Mathematical Analysis III*

Daniel Ueltschi

December 5, 2016

Contents

*These lecture notes are based on past notes by various lecturers at the University of Warwick, including Oleg Zaboronski and Sergey Nazarenko.

Part I — INTEGRATION

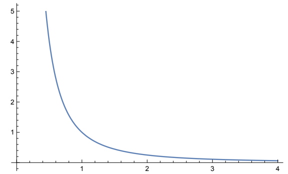
1 The challenges of a good definition of integration

Integration is a natural operation, it is the reverse of derivation. It is intuitively defined as "the area below the curve". If f is a nice function so that $g(t) = \int_a^t f(x) \mathrm{d}x$ makes sense, we have

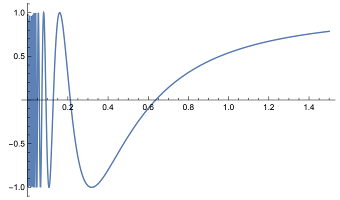
$$g'(t) = f(t). \tag{1.1}$$

The intuitive definition of integration becomes confusing when looking at irregular functions, such as

(a) $f(x) = 1/x^a$, with $a > 0$. Can we make sense of $\int_0^b \frac{\mathrm{d}x}{x^a}$ and $\int_b^\infty \frac{\mathrm{d}x}{x^a}$?



(b) $f(x) = \cos \frac{1}{x}$. What is $\int_0^b \cos \frac{1}{x} \mathrm{d}x$?



(c) Two variations of the previous example: $g(x) = x \cos \frac{1}{x}$ and $h(x) = \frac{1}{x} \cos \frac{1}{x}$. It is a good exercise to draw these functions!

(d) $f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{Q}, \\ 1 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$

It is natural to wonder whether one really wants to integrate such functions? There are good reasons to answer positively. As is well-known, $\mathbb{R}$ is a more convenient set of numbers than $\mathbb{Q}$. In a similar fashion, we will consider various spaces of functions and these spaces should be large enough so as to contain the limits of suitable Cauchy sequences. We are therefore seeking to define integration in a general fashion.

2 Step functions

We have just argued that integration needs to be defined very generally... but we now consider a restricted class of functions.

Definition 2.1. *A* **partition** P *of the interval* $[a,b]$ *is a set of numbers,* $P = \{p_0, p_1, \dots, p_k\}$ *such that* $a = p_0 < p_1 < \dots < p_{k-1} < p_k = b$*. Here,* $k \geq 1$ *is an arbitrary integer.*

If P, Q are two partitions and $Q \supset P$, one says that Q is a **refinement** of P.

Definition 2.2. *A function* $\varphi : [a,b] \to \mathbb{R}$ *is called a* **step function** *if there exists a partition* $P = \{p_0, \dots, p_k\}$ *of* $[a,b]$ *such that* φ *is constant on each subinterval* (p_{i-1}, p_i)*,* $1 \leq i \leq k$*.*

Notice that if φ is constant on each subinterval of P, and Q is a refinement of P, then φ is also constant on the subintervals of Q. We let $S[a,b]$ denote the set of all step functions on the interval $[a,b]$. It has the structure of a *vector space*:

Proposition 2.1. *If* $f, g \in S[a,b]$ *and* $\alpha, \beta \in \mathbb{R}$*, then* $\alpha f + \beta g \in S[a,b]$*.*

Proof. Let P a partition that is compatible with f, and Q a partition that is compatible with g. Then $P \cup Q$ is a refinement of both P and Q, where f, g are both constant on subintervals. It is then clear that $\alpha f + \beta g$ is constant on subintervals of $P \cup Q$, so it is a step function. □

It may be worth pointing out that many spaces of functions are vector spaces, such as the space $B[a,b]$ of bounded functions or the space $C[a,b]$ of continuous functions. Note also that

$$C[a,b] \subset B[a,b] \qquad \text{and} \qquad S[a,b] \subset B[a,b]. \tag{2.1}$$

We now define the integral of step functions.

Definition 2.3. (Integral of step functions) *Let* $\varphi \in S[a,b]$ *and* $P = \{p_0, \dots, p_k\}$ *a compatible partition; we let* φ_i *denote the value of* φ *in the interval* (p_{i-1}, p_i)*,* $i = 1, \dots, k$*. The integral is defined as*

$$\int_a^b \varphi(x) \mathrm{d}x = \sum_{i=1}^k \varphi_i (p_i - p_{i-1}).$$

Notice that the definition of the integral does not depend on the choice of the partition that is compatible with φ (why?). We also define the integral for $b < a$, by setting

$$\int_a^b \varphi(x) \mathrm{d}x = -\int_b^a \varphi(x) \mathrm{d}x. \tag{2.2}$$

Let us now point out important properties of integrals.

Proposition 2.2. (Additivity) *For any* $\varphi \in S[a,b]$ *and* $c \in (a,b)$*, we have*

$$\int_a^b \varphi(x) \mathrm{d}x = \int_a^c \varphi(x) \mathrm{d}x + \int_c^b \varphi(x) \mathrm{d}x.$$

This statement is rather obvious and no proof is needed.

Proposition 2.3. (Linearity) *For any* $\varphi, \psi \in S[a,b]$ *and* $\alpha, \beta \in (a,b)$*, we have*

$$\int_a^b \big(\alpha\varphi(x) + \beta\psi(x)\big) \mathrm{d}x = \alpha \int_a^b \varphi(x) \mathrm{d}x + \beta \int_a^b \psi(x) \mathrm{d}x.$$

Proof. Let $P = \{p_0, \dots, p_k\}$ be a partition that is compatible with φ, ψ (it is then compatible with $\alpha\varphi + \beta\psi$). Let φ_i, ψ_i be the values of φ, ψ on the interval (p_{i-1}, p_i). Then

$$\begin{aligned} \int_a^b \big(\alpha\varphi(x) + \beta\psi(x)\big) \mathrm{d}x &= \sum_{i=1}^k \big(\alpha\varphi_i + \beta\psi_i\big)(p_i - p_{i-1}) \\ &= \alpha \sum_{i=1}^k \varphi_i (p_i - p_{i-1}) + \beta \sum_{i=1}^k \psi_i (p_i - p_{i-1}) \\ &= \alpha \int_a^b \varphi(x) \mathrm{d}x + \beta \int_a^b \psi(x) \mathrm{d}x. \end{aligned} \tag{2.3}$$

□

Proposition 2.4. (Fundamental theorem of calculus for step functions) *Let* $\varphi \in S[a,b]$ *and* $P = \{p_0, \dots, p_k\}$ *a partition compatible with* φ*. Consider the function*

$$I : [a,b] \to \mathbb{R}$$
$$I(t) = \int_a^t \varphi(x) \mathrm{d}x.$$

Then
(a) $I(t)$ *is continuous on* $[a,b]$*.*
(b) I *is differentiable on* $\cup_{i=1}^k (p_{i-1}, p_i)$ *and* $I'(t) = \varphi(t)$*.*

Proof. Let φ_i be the value of φ on the interval (p_{i-1}, p_i), $i = 1, \dots, k$. Let $t \in (p_{j-1}, p_j)$; we have

$$I(t) = \sum_{i=1}^{j-1} \varphi_i (p_i - p_{i-1}) + \varphi_j (t - p_{j-1}). \tag{2.4}$$

This is of the form const + $\varphi_j t$. It is continuous and the derivative exists and is equal to $\varphi_j = \varphi(t)$. If $t = p_j$, we can check that $I(t) = \sum_{i=1}^{j} \varphi_i(p_i - p_{i-1})$, so it is continuous everywhere. □

3 Convergence of sequences of functions

The goal is to extend the definition of integrals to functions that can be approximated by step functions. For this, we need to define the notion of convergence for sequences of functions. This can be done in various ways — the variety being a strength of mathematical analysis. One way is to say that $\varphi_n \to f$ if for any $x \in [a,b]$, the sequence of numbers $(\varphi_n(x))_{n\geq 1}$ converges to $f(x)$. This is *pointwise convergence* which is an important and useful notion. But we now consider *uniform convergence*. It is best introduced with the help of the *supremum norm*, that turns the vector space of functions into a normed vector space.

Definition 3.1. *The* **supremum norm** *of the bounded function* $f \in B[a,b]$ *is*

$$\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|.$$

The notation is motivated by the family of norms $\|f\|_p = \left(\int_a^b |f(x)|^p \mathrm{d}x\right)^{1/p}$ (one needs $p \geq 1$ in order to satisfy the triangle inequality). Then $\|f\|_\infty = \lim_{p\to\infty} \|f\|_p$.

Notice that any step function $\varphi \in S[a,b]$ satisfies $\|\varphi\|_\infty < \infty$. We let $B[a,b]$ denote the set of *bounded* functions, i.e., functions that have finite supremum norm. It is a good exercise to verify that $\|\cdot\|_\infty$ satisfies all axioms of a norm; namely, it is

- positive: $\|f\|_\infty \geq 0$, and $\|f\|_\infty = 0$ implies that $f(x) = 0$ for all x;
- homogenous: $\|\alpha f\|_\infty = |\alpha|\,\|f\|_\infty$ for all functions f and all numbers α;
- triangle inequality: $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$ for all functions f, g.

Proposition 3.1. *For any* $\varphi \in S[a,b]$, *we have*

$$\Big|\int_a^b \varphi(x)\mathrm{d}x\Big| \leq \int_a^b |\varphi(x)|\mathrm{d}x \leq \|\varphi\|_\infty (b-a).$$

Proof. Let $P = \{p_0, \dots, p_k\}$ a partition that is compatible with φ, and φ_i the value of φ on the interval (p_{i-1}, p_i). Then

$$\begin{aligned} \Big|\int_a^b \varphi(x)\mathrm{d}x\Big| &= \Big|\sum_{i=1}^k \varphi_i (p_i - p_{i-1})\Big| \\ &\leq \sum_{i=1}^k |\varphi_i| (p_i - p_{i-1}) \qquad \Big(= \int_a^b |\varphi|\Big) \\ &\leq \sum_{i=1}^k \|\varphi\|_\infty (p_i - p_{i-1}) \\ &= \|\varphi\|_\infty (b-a). \end{aligned} \tag{3.1}$$

□

Recall that a function $f : A \to \mathbb{R}$, where A is a closed or open interval, is continuous at $x \in A$ if

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ such that } \forall y \in A \text{ with } |y-x| < \delta, \text{ we have } \big|f(x) - f(y)\big| < \varepsilon. \tag{3.2}$$

Here, $\delta = \delta(x,\varepsilon)$ depends on x and ε. The function f is continuous if is continuous at every point $x \in A$. For uniform continuity, we require that δ be independent of x.

Definition 3.2. *A function* $f : A \to \mathbb{R}$ *is* **uniformly continuous** *if*

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ such that } \forall x, y \in A \text{ with } |x-y| < \delta, \text{ we have } \big|f(x) - f(y)\big| < \varepsilon.$$

It is clear that uniform continuity implies continuity. The converse is not necessarily true, consider for instance the function $f(x) = 1/x$ on $(0,1]$. The lack of uniform continuity is due to the fast variation and the absence of limit at one end of the interval. This turns out to be the only possible difficulty. If the function is continuous on a *closed* interval, then it is uniformly continuous, as stated in the next lemma.

Lemma 3.2. *If* $f : [a,b] \to \mathbb{R}$ *is continuous, then it is uniformly continuous.*

Proof. We prove the contrapositive (i.e. instead of proving $A \implies B$, we prove non$B \implies$ non$A$). Let f a function on $[a,b]$ that is not uniformly continuous. Then

$$\exists \varepsilon > 0 \text{ such that } \forall \delta > 0, \exists x, y \in [a,b] \text{ with } |x-y| < \delta \text{ and } \big|f(x) - f(y)\big| > \varepsilon. \tag{3.3}$$

The strategy is to identify a point in $[a,b]$ where f is not continuous. Take $\delta_n = 1/n$; we get two sequences $(x_n)_{n\geq 1}$ and $(y_n)_{n\geq 1}$ such that $|x_n - y_n| < \frac{1}{n}$ and $|f(x_n) -$

$f(y_n)| > \varepsilon$. Since (x_n) is bounded, it has a convergent subsequence (x_{n_k}) by Bolzano-Weierstrass theorem. Let $u = \lim_{k\to\infty} x_{n_k}$. Since

$$|y_{n_k} - u| \leq |y_{n_k} - x_{n_k}| + |x_{n_k} - u| \to 0, \tag{3.4}$$

we also have $y_{n_k} \to u$. Further, $u \in [a, b]$ because the interval is closed (here we used the hypothesis that the interval is closed). We now have sequences (x_{n_k}), (y_{n_k}), numbers $\varepsilon > 0$ and $u \in [a, b]$, such that

- $\big|f(x_{n_k}) - f(y_{n_k})\big| > \varepsilon$;
- $x_{n_k} \to u$, $y_{n_k} \to u$.

Then f is not continuous at u, hence not continuous on $[a, b]$. □

We now introduce the class of functions for which integration will be defined.

Definition 3.3. *A function $f : [a, b] \to \mathbb{R}$ is* **regulated** *if there exists a sequence (φ_n) of step functions, $\varphi_n \in S[a, b]$, that converges uniformly to f, i.e. $\|f - \varphi_n\|_\infty \to 0$ as $n \to \infty$.*
Alternatively, f is regulated if for every $\varepsilon > 0$, there exists $\varphi \in S[a, b]$ such that $\|f - \varphi\|_\infty < \varepsilon$.

Let $R[a, b]$ denote the set of regulated functions on $[a, b]$. It is not hard to check that $R[a, b]$ is a vector space. Using the triangle inequality, we have

$$\|f\|_\infty = \|f - \varphi + \varphi\|_\infty \leq \|f - \varphi\|_\infty + \|\varphi\|_\infty < \infty, \tag{3.5}$$

so regulated functions are bounded. $R[a, b]$ contains all step functions, and also all continuous functions:

Proposition 3.3. *If $f : [a, b] \to \mathbb{R}$ is continuous, then it is regulated.*

Proof. If f is continuous, then it is uniformly continuous by Lemma 3.2. For every $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all $|x - y| < \delta$. Let $P = \{p_0, \dots, p_k\}$ be a partition of $[a, b]$ such that $p_i - p_{i-1} < \delta$ for all $i = 1, \dots, k$. Let $\varphi \in S[a, b]$ defined by

$$\varphi(x) = f(p_{i-1}) \quad \text{if } x \in [p_{i-1}, p_i), \quad \text{and } \varphi(b) = f(b). \tag{3.6}$$

We have $|f(x) - \varphi(x)| = |f(x) - f(p_{i-1})| < \varepsilon$ since $|x - p_{i-1}| < \delta$, so that $\|f - \varphi\|_\infty < \varepsilon$ indeed. □

Recall that a function is *piecewise continuous* if it is continuous at all points of the interval, except for a finite number of points. Using a partition that contains all points of discontinuity, we can extend Proposition 3.3 to certain piecewise continuous functions:

Corollary 3.4. *Let $f : [a, b] \to \mathbb{R}$ be a function and $P = \{p_0, \dots, p_k\}$ a partition such that f is continuous on each interval (p_{i-1}, p_i), and that it can be extended to a continuous function on $[p_{i-1}, p_i]$. Then f is regulated.*

Here, we invite the reader to think about functions on $[0, 1]$ that are *not* regulated. Can you conceive one that is bounded and continuous at every point save one? And can you prove the impossibility for a sequence of step functions to converge to this peculiar function?

We can now define the integral of regulated functions.

Definition 3.4. *The integral of $f \in R[a, b]$ is defined as*

$$\int_a^b f(x)\mathrm{d}x = \lim_{n\to\infty} \int_a^b \varphi_n(x)\mathrm{d}x,$$

where $(\varphi_n)_{n\geq 1}$ is any sequence of step functions that converges uniformly to f.

This definition raises two questions. Does the limit exists, and is it independent of the choice of the sequence of step functions? The answer is yes:

Proposition 3.5. *Let $f \in R[a, b]$ and $(\varphi_n)_{n\geq 1}, (\psi_n)_{n\geq 1}$ be sequences of step functions that converge uniformly to f. Then $\lim_{n\to\infty} \int_a^b \varphi_n$ and $\lim_{n\to\infty} \int_a^b \psi_n$ exist, and they are equal.*

Proof. Using Proposition 3.1, we have

$$\Big|\int_a^b \varphi_n(x)\mathrm{d}x - \int_a^b \varphi_m(x)\mathrm{d}x\Big| = \Big|\int_a^b \big(\varphi_n(x)\mathrm{d}x - \varphi_m(x)\big)\mathrm{d}x\Big| \leq \|\varphi_n - \varphi_m\|_\infty (b-a). \tag{3.7}$$

Since $\|\varphi_n - \varphi_m\|_\infty \leq \|\varphi_n - f\|_\infty + \|f - \varphi_m\|_\infty$, we have that for every $\varepsilon > 0$, there exists N such that $\|\varphi_n - \varphi_m\|_\infty < \frac{\varepsilon}{b-a}$ for all $m, n > N$. Then $\big|\int_a^b \varphi_n - \int_a^b \varphi_m\big| < \varepsilon$ for $m, n > N$, so that $\big(\int_a^b \varphi_n\big)_{n\geq 1}$ is a Cauchy sequence of numbers, and it therefore converges.

Finally, the two sequences (φ_n) and (φ_n) give the same integral:

$$\Big|\int_a^b \varphi_n(x)\mathrm{d}x - \int_a^b \psi_n(x)\mathrm{d}x\Big| \leq \|\varphi_n - \psi_n\|_\infty (b-a) \leq \big[\|\varphi_n - f\|_\infty + \|f - \psi_n\|_\infty\big](b-a) \to 0 \tag{3.8}$$

as $n \to \infty$. □

We established in Propositions 2.2 and 2.3 some properties of integrals in the case of step functions. These extend readily to integral of regulated functions.

Proposition 3.6. (Additivity) *For any $f \in R[a,b]$ and $c \in (a,b)$, we have*

$$\int_a^b f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x + \int_c^b f(x)\mathrm{d}x.$$

Proposition 3.7. (Linearity) *For any $f, g \in R[a,b]$ and $\alpha, \beta \in (a,b)$, we have*

$$\int_a^b \big(\alpha f(x) + \beta g(x)\big)\mathrm{d}x = \alpha \int_a^b fi(x)\mathrm{d}x + \beta \int_a^b g(x)\mathrm{d}x.$$

The proof is immediate, using that if $\varphi_n \to f$ and $\psi_n \to g$, then $\alpha\varphi_n + \beta\psi_n \to \alpha f + \beta g$. Then it follows from Proposition 2.3.

Proposition 3.8. *Let $f \in R[a,b]$ and $m, M \in \mathbb{R}$ such that $m \leq f(x) \leq M$ for all $x \in [a,b]$. Then*

$$m(b-a) \leq \int_a^b f(x)\mathrm{d}x \leq M(b-a).$$

Since $-\|f\|_\infty \leq f(x) \leq \|f\|_\infty$, this proposition also implies that $\left|\int_a^b f\right| \leq \|f\|_\infty(b-a)$.

Proof. For every n, there exists $\varphi_n \in S[a,b]$ such that $\|\varphi_n - f\|_\infty < \frac{1}{n}$. Then for all $x \in [a,b]$,

$$-m - \tfrac{1}{n} \leq f(x) - \tfrac{1}{n} \leq \varphi_n(x) \leq f(x) + \tfrac{1}{n} \leq M + \tfrac{1}{n}. \tag{3.9}$$

If $P = \{p_0, \dots, p_k\}$ is a partition compatible with φ_n, and $\varphi_{n,i}$ denotes the value of φ_n on the interval (p_{i-1}, p_i), then

$$-(m + \tfrac{1}{n})(b-a) \leq \sum_{i=1}^k \varphi_{n,i}(p_i - p_{i-1}) \leq (M + \tfrac{1}{n})(b-a), \tag{3.10}$$

The middle term is equal to $\int_a^b \varphi_n$. Taking the limit $n \to \infty$, we get the claim. □

Example: Consider the function $f(x) = x^\alpha$ on $[1,\infty)$. Let us check that f is uniformly continuous when $\alpha \in [0,1]$ but not when $\alpha > 1$. For $\alpha > 1$, observe that

$$f(x+\delta) - f(x) = (x+\delta)^\alpha - x^\alpha = x^\alpha\big[\big(1+\tfrac{\delta}{x}\big)^\alpha - 1\big] \geq x^\alpha\big[1 + \alpha\tfrac{\delta}{x} - 1\big] = \alpha\delta x^{\alpha-1}, \tag{3.11}$$

which diverges as $x \to \infty$. No uniform continuity in this case. For $\alpha \in [0,1]$, we use

$$0 \leq f(x+\delta) - f(x) = x^\alpha\big[\big(1+\tfrac{\delta}{x}\big)^\alpha - 1\big] \leq x^\alpha\big[1 + \alpha\tfrac{\delta}{x} - 1\big] = \alpha\delta x^{\alpha-1} \leq \alpha\delta, \tag{3.12}$$

uniformly in $x \in [0,\infty)$. Hence uniform continuity. The case $\alpha = 0$ needs special treatment, but it is trivial.

Recall the devil's staircase. Is it a regulated function? It is not a step function, and it is not clear whether it is continuous (it is, in fact). But it is definitely monotone increasing, which is enough.

Proposition 3.9. *If $f : [a,b] \to \mathbb{R}$ is monotone, then it is regulated.*

Proof. We can suppose that f is nondecreasing. Let $k \in \mathbb{N}$, and

$$q_i = f(a) + \tfrac{i}{k}\big(f(b) - f(a)\big), \tag{3.13}$$

where $i = 1, \dots, k$. Then let $p_i \in [a,b]$ be a number such that $f(x) \leq q_i$ if $x < p_i$ and $f(x) \geq q_i$ if $x > p_i$ (p_i is unique when f is increasing; it is not unique when f has flat parts). It is possible that $p_{i+1} = p_i$. We define

$$\varphi(x) = \begin{cases} q_i & \text{if } x \in (p_{i-1}, p_i) \text{ for some } p_{i-1} < p_i, \\ f(p_i) & \text{if } x = p_i. \end{cases} \tag{3.14}$$

Then φ is a step function and

$$0 \leq \varphi(x) - f(x) \leq \frac{f(b) - f(a)}{k}, \tag{3.15}$$

so that $\|\varphi - f\|_\infty$ is as small as we wish by choosing k large enough. □

Definition 3.5. *Let $f \in R[a,b]$, and for $x \in [a,b]$, define the function $F : [a,b] \to \mathbb{R}$ by*

$$F(x) = \int_a^x f(t)\mathrm{d}t.$$

F is called the **primitive**, *or* **indefinite integral**, *or* **antiderivative** *of f.*

Proposition 3.10. *The primitive F of any regulated function $f \in R[a,b]$ is continuous.*

Proof. We have

$$|F(x) - F(y)| = \Big|\int_a^x f(t)\mathrm{d}t - \int_a^y f(t)\mathrm{d}t\Big| = \Big|\int_y^x f(t)\mathrm{d}t\Big| \leq \|f\|_\infty |x-y|. \tag{3.16}$$

Then for every $\varepsilon > 0$, we can choose $\delta = \varepsilon/\|f\|_\infty$ to get $|F(x) - F(y)| < \varepsilon$ for all x, y with $|x - y| < \delta$. □

We can now formulate a version of the fundamental theorem of calculus.

Theorem 3.11. *Let $f \in R[a,b]$, and assume that f is continuous at $c \in (a,b)$. Then $F(x) = \int_a^x f$ is differentiable at c, and $F'(c) = f(c)$.*

Proof. We have $F(c+h) - F(c) = \int_c^{c+h} f(t) \mathrm{d}t$. Since f is continuous at c, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $f(c) - \varepsilon \leq f(t) \leq f(c) + \varepsilon$ for $|t - c| < \delta$. Then

$$h\big(f(c) - \varepsilon\big) \leq \int_c^{c+h} f(t) \mathrm{d}t \leq h\big((f)c) + \varepsilon\big). \tag{3.17}$$

Consequently,

$$-\varepsilon \leq \frac{F(c+h) - F(c)}{h} - f(c) \leq \varepsilon \tag{3.18}$$

for all $|h| < \delta$. This is precisely the meaning of $\lim_{h\to 0} \frac{F(c+h)-F(c)}{h} = f(c)$. □

If $f \in C[a,b]$, it is regulated by Proposition 3.3, and Theorem 3.11 shows that its primitive is differentiable at any point of (a,b). Another useful version of the fundamental theorem of calculus is the following.

Theorem 3.12. *Let $f \in C[a,b]$ and F a differentiable function on $[a,b]$ such that $F'(x) = f(x)$ for all $x \in (a,b)$. Then*

$$\int_a^b f(x) \mathrm{d}x = F(b) - F(a).$$

Proof. Consider the function g on $[a,b]$,

$$g(x) = F(x) - \int_a^x f(t) \mathrm{d}t. \tag{3.19}$$

This function is differentiable and $g'(x) = F'(x) - f(x) = 0$. This implies that g is constant (this follows from the mean value theorem), so that $g(a) = g(b)$, i.e.,

$$F(a) = F(b) - \int_a^b f(t) \mathrm{d}t. \tag{3.20}$$

□

4 Methods of integration

The main two methods are the integration by parts and integration by substitution. They both allow to replace the integral of a function by that of another function which is hopefully simpler. The goal of this section is to establish the validity of the corresponding formulæ.

We define the product of two functions by $(fg)(x) = f(x)g(x)$.

Proposition 4.1. *If $f, g \in R[a,b]$, then $fg \in R[a,b]$.*

Proof. First, let us note that $\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty$. Let (φ_n), (ψ_n) be regulated functions such that $\varphi_n \to f$ and $\psi_n \to g$ uniformly. Then $\varphi_n \psi_n$ is a step function for all n, and

$$\begin{aligned} \|fg - \varphi_n \psi_n\|_\infty &\leq \|fg - \varphi_n g\|_\infty + \|\varphi_n g - \varphi_n \psi_n\|_\infty \\ &\leq \|f - \varphi_n\|_\infty \|g\|_\infty + \|\varphi_n\|_\infty \|g - \psi_n\|_\infty. \end{aligned} \tag{4.1}$$

Since $\|\varphi_n\|_\infty \to \|f\|_\infty$, it is bounded. Then everything tends to 0 as $n \to \infty$, so fg is indeed equal to the limit of a sequence of step functions. □

Theorem 4.2. (Integration by parts) *Let f, g be continuously differentiable functions on $[a,b]$. Then*

$$\int_a^b f'g = fg\Big|_a^b - \int_a^b fg'.$$

It is perhaps worth clarifying the notation used in the theorem; written explicitly, the equation is

$$\int_a^b f'(x)g(x) \mathrm{d}x = f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x) \mathrm{d}x. \tag{4.2}$$

Proof. Recall that $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$. By Theorem 3.12 (the second version of the fundamental theorem of calculus), we have

$$(fg)(b) - (fg)(a) = \int_a^b (fg)'(x) \mathrm{d}x = \int_a^b f'g + \int_a^b fg'. \tag{4.3}$$

Rearranging this identity, we get the formula of integration by parts. □

This theorem is amazingly useful and versatile. Let us describe a nice application to the Gamma function $\Gamma : [0,\infty) \to [0,\infty)$, defined by

$$\Gamma(x) = \int_0^\infty t^{x-1} \, \mathrm{e}^{-t} \, \mathrm{d}t. \tag{4.4}$$

Notice that $[0, \infty)$ is not a closed interval, and Γ is not regulated around 0 when $x < 1$. A correct definition for the integral above is

$$\Gamma(x) = \lim_{a \to 0+} \lim_{b \to \infty} \int_a^b t^{x-1} \, \mathrm{e}^{-t} \, \mathrm{d}t. \tag{4.5}$$

One can check that the limits exist, and that it can be taken in any order. We use Theorem 4.2 with $f'(t) = \mathrm{e}^{-t}$ and $g(t) = t^{x-1}$. Then $f(t) = -\mathrm{e}^{-t}$ and $g'(t) = (x-1)t^{x-2}$, and we obtain

$$\begin{aligned} \Gamma(x) &= \int_0^\infty \mathrm{e}^{-t} \, t^{x-1} \mathrm{d}t \\ &= -\mathrm{e}^{-t} \, t^{x-1} \Big|_0^\infty + (x-1) \int_0^\infty \mathrm{e}^{-t} \, t^{x-2} \mathrm{d}t \\ &= (x-1)\Gamma(x-2). \end{aligned} \tag{4.6}$$

The first term in the middle equation is equal to $\lim_{a \to 0+} \lim_{b \to \infty} (-\mathrm{e}^{-t} \, t^{x-1}) \big|_a^b$, which is 0 if $x > 1$. This shows that for all $x > 0$, we have

$$\Gamma(x+1) = x\Gamma(x). \tag{4.7}$$

Since $\Gamma(1) = \int_0^1 \mathrm{e}^{-t} \, \mathrm{d}t = -\mathrm{e}^{-t} \big|_0^\infty = 1$, we obtain $\Gamma(n+1) = n!$ for integer n. The Gamma function generalises factorials to real numbers!

We now discuss the method of substitution. It is based on the chain rule for derivatives.

Theorem 4.3. *Let $f \in C[a,b]$ and $g : [c,d] \to [a,b]$ be differentiable. Then*

$$\int_c^d f\big(g(x)\big) g'(x) \mathrm{d}x = \int_{g(c)}^{g(d)} f(t) \mathrm{d}t.$$

Note that the above formula can also be written as

$$\int_a^b f(x) \mathrm{d}x = \int_{g^{-1}(a)}^{g^{-1}(b)} f\big(g(t)\big) g'(t) \mathrm{d}t, \tag{4.8}$$

where $g^{-1}(a)$ is equal to any value $y \in [c,d]$ such that $g(y) = a$.

Proof. The claim follows from the first version of the fundamental theorem of calculus, Theorem 3.11, and from the chain rule for derivatives. Let $F(x) = \int_a^x f$. Then F is differentiable with $F'(x) = f(x)$, and

$$\int_{g(c)}^{g(d)} f(t) \mathrm{d}t = F\big(g(d)\big) - F\big(g(c)\big). \tag{4.9}$$

Further, we have

$$\frac{\mathrm{d}}{\mathrm{d}x} F\big(g(x)\big) = F'\big(g(x)\big) g'(x) = f\big(g(x)\big) g'(x), \tag{4.10}$$

so that

$$\int_c^d f\big(g(x)\big) g'(x) \mathrm{d}x = F\big(g(x)\big)\Big|_c^d. \tag{4.11}$$

This is identical to the right side of Eq. (4.9). ☐

Let us review a few examples.

(a) Consider the integral $\int_a^b \cos(x^2) x \mathrm{d}x$. We take $f(x) = \cos x$ and $g(x) = x^2$, and Theorem 4.3 shows that the integral is equal to $\frac{1}{2} \int_{a^2}^{b^2} \cos t \, \mathrm{d}t = \frac{1}{2}\big(\sin b^2 - \sin a^2\big)$.

We can also proceed more directly and write $\int_a^b \cos(x^2) x \mathrm{d}x = \frac{1}{2} \int_a^b \Big(\frac{\mathrm{d}}{\mathrm{d}x} \sin(x^2)\Big) \mathrm{d}x$, which obviously gives the same result.

(b) $\displaystyle\int_a^b \sin x \cos x \mathrm{d}x = \frac{1}{2} \int_a^b \Big(\frac{\mathrm{d}}{\mathrm{d}x} \sin^2 x\Big) \mathrm{d}x = \frac{1}{2}\big(\sin^2 b - \sin^2 a\big).$

(c) $\displaystyle\int_0^\pi \mathrm{e}^{\sin x} \cos x \, \mathrm{d}x = \Big(\frac{\mathrm{d}}{\mathrm{d}x} \mathrm{e}^{\sin x}\Big) \mathrm{d}x = \mathrm{e}^{\sin x} \Big|_0^\pi = 0.$

If you find the answer 0 to be disappointing, evaluate the integral from 0 to $\pi/2$ instead!

(d) $\displaystyle\int_1^{\mathrm{e}} \frac{1}{x} \log^2 x \, \mathrm{d}x = \frac{1}{3} \int_1^{\mathrm{e}} \Big(\frac{\mathrm{d}}{\mathrm{d}x} \log^3 x\Big) \mathrm{d}x = \frac{1}{3}.$

It is natural to approximate an integral by a direct sum. It is tempting to decompose the interval $[a,b]$ into smaller intervals, and to approximate the function by a constant within each interval. Namely, given $n \in \mathbb{N}$, let $a_i = a + i\frac{b-a}{n}$, $i = 1, \dots, n$. The question is whether $\sum_{i=1}^n \frac{b-a}{n} f(a_i)$ is a good approximation for f. We check that this converges to $\int_a^b f$ as $n \to \infty$.

Proposition 4.4. *Let $f \in R[a,b]$. Then*

$$\lim_{n \to \infty} \sum_{i=1}^n \frac{b-a}{n} f(a_i) = \int_a^b f(t) \mathrm{d}t.$$

Proof. We first prove it for step function, then we use a standard continuity argument. If $\varphi \in S[a,b]$, there exists $P = \{p_0, \dots, p_k\}$ such that φ is constant on (p_{i-1}, p_i). Then

$$\Big| \int_a^b \varphi(t)\mathrm{d}t - \sum_{i=1}^n \frac{b-a}{n}\varphi(a_i) \Big| \leq 2k\frac{b-a}{n}\|\varphi\|_\infty. \tag{4.12}$$

Indeed, the differences occur only at the points of discontinuity of φ, and the maximum jump is $2\varphi\|_\infty$. This goes to 0 as $n \to \infty$, for any fixed φ; the claim holds therefore for step functions.

If $f \in R[a,b]$, then for any $\varepsilon > 0$, there exists $\varphi \in S[a,b]$ such that $\|f-\varphi\|_\infty < \varepsilon$. Then

$$\begin{aligned} \Big| \sum_{i=1}^n \frac{b-a}{n} f(a_i) - \int_a^b f(t)\mathrm{d}t \Big| \leq & \Big| \sum_{i=1}^n \frac{b-a}{n} f(a_i) - \sum_{i=1}^n \frac{b-a}{n}\varphi(a_i) \Big| \\ & + \Big| \sum_{i=1}^n \frac{b-a}{n}\varphi(a_i) - \int_a^b \varphi(t)\mathrm{d}t \Big| \\ & + \Big| \int_a^b \varphi(t)\mathrm{d}t - \int_a^b f(t)\mathrm{d}t \Big|. \end{aligned} \tag{4.13}$$

The first and third terms of the right side are less than $\|f-\varphi\|_\infty(b-a)$. The second term vanishes in the limit $n \to \infty$. We can make the right side as small as we wish by first choosing ε small enough, then n large enough. □

5 Characterisation of regulated functions

A natural question is whether a given function is regulated. So far, we have seen that:

- Step functions are regulated (this immediately follows from the definition of regulated functions).
- Continuous functions on closed intervals are regulated (Proposition 3.3).
- Piecewise continuous functions, with some extra conditions, are regulated.
- Monotone functions on closed intervals are regulated (Proposition 3.9).
- A function such as $f(x) = \sin\frac{1}{x}$ on $(0,1]$ (and $f(0)$ given some value) is not regulated.

It turns out that a rather simple criterion holds true, namely, that left and right limits of the function exist at every point. We use the notation

$$f(x+) = \lim_{y \to x+} f(y) \qquad \text{and} \qquad f(x-) = \lim_{y \to x-} f(y). \tag{5.1}$$

Notice that $f(x\pm)$ may or may not exist.

Proposition 5.1. *Let $f : [a,b] \to \mathbb{R}$. It is regulated iff for every $x \in (a,b)$, $f(x+)$ and $f(x-)$ exist, and $f(a+)$ and $f(b-)$ exist.*

This is a beautifully explicit criterion; the proof is not easy and we treat both implications separately.

Proof that f regulated implies that left and right limits exist: Let $x \in (a,b]$ and $x_n \to x-$. Let $\varphi_k \in S[a,b]$ such that $\|f-\varphi\|_\infty < \frac{1}{k}$. For any k, there exists δ_k such that φ_k is constant on $(x-\delta_k, x)$. Then there exists N_k such that $|x_n - x| < \delta_k$ if $n > N_k$, and

$$\big| f(x_m) - f(x_n) \big| < \tfrac{2}{k} \tag{5.2}$$

for all $m, n > N_k$. So $(f(x_n))_{n\geq 1}$ is Cauchy, hence convergent. Let $f(x-)$ denote its limit.

If (x_n') is another sequence that converges to $x-$, we can consider the mixed sequence $(x_n") = (x_1, x_1', x_2, x_2', \dots)$. We have $x_n" \to x-$, so $f(x_n")$ is Cauchy by the above argument, hence convergent. All its subsequences converge to the same limit, so $f(x_n') \to f(x-)$ as well. □

Proof that the existence of left and right limits implies that f is regulated: The idea is to consider the set A_ε, defined for $\varepsilon > 0$ by

$$A_\varepsilon = \Big\{ c \in [a,b] : \exists \varphi \in S[a,b] \text{ such that } \big\| (f-\varphi)\big|_{[a,c]} \big\|_\infty < \varepsilon \Big\}. \tag{5.3}$$

Here, $f\big|_{[a,c]}$ denotes the restriction of f to the interval $[a,c]$. The set A_ε is the maximal interval that contains a, and such that f can be approximated by a step function. The goal is to show that $A_\varepsilon = [a,b]$ for any $\varepsilon > 0$.

(i) $A_\varepsilon \neq \emptyset$: Since $f(a+)$ exists, there exists $\delta > 0$ such that $|f(x) - f(a+)| < \varepsilon$ for all $x \in (a, a+\delta)$. We can take

$$\varphi(x) = \begin{cases} f(a) & \text{if } x = a, \\ f(a+) & \text{if } x \in (a, a+\delta), \\ f(a+\delta) & \text{if } x \in [a+\delta, b]. \end{cases} \tag{5.4}$$

We have $\varphi \in S[a,b]$ and $\big\| (\varphi - f)\big|_{[a,a+\delta]} \big\|_\infty < \varepsilon$, so A_ε contains the interval $[a, a+\delta]$.

(ii) $\sup A_\varepsilon = b$: Assume that $\sup A_\varepsilon = c < b$. There exists $\varphi \in S[a,b]$ such that $\big\| (\varphi - f)\big|_{[a,c]} \big\|_\infty < \varepsilon$. Since $f(c+)$ exists, we can repeat the previous argument and get an approximating step function φ' on $[c, c+\delta]$ for some $\delta > 0$. Combining φ and φ', we get an approximation of f on $[a, c+\delta]$, so that $\sup A_\varepsilon > c$, contradiction.

(iii) $b \in A_\varepsilon$: Since $f(b-)$ exists, we can repeat the argument in (i) to get an approximating step function on $[b-\delta, b]$ for some $\delta > 0$. There is also an approximating step function on $[a, b-\delta]$ by (i) and (ii). They can be combined to give an approximation on $[a, b]$.

Then $A_\varepsilon = [a, b]$ for any $\varepsilon > 0$, so $f \in R[a, b]$. □

6 Improper and Riemann integrals

There exist interesting functions that are not regulate, but whose integrals are well-defined. An example is the Gamma function, $\gamma(x) = \int_0^\infty t^{x-1} \mathrm{e}^{-t} \, \mathrm{d}t$; the integrant is not always regulated at 0 (if $x < 1$) and the interval is infinite. Let us consider a few examples.

(a) $\int_0^1 x^{-s} \mathrm{d}x$, $s > 0$. The function is not regulated; we can calculate explicitly; if $s \neq 1$,

$$\int_a^1 x^{-s} \mathrm{d}x = \tfrac{1}{1-s} x^{1-s} \Big|_a^1 = \tfrac{1}{1-s}(1 - a^{1-s}). \tag{6.1}$$

As $a \to 0+$, the latter converges iff $s \leq 1$. We would like to declare that $\int_0^1 x^{-s} \mathrm{d}x = \frac{1}{1-s}$ for $s \in [0, 1)$.

(b) $\int_0^1 x^{-1} \mathrm{d}x$, which is the case $s = 1$ of item (a). We have

$$\int_a^1 x^{-1} \mathrm{d}x = \log x \Big|_a^1 = -\log a. \tag{6.2}$$

This tends to ∞ as $a \to 0+$, so this integral should not converge.

(c) $\int_1^\infty x^{-s} \mathrm{d}x$. The interval is infinite, but we can consider $\int_1^b x^{-s}$ and attempt to let $b \to \infty$. If $s \neq 1$, we get

$$\int_1^b x^{-s} \mathrm{d}x = \tfrac{1}{1-s}(b^{1-s} - 1). \tag{6.3}$$

This converges (to $\frac{1}{s-1}$) iff $s > 1$.

(d) $\int_1^\infty x^{-1} \mathrm{d}x$. Since $\int_1^b x^{-1} \mathrm{d}x = \log b$, this tends to ∞ as $b \to \infty$, so the integral does not converge.

Definition 6.1. *Let A be an interval on $\mathbb{R}$ (it can be open, closed, neither, finite, infinite) and f a function $A \to \mathbb{R}$ that is regulated on all closed intervals $[c, d] \subset A$. Let $a = \inf A$ and $b = \sup A$. If $\lim_{c \to a+} \int_c^e f(x) \mathrm{d}x$ exists for some $e \in (a, b)$, and $\lim_{d \to b-} \int_e^d f(x) \mathrm{d}x$ exists for some $e \in (a, b)$, then the* **improper integral** *$\int_A f(x) \mathrm{d}x$ exists, and is defined to be*

$$\int_A f(x) \mathrm{d}x = \lim_{c \to a+} \int_c^e f(x) \mathrm{d}x + \lim_{d \to b-} \int_e^d f(x) \mathrm{d}x.$$

It is not hard to verify that $\int_A f$ does not depend on the midpoint $e \in (a, b)$.

Let us discuss a few examples.

(a) $f(x) = \frac{1}{\sqrt{x}} \cos \sqrt{x}$ on $(0, \infty)$. The primitive is $F(x) = 2 \sin \sqrt{x}$, and $\int_a^b f(x) \mathrm{d}x = 2 \sin \sqrt{b} - 2 \sin \sqrt{a}$. We can take the limit $a \to 0+$, but not the limit $b \to \infty$. Then $\int_0^1 f(x) \mathrm{d}x$ exists as an improper integral, but not $\int_1^\infty f(x) \mathrm{d}x$.

(b) Can we modify the function above so that the improper integral exists? Consider $f(x) = x^{-\frac{1}{2} - s} \cos \sqrt{x}$ where s is a fixed parameter. For which s does the improper integral exists?

We can gain insight using integration by parts; namely,

$$\int_a^b \frac{1}{x^s} \frac{1}{\sqrt{x}} \cos \sqrt{x} \, \mathrm{d}x = \frac{2 \sin \sqrt{x}}{x^s} \Big|_a^b + s \int_a^b \frac{2 \sin \sqrt{x}}{x^{s+1}} \mathrm{d}x. \tag{6.4}$$

In order to take the limit $a \to 0+$, the first term of the right side gives the condition $s \leq \frac{1}{2}$; the second term gives $s < \frac{1}{2}$. As for the limit $b \to \infty$, the first term gives the condition $s > 0$; this also implies that the second term converges. Thus both limits exist if $s \in (0, \frac{1}{2})$.

(c) If we consider an integral of a function such as $f(x) = |x-1|^{-1/2}$, we can decompose it as

$$\int_0^2 \frac{\mathrm{d}x}{|x-1|^{1/2}} = \int_0^1 \frac{\mathrm{d}x}{|x-1|^{1/2}} + \int_1^2 \frac{\mathrm{d}x}{|x-1|^{1/2}}, \tag{6.5}$$

and use the notion of improper integrals for each terms (this integral is then equal to 4).

We now discuss the integral in the sense of Riemann. The idea is to find good upper and lower bounds of the integral using step functions.

Definition 6.2. *Let $f : [a, b] \to \mathbb{R}$, and define*

- *The* **upper sum** $U_f = \inf\big\{\int_a^b \varphi : \varphi \in S[a, b]$ *and* $\varphi(x) \geq f(x) \; \forall x \in [a, b]\big\}$.
- *The* **lower sum** $L_f = \sup\big\{\int_a^b \varphi : \varphi \in S[a, b]$ *and* $\varphi(x) \leq f(x) \; \forall x \in [a, b]\big\}$.

We say that f is **Riemann-integrable** *iff $U_f = L_f$, in which case we define $\int_a^b f = U_f = L_f$.*

How does Riemann integration compare with the integral of regulated functions? It turns out to be more general.

Proposition 6.1. *If $f \in R[a,b]$, then it is Riemann-integrable, and both definitions give the same value.*

Proof. Since $f \in R[a,b]$, for any n there exists $\varphi_n \in S[a,b]$ such that

$$\varphi_n(x) - \tfrac{1}{n} \leq f(x) \leq \varphi_n(x) + \tfrac{1}{n}, \tag{6.6}$$

for all $x \in [a,b]$. A shifted step function is still a step function, so we get

$$\int_a^b \varphi_n - \tfrac{b-a}{n} \leq L_f \leq U_f \leq \int_a^b \varphi_n + \tfrac{b-a}{n}. \tag{6.7}$$

Letting $n \to \infty$, we see that $L_f = U_f$, and they are both equal to $\int_a^b f$. □

Let us look at some examples.

(a) $f(x) = \frac{1}{\sqrt{x}}$ on $[0,1]$ is still not Riemann, since there exists no step function that is larger than f.

(b) On the interval $[0,1]$, let $f(x) = \begin{cases} 1 & \text{if } x = 2^{-n} \text{ for some } n = 0,1,2,... \\ 0 & \text{otherwise.} \end{cases}$ We observe that $f(0+)$ does not exist, so it is not regulated. But we can use

$$\varphi_N(x) = \begin{cases} 1 & \text{if } x \leq 2^{-N}, \\ f(x) & \text{if } x > 2^{-N}, \end{cases} \tag{6.8}$$

to get $U_f \leq \int_0^1 \varphi_N = 2^{-N}$, hence $U_f \leq 0$. Using the 0 step function, we get $L_f \geq 0$. Then f is Riemann-integrable and its integral is zero.

(c) The function $\sin \frac{1}{x}$ on $[0,1]$ is Riemann-integrable, although it is not regulated.

(d) The function $f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}, \\ 0 & \text{if } x \in \mathbb{Q}, \end{cases}$ on the interval $[0,1]$ is not Riemann-integrable, and not regulated. Its integral cannot be defined as improper integral either. It is **Lebesgue-integrable**, though, and $\int_0^1 f = 1$.

7 Uniform and pointwise convergence

"Analysis is the art to take limits". Sequences of functions can converge in several different ways, and this diversity should be embraced as to make the theory more interesting and far-reaching. We have already seen the notion of *uniform convergence*: $f_n \to f$ uniformly if $\|f_n - f\|_\infty \to 0$ as $n \to \infty$. Another useful notion is *pointwise convergence*:

Definition 7.1. *Let $A \subset \mathbb{R}$ and $f_1, f_2, \dots$ be functions $A \to \mathbb{R}$. We say that $(f_n)_{n\geq 1}$* **converges** *to f* **pointwise** *if $f_n(x) \to f(x)$ for every $x \in A$.*

It is worth comparing pointwise and uniform convergence.

- $f_n \to f$ pointwise means that for all $x \in A$, for all $\varepsilon > 0$, there exists $N_{x,\varepsilon}$ such that $|f_n(x) - f(x)| < \varepsilon$ for all $n > N_{x,\varepsilon}$.
- $f_n \to f$ uniformly means that $\forall \varepsilon > 0$, there exists N_ε such that $|f_n(x) - f(x)| < \varepsilon$ for all $x \in A$ and all $n > N_{x,\varepsilon}$.

Uniform convergence implies pointwise convergence, but the converse is not always true. Here are some examples (please draw them!).

(a) On the interval $[0,1]$, $f_n(x) = \begin{cases} 1 - nx & \text{if } x \in [0, \frac{1}{n}], \\ 0 & \text{if } x \in (\frac{1}{n}, 1]. \end{cases}$ It is not hard to check that the pointwise limit f exists and is given by $f(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{if } x \in (0,1]. \end{cases}$ Convergence is not uniform, since $\|f_n - f\|_\infty = 1$ for all n, so it does not tend to 0.

(b) On the interval $[0,1]$, $f_n(x) = x^n$. The pointwise limit f exists and is given by $f(x) = \begin{cases} 0 & \text{if } x \in [0,1), \\ 1 & \text{if } x = 1. \end{cases}$ Convergence is not uniform, since we again have $\|f_n - f\|_\infty = 1$ for all n.

(c) Let g be a (non identically zero) function $\mathbb{R} \to \mathbb{R}$ such that $g(x) = 0$ for $|x| > 1$, and let $f_n(x) = g(x - n)$. We have $\lim_{n\to\infty} f_n(x) = 0$ for all x, but $\|f_n\|_\infty = \|g\|_\infty$, which does not depend on n and does not tend to 0. Convergence is then pointwise but not uniform.

Recall that a *metric space* (i.e. a set with a notion of distance between its elements) is *complete* if every Cauchy sequence converges within the space. We know that $(\mathbb{R}, |\cdot|)$ is complete.

Proposition 7.1. *The space of regulated function with the sup norm, $(R[a,b], \|\cdot\|_\infty)$, is complete.*

Proof. Let $(f_n)_{n\geq 1}$ be a sequence of regulated functions, that is Cauchy with respect to the sup norm. We show that (a) (f_n) converges to some function f; (b) convergence is uniform; (c) the limit function f is regulated.

(a) f_n converges pointwise. For every fixed x, we have $|f_m(x) - f_n(x)| \leq \|f_m - f_n\|_\infty$, so $(f_n(x))_{n\geq 1}$ is a Cauchy sequence of real numbers. It converges, and we define $(f(x) = \lim_{n\to\infty} f_n(x)$ for each $x \in [a,b]$.

(b) $f_n \to f$ uniformly. For every $\varepsilon > 0$, there exists N_ε such that

$$f_m(x) - \varepsilon \le f_n(x) \le f_m(x) + \varepsilon, \tag{7.1}$$

for all x and all $m, n > N_\varepsilon$. Taking the limit $m \to \infty$, we get

$$f(x) - \varepsilon \le f_n(x) \le f(x) + \varepsilon, \tag{7.2}$$

for all x and all $n > N_\varepsilon$. Then $\|f_n - f\|_\infty \to 0$ indeed.

(c) $f \in R[ab]$. Let $\varphi_{n,m}$ be step functions such that $\varphi_{n,m} \to f_n$ as $m \to \infty$, uniformly. Then there exists m_n such that $\|\varphi_{n,m_n} - f_n\|_\infty < \frac{1}{n}$. The sequence (φ_{n,m_n}) converges uniformly to f:

$$\|\varphi_{n,m_n} - f\|_\infty \le \|\varphi_{n,m_n} - f_n\|_\infty + \|f_n - f\|_\infty \tag{7.3}$$

and the right side goes to 0 as $n \to \infty$, so f is regulated. □

Next, a great theorem of analysis.

Theorem 7.2. *Let A be an arbitrary interval in $\mathbb{R}$ (open, closed, neither, finite, or infinite) and $C(A)$ be the space of continuous functions $A \to \mathbb{R}$. Then $(C(A), \|\cdot\|_\infty)$ is complete.*

Proof. Let $(f_n)_{n\ge1}$ be a Cauchy sequence of functions in $C(A)$ (with respect with the sup norm). Proceeding exactly as in the previous proof, we get the existence of a function $f : A \to \mathbb{R}$ such that $\|f_n - f\|_\infty \to 0$. There remains to verify that f is continuous.

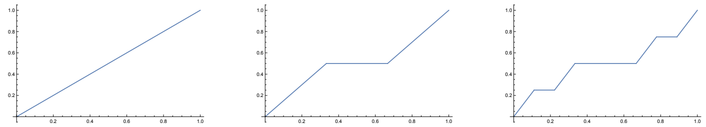
Let $x \in A$. For every $\varepsilon > 0$, there exists f_n in the above sequence such that $\|f_n - f\|_\infty < \frac{\varepsilon}{3}$. Further, there exists $\delta > 0$ such that $|f_n(y) - f_n(x)| < \frac{\varepsilon}{3}$ for all $y \in A$ with $|y - x| < \delta$. Then

$$|f(y) - f(x)| \le |f(y) - f_n(y)| + |f_n(y) - f_n(x)| + |f_n(x) - f(x)| \le \varepsilon, \tag{7.4}$$

for all $y \in A$ with $|y - x| < \delta$. This shows that f is indeed continuous. □

The devil's staircase function is a beautiful application of Theorem 7.2, which can be used to establish that it is continuous. We can define the devil's staircase as the limit of functions $f_n : [0,1] \to [0,1]$, where $f_0(x) = x$, and

$$f_{n+1}(x) = \begin{cases} \frac12 f_n(3x) & \text{if } x \in [0, \frac13), \\ \frac12 & \text{if } x \in [\frac13, \frac23], \\ \frac12 + \frac12 f_n(3x-2) & \text{if } x \in (\frac23, 1]. \end{cases} \tag{7.5}$$



The functions f_0, f_1, f_2 are illustrated above. At each iteration, the oblique straight lines are replaced by two oblique lines and a flat line in the centre. It is apparent that all f_n are continuous. Further,

$$\begin{aligned} \|f_{n+1} - f_n\|_\infty &= \max\Big\{ \sup_{x\in[0,\frac13)} |f_{n+1}(x) - f_n(x)|, \sup_{x\in[\frac13,\frac23]} |f_{n+1}(x) - f_n(x)|, \sup_{x\in(\frac23,1]} |f_{n+1}(x) - f_n(x)| \Big\} \\ &= \sup_{x\in[0,\frac13)} |f_{n+1}(x) - f_n(x)| \\ &= \sup_{x\in[0,\frac13)} |\tfrac12 f_n(3x) - \tfrac12 f_{n-1}(3x)| \\ &= \tfrac12 \|f_n - f_{n-1}\|_\infty. \end{aligned} \tag{7.6}$$

Iterating, we get $\|f_{n+1} - f_n\|_\infty \le 2^{-n}\|f_1 - f_0\|_\infty$. It follows that $(f_n)_{n\ge1}$ is a Cauchy sequence of continuous functions (with respect to the sup norm), so it converges to a continuous function by Theorem 7.2.

Notice that f is a continuous function that is constant on intervals whose total length is equal to 1!

8 Functions of two variables

We consider functions $f : D \to \mathbb{R}$ where $D = [a,b] \times [c,d]$, and explore the properties of objects such as $\int_a^b f(x,t)\mathrm{d}x$, $\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b f(x,t)\mathrm{d}x$, and $\int_c^d \big[\int_a^b f(x,t)\mathrm{d}x\big]\mathrm{d}t$. First, let us recall the notion of continuity.

Definition 8.1. *Let $f : D \to \mathbb{R}$, where D is an arbitrary set in $\mathbb{R}^2$.*

- *f is* **continuous** *at $(x_0, t_0) \in D$ if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x,t) - f(x_0,t_0)| < \varepsilon$ for all $(x,t) \in D$ with $|x - x_0| + |t - t_0| < \delta$.*

- *f is* **uniformly continuous** *if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x,t) - f(x',t')| < \varepsilon$ for all $(x,t), (x',t') \in D$ with $|x - x'| + |t - t'| < \delta$.*

As in the one-variable case (Lemma 3.2), one can prove that continuity implies uniform continuity if D is a closed domain.

Lemma 8.1. *Assume that f is uniformly continuous on D, and let $I : [c,d] \to \mathbb{R}$ be defined by $I(t) = \int_a^b f(x,t)\mathrm{d}x$. Then I is uniformly continuous on the interval $[c,d]$.*

Proof. We have $I(s) - I(t) = \int_a^b \big[f(x,s) - f(x,t)\big]\mathrm{d}x$. We know that for every $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x,s) - f(y,t)| < \varepsilon$ whenever $|x-y| + |s-t| < \delta$. For this δ, we have $|I(s) - I(t)| < \varepsilon(b-a)$ whenever $|s-t| < \delta$, so I is uniformly continuous indeed. □

Next, we exchange derivative and integral.

Proposition 8.2. *Assume that $f(x,t)$ and $\frac{\partial f}{\partial t}(x,t)$ are continuous on $[a,b] \times [c,d]$. Then for all $t \in (c,d)$, we have*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_a^b f(x,t)\,\mathrm{d}x = \int_a^b \frac{\partial f}{\partial t}(x,t)\,\mathrm{d}x.$$

Proof. Let $F(t) = \int_a^b f(x,t)\mathrm{d}x$ and $G(t) = \int_a^b \frac{\partial f}{\partial t}(x,t)\mathrm{d}x$. (Since f and $\frac{\partial f}{\partial t}$ are continuous, these integrals exist.) We have

$$\begin{aligned}\frac{F(t+h) - F(t)}{h} - G(t) &= \int_a^b \Big[\frac{f(x,t+h) - f(x,t)}{h} - \frac{\partial f}{\partial t}(x,t)\Big]\mathrm{d}x \\ &= \int_a^b \Big[\frac{\partial f}{\partial t}(x,\tau) - \frac{\partial f}{\partial t}(x,t)\Big]\mathrm{d}x.\end{aligned} \tag{8.1}$$

We used the mean-value theorem for the last line, and τ is a number that depends on x,t,h, which satisfies $\tau \in [t,t+h]$. Since $\frac{\partial f}{\partial t}$ is uniformly continuous, the last integrant vanishes as $h \to 0$. It follows that F is differentiable, and its derivative is G. □

Next, a version of Fubini theorem, which allows to exchange the order of integration.

Theorem 8.3. *Assume that $f : [a,b] \times [c,d] \to \mathbb{R}$ is continuous. Then*

$$\int_a^b \Big[\int_c^d f(x,t)\mathrm{d}t\Big]\mathrm{d}x = \int_c^d \Big[\int_a^b f(x,t)\mathrm{d}x\Big]\mathrm{d}t.$$

Physicists have another way to denote multiple integrals: The above identity is written $\int_a^b \mathrm{d}x \int_c^d \mathrm{d}t\, f(x,t) = \int_c^d \mathrm{d}t \int_a^b \mathrm{d}x\, f(x,t)$. This is less elegant but more convenient, and we will often use it.

Proof. Let

$$F(y) = \int_a^y \mathrm{d}x \int_c^d \mathrm{d}t\, f(x,t) - \int_c^d \mathrm{d}t \int_a^y \mathrm{d}x\,(f(x,t). \tag{8.2}$$

We show that $F(y) = 0$ for all $y \in [a,b]$. It is clear that $F(a) = 0$, and using the fundamental theorem of calculus, we have

$$F'(y) = \int_c^d \mathrm{d}t\, f(y,t) - \frac{\mathrm{d}}{\mathrm{d}y} \int_c^d \mathrm{d}t \int_a^y \mathrm{d}x\, f(x,t). \tag{8.3}$$

The second term of the right side is of the form $\frac{\mathrm{d}}{\mathrm{d}y} \int_c^d g(y,t)\mathrm{d}t$, with $g(y,t) = \int_a^y f(x,t)\mathrm{d}x$. The function g is continuous in y,t, and its derivative $\frac{\partial g}{\partial y} = f(y,t)$ is also continuous. Then we can use Proposition 8.2 to exchange derivative and integral. We get

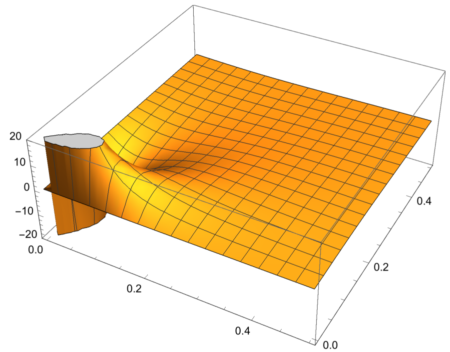
$$F'(y) = \int_c^d \mathrm{d}t\, f(y,t) - \int_c^d \mathrm{d}t\, f(y,t) = 0. \tag{8.4}$$

We have shown that F is constant, hence 0. □

In most cases, one can exchange the order of integration. But here is a counterexample: On the domain, $[0,2] \times [0,1]$, let

$$f(x,y) = \begin{cases} \frac{xy(x^2-y^2)}{(x^2+y^2)^3} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } x = y = 0. \end{cases} \tag{8.5}$$

This function behaves badly at $(0,0)$, as it diverges both to $+\infty$ and $-\infty$; see the illustration.



Integrals can be calculated explicitly (by substitution or integration by parts) and we find

$$\begin{aligned}\int_0^2 \mathrm{d}x \int_0^1 \mathrm{d}y\, f(x,y) &= \frac{1}{5}, \\ \int_0^1 \mathrm{d}y \int_0^2 \mathrm{d}x\, f(x,y) &= -\frac{1}{20}.\end{aligned} \tag{8.6}$$

Here, order of integration matters!

We now discuss the exchange of limits and derivatives: If $f_n \to f$, does f'_n converge to f'? An equivalent question is whether $(C^1[a,b], \|\cdot\|_\infty)$ is complete? Here, $C^1[a,b]$ denotes the space of continuously differentiable functions on the interval $[a,b]$. Let us look at an example.

On the interval $[-1,1]$, consider $f_n(x) = \sqrt{x^2 + \frac{1}{n}}$. It is clear that $f_n(x) \to |x|$ pointwise. Further, we have

$$|f_n(x) - f(x)| = \frac{|f_n(x)^2 - f(x)^2|}{|f_n(x) + f(x)|} = \frac{\frac{1}{n}}{\sqrt{x^2 + \frac{1}{n}} + |x|} \leq \frac{1}{\sqrt{n}} \tag{8.7}$$

for all x, so that $\|f_n - f\|_\infty \to 0$ as $n \to \infty$. The functions f_n are clearly differentiable (with $f'_n(x) = (x^2 + \frac{1}{n})^{-1/2} x$), they converge uniformly, but the limit $f(x) = |x|$ is not differentiable. This shows that $C^1[-1,1]$ is not complete with respect to the sup norm.

But there are many cases where the limit is differentiable.

Theorem 8.4. *Assume that $(f_n)_{n\geq 1}$ is a sequence of functions in $C^1[a,b]$ such that (f_n) and (f'_n) are Cauchy with respect to the sup norm. Then there exists $f \in C^1[a,b]$ such that $f_n \to f$ and $f'_n \to f'$, uniformly.*

Proof. By Theorem 7.2 ($C[a,b]$ is complete), there exist continuous functions f and g such that $f_n \to f$ and $f'_n \to g$ uniformly. There remains to check that f is differentiable, and that $f' = g$. We have

$$\int_a^x g = \int_a^x \lim_n f'_n = \lim_n \int_a^x f'_n = \lim_n \big(f_n(x) - f_n(a)\big) = f(x) - f(a). \tag{8.8}$$

Then $f(x) = f(a) + \int_a^x g$, and the fundamental theorem of calculus (Theorem 3.11) shows that f' exists and it is equal to g. □

9 Series of functions

Infinite sums of functions are mathematically interesting and they have important applications. Key examples include Taylor series, $f(x) = \sum_{k=0}^\infty \frac{f^{(k)}(a)}{k!}(x-a)^k$, and Fourier series, $f(x) = \sum_{k=0}^\infty (a_k \cos kx + b_k \sin kx)$. Recall that a series of numbers $\sum_k a_k$ converges iff the sequence of partial sums, $(\sum_{k=0}^n a_k)_{n\geq 1}$, converges as $n \to \infty$. This notion extends straightforwardly to series of functions.

Definition 9.1. *Let A be an arbitrary interval in $\mathbb{R}$, and let $(f_k)_{k\geq 0}$ be functions $A \to \mathbb{R}$.*

- *The series of functions $\sum_{k=0}^\infty f_k$* **converges pointwise** *if $\lim_{n\to\infty} \sum_{k=0}^n f_k(x)$ exists for each $x \in A$.*
- *The series of functions $\sum_{k=0}^\infty f_k$* **converges uniformly** *if there exists a function $S : A \to \mathbb{R}$ such that*

$$\lim_{n\to\infty} \Big\| \sum_{k=0}^n f_k - S \Big\|_\infty = 0.$$

We have seen that, if $g_n \to g$ uniformly, then $\int_a^b g_n \to \int_a^b g$. This immediately implies that, if $\sum_k f_k$ is a series of functions that converges uniformly, we have

$$\sum_{k=0}^\infty \int_a^b f_k(x)\,\mathrm{d}x = \int_a^b \Big(\sum_{k=0}^\infty f_k(x) \Big)\,\mathrm{d}x. \tag{9.1}$$

If all functions f_k are continuous and the series $\sum_k f_k$ converges uniformly, it follows from Theorem 7.2 that the limiting function is continuous. And if f_k are differentiable functions and $\sum_k f_k$, $\sum_k f'_k$ converge uniformly, the limiting function is also differentiable by Theorem 8.4.

Now, a simple criterion for the uniform convergence of series of functions. It is enough for many applications.

Proposition 9.1. *Let $(f_k)_{k\geq 0}$ be functions $A \to \mathbb{R}$, and assume that $\sum_{k=0}^\infty \|f_k\|_\infty < \infty$. Then $\sum_{k=0}^\infty f_k$ converges uniformly.*

Proof. For any $m < n$, we have

$$\Big\| \sum_{k=0}^m f_k - \sum_{k=0}^n f_k \Big\|_\infty = \Big\| \sum_{k=m+1}^n f_k \Big\|_\infty \leq \sum_{k=m+1}^\infty \|f_k\|_\infty. \tag{9.2}$$

The latter tends to 0 as $m, n \to \infty$. Then $(\sum_{k=0}^n f_k)_{n\geq 1}$ is a Cauchy sequence of functions with respect to the sup norm, so its pointwise limit exists and $\sum_{k=0}^n f_k$ converges uniformly to it. □

Jean Baptiste Joseph Fourier (1768–1830)

Joseph Fourier hesitated between religion and mathematics, and became an ardent revolutionary by 1793. He was arrested in 1794 during the Terror, but he escaped the guillotine thanks to political changes (Robespierre was guillotined on 28 July 1794). He studied at École Normale with Lagrange, Laplace, Monge, and was arrested and freed again in 1795.
In 1798, he joined Napoléon in his invasion (in French: "expédition") of Egypt. It started well, until Nelson destroyed the French fleet in the Battle of the Nile (1 August 1798). Napoléon returned to Paris in 1799, but Fourier and the rest of the expeditionary force remained there until 1801.
Back in France, he became Prefect of Isère (Grenoble) and he worked on the *Description of Egypt*. Between 1804 and 1807, Fourier wrote *On the propagation of heat in solid bodies*. This work was controversial at the time because of issues with Fourier series.
When Napoléon marched through Grenoble in the "hundred days", Fourier fled instead of welcoming him. He was nonetheless nominated Prefect of the Rhône. Shortly after this, Waterloo, and the end of Napoléon's epopee.
In 1822 he published his essay *Théorie analytique de la chaleur*, which Lord Kelvin described as "a great mathematical poem".

Let us now discuss Fourier series. The goal is to express $f : [0, 2\pi] \to \mathbb{R}$ as a series of sines and cosines. The main result is the following.

Theorem 9.2. *Let $(a_k)_{k\geq 0}$ and $(b_k)_{k\geq 0}$ be sequences of numbers such that $\sum_k |a_k| < \infty$ and $\sum_k |b_k| < \infty$. Then*

- *The series $\displaystyle\sum_{k=0}^{\infty} \big(a_k \cos kx + b_k \sin kx\big)$ converges uniformly on $\mathbb{R}$.*
- *The limiting function, $f(x)$, is continuous and 2π-periodic, i.e. $f(x + 2\pi) = f(x)$.*
- *We have $a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\mathrm{d}x$, and for all $k = 1, 2, 3, \ldots$,*

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(x) \cos kx \,\mathrm{d}x,$$
$$b_k = \frac{1}{\pi}\int_0^{2\pi} f(x) \sin kx \,\mathrm{d}x.$$

This theorem can be used in reverse order: Given a regulated function f on $[0, 2\pi]$, one can define the Fourier coefficients (a_k) and (b_k); if the latter are absolutely summable, the function f can be written as a convergent Fourier series.

Proof. Let $f_k = a_k \cos kx + b_k \sin kx$. Then $\|f_k\|_\infty < |a_k| + |b_k|$, and Proposition 9.1 implies that the Fourier series converges uniformly. Let $f = \sum_k f_k$ denote the limiting function.

Since f is the uniform limit of continuous functions, it is continuous by Theorem 7.2. 2π-periodicity is immediate, since $\cos kx$ and $\sin kx$ are 2π-periodic functions for all $k \in \mathbb{N}$.

The formulæ for the coefficients is not hard to establish, knowing that integrals can be interchanged with series that converge uniformly.

$$\begin{aligned}\frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\,\mathrm{d}x &= \frac{1}{\pi}\int_0^{2\pi}\sum_{p=0}^{\infty}\big(a_p\cos px + b_p \sin px\big)\cos kx\,\mathrm{d}x \\ &= \frac{1}{\pi}\sum_{p=0}^{\infty} a_p \int_0^{2\pi}\cos px\cos kx\,\mathrm{d}x + \frac{1}{\pi}\sum_{p=0}^{\infty} b_p\int_0^{2\pi}\sin px \cos kx\,\mathrm{d}x.\end{aligned} \tag{9.3}$$

One can check that

$$\begin{aligned}&\int_0^{2\pi}\cos px\cos kx\,\mathrm{d}x = \pi\delta_{p,k} \qquad \text{if } (p,k) \neq (0,0), \\ &\int_0^{2\pi}\sin px\cos kx\,\mathrm{d}x = 0,\end{aligned} \tag{9.4}$$

where $\delta_{p,k} = 1$ if $p = k$, and 0 otherwise, is called "Kronecker's symbol". The last line in Eq. (9.3) is then equal to a_k. The case of b_k is similar. □

If a function f is not continuous on $[0, 2\pi]$ (or if $f(0) \neq f(2\pi)$), its Fourier coefficients cannot be absolutely summable — that leads otherwise to a contradiction. If f is continuous, the situation is not clear. But if we assume f to be twice differentiable, it is definitely enough:

Proposition 9.3. *Assume that $f \in C^2[0, 2\pi]$, and also that $f(0) = f(2\pi)$ and $f'(0) = f'(2\pi)$. Then*

$$|a_k|, |b_k| \leq \frac{2\|f''\|_\infty}{k^2}.$$

Proof. We only prove that $|a_k| \leq \frac{2\|f''\|_\infty}{k^2}$, the other case being similar. Integrating by parts, and observing that the boundary terms vanish, we have

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_0^{2\pi} f(x)\,\cos kx\,\mathrm{d}x \\ &= \frac{1}{\pi}\Big[\frac{1}{k}f(x)\sin kx\Big|_0^{2\pi} - \frac{1}{k}\int_0^{2\pi} f'(x)\,\sin kx\,\mathrm{d}x\Big] \\ &= -\frac{1}{\pi k}\Big[-\frac{1}{k}f'(x)\cos kx\Big|_0^{2\pi} + \frac{1}{k}\int_0^{2\pi} f''(x)\,\cos kx\,\mathrm{d}x\Big] \\ &= -\frac{1}{\pi k^2}\int_0^{2\pi} f''(x)\,\cos kx\,\mathrm{d}x. \end{aligned} \tag{9.5}$$

Then $|a_k| \leq \frac{1}{\pi k^2}\int_0^{2\pi}|f''(x)|\,\mathrm{d}x \leq \frac{2}{k^2}\|f''\|_\infty$. □

Let us make use of Fourier series in order to study the heat equation. It describes the evolution of temperature in a metallic rod. If $T(x)$ denotes the temperature at x, the heat flux between x and $x+h$ is proportional to $T(x)-T(x+h)$. The change of temperature at x is then proportional to $T(x)-T(x+h)+T(x)-T(x-h)$. This suggests that the time derivative of T is proportional to the second derivative with respect to space.

We consider here a ring of perimeter 2π. Let $T(x,t)$ the temperature at $x \in [0,2\pi)$ and time $t \geq 0$. Its evolution satisfies the **heat equation**

$$\frac{\partial}{\partial t}T(x,t) = \frac{\partial^2}{\partial x^2}T(x,t). \tag{9.6}$$

The initial condition is give by the function $T_0(x)$. Let us assume that T_0 is sufficiently smooth (e.g. C^2) so that it is given by a Fourier series, $T_0(x) = \sum_{k=0}^\infty (a_k \cos kx + b_k \sin kx)$. We also assume that at all times, we have

$$T(x,t) = \sum_{k=0}^\infty \big(a_k(t)\cos kx + b_k(t)\sin kx\big). \tag{9.7}$$

Assuming we can differentiate inside the sums, the time first- and space second derivatives are

$$\begin{aligned} \frac{\partial}{\partial t}T(x,t) &= \sum_{k=0}^\infty \big(\dot a_k(t)\cos kx + \dot b_k(t)\sin kx\big), \\ \frac{\partial^2}{\partial x^2}T(x,t) &= \sum_{k=0}^\infty \big(-k^2 a_k(t)\cos kx - k^2 b_k(t)\sin kx\big). \end{aligned} \tag{9.8}$$

Here, $\dot a_k(t), \dot b_k(t)$ denote the time derivatives. Equating all terms, we get ordinary differential equations for each modes, namely

$$\dot a_k(t) = -k^2 a_k(t), \qquad a_k(0) = a_k, \tag{9.9}$$

and similarly for b_k. The solutions are $a_k(t) = a_k\,\mathrm{e}^{-k^2 t}$. Let us then *define*

$$T(x,t) = \sum_{k=0}^\infty \big(a_k\,\mathrm{e}^{-k^2t}\,\cos kx + b_k\,\mathrm{e}^{-k^2t}\,\sin kx\big). \tag{9.10}$$

We can now proceed backwards and check that this series converges absolutely, so it defines a function that indeed solves the heat equation. This establishes existence! It is possible to prove uniqueness, although this is more difficult and we leave it aside. Of special interest is to study the solution above. As the time increases, the Fourier coefficients become smaller, it converges better and the function becomes smoother. In the limit $t \to \infty$, we see that $a_k(t), b_k(t) \to 0$ for all $k \geq 1$, and the function converges to the constant function a_0, which is equal to the average initial temperature, $a_0 = \frac{1}{2\pi}\int_0^{2\pi} T_0(x)\,\mathrm{d}x$.

Part II — NORMS AND INNER PRODUCTS

The reader will have noticed the similarities between the absolute values of numbers, and the sup norm of functions. Both allow to define convergence and Cauchy sequences. A shared property is the triangle inequality which was used over and over again. It is worth formalising this structure for general vector spaces.

10 Normed vector spaces

Recall that a vector space (over $\mathbb{R}$) is a set where elements can be added with one another, and multiplied by numbers.

Definition 10.1. *Let V a vector space over $\mathbb{R}$. A function $\|\cdot\| : V \to \mathbb{R}$ is a* **norm** *if*

(i) it is positive: $\|v\| \geq 0$ for all $v \in V$, and $\|v\| = 0$ iff $v = 0$;

(ii) it is homogeneous: $\|\alpha v\| = |\alpha|\,\|v\|$ for all $\alpha \in \mathbb{R}$ and $v \in V$;

(iii) it satisfies the triangle inequality: $\|u+v\| \leq \|u\|+\|v\|$ for all $u, v \in V$.

A vector space equipped with a norm is a **normed space**.

Notice that (ii) and (iii) imply that $\|v\| \geq 0$: We have $0 = \|0\| = \|v - v\| \leq \|v\| + \|-v\| = 2\|v\|$, for all $v \in V$.

Special cases of normed vector spaces are $(\mathbb{R}, |\cdot|)$ or $(\mathbb{C}, |\cdot|)$. More interesting are norms on $V = \mathbb{R}^n$; vectors are denoted $x = (x_1, \dots, x_n)$. In cases (a) and (b), it is not hard to check that all axioms for a norm are satisfied.

(a) The sup norm $\|x\|_\infty = \max_{i=1,\dots,n} |x_i|$.

(c) The 1-norm $\|x\|_1 = \sum_{i=1}^n |x_i|$.

(d) The *Euclidean norm* $\|x\|_2 = \big(\sum_{i=1}^n |x_i|^2\big)^{1/2}$. This norm is induced by an inner product and the triangle inequality follows from the Cauchy-Schwarz inequality, as will be seen later.

(e) More generally, the p-norm $\|x\|_p = \big(\sum_{i=1}^n |x_i|^p\big)^{1/p}$. The triangle inequality holds for $p \geq 1$ and is known as *Minkowski inequality*:

$$\Big(\sum_{i=1}^n |x_i + y_i|^p\Big)^{1/p} \leq \Big(\sum_{i=1}^n |x_i|^p\Big)^{1/p} + \Big(\sum_{i=1}^n |y_i|^p\Big)^{1/p}. \tag{10.1}$$

The proof of this inequality is a bit difficult.

It is a good exercise to show that for all $x \in \mathbb{R}^n$, we have $\lim_{p\to\infty} \|x\|_p = \|x\|_\infty$.

Definition 10.2. *Let $(V, \|\cdot\|)$ be a normed space.*

- *A sequence $(v_n)_{n\geq 1}$ in V* **converges** *to $v \in V$ iff $\lim_{n\to\infty} \|v - v_n\| = 0$.*
- *A sequence $(v_n)_{n\geq 1}$ in V is* **a Cauchy sequence** *if for every $\varepsilon > 0$, there exists N_ε such that $\|v_m - v_n\| < \varepsilon$ for all $m, n > N_\varepsilon$.*
- *A normed vector space is* **complete** *iff every Cauchy sequence converges; a complete normed space is called a* **Banach space**.

Stefan Banach
(1892–1945)

Stefan Banach is one of the most important 20th century mathematicians. He was born in Kraków and he moved in 1910 to Lemberg (then in the Habsburg empire) in order to start university studies. He believed then that there was nothing new to discover in mathematics, so he studied engineering. He spent World War I in Kraków, where he met Hugo Steinhaus, who became a close friend and collaborator. He moved back to Lwów (now in Poland) in 1920. With other mathematicians, he founded the Lwów School of Mathematics, headquartered at the Scottish Café.
In 1939–41 Lwów was occupied by the Soviet Union and Banach became a corresponding member of the Academy of Sciences of Ukraine. But after the Nazi invasion in 1941, universities in Lwów were closed and Banach avoided deportation by working in a typhus research institute. The Soviets reoccupied Lwów in 1944. As other Poles, Banach was preparing to leave the town, now called Lviv, but he died of lung cancer in 1945.
Banach essentially founded linear analysis, and he obtained such fundamental results as the Hahn–Banach, Banach–Steinhaus, and Banach–Alaoglu theorems. The Banach–Tarski paradox is a mathematically rigorous theorem that relies on the axiom of choice and that states that a three-dimensional unit ball can be cut in finitely-many pieces, these pieces moved and rotated, and reassembled so as to give two unit balls. This puzzling result raises questions about the validity of the axiom of choice.

One easily proves that if $\lim_n v_n$ exists, it is unique: If $\|v_n - v\| \to 0$ and $\|v_n - v'\| \to 0$, then $\|v - v'\| \leq \|v - v_n\| + \|v_n - v'\| \to 0$; then $\|v - v'\| = 0$, so that $v = v'$.

Notice that $(\mathbb{R}, |\cdot|)$ is a Banach space, and also $(\mathbb{R}^n, \|\cdot\|_p)$ for $p \in [1, \infty]$. The space of step functions with the sup norm, $(S[a,b], \|\cdot\|_\infty)$, is not complete, hence not

Banach. But $(R[a,b], \|\cdot\|_\infty)$ and $(C[a,b], \|\cdot\|_\infty)$ are Banach spaces, by Proposition 7.1 and Theorem 7.2 respectively.

Definition 10.3. *Two norms $\|\cdot\|$ and $\|\cdot\|'$ on the vector space V are* **equivalent** *is there exist $k, K > 0$ such that for all $v \in V$,*

$$k\|v\| \leq \|v\|' \leq K\|v\|.$$

The motivation for this notion is that equivalent norms give the same notion of convergence (they induce the same *topology*). Indeed, if $(v_n)_{n\geq 1}$ satisfies $\|v_n - v\| \to 0$, then $\|v_n - v\|' \to 0$ for all norms $\|\cdot\|'$ that are equivalent to $\|\cdot\|$. Notice also that if $\|\cdot\|$ and $\|\cdot\|'$ are equivalent, and $\|\cdot\|'$ and $\|\cdot\|''$ are equivalent, then $\|\cdot\|$ and $\|\cdot\|''$ are also equivalent. Further, if $(V, \|\cdot\|)$ is Banach, and $\|\cdot\|$ and $\|\cdot\|'$ are equivalent, then $(V, \|\cdot\|')$ is also Banach.

Now comes a disappointing result.

Proposition 10.1. *All norms on $\mathbb{R}^n$ are equivalent.*

Proof. We show that all norms are equivalent to $\|\cdot\|_\infty$. Let $(e_j)_{j=1}^n$ denote the usual basis of $\mathbb{R}^n$, where $x = (x_1, \dots, x_n) = \sum_{j=1}^n x_j e_j$. Then

$$\|x\| = \Big\|\sum_{j=1}^n x_j e_j\Big\| \leq \sum_{j=1}^n |x_j|\,\|e_j\| \leq \|x\|_\infty \sum_{j=1}^n \|e_j\|. \tag{10.2}$$

We can define $K = \sum_{j=1}^n \|e_j\|$; this number is independent of x, and $\|x\| \leq K\|x\|_\infty$ for all x.

Next, we define

$$k = \inf\Big\{\frac{\|x\|}{\|x\|_\infty} : x \in \mathbb{R}^n \setminus \{0\}\Big\}. \tag{10.3}$$

Since $k \leq \frac{\|x\|}{\|x\|_\infty}$, we have $k\|x\|_\infty \leq \|x\|$ for all x; this is what we want to show, but we need to check that $k \neq 0$. By rescaling, we have that

$$k = \inf\big\{\|x\| : x \in \mathbb{R}^n \text{ and } \|x\|_\infty = 1\big\}. \tag{10.4}$$

We proceed *ab absurdo*. If $k = 0$, there exists a sequence $(x_\ell)_{\ell\geq 1}$ such that $\|x\|_\infty = 1$ for all ℓ, and $\|x_\ell\| \to 0$ as $\ell \to \infty$.

There exists a subsequence $(x_{\ell_m})_{m\geq 1}$ that converges with respect to the $\|\cdot\|_\infty$. Indeed, by Bolzano-Weierstrass, there exists a subsequence of (x_ℓ) such that the first components converge; again by Bolzano-Weierstrass, there exists a subsequence of the subsequence such that the second components converge (the first components still converge). We can continue to extract subsequences until all components converge. It is not hard to check that the resulting subsequence converges with respect to $\|\cdot\|_\infty$.

Let $x \in \mathbb{R}^n$ denote the limit of the subsequence (x_{ℓ_m}). The triangle inequality implies that

$$\Big|\|x\|_\infty - \|x_{\ell_m}\|_\infty\Big| \leq \|x - x_{\ell_m}\|_\infty. \tag{10.5}$$

This tends to 0 as $m \to \infty$, so that $\|x\|_\infty = 1$ since $\|x_\ell\|_\infty = 1$ for all ℓ. But we also have

$$\Big|\|x\| - \|x_{\ell_m}\|\Big| \leq \|x - x_{\ell_m}\| \leq K\|x - x_{\ell_m}\| \to 0, \tag{10.6}$$

where we used the bound that we have already established. Since $\|x_\ell\| \to 0$, this implies that $\|x\| = 0$, hence $x = 0$ and $\|x\|_\infty = 0$, contradiction. The conclusion is that $k > 0$. □

Since all finite-dimensional vector spaces are isomorphic to $\mathbb{R}^n$ for some n, this proposition shows that all norms are equivalent in finite-dimensional spaces. Things are much more interesting in infinite-dimensional spaces, such as spaces of sequences.

Given $x = (x_k)_{k\geq 0}$, we define

- $\|x\|_p = \Big(\sum_{k=0}^\infty |x_k|^p\Big)^{1/p}$ and $\ell_p = \big\{(x_k)_{k\geq 0} : \|x\|_p < \infty\big\}$.
- $\|x\|_\infty = \sup_{k\in\mathbb{N}} |x_k|$ and $\ell_\infty = \big\{(x_k)_{k\geq 0} : \|x\|_\infty < \infty\big\}$. This is the space of all bounded sequences.

One can check that $(\ell_p, \|\cdot\|_p)$ are normed vector spaces for $p \in [0, \infty]$. (Minkowski inequality holds for $p \geq 1$, and it is not hard to prove for $p = 1, 2, \infty$.) Notice also that $\ell_p \subset \ell_q$ when $p < q$ (exercise!), and that ℓ_p is a Banach space for all $p \in [1, \infty]$.

The $\|\cdot\|_p$ norms are *not* equivalent; indeed, for $p < q$ there exist sequences $(x^{(n)})_{n\geq 1}$ of elements of ℓ_p such that $\|x^{(n)}\|_p > \varepsilon > 0$ for all n, but $\|x^{(n)}\|_q \to 0$. For instance, consider

$$x_k^{(n)} = \begin{cases} (k \log n)^{-1/p} & \text{if } 1 \leq k \leq n, \\ 0 & \text{otherwise.} \end{cases} \tag{10.7}$$

Then $\|x^{(n)}\|_p = \big(\frac{1}{\log n}\sum_{k=1}^n \frac{1}{k}\big)^{1/p}$, which tends to 1 as $n \to \infty$; but $\|x^{(n)}\|_q = (\log n)^{-1/p}\big(\sum_{k=1}^n k^{-q/p}\big)^{1/q}$. Since $q/p > 1$, the sum is bounded uniformly in n, so the prefactor makes it go to 0 as $n \to \infty$.

11 Inner products

Definition 11.1. *An* **inner product** *(or* **scalar product***) on the vector space V is a function $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ such that*

(i) it is symmetric: $\langle x,y\rangle = \langle y,x\rangle$ for all $x,y \in V$;

(ii) it is linear: $\langle x, \alpha y + \beta z\rangle = \alpha\langle x,y\rangle + \beta\langle x,z\rangle$ for all $\alpha,\beta \in \mathbb{R}$ and $x,y \in V$;

(iii) it satisfies $\langle x,x\rangle \geq 0$ for all $x \in V$, and $\langle x,x\rangle = 0$ implies that $x = 0$.

This definition is for *real* inner products.[1]

Definition 11.2. *The* **induced norm** *of an inner product is*

$$\|x\| = \sqrt{\langle x,x\rangle}.$$

A vector space with an inner product, which is complete with respect to the induced norm, is called a **Hilbert space**.

Theorem 11.1.

(a) The induced norm is a norm.

(b) The inner product satisfies **Cauchy–Буняко́вский-Schwarz inequality***:*

$$\big|\langle x,y\rangle\big| \leq \|x\|\,\|y\|.$$

(c) The induced norm satisfies the **parallelogram identity***:*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Proof. (a) and (c) are easy and are left as an exercise.

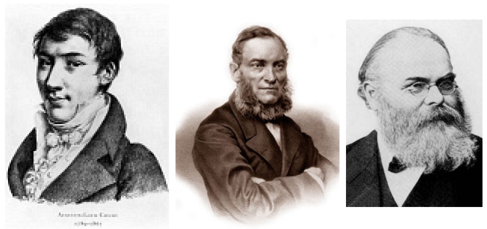
For (b), we use the one inequality at hand, applied to $x + \alpha y$:

$$0 \leq \langle x+\alpha y, x+\alpha y\rangle = \|x\|^2 + 2\alpha\langle x,y\rangle + \alpha^2\|y\|^2. \tag{11.1}$$

This is nonnegative for all $\alpha \in \mathbb{R}$, so this quadratic polynomial cannot have two zeroes. This means that the discriminant is nonpositive, that is, $4\langle x,y\rangle^2 - 4\|x\|^2\|y\|^2 \leq 0$. This is precisely the Cauchy-Schwarz inequality. □

[1] In quantum mechanics, one needs complex vector spaces where inner products satisfy $\langle x,y\rangle = \overline{\langle y,x\rangle}$; they are antilinear in the first variable and linear in the second.

Cauchy–Bunyakovsky–Schwarz inequality



The full setting and the inequality came only progressively. Augustin-Louis Cauchy (1789–1857) obtained it for sums in 1821. Ви́ктор Яковлевич Буняко́вский (Viktor Yakovlevich Bunyakovsky, 1804–1889) extended it to integrals in 1859. Hermann Amandus Schwarz (1843–1921) proposed a modern proof in 1888.

Inner product spaces enjoy many more properties than normed spaces, so it helps to know whether a given norm is the induced norm for some inner product. It follows from Theorem 11.1 (c) that a necessary condition is that the norm satisfies the parallelogram identity. It turns out that it is a sufficient condition.

Proposition 11.2. *If a norm satisfies the parallelogram identity, then there exists an inner product such that it is the induced norm. The inner product is given by*

$$\begin{aligned}\langle x,y\rangle &= \tfrac12\big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\big)\\ &= \tfrac14\big(\|x+y\|^2 - \|x-y\|^2\big).\end{aligned}$$

The two forms of the inner product are identical because of the parallelogram identity.[2]

Proof. The proofs of properties (i) and (iii) of the inner product are immediate, but linearity is harder to establish than could have perhaps been expected. We start with

$$\|x+y+z\|^2 - \|x-y-z\|^2 = \|x+y+z\|^2 + \|x+y-z\|^2 - \|x-z+y\|^2 - \|x-z-y\|^2. \tag{11.2}$$

Using the parallelogram identity, we then get

$$\begin{aligned}\|x+y+z\|^2 - \|x-y-z\|^2 &= 2\|x+y\|^2 + 2\|z\|^2 - 2\|x-z\|^2 - 2\|y\|^2\\ &= 2\|x+y\|^2 + 2\|y\|^2 - 2\|x-y\|^2 - 2\|z\|^2.\end{aligned} \tag{11.3}$$

[2] This proposition remains true for complex vector spaces, but the inner product is given by the *polarisation identity*

$$\langle x,y\rangle = \tfrac14\big(\|x+y\|^2 - \|x-y\|^2 - \mathrm{i}\|x+\mathrm{i}y\|^2 + \mathrm{i}\|x-\mathrm{i}y\|^2\big).$$

The proof is nearly identical.

The second equation is obtained by exchanging y and z. Averaging the two right sides, we get

$$\|x+y+z\|^2 - \|x-y-z\|^2 = \|x+y\|^2 - \|x-y\|^2 + \|x+z\|^2 - \|x-z\|^2. \tag{11.4}$$

We have just proved that for all $x, y, z \in V$,

$$\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle. \tag{11.5}$$

There remains to prove that $\langle x, \alpha y\rangle = \alpha\langle x, y\rangle$. Eq. (11.5) proves it for $\alpha \in \mathbb{N}$. Since $\langle x, -y\rangle = -\langle x, y\rangle$, this holds for $\alpha \in \mathbb{Z}$. Further, by the homogeneity of norms, we have

$$\langle \alpha x, \alpha y\rangle = \tfrac{1}{4}\big(\|\alpha(x+y)\|^2 - \|\alpha(x-y)\|^2\big) = \alpha^2\langle x, y\rangle. \tag{11.6}$$

Now let $\alpha = p/q$ with $p, q \in \mathbb{N}$. Using both equations above, we have

$$\langle x, \alpha y\rangle = \tfrac{1}{q^2}\langle qx, qy\rangle = \tfrac{p}{q}\langle x, y\rangle. \tag{11.7}$$

So linearity holds for all $\alpha \in \mathbb{Q}$. A continuity argument extends it to all $\alpha \in \mathbb{R}$. Let α_n be rational numbers that converge to α; since the norm is continuous, we have

$$\begin{aligned}\langle x, \alpha y\rangle &= \tfrac{1}{4}\big(\|x+\alpha y\|^2 - \|x-\alpha y\|^2\big)\\ &= \lim_{n\to\infty} \tfrac{1}{4}\big(\|x+\alpha_n y\|^2 - \|x-\alpha_n y\|^2\big)\\ &= \lim_{n\to\infty} \langle x, \alpha_n y\rangle\\ &= \lim_{n\to\infty} \alpha_n\langle x, y\rangle\\ &= \langle x, y\rangle.\end{aligned} \tag{11.8}$$

We used linearity with respect to rational numbers, which we checked in Eq. (11.7). □

12 Linear maps (operators)

Functional analysis (also called linear analysis) is the domain of mathematics that studies vector spaces or arbitrary dimensions and linear maps between these spaces. This field originates from quantum mechanics, but it has much broader applications.

A **linear map** (or **operator**) between vector spaces V and V' is a map $T : V \to V'$ that satisfies

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y), \tag{12.1}$$

for all numbers $\alpha, \beta \in \mathbb{R}$ and vectors $x, y \in V$. It is customary to write Tx instead of $T(x)$. Linear maps are quite peculiar, but they include many interesting examples:

- Let $V = \mathbb{R}^m$ and $V' = \mathbb{R}^n$, and T an $n \times m$ matrix. Then

$$Tx = \begin{pmatrix} T_{11} & \cdots & T_{1m} \\ \vdots & & \vdots \\ T_{n1} & \cdots & T_{nm}\end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_m\end{pmatrix}, \tag{12.2}$$

  which gives a vector with n components.
- A differential operator: the map $T : C^1[a,b] \to C[a,b]$ where $Tf = f'$.
- An integral operator: the map $T : C[a,b] \to C[a,b]$ where the image Tf is the function

$$(Tf)(x) = \int_a^x f(t)\,\mathrm{d}t, \tag{12.3}$$

  with $x \in [a,b]$.

Definition 12.1. *Let $(V, \|\cdot\|)$ and $(V', \|\cdot\|')$ be normed spaces. The* **operator norm** *of the linear map $T : V \to V'$ is defined as*

$$\|T\| = \sup_{x\in V, x\neq 0} \frac{\|Tx\|'}{\|x\|} = \sup_{x\in V, \|x\|=1} \|Tx\|'.$$

We say that T is **bounded** *if $\|T\| < \infty$.*

The reader should verify that both supremums above give the same value. It follows from the definition that $\|Tx\|' \leq \|T\|\,\|x\|$, a useful inequality. Before we state the next proposition, let us recall that a map f between normed space $(V, \|\cdot\|)$ and $(V', \|\cdot\|')$ is **continuous** at $y \in V$ if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $\|f(x) - f(y)\|' < \varepsilon$ for all $x \in V$ such that $\|x - y\| < \delta$. Here, the map f is not necessarily linear.

Linear maps are continuous iff they are bounded.

Proposition 12.1. *Let $(V, \|\cdot\|)$ and $(V', \|\cdot\|')$ be normed spaces and $T : V \to V'$ be a linear map. The following are equivalent:*

(a) T is continuous at 0 (here, $0 \in V$).

(b) T is continuous.

(c) $\|T\| < \infty$.

Proof. It is clear that (b) implies (a). The converse implication follows from linearity: Continuity at 0 implies that for all $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\|Tx - Ty\|' = \|T(x-y)\|' < \varepsilon \tag{12.4}$$

whenever $\|x-y\| < \delta$. This means that T is also continuous at y.

Next, we show that (a) implies (c). For $x \in V$ with $\|x\| = 1$, we have

$$\|Tx\|' = \tfrac{1}{\delta}\|T(\delta x)\|' < \tfrac{\varepsilon}{\delta}. \tag{12.5}$$

The upper bound holds for all x (such that $\|x\| = 1$), so that $\|T\| < \frac{\varepsilon}{\delta} < \infty$.

Finally, (c) implies (a): We choose $\delta = \varepsilon / \|T\|$, so that for all $\|x\| < \delta$, we have

$$\|Tx\|' \leq \|T\|\,\|x\| < \|T\| \tfrac{\varepsilon}{\|T\|} = \varepsilon. \tag{12.6}$$

Then T is continuous at 0 indeed. □

One easily checks that all matrices have finite operator norms. But differential operators are typically unbounded. Consider the normed spaces $V = (C^1[0,1], \|\cdot\|_\infty)$ and $V' = (C[0,1], \|\cdot\|_\infty)$, and the operator $Tf = f'$. Let $f_n(x) = x^n$; then $f_n'(x) = nx^{n-1}$, so that $\|f_n\|_\infty = 1$ and $\|f_n'\|_\infty = n$. Then

$$\|T\| \geq \frac{\|f_n'\|_\infty}{\|f_n\|_\infty} = n. \tag{12.7}$$

This is true for any n, so that $\|T\| = \infty$.

Let us now look at the integral operator mentioned above. Let $V = (C[a,b], \|\cdot\|_\infty)$ and $T : V \to V$ defined by $(Tf)(x) = \int_a^x f(t)\,\mathrm{d}t$. We have

$$\|T\| = \sup_{\|f\|_\infty = 1} \|Tf\| = \sup_{\|f\|_\infty=1} \sup_{x\in[a,b]} \Big| \int_a^x f(t)\,\mathrm{d}t \Big| = b - a. \tag{12.8}$$

Indeed, the right side is an upper bound because $|f(t)| \leq 1$; it is also a lower bound by choosing the constant function $f(t) = 1$ for all $t \in [a,b]$.

13 Open and closed sets

The notion of open and closed sets is rather technical and its beauty is very well hidden. But this cannot be avoided without prejudice, so we discuss it in this section.

Definition 13.1. *Let $(V, \|\cdot\|)$ be a normed space.*

- *The* **open ball** *$B(x,\delta)$ of radius $\delta > 0$ centred at $x \in V$ is*

  $$B(x,\delta) = \big\{ y \in V : \|x-y\| < \delta \big\}.$$

- *A subset $U \subset V$ is* **open** *if for every $x \in U$, there exists $\delta > 0$ such that $B(x,\delta) \subset U$.*
- *A subset $U \subset V$ is* **closed** *if its complement $U^{\mathrm{c}} = V \setminus U$ is open.*

The set $\overline{B}(x,\delta) = \big\{ y \in V : \|x-y\| \leq \delta \big\}$ is called the **closed ball** of radius δ centred on x. Heuristically, open sets are sets that do not contain their boundary, while closed sets contain their boundary.

Let us review some examples.

(a) On $\mathbb{R}$, the interval $(a,b) = \{x \in \mathbb{R} : a < x < b\}$ is open; $[a,b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ is closed; $(a,b] = \{x \in \mathbb{R} : a < x \leq b\}$ is neither closed nor open. $\mathbb{R}$ is both closed and open.

(b) On $\mathbb{R}^2$, it is a good exercise to draw the balls of radius 1, centred at 0, with norms $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$.

(c) Consider a function $f \in C[a,b]$, and draw a few other continuous functions that belong to $B(f,\delta)$ with respect to $\|\cdot\|_\infty$. What about other norms, such as $\|\cdot\|_1$?

The property of being open (or closed) depends on the norm; if two norms are equivalent, they yield the same open and closed sets.

Let $f : V \to V'$ be a map between arbitrary sets V and V'; for $U \subset V'$, we denote

$$f^{-1}(U) = \big\{ x \in V : f(x) \in U \big\}. \tag{13.1}$$

This definition makes sense even when f is not invertible. This and the notion of open sets allow for an abstract but surprisingly useful characterisation of continuity.

Proposition 13.1. *A map f between normed spaces $(V, \|\cdot\|)$ and $(V', \|\cdot\|')$ is continuous iff $f^{-1}(U)$ is open for any open $U \subset V'$.*

It is a good idea to illustrate this proposition by looking at continuous and discontinuous functions $\mathbb{R} \to \mathbb{R}$.

In mathematics, the most general notion of continuity involves a *topology*, which is a collection of open sets; the property above is then used as definition of continuity. This allows to work with objects that do not fit in a normed space. A practical application of this proposition is that compositions of continuous maps are continuous, as stated in the next corollary. Recall the notation $(g \circ f)(x) = g(f(x))$.

Corollary 13.2. *If $f : V \to V'$ and $g : V' \to V''$ are continuous, then $g \circ f : V \to V$" is continuous.*

The proof is immediate, as $g \circ f$ maps back open sets to open sets.

Proof of Proposition 13.1. Assume that f is continuous, and let $U \subset V'$ be open. We show that $f^{-1}(U)$ is open. Let $x \in f^{-1}(U)$; then $f(x) \in U$, and there exists $\varepsilon > 0$ such that $B(f(x),\varepsilon) \subset U$.

Since f is continuous, there exists $\delta > 0$ such that $y \in B(x,\delta) \implies f(y) \in B(f(x),\varepsilon) \subset U$. Then $y \in f^{-1}(U)$. This shows that $B(x,\delta) \subset f^{-1}(U)$, so we have proved that $f^{-1}(U)$ is open.

We now assume that $f^{-1}(U)$ is open for any open $U \subset V'$. Let $x \in V$ and $\varepsilon > 0$. Since $f^{-1}(B(f(x), \varepsilon))$ is open, it contains the ball $B(x, \delta)$ for some $\delta > 0$. Then $f(B(x, \delta)) \subset B(f(x), \varepsilon)$, so f is continuous at x. □

Here is a useful characterisation of closed sets, that indeed suggests that they include boundary points.

Proposition 13.3. *A subset U of the normed space $(V, \|\cdot\|)$ is closed iff $x_n \in U$ for all n, and $x_n \to x$, imply that $x \in U$.*

Then $(0, 1] \in \mathbb{R}$ is not closed since the sequence $(\frac{1}{n})_{n \geq 1}$ converges to 0, but 0 does not belong to $(0, 1]$.

Proof. We prove that: U not closed $\Leftrightarrow$ there exists (x_n) in U with $x_n \to x \notin U$.
$\Rightarrow$: If U is not closed, $V \setminus U$ is not open, so there exists $x \in V \setminus U$ such that $B(x, \delta) \not\subset V \setminus U$ for all $\delta > 0$. It follows that for all n, there exists $x_n \in U$ such that $\|x_n - x\| < \frac{1}{n}$, so $x_n \to x$.
$\Leftarrow$: Let (x_n) in U such that $x_n \to x \notin U$. For any $\delta > 0$, the ball $B(x, \delta)$ contains x_n for n large enough; but $x_n \notin V \setminus U$, so $V \setminus U$ is not open, and U is not closed. □

14 The contraction mapping theorem

Definition 14.1. *Let $(V, \|\cdot\|)$ be normed space and $U \subset V$ a closed subset. A map $f : U \to U$ is a* **contraction** *on U if there exists $\eta < 1$ such that*

$$\big\|f(x) - f(y)\big\| \leq \eta \|x - y\|$$

for all $x, y \in U$.

Notice that it is not enough that $\|f(x) - f(y)\| < \|x - y\|$ for all $x, y \in U$; one needs the constant $\eta < 1$. It is worth pointing out that contractions are continuous maps.

A vector $x \in V$ is a **fixed point** of the function $f : V \to V$ if $f(x) = x$.

Theorem 14.1. (Contraction mapping theorem)
Let V be a Banach space; $U \subset V$ a closed subset; f a contraction on U. Then f **has** *a* **unique** *fixed point in U.*
Further, starting from any $x_0 \in U$, the sequence (x_n) defined inductively by $x_n = f(x_{n-1})$, $n \geq 1$, converges to the fixed point.

This result is also called the **Banach fixed point theorem**. We state it for Banach spaces, but the space does not need to be linear — instead of a norm, a *distance* is enough. We need completeness, though.

Proof. We start by verifying that the sequence (x_n) of the theorem is Cauchy. We have

$$\|x_{n+1} - x_n\| = \|f(x_n) - f(x_{n-1})\| \leq \eta \|x_n - x_{n-1}\| \leq \cdots \leq \eta^n \|x_1 - x_0\|. \tag{14.1}$$

If $m < n$,

$$\|x_m - x_n\| \leq \sum_{k=m}^{n-1} \|x_k - x_{k+1}\| \leq \sum_{k=m}^{\infty} \eta^k \|x_1 - x_0\| = \frac{\eta^m}{1 - \eta} \|x_1 - x_0\|. \tag{14.2}$$

We see that for every $\varepsilon > 0$, we can find N large enough so that $\|x_m - x_n\| < \varepsilon$ for all $m, n > N$.

The sequence (x_n) converges since V is complete; let x denote the limit. It belongs to U since U is closed, and it is fixed point by continuity of f:

$$f(x) = \lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} x_{n+1} = x. \tag{14.3}$$

Finally, the fixed point is unique: If x, z are two fixed points, we have

$$\|x - z\| = \|f(x) - f(z)\| \leq \eta \|x - z\|, \tag{14.4}$$

which implies that $\|x - z\| = 0$. □

The proof shows that convergence of the sequence (x_n) is exponentially fast.

Let us illustrate the theorem with a few applications. First, consider the equation $\mathrm{e}^{-ax} = x$ with the parameter $a \in (0, 1)$. We are looking for solution(s) $x \in [0, \infty)$. Let $f(x) = \mathrm{e}^{-ax}$; by the mean-value theorem,

$$|\,\mathrm{e}^{-ax} - \mathrm{e}^{-ay}\,| = |f'(\xi)|\,|x - y| \tag{14.5}$$

with ξ between x and y. Then $|f'(\xi)| = a\,\mathrm{e}^{-a\xi} \leq a$, so that $|f(x) - f(y)| < a|x - y|$, and f is a contraction. This proves that there exists a unique solution to the above equation.

Integral equations. Consider $C[a, b]$ with the $\|\cdot\|_\infty$ norm. We are given functions $k : [a, b] \times [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$, and we consider the *Fredholm integral equation of the 2nd kind*:

$$f(x) - \int_a^b k(x, y) f(y) \mathrm{d}y = g(x). \tag{14.6}$$

Introducing the operator $T : C[a, b] \to C[a, b]$ by

$$(Tf)(x) = g(x) + \int_a^b k(x, y) f(y) \mathrm{d}y, \tag{14.7}$$

we see that the equation above is the fixed point equation $Tx = x$. We seek sufficient conditions for the existence of a unique solution. We have

$$\begin{aligned}\|Tf - Tg\|_\infty &= \sup_{x\in[a,b]} \Big| \int_a^b k(x,y)\big(f(y) - g(y)\big)\,\mathrm{d}y \Big| \\ &\leq \sup_{x\in[a,b]} \int_a^b |k(x,y)|\big|f(y) - g(y)\big|\mathrm{d}y \\ &\leq \|f - g\|_\infty \sup_{x\in[a,b]} \int_a^b |k(x,y)|\,\mathrm{d}y.\end{aligned} \tag{14.8}$$

We see that T is a contraction if

$$\sup_{x\in[a,b]} \int_a^b |k(x,y)|\,\mathrm{d}y < 1. \tag{14.9}$$

The contraction mapping theorem implies the existence of a unique fixed point.

It is worth doing the same exercise with norm $\|\cdot\|_1$ instead of $\|\cdot\|_\infty$. We find that T is a contraction provided

$$\int_a^b \Big(\sup_{y\in[a,b]} |k(x,y)| \Big)\,\mathrm{d}x < 1. \tag{14.10}$$

This condition is distinct from the one before, so we can solve other equations. But a non-trivial problem occurs here, namely that the space $(C[a,b], \|\cdot\|_1)$ is not complete; it must first be completed, in a similar way that $\mathbb{R}$ is the completion of $\mathbb{Q}$, so the contraction mapping theorem can be applied.

Ordinary differential equations. Consider an ordinary differential equation of the form

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= F\big(t, x(t)\big), \\ x(t_0) &= x_0.\end{aligned} \tag{14.11}$$

Here, x is a function on $\mathbb{R}$, $x_0 \in \mathbb{R}$ is the initial conditions, and F is a function $\mathbb{R}^2 \to \mathbb{R}$. We are looking for sufficient conditions that guarantees the existence of a differentiable function x that solves this equation. We use the contraction mapping theorem.

It helps to use an integral equation; assuming that x is a solution to the equation above, we can integrate and we get

$$x(t) - x(t_0) = \int_{t_0}^t F\big(s, x(s)\big)\,\mathrm{d}s, \tag{14.12}$$

or $x(t) = x(t_0) + \int_{t_0}^t F\big(s, x(s)\big)\,\mathrm{d}s$. Conversely, if x solves the integral equation, then it solves the differential equation by the fundamental theorem of calculus (Theorem 3.11; we should assume here that F is continuous).

Let T denote the **Picard operator**

$$\begin{aligned}T : &C[t_0 - \delta, t_0 + \delta] \to C[t_0 - \delta, t_0 + \delta], \\ &x \mapsto x_0 + \int_{t_0}^t F\big(s, x(s)\big)\,\mathrm{d}s.\end{aligned} \tag{14.13}$$

The integral equation becomes $Tx = x$, so we are looking for fixed points of T. We equip $C[t_0 - \delta, t_0 + \delta]$ with the sup norm (so it is complete indeed). Let us assume that F is **Lipschitz-continuous** with respect to the second variable, namely that

$$|F(s,t) - F(s,u)| \leq L|t - u|, \tag{14.14}$$

for a constant L that does not depend on s, t, u. Then

$$\begin{aligned}\big|Tx(t) - Ty(t)\big| &= \Big| \int_{t_0}^t \Big(F\big(s, x(s)\big) - F\big(s, y(s)\big) \Big)\,\mathrm{d}s \Big| \\ &\leq \int_{t_0}^t L\big|x(s) - y(s)\big|\,\mathrm{d}s \\ &\leq L|t - t_0|\,\|x - y\|_\infty.\end{aligned} \tag{14.15}$$

This shows that $\|Tx - Ty\|_\infty \leq L\delta\,\|x - y\|_\infty$, so T is a contraction if $L\delta < 1$. Hence the existence of a unique solution. We have just proved the **Picard-Lindelöf theorem**.

Theorem 14.2. *Let $D = [t_0 - a, t_0 + a] \times [x_0 - b, x_0 + b]$. Assume that $F : D \to \mathbb{R}$ is continuous in the first variable and Lipschitz-continuous in the second variable. Then there exists $\delta > 0$ such that the differential equation $\frac{\mathrm{d}x}{\mathrm{d}t} = F(t, x(t))$* **has** *a* **unique** *solution $x \in C^1[t_0 - \delta, t_0 + \delta]$.*

It is worth pointing out that the conclusion of this theorem is not obvious. For instance, the equation $\frac{\mathrm{d}x}{\mathrm{d}t} = \sqrt{x(t)}$ has two solutions with $t \in [0,1]$: $x \equiv 0$ and $x(t) = \frac{1}{4}t^2$. The theorem does not apply because the function $F(t) = \sqrt{t}$ is not Lipschitz.

Newton-Raphson method. This is a method that helps to solve the equation $f(x) = 0$, where f is a function $[a,b] \to \mathbb{R}$. The idea is to start with a reasonable guess, x_0, and to apply the following iteration:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{14.16}$$

Then we hope that x_n tends to the solution as $n \to \infty$. Draw a picture to see why!

The method does not always work. A counter-example is the function $g(x) = x^{1/3}$, with $x_0 \neq 0$. One gets $x_{n+1} = -2x_n$, which does not define a convergent sequence.

We assume that $f \in C^2(\mathbb{R})$, that $g'(x) \neq 0$ for all x, and that

$$\eta = \sup_{x \in \mathbb{R}} \frac{|f(x) f''(x)|}{f'(x)^2} < 1. \tag{14.17}$$

As we shall see, it is enough to guarantee that the equation $f(x) = 0$ has a unique solution, and that x_n converges to it.

We work with the Banach space $(\mathbb{R}, |\cdot|)$. Let us introduce

$$g(x) = x - \frac{f(x)}{f'(x)}. \tag{14.18}$$

The fixed point equation $g(x) = x$ is equivalent to $f(x) = 0$. By the mean-value theorem, there exists ξ such that

$$\big|g(x) - g(y)\big| = |g'(\xi)|\,|x-y| = \left|\frac{f(\xi)f''(\xi)}{f'(\xi)^2}\right| |x-y| \leq \eta |x-y|. \tag{14.19}$$

It follows that the equation has a unique solution, and that the sequence $(x_n)_{n\geq 0}$ converges exponentially fast to it. In fact, it can be checked that convergence is much faster: Using Taylor expansion with remainder, we have

$$\begin{aligned} g(x+a) &= x + a - \frac{f(x+a)}{f'(x+a)} \\ &= x + a - \frac{f(x) + a f'(x) + \frac12 a^2 f''(\xi_1)}{f'(x) + a f''(\xi_2)} \\ &= x + a \frac{f'(x) + a f''(\xi_2)}{f'(x) + a f''(\xi_2)} - a \frac{f'(x) + \frac12 a f''(\xi_1)}{f'(x) + a f''(\xi_2)} \\ &= x + a^2 \frac{f''(\xi_2) - \frac12 f''(\xi_1)}{f'(x) + a f''(\xi_2)}. \end{aligned} \tag{14.20}$$

Here, ξ_1, ξ_2 are close to x, $|\xi_{1,2} - x| < |a|$. Let

$$C = \sup_{|y_1|,|y_2|<|a|} \frac{f''(x+y_2) - \frac12 f''(x+y_1)}{f'(x) + a f''(x+y_2)}. \tag{14.21}$$

Then $|g(x+a) - x| \leq C a^2$. Let us now look at the sequence (x_n) defined by $x_{n+1} = g(x_n)$, and let $a_n = |x - x_n|$. We have just shown that $a_{n+1} \leq C a_n^2$. Let us now assume that $a_0 \leq C^{-1}\varepsilon$ for some $\varepsilon < 1$ (since $x_n \to x$, we can choose n large enough so that $|x_n - x| < C^{-1}\varepsilon$, and re-start the sequence here). It is easy to prove by induction that

$$a_n < C^{-1} \varepsilon^{2^n} \tag{14.22}$$

for all $n \geq 0$. Indeed, this is true for $n = 0$, and

$$a_{n+1} \leq C \Big(C^{-1} \varepsilon^{2^n} \Big)^2 = C^{-1} \varepsilon^{2^{n+1}}. \tag{14.23}$$

Notice that ε can be taken to be arbitrarily small! Taking $\varepsilon = \frac{1}{10}$, convergence is so fast that the number of correct digits *doubles* at every step!

Jacobi algorithm. Our last application of the contraction mapping theorem is a method that allows to compute the solution to the linear algebra equation

$$Ax = b, \tag{14.24}$$

where x, b are vectors in $\mathbb{R}^n$, and A is an $n \times n$ matrix. Assuming that A is invertible, the solution is $x = A^{-1}b$, but computation of A^{-1} is tricky. Jacobi algorithm consists of decomposing $A = D + R$, where D contains the diagonal of A and R the off-diagonal terms:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & \\ \vdots & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix} + \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{pmatrix}. \tag{14.25}$$

We assume that all diagonal terms differ from zero. The matrix D^{-1} is easy to compute, since

$$D^{-1} = \begin{pmatrix} a_{11}^{-1} & 0 & \cdots & 0 \\ 0 & a_{22}^{-1} & & \\ \vdots & & \ddots & \\ 0 & & & a_{nn}^{-1} \end{pmatrix}. \tag{14.26}$$

We now introduce the map $f : \mathbb{R}^n \to \mathbb{R}^n$ by

$$f(x) = D^{-1}(b - Rx). \tag{14.27}$$

The fixed point equation $f(x) = x$ amounts to $Dx = b - R(x)$, which is equivalent to $(D+R)x = b$, i.e. $Ax = b$. Let $\|\cdot\|$ be any norm on $\mathbb{R}^n$; recall that it induces the operator norm on matrices $\|M\| = \sup_{\|x\|=1} \|Mx\|$. Then we have

$$\|f(x) - f(y)\| = \|D^{-1}R(x-y)\| \leq \|D^{-1}R\|\,\|x-y\|. \tag{14.28}$$

We make the new assumption $\|D^{-1}R\| < 1$ so that f is a contraction. This allows to calculate the first terms of the inductive sequence $x_{n+1} = f(x_n)$, which gives a good approximation to the solution of Eq. (14.24).

The method works for matrices A such that $\|D^{-1}R\| < 1$; it may help to choose the norm wisely, so as to make $\|D^{-1}R\|$ as small as possible.

Index