{j_l}^l) + d(x_{j_l}^l, x_{j_k}^k) \le \frac{1}{l} + d(x_{j_l}^l, x_{j_k}^k) \quad \text{for } k \ge j_l \\ &\to 0 \quad \text{as } k, l \to \infty \quad \text{(recall (2-10)).} \end{aligned}$$

The assertions on J are easily verified. □

This means that every metric space that is not complete can be extended to a complete space. Examples of completions are the space of Lebesgue integrable functions in Appendix A3 and the Sobolev spaces in 3.27.

E2 Exercises

E2.1 Open and closed sets. If $(X, \mathcal{T})$ is a topological space, then it holds for $A \subset X$ that:

(1) $X \setminus \operatorname{clos}(A) = \operatorname{intr}(X \setminus A)$.

(2) $\operatorname{intr}(A)$ is open, and $\operatorname{clos}(A)$ is closed.

(3) $A \in \mathcal{T} \iff A = \operatorname{intr}(A)$.

(4) $X \setminus A \in \mathcal{T} \iff A = \operatorname{clos}(A)$.

Solution (1). From the negation of the definition of a closure in 2.11 it follows for $x \in X \setminus \operatorname{clos}(A)$ that there exists an $U \in \mathcal{T}$ with $x \in U$ and $U \cap A = \emptyset$. This means $U \subset X \setminus A$ and $U \in \mathcal{T}$ with $x \in U$, and this is the definition of a point $x \in \operatorname{intr}(X \setminus A)$. □

Solution (2). Let $\mathcal{T}' := \{U \in \mathcal{T};\ U \subset A,\ U \cap \operatorname{intr}(A) \ne \emptyset\}$. On recalling the definition of the interior of A we then have that

$$\operatorname{intr}(A) \subset V := \bigcup_{U \in \mathcal{T}'} U \in \mathcal{T}\,.$$

Moreover, $x \in U \in \mathcal{T}'$ implies that $U \in \mathcal{T}$ and $x \in U \subset A$, and so $x \in \operatorname{intr}(A)$. Hence, $\operatorname{intr}(A) = V \in \mathcal{T}$. The second claim now follows from (1). □

Solution (3). If $A \in \mathcal{T}$ and $x \in A$, then $x \in U := A$ with $U \in \mathcal{T}$, and so $A \subset \operatorname{intr}(A) \subset A$. Conversely, $A = \operatorname{intr}(A) \in \mathcal{T}$ by (2). □

Solution (4). Follows from (3) on noting (1). □

E2.2 Distance and neighbourhoods. Let (X, d) be a metric space and $A \subset X$. Then:

(1) $\operatorname{dist}(\bullet, A)$ is a Lipschitz continuous function with Lipschitz constant ≤ 1, where equality holds if $X \setminus \overline{A}$ is nonempty.

(2) The neighbourhoods $\mathrm{B}_r(A)$ for $r > 0$ are open sets. In particular, all balls $\mathrm{B}_r(x)$ for $x \in X$ and $r > 0$ are open.

(3) For $r_1, r_2 > 0$, one has $\mathrm{B}_{r_1}(\mathrm{B}_{r_2}(A)) \subset \mathrm{B}_{r_1+r_2}(A)$, and equality holds if X is a normed space.

Solution (1). Let $x, y \in X$. Given $\varepsilon > 0$ choose $a \in A$ such that $d(x, a) \leq \operatorname{dist}(x, A) + \varepsilon$. On employing the triangle inequality it then follows that

$$\operatorname{dist}(y, A) - \operatorname{dist}(x, A) \leq d(y, a) - d(x, a) + \varepsilon \leq d(y, x) + \varepsilon\,.$$

A symmetry argument then yields that

$$|\operatorname{dist}(y, A) - \operatorname{dist}(x, A)| \leq d(x, y)\,.$$

This corresponds to the definition of Lipschitz continuity in 3.7 with Lipschitz constant ≤ 1. If $x \in X \setminus \overline{A}$, then $\mathrm{B}_\varepsilon(x) \cap A = \emptyset$ for an $\varepsilon > 0$, and hence $\operatorname{dist}(x, A)$ is positive. Now choose for every $\varepsilon > 0$ a $y \in A$ such that $d(x, y) \leq (1 + \varepsilon)\operatorname{dist}(x, A)$. It follows that

$$|\operatorname{dist}(y, A) - \operatorname{dist}(x, A)| = \operatorname{dist}(x, A) \geq \frac{1}{1 + \varepsilon} d(x, y)\,,$$

which shows that the Lipschitz constant is equal to 1. □

Solution (2). Let $x \in \mathrm{B}_r(A)$ and $\delta := r - \operatorname{dist}(x, A) > 0$. If $y \in \mathrm{B}_\delta(x)$, then, by (1),

$$\operatorname{dist}(y, A) \leq \operatorname{dist}(x, A) + d(x, y) < \operatorname{dist}(x, A) + \delta = r\,,$$

and so $\mathrm{B}_\delta(x) \subset \mathrm{B}_r(A)$. □

Solution (3). Let $x \in \mathrm{B}_{r_1}(\mathrm{B}_{r_2}(A))$, i.e. $\operatorname{dist}(x, \mathrm{B}_{r_2}(A)) < r_1$. Then there exists a $y \in \mathrm{B}_{r_2}(A)$ with $d(x, y) < r_1$. It follows from (1) that

$$\operatorname{dist}(x, A) \leq \operatorname{dist}(y, A) + d(x, y) < r_2 + r_1\,.$$

Now let X be a normed space and $x \in \mathrm{B}_{r_1+r_2}(A)$. Then there exists a $y \in A$ with $\|x - y\| < r_1 + r_2$. It follows for

$$z := (1 - s)x + sy\,, \quad s := \frac{r_2}{r_1 + r_2}\,,$$

that

$$\|z - y\| = (1-s)\|x-y\| < r_1 \quad \text{and} \quad \|x - z\| = s\|x-y\| < r_2\,,$$

and so $x \in \mathrm{B}_{r_2}(\mathrm{B}_{r_1}(A))$. □

E2.3 Construction of metrics. Let $\psi : [0,\infty[\; \to [0,\infty[$ be a continuously differentiable strictly monotone function with $\psi(0) = 0$ and nonincreasing derivative ψ'. Then

$$d \text{ is a metric on } X \quad \Longrightarrow \quad \psi \circ d \text{ is a metric on } X\,.$$

Example:

$$\psi(t) := \frac{t}{1+t}\,.$$

Solution. We have to verify the metric axioms in 2.6 for $\psi \circ d$. The axiom (M1) is satisfied, since

$$\psi(d(x,y)) = 0 \quad \Longleftrightarrow \quad d(x,y) = 0 \quad \Longleftrightarrow \quad x = y\,.$$

The axiom (M3) follows from

$$\begin{aligned}\psi(d(x,y)) \le \psi(d(x,z) + d(z,y)) &= \psi(d(x,z)) + \int_0^{d(z,y)} \psi'(d(x,z)+t)\,\mathrm{d}t \\ &\le \psi(d(x,z)) + \int_0^{d(z,y)} \psi'(t)\,\mathrm{d}t = \psi(d(x,z)) + \psi(d(z,y))\,.\end{aligned}$$

□

E2.4 Convergence. Prove the assertions on convergence in 2.17.

Proof 2.17(1). Assume that $f(s) \to y_1$ and $f(s) \to y_2$ in Y as $\varphi(s) \to x_0$ with $y_1 \ne y_2$. As Y is a Hausdorff space, there exist $y_1 \in V_1 \in \mathcal{T}_Y$ and $y_2 \in V_2 \in \mathcal{T}_Y$ such that $V_1 \cap V_2 = \emptyset$. However, the definition of convergence yields that there exists a $U_1 \in \mathcal{T}_X$ such that $x_0 \in U_1$ and $f\big(\varphi^{-1}(U_1)\big) \subset V_1$, and then a $U_2 \in \mathcal{T}_X$ such that $x_0 \in U_2 \subset U_1, \varphi^{-1}(U_2) \ne \emptyset$ and $f\big(\varphi^{-1}(U_2)\big) \subset V_2$. As $U_2 \subset U_1$ it follows that $f\big(\varphi^{-1}(U_2)\big) \subset V_2 \cap V_1 = \emptyset$, and so $\varphi^{-1}(U_2) = \emptyset$, which is a contradiction. □

Proof 2.17(2). For $x_0 \in U_0 \in \mathcal{T}_X$ the definition of convergence gives that $\varphi^{-1}(U_0) \ne \emptyset$, i.e. $\varphi(S) \cap U_0 \ne \emptyset$, and so $x_0 \in \operatorname{clos}(\varphi(S))$. In addition it follows from the definition of convergence that for $y_0 \in V_0 \in \mathcal{T}_Y$ there exists an $s \in S$ with $f(s) \in V_0$, and so $y_0 \in \operatorname{clos}(f(S))$. □

Proof 2.17(3). Choosing $U_0 = X$ and $V_0 = V$ yields convergence in 2.17(3). Conversely, set $V = V_0$. Then if $x_0 \in U \in \mathcal{T}_X$ with $f(U \cap S) \subset V$ as in 2.17(3), it holds for $\widetilde{U} = U \cap U_0$ that

$$\widetilde{U} \cap S \ne \emptyset \text{ (since } x_0 \in \operatorname{clos}(S)) \text{ and } f(\widetilde{U} \cap S) \subset V_0\,.$$

□

Proof 2.17(4). Let $y_0 \in V \in \mathcal{T}_Y$ and then U as in 2.17(3). It follows from $x_0 \in U \cap S$ that $f(x_0) \in V$. As Y is a Hausdorff space, this implies that $f(x_0) = y_0$. □

Proof 2.17(5). Let f be continuous, and $V \in \mathcal{T}_Y$ with $x_0 \in f^{-1}(V)$. Since f is continuous at x_0, there exists a $U \in \mathcal{T}_X$ such that $x_0 \in U$ and $f(U) \subset V$, i.e. $x_0 \in U \subset f^{-1}(V)$. Hence $f^{-1}(V) \in \mathcal{T}_X$. Conversely, let $x_0 \in X$ and $f(x_0) \in V \in \mathcal{T}_Y$. Then $x_0 \in U := f^{-1}(V) \in \mathcal{T}_X$, which proves the continuity of f in x_0. □

E2.5 Examples of continuous maps.

(1) Let $\mathcal{T}_1$, $\mathcal{T}_2$ be two topologies on X. Then the ***identity*** $\mathrm{Id} : X \to X$, defined by $\mathrm{Id}(x) := x$, is a continuous map from $(X, \mathcal{T}_2)$ to $(X, \mathcal{T}_1)$ if and only if $\mathcal{T}_2$ is stronger than $\mathcal{T}_1$.

(2) If (X, d) is a metric space, then $d : X \times X \to \mathbb{R}$ is continuous.

(3) If $(X, \|\bullet\|)$ is a normed space, then the norm is a continuous map from X to $\mathbb{R}$.

(4) Let $(\bullet_1 , \bullet_2)$ be a scalar product on the $\mathbb{K}$-vector space X, let $\|\bullet\|$ be the corresponding induced norm and consider the normed space $(X, \|\bullet\|)$. Then the scalar product is a continuous map from $X \times X$ to $\mathbb{K}$.

Solution (2). Use E2.2(1). □

Solution (3). This follows from (2) and the definition of the induced metric in 2.6. □

Solution (4). Employ the Cauchy-Schwarz inequality 2.2(2). □

E2.6 Completeness of Euclidean space. The set $\mathbb{K}^n$ is complete with respect to all of the metrics given in 2.5 and 2.8.

Solution. First show the completeness with respect to the ∞-norm in 2.5: If $\left(x^k\right)_{k \in \mathbb{N}}$ is a Cauchy sequence with respect to this norm, $x^k = \left(x_i^k\right)_{i=1,\dots,n}$, then $\left|x_i^k - x_i^l\right| \le \left\|x^k - x^l\right\|_\infty$, and so $\left(x_i^k\right)_{k \in \mathbb{N}}$ are Cauchy sequences in $\mathbb{K}$, which means that there exist $x_i = \lim_{k \to \infty} x_i^k$ in $\mathbb{K}$ (because $\mathbb{R}$ and $\mathbb{C}$ are complete, with the completeness of the latter following from that of $\mathbb{R}^2$, which is shown here). Hence $\left|x_i^k - x_i\right| \to 0$ as $k \to \infty$ for every $i \in \{1, \dots, n\}$, which implies that $\left\|x^k - x\right\|_\infty \to 0$ as $k \to \infty$.

The completeness with respect to the other metrics then follows from the results in 2.16. □

E2.7 Incomplete function space. Let $I := [a, b] \subset \mathbb{R}$ be an interval with $a < b$, and for $n \in \mathbb{N}$ let

$$\mathcal{P}_n := \{f : I \to \mathbb{R};\ f \text{ is a polynomial of degree } \le n\}.$$

Then $\mathcal{P} := \bigcup_{n\in\mathbb{N}} \mathcal{P}_n$ equipped with

$$\|f\|_\infty := \sup_{x\in I} |f(x)| \quad \text{for } f \in \mathcal{P}$$

is a normed space that is not complete.

Solution. The norm axioms are easily verified. Setting

$$f(x) := \mathrm{e}^x = \sum_{i=1}^{\infty} \frac{1}{i!}x^i\,, \quad f_n(x) := \sum_{i=1}^{n} \frac{1}{i!}x^i$$

we have that

$$\sup_{x\in I} |f_n(x) - f(x)| \to 0 \quad \text{as } n \to \infty.$$

Hence $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{P}$. If $g = \lim_{n\to\infty} f_n$ existed in $\mathcal{P}$, it would follow that $|f_n(x) - g(x)| \le \|f_n - g\|_\infty \to 0$ as $n \to \infty$ for all $x \in I$, and so $g = f \notin \mathcal{P}$, which is a contradiction. □

E2.8 On completeness. Let (X, d) be a metric space. Then:

(1) If (X, d) is complete and $Y \subset X$ is closed, then (Y, d) is also a complete metric space.

(2) If $Y \subset X$ and (Y, d) is complete, then Y is closed in X (as a subset of the metric space (X, d)).

Solution (1). If $\left(x^k\right)_{k\in\mathbb{N}}$ is a Cauchy sequence in Y, then it is also a Cauchy sequence in X. The completeness of X yields that it has a limit $x \in X$. As Y is closed it follows that $x \in Y$. □

Solution (2). Let $\left(x^k\right)_{k\in\mathbb{N}}$ be a sequence in Y converging in X to $x \in X$. Since Y is equipped with the metric d, it is a Cauchy sequence in Y. The completeness of Y yields that it has a limit $y \in Y$. Now y must also be the limit of the sequence in X, and so $x = y \in Y$. □

E2.9 Hausdorff distance between sets. Let (X, d) be a metric space and

$$\mathcal{A} := \{A \subset X\,;\; A \text{ is nonempty, bounded and closed}\}\,.$$

The ***Hausdorff distance*** between $A_1 \in \mathcal{A}$ and $A_2 \in \mathcal{A}$ is defined by

$$d_H(A_1, A_2) := \inf\{\varepsilon > 0\,;\; A_1 \subset \mathrm{B}_\varepsilon(A_2) \text{ and } A_2 \subset \mathrm{B}_\varepsilon(A_1)\}\,.$$

Then d_H is a metric on $\mathcal{A}$, and for $A, B \in \mathcal{A}$ we have

$$\begin{aligned} d_H(A, B) &= \max\left(\sup_{a\in A} \mathrm{dist}(a, B),\; \sup_{b\in B} \mathrm{dist}(b, A)\right) \\ &= \sup_{x\in M} |\mathrm{dist}(x, A) - \mathrm{dist}(x, B)| \end{aligned}$$

for any set M with $A \cup B \subset M \subset X$.

Solution. If $d_H(A_1, A_2) = 0$, then

$$A_1 \subset \bigcap_{\varepsilon>0} \mathrm{B}_\varepsilon(A_2) = \overline{A_2} = A_2\,,$$

and similarly $A_2 \subset A_1$. Moreover, d_H is symmetric by definition. Given $A_1, A_2, A_3 \in \mathcal{A}$ and $\delta > 0$, there exist numbers $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ such that

$$\begin{aligned}\varepsilon_1 &\le d_H(A_1, A_2) + \delta,\ A_1 \subset \mathrm{B}_{\varepsilon_1}(A_2)\,,\ A_2 \subset \mathrm{B}_{\varepsilon_1}(A_1)\,,\\ \varepsilon_2 &\le d_H(A_2, A_3) + \delta,\ A_2 \subset \mathrm{B}_{\varepsilon_2}(A_3)\,,\ A_3 \subset \mathrm{B}_{\varepsilon_2}(A_2)\,.\end{aligned}$$

By E2.2(3),

$$\begin{aligned}A_1 &\subset \mathrm{B}_{\varepsilon_1}(\mathrm{B}_{\varepsilon_2}(A_3)) \subset \mathrm{B}_{\varepsilon_1+\varepsilon_2}(A_3)\ ,\\ A_3 &\subset \mathrm{B}_{\varepsilon_2}(\mathrm{B}_{\varepsilon_1}(A_1)) \subset \mathrm{B}_{\varepsilon_1+\varepsilon_2}(A_1)\ ,\end{aligned}$$

and hence

$$d_H(A_1, A_3) \le \varepsilon_1 + \varepsilon_2 \le d_H(A_1, A_2) + d_H(A_2, A_3) + 2\delta\,.$$

This shows that d_H defines a metric.

Now let $A, B \in \mathcal{A}$, $d := d_H(A, B)$ and

$$\begin{aligned}d_{\max} &:= \max\left(\sup_{a\in A}\operatorname{dist}(a, B),\ \sup_{b\in B}\operatorname{dist}(b, A)\right),\\ d_{\sup} &:= \sup_{x\in M} |\operatorname{dist}(x, A) - \operatorname{dist}(x, B)|\,.\end{aligned}$$

Then $d_{\sup} \ge d_{\max}$, on noting that

$$d_{\sup} \ge \sup_{x\in B} |\operatorname{dist}(x, A) - 0|\,,$$

and applying a symmetry argument. Moreover, $d_{\max} \ge d$, as for $\delta > 0$ we have that

$$B \subset \mathrm{B}_{d_{\max}+\delta}(A)\ ,$$

and hence, by a symmetry argument, that $d_{\max} + \delta \ge d$. Furthermore, $d \ge d_{\max}$, since $B \subset \mathrm{B}_\varepsilon(A)$ and $A \subset \mathrm{B}_\varepsilon(B)$ implies that

$$\operatorname{dist}(b, A) < \varepsilon \quad \text{for } b \in B \quad \text{and} \quad \operatorname{dist}(a, B) < \varepsilon \quad \text{for } a \in A\,,$$

and so $d_{\max} \le \varepsilon$. Finally, $d_{\max} \ge d_{\sup}$, because for $x \in X$ and $\delta > 0$, there exists a $b \in B$ such that

$$\operatorname{dist}(x, B) \ge d(x, b) - \delta\,.$$

Thanks to E2.2(1),

$$\operatorname{dist}(x, A) - \operatorname{dist}(x, B) \le \operatorname{dist}(x, A) - d(x, b) + \delta \le \operatorname{dist}(b, A) + \delta\,,$$

and hence, by a symmetry argument, $d_{\sup} \le d_{\max} + \delta$. □