14. Metric space convergence

We are ready at last to commence the discussion of topics that should be familiar from calculus (which can be thought of as a non-technical, or practical-use oriented, version of analysis). We start with limits of sequences.

14.1 Convergence of real-valued sequences.

The concept of *convergence* is fundamental for analysis. We will first discuss it in the context of convergence of sequences. Recall that the notion of a sequence was introduced already in Definition 12.8 where we defined an A-valued sequence $\{x_n\}_{n\in\mathbb{N}}$ to be a function $\mathbb{N} \to A$ with $\mathrm{Dom}(f) = \mathbb{N}$ whose value at n is x_n. We will largely suppress this technical interpretation in what follows and think of a sequence intuitively as a line-up of objects indexed by the naturals.

Consider the following example of a sequence $\{a_n\}_{n\in\mathbb{N}}$ taking values in rationals which is defined recursively as

$$a_0 := 1 \;\wedge\; \forall n \in \mathbb{N}\colon\; a_{n+1} := 3 - \frac{1}{a_n} \tag{14.1}$$

It is easy to evaluate a couple of first terms,

$$a_0 = 1,\, a_1 = 2,\, a_3 = 2.5,\, a_4 = 2.6,\, a_5 = 2.615\dots \tag{14.2}$$

It appears that the values rise with the rising index, albeit not above the value 3. And, indeed, we easily prove:

Lemma 14.1 $\forall n \in \mathbb{N}\colon\; 1 \leqslant a_n \leqslant 3 \;\wedge\; a_n < a_{n+1}$

Proof. Let $P_n := 1 \leqslant a_n \leqslant 3 \;\wedge\; a_n < a_{n+1}$. Noting that $1 = a_0 < a_1 = 2 \leqslant 3$ we get that P_0 is TRUE. Assuming P_n, we have $\frac{1}{a_{n+1}} < \frac{1}{a_n}$ and so

$$a_{n+2} = 3 - \frac{1}{a_{n+1}} < 3 - \frac{1}{a_n} = a_{n+1} \tag{14.3}$$

Since $\frac{1}{a_n} \geqslant 0$, last equality also gives $a_{n+1} \leqslant 3$ (in fact, as $\frac{1}{a_n} \geqslant \frac{1}{3}$ we even have $a_{n+1} \leqslant 2.666\dots$) while the fact that $a_n \geqslant 1$ implies $1 \leqslant a_n \leqslant a_{n+1}$. Hence $P_n \Rightarrow P_{n+1}$ and so the claim holds by induction. □

We have thus verified that $\{a_n\}_{n\in\mathbb{N}}$ conforms to:

Definition 14.2 *Let $(A, \leqslant)$ be a partially ordered set. An sequence $\{x_n\}_{n\in\mathbb{N}}$ taking values in A is then said to be*

- non-decreasing *if* $\forall n \in \mathbb{N}\colon\; x_n \leqslant x_{n+1}$ *and* strictly increasing *if* $\forall n \in \mathbb{N}\colon\; x_n < x_{n+1}$
- non-increasing *if* $\forall n \in \mathbb{N}\colon\; x_{n+1} \leqslant x_n$ *and* strictly decreasing *if* $\forall n \in \mathbb{N}\colon\; x_{n+1} < x_n$

Such sequences are generally referred to as monotone*.*

We remark that the words *increasing*, resp., *decreasing* are used as colloquial equivalents of non-decreasing, resp., non-increasing, but the uncertainty which of "$\leqslant$" or "$<$" is meant makes them less desirable when precision is of concern.

Another observation that may be gleaned from (14.2) is that the values of the sequence are getting closer and closer together, and perhaps even approach a "limit" point. Here are the precise definitions of these intuitive terms:

Definition 14.3 (Cauchy sequence in $\mathbb{R}$) *A real-valued sequence $\{x_n\}_{n\in\mathbb{N}}$ is said to be Cauchy if*

$$\forall k \in \mathbb{N}\,\exists n_0 \in \mathbb{N}\,\forall m, n \in \mathbb{N}\colon\;\; n, m \geqslant n_0 \;\Rightarrow\; |x_m - x_n| < \frac{1}{k+1} \tag{14.4}$$

Definition 14.4 (Limit of $\mathbb{R}$-valued sequence) *An real-valued sequence $\{x_n\}_{n\in\mathbb{N}}$ is said to have a limit, or converges, if*

$$\exists L \in \mathbb{R}\,\forall k \in \mathbb{N}\,\exists n_0 \in \mathbb{N}\,\forall m, n \in \mathbb{N}\colon\;\; n \geqslant n_0 \;\Rightarrow\; |x_n - L| < \frac{1}{k+1} \tag{14.5}$$

Any such L is then called a limit of $\{x_n\}_{n\in\mathbb{N}}$. We abbreviate (14.5) *as $x_n \to L$.*

Indeed, we readily prove:

Lemma 14.5 *The sequence $\{a_n\}_{n\in\mathbb{N}}$ from* (14.1) *is Cauchy.*

Proof. Let $n \in \mathbb{N}$. A calculation shows

$$a_{n+2} - a_{n+1} = \Bigl(3 - \frac{1}{a_{n+1}}\Bigr) - \Bigl(3 - \frac{1}{a_n}\Bigr) = \frac{1}{a_n} - \frac{1}{a_{n+1}} = \frac{a_{n+1} - a_n}{a_n a_{n+1}} \tag{14.6}$$

Taking absolute values and noting that $a_n \geqslant 1$ but $a_{n+1} \geqslant a_1 = 2$ gives

$$|a_{n+2} - a_{n+1}| \leqslant \frac{|a_{n+1} - a_n|}{a_n a_{n+1}} \leqslant \frac{1}{2}|a_{n+1} - a_n| \tag{14.7}$$

We then use induction to verify that, for all $n \in \mathbb{N}$,

$$|a_{n+1} - a_n| \leqslant 2^{-n}|a_1 - a_0| = 2^{-n} \tag{14.8}$$

and then, for all $m, n \geqslant N$,

$$|a_m - a_n| \leqslant 2^{-N+1} \tag{14.9}$$

Since the right-hand side is decreasing in N, it then suffices to show that, for each $k \in \mathbb{N}$ there is $N \in \mathbb{N}$ such that $2^{-N+1} \leqslant \frac{1}{k+1}$. Since $\forall n \in \mathbb{N}\colon n + 1 \leqslant 2^n$ as verified again by induction, it suffices to choose $N := k + 1$. □

A very similar argument also gives:

Lemma 14.6 *The sequence $\{a_n\}_{n\in\mathbb{N}}$ from* (14.1) *is convergent with limit $L := \frac{1+\sqrt{5}}{2}$.*

Proof. We start by explaining where the precise value of L comes from. If we already know that the sequence is convergent, both a_n and a_{n+1} are close to L for all n large. Replacing these values by L in the recursive formula produces the equation $L = 3 - \frac{1}{L}$. This is a quadratic equation whose only positive solution is $\frac{1+\sqrt{5}}{2}$.

We now convert this argument to a proof that L is a limit of $\{a_n\}_{n\in\mathbb{N}}$. Suppose that L satisfies $L = 3 - \frac{1}{L}$. Then

$$L - a_{n+1} = \Bigl(3 - \frac{1}{L}\Bigr) - \Bigl(3 - \frac{1}{a_n}\Bigr) = \frac{1}{a_n} - \frac{1}{L} = \frac{L - a_n}{a_n L} \tag{14.10}$$

Taking absolute value and using that $L \geqslant \frac{3}{2}$ gives that

$$|L - a_{n+1}| \leqslant \frac{2}{3}\,|L - a_n| \tag{14.11}$$

Assume that $L \geqslant 3/2$. Then we again readily prove by induction that, for all $n \in \mathbb{N}$,

$$|L - a_n| \leqslant \left(\frac{2}{3}\right)^n |L - a_0| \leqslant \left(\frac{2}{3}\right)^n |L - 1| \tag{14.12}$$

where we also used that $L \leqslant 2$. In order to complete the proof, we need a lemma whose proof (based on Archimedean principle) we leave to the reader:

Lemma 14.7 $\forall x, y \in \mathbb{R}\colon\ x, y > 0 \ \wedge\ x < 1 \ \Rightarrow\ \exists n \in \mathbb{N}\colon x^n < y$

Indeed, given any $k \in \mathbb{N}$ and setting $y := \frac{1}{k+1}|L-1|^{-1}$ and $x := \frac{2}{3}$, this lemma gives us $n_0 \in \mathbb{N}$ such that the right-hand side of (14.12) is less than $\frac{1}{k+1}$ for all $n \geqslant n_0$. □

Some remarks are in order. First, note that while $\{a_n\}_{n\in\mathbb{N}}$ is $\mathbb{Q}$-valued, the limit L is not rational. It thus follows that the sequence is Cauchy even as as $\mathbb{Q}$-valued sequence, but it is not convergent in $\mathbb{Q}$. This is because being Cauchy is an *intrinsic* property of the sequence while being convergent depends also on the ambient space in which the sequence is immersed. We will return to this question more systematically later.

Second, the above procedure for controlling limit behavior of recursively defined sequences works quite generally. Indeed, if $f\colon \mathbb{R} \to \mathbb{R}$ is a given function, we can define the sequence by $\forall n \in \mathbb{N}\colon a_{n+1} := f(a_n)$ starting from the initial value a_0. The limit L, if such exist, will typically be a solution to the equation $L = f(L)$, meaning that L is a *fixed point* of f. An example of this for $f(x) := \frac{1}{1+x}$ appears as a homework exercise.

14.2 Metric spaces.

Having digested the above notions in the context of real-valued sequences, we now turn to their generalizations beyond the reals. Here we note that what made Definitions 14.3 and 14.4 work was that in $\mathbb{R}$ we have a natural notion of *closeness*. Indeed, we say that x and y are close if $|x - y|$ is small. We thus put forward:

Definition 14.8 (Metric space) *A* metric space *is a pair (X, ϱ), where X is a set and $\varrho\colon X \times X \to \mathbb{R}$ is a function that obeys:*

(1) *(positivity)* $\forall x, y \in X\colon\ \varrho(x,y) \geqslant 0 \ \wedge\ (\varrho(x,y) = 0 \Leftrightarrow x = y)$
(2) *(symmetry)* $\forall x, y \in X\colon\ \varrho(x,y) = \varrho(y,x)$
(3) *(triangle inequality)* $\forall x, y, z \in X\colon\ \varrho(x,y) \leqslant \varrho(x,z) + \varrho(z,x)$

We call any such ϱ a metric *on X.*

The triangle inequality expresses the intuitive fact that the passage from x to y via z will be at least as long as the shortest possible way. Metric thus axiomatizes the intuitive notion of *distance*, with both words used synonymously in practice.

Using the definition of absolute value it is fairly easy to check that

$$\varrho(x,y) := |x - y| \tag{14.13}$$

defines a metric on $\mathbb{R}$ (and also on any subset thereof; in particular, on $\mathbb{Q}$). However, as asked to show in the homework, also

$$\varrho'(x,y) := \Big|\frac{x}{1+|x|} - \frac{y}{1+|y|}\Big| \tag{14.14}$$

is a metric on $\mathbb{R}$.

The metric (14.13) finds a number of possible generalizations in $\mathbb{R}^d$. Arguably the most natural of these is the *Euclidean metric* which, for points $x = (x_1,\dots,x_d)$ and $y = (y_1,\dots,y_d)$ is given by

$$\varrho_2(x,y) := \Big(\sum_{i=1}^d (x_i - y_i)^2\Big)^{1/2} \tag{14.15}$$

This is linked to the one-dimensional case by the fact that $\varrho_2(x,y)$ is the length of a straight line segment between x and y as measured by the metric (14.13).

However, the Euclidean metric is not the only metric on $\mathbb{R}^d$ that is linked to (14.13). One other such metric is the ∞*-metric*

$$\varrho_\infty(x,y) := \max_{i=1,\dots,d} |x_i - y_i| \tag{14.16}$$

which correspond to the largest difference of the coordinates of the two points. Here the triangle inequality is verified by noting that, for all $x,y,z \in \mathbb{R}^d$ and $i = 1,\dots,d$,

$$|x_i - y_i| \leqslant |x_i - z_i| + |z_i - y_i| \leqslant \varrho(x,z) + \varrho(z,y) \tag{14.17}$$

This metric is sometimes called the chessboard distance, because it corresponds to the minimal number of moves a king needs to get from one point to another on a chessboard.

The metrics ϱ_2 and ϱ_∞ happen to be special instances from a whole one-parameter family of *p-metrics* on $\mathbb{R}^d$. These are indexed by a real-valued parameter $p \in [1,\infty]$ and, for p finite, they are given by

$$\varrho_p(x,y) := \Big(\sum_{i=1}^d |x_i - y_i|^p\Big)^{1/p} \tag{14.18}$$

To see that the ∞-metric belongs to this family, we check that $\varrho_p(x,y)$ converges to $\varrho_\infty(x,y)$ as $p \to \infty$. (This requires concepts we have not yet talked about in detail.) The $p = 1$ case is sometimes referred to as the Manhattan distance, or taxicab distance, as it corresponds to total distance traveled via a square grid.

In order to prove that ϱ_p are actually metrics (for $p \in [1,\infty]$) we use the fact that ϱ_p actually arises from a norm. The latter is a concept defined for all vector spaces:

Definition 14.9 (Norm) *Given a vector space V over a field F such that $F = \mathbb{R} \vee F = \mathbb{C}$, a function $\|\cdot\|\colon V \to \mathbb{R}$ is a* norm *if it satisfies the following requirements:*

(1) *(positivity)* $\forall u \in V\colon\ 0 \leqslant \|u\| \wedge (\|u\| = 0 \Leftrightarrow u = 0)$
(2) *(homogeneity)* $\forall u \in V\ \forall \lambda \in F\colon\ \|\lambda u\| = |\lambda|\|u\|$
(3) *(triangle inequality)* $\forall u,v \in V\colon\ \|u+v\| \leqslant \|u\| + \|v\|$.

A vector space V endowed with a norm $\|\cdot\|$ is called a normed space.

We then check:

Lemma 14.10 *Let $\|\cdot\|$ be a norm on vector space V. Then $\varrho\colon V\times V\to\mathbb{R}$ defined by*

$$\varrho(x,y) := \|x-y\| \tag{14.19}$$

is a metric on V.

Proof. The positivity of ϱ follows from the positivity of $\|\cdot\|$ while homogeneity of the latter implies symmetry via $\|x-y\| = \|(-1)(y-x)\| = \|y-x\|$. The triangle inequality for the norm along with the rewrite $x-y = (x-z)+(z-y)$ then gives the triangle inequality for ϱ. □

Noting that $\varrho_p(x,y) = \|x-y\|_p$, for $\|\cdot\|_p$ defined as in the next claim, all we just need to prove:

Proposition 14.11 (p-norms on $\mathbb{R}^d$) *Let $d\in\mathbb{N}$ be such that $d\geqslant 1$ and $p\in\mathbb{R}$ obey $p\geqslant 1$. For $x=(x_1,\dots,x_d)\in\mathbb{R}^d$ let*

$$\|x\|_p := \Bigl(\sum_{i=1}^d |x_i|^p\Bigr)^{1/p} \tag{14.20}$$

Then $\|\cdot\|_p$ is a norm on $\mathbb{R}^d$.

We will not give a proof but only comment that its most difficult part (the triangle inequality) is deduced from the so called *Minkowski inequality* which is itself proved by way of the *Hölder inequality*. When $p=2$, the latter reduces to the *Cauchy-Schwarz inequality* which can be proved easily without much calculus.

As our last example of a metric we introduce the concept of *discrete metric* which is defined for any non-empty set A by

$$\varrho(x,y) := \begin{cases} 0, & \text{if } x=y,\\ 1, & \text{if } x\neq y, \end{cases} \tag{14.21}$$

This metric is not very useful in practice but is very good for theory building as it provides an easy test case for various facts about metric spaces.

14.3 Sequences in metric spaces.

We are now ready to go back to sequences taking values in a metric space and generalize the notions introduced in Definitions 14.3–14.4 as

Definition 14.12 (Cauchy and convergent sequences) *Let (X,ϱ) be a metric space. We say that an X-valued sequence $\{x_n\}_{n\in\mathbb{N}}$ in is*

- Cauchy *if* $\forall\epsilon>0\,\exists n_0\geqslant 0\,\forall n,m\geqslant n_0\colon\ \varrho(x_n,x_m)<\epsilon$
- convergent *if* $\exists z\in X\,\forall\epsilon>0\,\exists n_0\geqslant 0\,\forall n\geqslant n_0\colon\ \varrho(x_n,z)<\epsilon$

We call any such z a limit of $\{x_n\}_{n\in\mathbb{N}}$ and write $x_n\to z$ in this case.

We now observe a couple of general facts:

Lemma 14.13 (Uniqueness of the limit) *Any sequence has at most one limit. More precisely, for any metric space (X,ϱ) and any X-valued sequence $\{x_n\}_{n\in\mathbb{N}}$,*

$$\forall z,\tilde z\in X\colon\ x_n\to z\ \wedge\ x_n\to\tilde z\ \Rightarrow\ z=\tilde z \tag{14.22}$$

Proof. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in a metric space (X,ϱ) and $z,\tilde{z}\in X$ be such that $x_n\to z$ and $x_n\to\tilde{z}$. Assuming $z\neq\tilde{z}$, we have $\epsilon := \frac{1}{2}\varrho(z,\tilde{z})>0$. Then, by $x_n\to z$, there is n_0 be such that $\varrho(x_n,z)<\epsilon$ for $n\geqslant n_0$ and, by $x_n\to\tilde{z}$ there is $\tilde{n}_0$ such that $\varrho(x_n,\tilde{z})<\epsilon$ for $n\geqslant\tilde{n}_0$. Taking $n:=\max\{n_0,\tilde{n}_0\}$ we then have

$$\varrho(z,\tilde{z})\leqslant\varrho(x_n,z)+\varrho(x_n,\tilde{z})<\epsilon+\epsilon=2\epsilon=\varrho(z,\tilde{z}) \tag{14.23}$$

which (due to "$<$") is impossible. Hence $\varrho(z,\tilde{z})=0$ and thus $z=\tilde{z}$ as claimed. □

The notation $\lim_{n\to\infty}x_n$ is often used to denote the (unique) limit of the sequence $\{x_n\}_{n\in\mathbb{N}}$. Note that the existence of the limit is implicit in this notation — we would not write this if the limit did not exist.

The two notions from Definition 14.12 are closely related:

Lemma 14.14 (Convergent implies Cauchy) *If $x_n\to x$ then $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy.*

Proof. Fix $\epsilon>0$ and let $n_0\in\mathbb{N}$ be such that for all $n\geqslant n_0$ we have $\varrho(x_n,x)<\epsilon/2$. By the triangle inequality we then get

$$\forall m,n\geqslant n_0\colon\ \varrho(x_m,x_n)\leqslant\varrho(x_m,x)+\varrho(x,x_n)<\epsilon/2+\epsilon/2=\epsilon. \tag{14.24}$$

This yields (2.4) and so $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy. □

For spaces with discrete metric we get even a full characterization of convergent and Cauchy sequences:

Lemma 14.15 (Cauchy/convergent sequences in discrete metric) *Consider any set X endowed with the discrete metric ϱ. Then for any sequence $\{x_n\}_{n\in\mathbb{N}}$ of points in X,*

$$\{x_n\}_{n\in\mathbb{N}}\text{ is Cauchy}\quad\Leftrightarrow\quad\exists n\in\mathbb{N}\,\forall m\geqslant n\colon\ x_m=x_n. \tag{14.25}$$

Every Cauchy sequence is thus eventually constant and all Cauchy sequences converge.

Proof. Since ϱ takes values 0 and 1, letting $\epsilon:=1/2$ in (2.4) forces $\varrho(x_n,x_m)=0$ once $m,n\geqslant n_0$. The positivity axiom for d then shows that $x_n=x_m$ for all $m,n\geqslant n_0$. Eventually constant sequences are always convergent. □

The last example is of course very special; the notion of being Cauchy is actually weaker than being convergent. As shown in a homework exercise, both notions depend on which metric we consider. Indeed, the $\mathbb{R}$-valued sequence $x_n:=n$ is not Cauchy under the Euclidean metric (14.13) but it is Cauchy yet not convergent under the metric (14.14). We will spend considerable time discussing these aspects further.

As our final note, we observe that while the sequence $x_n:=(-1)^n$ is neither convergent nor even Cauchy, restricting to even numbered indices gives us a constant, and thus convergent sequence. This naturally leads to:

Definition 14.16 *Given a sequence $\{x_n\}_{n\in\mathbb{N}}$ its* subsequence *is any sequence of the form $\{x_{n_k}\}_{k\in\mathbb{N}}$ where $\{n_k\}_{k\in\mathbb{N}}$ is a strictly increasing sequence taking values in $\mathbb{N}$.*

Informally, a subsequence arises via a selection of some values in the line-up of the whole sequence. This selection is usually driven by the desire to refine possible behaviors as the index becomes large. The following lemma was left to homework:

Lemma 14.17 *Let $\{x_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence. If $\{x_n\}_{n\in\mathbb{N}}$ admits a convergent subsequence, $x_{n_k} \to z$, then $\{x_n\}_{n\in\mathbb{N}}$ is convergent and $x_n \to z$.*

It follows that a non-convergent Cauchy sequence admits no convergent subsequence. A similar criterion for convergent sequences is given in:

Lemma 14.18 (Characterization of convergent sequences) *Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in a metric space (X, ϱ). Then the following are equivalent:*

(1) *$\{x_n\}_{n\in\mathbb{N}}$ is convergent,*
(2) *there exists $z \in X$ such that every subsequence of $\{x_n\}_{n\in\mathbb{N}}$ contains a subsequence that converges to z.*

We leave the proof of this equivalence, which helps decide whether an abstract notion of convergence can possibly arise from a metric, to the reader.