Chapter 1

# Reminders: convergence of sequences and series

This first chapter, which is quite elementary, is essentially a survey of the notion of convergence of sequences and series. Readers who are very confortable with this concept may start reading the next chapter.

However, although the mathematical objects we discuss are well known in principle, they have some unexpected properties. We will see in particular that the order of summation may be crucial to the evaluation of the series, so that changing the order of summation may well change its sum.

We start this chapter by discussing two physical problems in which a limit process is *hidden*. Each leads to an apparent paradox, which can only be resolved when the underlying limit is explicitly brought to light.

## 1.1 The problem of limits in physics

### 1.1.a Two paradoxes involving kinetic energy

#### First paradox

Consider a truck with mass $m$ driving at constant speed $v = 60$ mph on a perfectly straight stretch of highway (think of Montana). We assume that, friction being negligible, *the truck uses no gas to remain at constant speed.*

On the other hand, to accelerate, it must use (or guzzle) $\ell$ gallons of gas in order to increase its kinetic energy by an amount of 1 joule. This assumption, although it is imperfect, is physically acceptable because each gallon of gas yields the same amount of energy.

So, when the driver decides to increase its speed to reach $v' = 80$ mph, the quantity of gas required to do so is equal to the difference of kinetic energy, namely, it is

$$\ell(E_c' - E_c) = \frac{1}{2}\ell m(v'^2 - v^2) = \frac{1}{2}\ell m(6\,400 - 3\,600) = 1\,400 \times \ell m.$$

With $\ell \cdot m = \frac{1}{10\,000}\,\mathrm{J} \cdot \mathrm{mile}^{-2} \cdot \mathrm{h}^2$, say, this amounts to 0.14 gallon. Jolly good.

Now, let us watch the same scene of the truck accelerating, as observed by a highway patrolman, initially driving as fast as the truck $w = v = 60$ mph, but with a motorcycle which is unable to go faster.

The patrolman, having his college physics classes at his fingertips, argues as follows: "in my own galilean reference frame, the relative speed of the truck was previously $v^* = 0$ and is now $v^{*\prime} = 20$ mph. To do this, the amount of gas it has guzzled is equal to the difference in kinetic energies:

$$\ell(E_c^{*\prime} - E_c^*) = \frac{1}{2}\ell m\Big((v^{*\prime})^2 - (v^*)^2\Big) = \frac{1}{2}\ell m(400 - 0) = 200 \times \ell m,$$

or around 0.02 gallons."

*There is here a clear problem*, and one of the two observers must be wrong. Indeed, the galilean relativity principle states that all galilean reference frames are equivalent, and computing kinetic energy in the patrolman's reference frame is perfectly legitimate.

How is this paradox resolved?

We will come to the solution, but first here is another problem. The reader, before going on to read the solutions, is earnestly invited to think and try to solve the problem by herself.

### Second paradox

Consider a highly elastic rubber ball in free fall as we first see it. At some point, it hits the ground, and we assume that *this is an elastic shock*.

Most high-school level books will describe the following argument: "assume that, at the instant $t = 0$ when the ball hits the ground, the speed of the ball is $v_1 = -10\ \mathrm{m{\cdot}s^{-1}}$. Since the shock is elastic, there is conservation of total energy before and after. Hence the speed of the ball after the rebound is $v_2 = -v_1$, or simply $+10\ \mathrm{m{\cdot}s^{-1}}$ going up."

This looks convincing enough. But it is not so impressive if seen from the point of view of an observer who is also moving down at constant speed $v_{\text{obs}} = v_1 = -10\ \mathrm{m{\cdot}s^{-1}}$. For this observer, the speed of the ball before the shock is $v_1^* = v_1 - v_{\text{obs}} = 0\ \mathrm{m{\cdot}s^{-1}}$, so it has zero kinetic energy. However, after rebounding, the speed of the ball is $v_2^* = v_2 - v_{\text{obs}} = 20\ \mathrm{m{\cdot}s^{-1}}$, and therefore it has nonzero kinetic energy! With the analogue of the reasoning above, one should still have found $v_2^* = v_1^* = 0$ (should the ball go through the ground?)

So there is something fishy in this argument also. It is important to remember that *the fact that the right answer is found in the first case does not imply that the argument that leads to the answer is itself correct.*

Readers who have solved the first paradox will find no difficulty in this second one.

## Paradoxes resolved

Kinetic energy is of course not the same in every reference frame. But this is not so much the kinetic energy we are interested in; rather, we want the difference before and after the event described.

Let's go back to elementary mechanics. What happens, in two distinct reference frames, to a system of $N$ solid bodies with initial speed $\boldsymbol{v}_i$ ($i = 1, \ldots, N$) and final speed $\boldsymbol{v}'_i$ after some shock?

In the first reference frame, the difference of kinetic energy is given by

$$\Delta E_c = \sum_{i=1}^{N} m_i (\boldsymbol{v}_i'^2 - \boldsymbol{v}_i^2).$$

In a second reference frame, with relative speed $\boldsymbol{w}$ with respect to the first, the difference is equal to

$$\begin{aligned}\Delta E_c^* &= \sum_{i=1}^{N} m_i \Big( (\boldsymbol{v}'_i - \boldsymbol{w})^2 - (\boldsymbol{v}_i - \boldsymbol{w})^2 \Big) \\ &= \sum_{i=1}^{N} m_i (\boldsymbol{v}_i'^2 - \boldsymbol{v}_i{}^2) - 2\boldsymbol{w} \cdot \left( \sum_{i=1}^{N} m_i (\boldsymbol{v}'_i - \boldsymbol{v}_i) \right) = \Delta E_c - 2\boldsymbol{w} \cdot \Delta \boldsymbol{P},\end{aligned}$$

(we use $*$ as exponents for any physical quantity expressed in the new reference frame), so that $\Delta E_c^* = \Delta E_c$ as long as *the total momentum is preserved* during the shock, in other words if $\Delta \boldsymbol{P} = \boldsymbol{0}$.

In the case of the truck and the patrolman, we did not really take the momentum into account. In fact, the truck can accelerate *because it "pushes back" the whole earth behind it!*

So, let us take up the computation with a terrestrial mass $M$, which is large but not infinite. We will take the limit $[M \to \infty]$ at the very end of the computation, and more precisely, we will let $[M/m \to \infty]$.

At the beginning of the "experiment," in the terrestrial reference frame, the speed of the truck is $\boldsymbol{v}$. At the end of the experiment, the speed is $\boldsymbol{v}'$. Earth, on the other hand, has original speed $\boldsymbol{V} = \boldsymbol{0}$, and if one remains in the same galilean reference frame, final speed $\boldsymbol{V}' = \frac{m}{M}(\boldsymbol{v} - \boldsymbol{v}')$ (because of conservation of total momentum).[1] The kinetic energy of the system at the beginning is then $\frac{1}{2}m\boldsymbol{v}^2$ and at the end it is $\frac{1}{2}m\boldsymbol{v}'^2 + \frac{1}{2}M\boldsymbol{V}'^2$. So, the difference is given by

$$\Delta E_c = \frac{1}{2}m(\boldsymbol{v}'^2 - \boldsymbol{v}^2) + \frac{1}{2}\frac{m^2}{M}(\boldsymbol{v} - \boldsymbol{v}')^2 = \frac{1}{2}m(\boldsymbol{v}'^2 - \boldsymbol{v}^2)\left[1 + \mathrm{O}\left(\frac{m}{M}\right)\right].$$

This is the amount of gas involved! So we see that, up to negligible terms, the first argument gives the right answer, namely, 0.14 gallons.

We now come back to the patrolman's frame, moving with speed $\boldsymbol{w}$ with respect to the terrestrial frame. The initial speed of the truck is $\boldsymbol{v}^* = \boldsymbol{v} - \boldsymbol{w}$, and the final speed is $\boldsymbol{v}'^* = \boldsymbol{v}' - \boldsymbol{w}$. The Earth has initial speed $\boldsymbol{V}^* = -\boldsymbol{w}$ and final speed $\boldsymbol{V}'^* = -\boldsymbol{w} + \frac{m}{M}{}'(\boldsymbol{v} - \boldsymbol{v}')$. The difference is now:

$$\begin{aligned}
\Delta E_c^* &= \frac{1}{2}m\left(\boldsymbol{v}'^{*2} - \boldsymbol{v}^{*2}\right) + \frac{1}{2}M\left(\boldsymbol{V}'^{*2} - \boldsymbol{V}^{*2}\right) \\
&= \frac{1}{2}m(\boldsymbol{v}' - \boldsymbol{w})^2 - \frac{1}{2}m(\boldsymbol{v} - \boldsymbol{w})^2 + \frac{1}{2}M\left[\frac{m}{M}(\boldsymbol{v} - \boldsymbol{v}') - \boldsymbol{w}\right]^2 - \frac{1}{2}M\boldsymbol{w}^2 \\
&= \frac{1}{2}m\boldsymbol{v}'^2 - \frac{1}{2}m\boldsymbol{v}^2 + m(\boldsymbol{v} - \boldsymbol{v}')\cdot\boldsymbol{w} + \frac{1}{2}\frac{m^2}{M}(\boldsymbol{v} - \boldsymbol{v}')^2 - m(\boldsymbol{v} - \boldsymbol{v}')\cdot\boldsymbol{w}, \\
\Delta E_c^* &= \Delta E_c.
\end{aligned}$$

Hence the difference of kinetic energy is preserved, as we expected. So even in this other reference frame, a correct computation shows that the quantity of gas involved is the same as before.

The patrolman's mistake was to forget the *positive* term $-m(\boldsymbol{v} - \boldsymbol{v}')\cdot\boldsymbol{w}$, corresponding to the difference of kinetic energy of the Earth in its galilean frame. This term does not tend to 0 as $[M/m \to \infty]$ !

*From the point of view of the patrolman's frame, 0.02 gallons are needed to accelerate the truck, and the remaining 0.12 gallons are needed to accelerate the Earth!*

We can summarize this as a table, where T is the truck and E is the Earth.

| | Initial speed | Final speed | $E_c$ init. | $E_c$ final | $\Delta E_c$ |
|---|---|---|---|---|---|
| T | $\boldsymbol{v}$ | $\boldsymbol{v}'$ | $\frac{1}{2}m\boldsymbol{v}^2$ | $\frac{1}{2}m\boldsymbol{v}'^2$ | $\frac{m}{2}(\boldsymbol{v}'^2 - \boldsymbol{v}^2)$ |
| E | $\boldsymbol{0}$ | $\frac{m}{M}(\boldsymbol{v} - \boldsymbol{v}')$ | 0 | $\frac{1}{2}\frac{m^2}{M}(\boldsymbol{v} - \boldsymbol{v}')^2$ | $+\frac{m^2}{2M}(\boldsymbol{v} - \boldsymbol{v}')^2$ |
| T* | $\boldsymbol{v} - \boldsymbol{w}$ | $\boldsymbol{v}' - \boldsymbol{w}$ | $\frac{1}{2}m(\boldsymbol{v} - \boldsymbol{w})^2$ | $\frac{1}{2}m(\boldsymbol{v}' - \boldsymbol{w})^2$ | $\frac{m}{2}(\boldsymbol{v}'^2 - \boldsymbol{v}^2)$ |
| E* | $-\boldsymbol{w}$ | $\frac{m}{M}(\boldsymbol{v} - \boldsymbol{v}') - \boldsymbol{w}$ | $\frac{1}{2}M\boldsymbol{w}^2$ | $\frac{M}{2}\left(\frac{m}{M}(\boldsymbol{v} - \boldsymbol{v}') - \boldsymbol{w}\right)^2$ | $+\frac{m^2}{2M}(\boldsymbol{v} - \boldsymbol{v}')^2$ |

[1] Note that the terrestrial reference frame is then not galilean, since the Earth started "moving" under the truck's impulsion.

(Illustration © Craig THOMPSON 2001)

Fig. 1.1 – Romeo, Juliet, and the boat on a lake.

The second paradox is resolved in the same manner: the Earth's rebound energy must be taken into account after the shock with the ball.

The interested reader will find another paradox, relating to optics, in Exercise 1.3 on page 43.

### 1.1.b Romeo, Juliet, and viscous fluids

Here is an example in mechanics where a function $f(x)$ is defined on $[0, +\infty[$, but $\lim_{x\to 0^+} f(x) \neq f(0)$.

Let us think of a summer afternoon, which Romeo and Juliet have dedicated to a pleasant boat outing on a beautiful lake. They are sitting on each side of their small boat, immobile over the waters. Since the atmosphere is conducive to charming murmurs, Romeo decides to go sit by Juliet.

Denote by $M$ the mass of the boat and Juliet together, $m$ that of Romeo, and $L$ the length of the walk from one side of the boat to the other (see Figures 1.1 and 1.2).

Two cases may be considered: one where the friction of the boat on the lake is negligible (a perfect fluid), and one where it is given by the formula $f = -\eta v$, where $f$ is the force exerted by the lake on the boat, $v$ the speed of the boat on the water, and $\eta$ a viscosity coefficient. We consider the problem only on the horizontal axis, so it is *one-dimensional.*

We want to compute how far the boat moves

1. in the case $\eta = 0$;

2. in the case $\eta \neq 0$.

Let $\ell$ be this distance.

The first case is very easy. Since no force is exerted in the horizontal plane on the system "boat + Romeo + Juliet," the center of gravity of this system

(Illustration © Craig THOMPSON 2001)

Fig. 1.2 – Romeo moved closer.

does not move during the experiment. Since Romeo travels the distance $L$ relative to the boat, it is easy to deduce that the boat must cover, in the opposite direction, the distance

$$\ell = \frac{m}{m+M} L.$$

In the second case, let $x(t)$ denote the positive of the boat and $y(t)$ that of Romeo, relative to the Earth, not to the boat. The equation of movement for the center of gravity of the system is

$$M\ddot{x} + m\ddot{y} = -\eta\dot{x}.$$

We now integrate on both sides between $t = 0$ (before Romeo starts moving) and $t = +\infty$. Because of the friction, we know that as $[t \to +\infty]$, the speed of the boat goes to 0 (hence also the speed of Romeo, since he will have been long immobile with respect to the boat). Hence we have

$$(M\dot{x} + m\dot{y})\Big|_0^{+\infty} = 0 = -\eta\big(x(+\infty) - x(0)\big)$$

or $\eta\ell = 0$. Since $\eta \neq 0$, we have $\ell = 0$, whichever way Romeo moved to the other side. In partcular, if we take the limit when $\eta \to 0$, hoping to obtain the nonviscous case, we have:

$$\lim_{\substack{\eta\to 0\\ \eta>0}} \ell(\eta) = 0 \qquad \text{hence} \qquad \lim_{\substack{\eta\to 0\\ \eta>0}} \ell(\eta) \neq \ell(0).$$

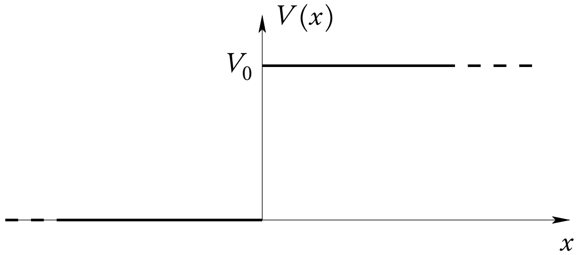



Fig. 1.3 — Potential wall $V(x) = V_0 H(x)$.

Conclusion: *The limit of a viscous fluid to a perfect fluid is singular.* It is not possible to formally take the limit when the viscosity tends to zero to obtain the situation for a perfect fluid. In particular, it is easier to model flows of nonviscous perfect fluids by "real" fluids which have large viscosity, because of turbulence phenomena which are more likely to intervene in fluids with small viscosity. The interested reader can look, for instance, at the book by Guyon, Hulin, and Petit [44].

**Remark 1.1** The exact form $f = -\eta v$ of the friction term is crucial in this argument. If the force involves additional (nonlinear) terms, the result is completely different. Hence, if you try to perform this experiment in practice, it will probably not be conclusive, and the boat is not likely to come back to the same exact spot at the end.

### 1.1.c Potential wall in quantum mechanics

In this next physical example, there will again be a situation where we have a limit $\lim_{x\to 0} f(x) \neq f(0)$; however, the singularity arises here in fact because of a *second* variable, and the true problem is that we have a double limit which *does not commute*: $\lim_{x\to 0}\lim_{y\to 0} f(x,y) \neq \lim_{y\to 0}\lim_{x\to 0} f(x,y)$.

The problem considered is that of a quantum particle arriving at a potential wall. We look at a one-dimensional setting, with a potential of the type $V(x) = V_0 H(x)$, where $H$ is the Heaviside function, that is, $H(x) = 0$ if $x < 0$ and $H(x) = 1$ if $x > 0$. The graph of this potential is represented in Figure 1.3.

A particle arrives from $x = -\infty$ in an energy state $E > V_0$; part of it is transmitted beyond the potential wall, and part of it is reflected back. We are interested in the reflection coefficient of this wave.

The incoming wave may be expressed, for negative values of $x$, as the sum of a progressive wave moving in the direction of increasing $x$ and a reflected wave. For positive values of $x$, we have a transmitted wave in the direction of increasing $x$, but no component in the other direction. According to the Schrödinger equation, the wave function can therefore be written in the form

$$\varphi(x,t) = \psi(x)\, f(t),$$

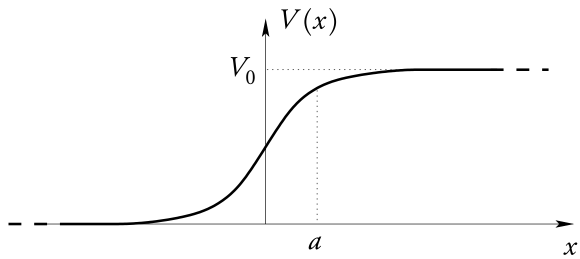



Fig. 1.4 — "Smoothed" potential $V(x) = V_0/(1 + \mathrm{e}^{-x/a})$, with $a > 0$.

where

$$\psi(x) = \begin{cases} \mathrm{e}^{\mathrm{i}kx} + B\,\mathrm{e}^{-\mathrm{i}kx} & \text{if } x < 0, \text{ with } k \stackrel{\text{def}}{=} \frac{\sqrt{2mE}}{\hbar}, \\ A\,\mathrm{e}^{\mathrm{i}k'x} & \text{if } x > 0, \text{ with } k' \stackrel{\text{def}}{=} \frac{\sqrt{2m(E-V_0)}}{\hbar}. \end{cases}$$

The function $f(t)$ is only a time-related phase factor and plays no role in what follows. The reflection coefficient of the wave is given by the ratio of the currents associated to $\psi$ and is given by $R = 1 - \frac{k'}{k}\,|A|^2$ (see [20, 58]). There remains to find the value of $A$. To find it, it suffices to write the equation expressing the continuity of $\psi$ and $\psi'$ at $x = 0$, Since $\psi(0^+) = \psi(0^-)$, we have $1 + B = A$. And since $\psi'(0^+) = \psi'(0^-)$, we have $k(1 - B) = k'A$, and we deduce that $A = 2k/(k + k')$. The reflection coefficient is therefore equal to

$$R = 1 - \frac{k'}{k}\,|A|^2 = \left(\frac{k - k'}{k + k'}\right) = \left(\frac{\sqrt{E} - \sqrt{E - V_0}}{\sqrt{E} + \sqrt{E - V_0}}\right)^2. \qquad \text{(1.1)}$$

Here comes the surprise: this expression (1.1) is independant of $\hbar$. In particular, the limit as $[\hbar \to 0]$ (which defines the "classical limit") yields a nonzero reflection coefficient, although we know that in classical mechanics a particle with energy $E$ does not reflect against a potential wall with value $V_0 < E$![2] So, displaying explicitly the dependency of $R$ on $\hbar$, we have:

$$\lim_{\substack{\hbar \to 0 \\ \hbar \neq 0}} R(\hbar) \neq 0 = R(0).$$

In fact, we have gone a bit too fast. We take into account the physical aspects of this story: the "classical limit" is certainly not the same as brutally writing "$\hbar \to 0$." Since Planck's constant is, as the name indicates, just a constant, this makes no sense. To take the limit $\hbar \to 0$ means that one arranges for the quantum dimensions of the system to be much smaller than all other dimensions. Here the quantum dimension is determined by the de

[2] The particle goes through the obstacle with probability 1.

Broglie wavelength of the particle, that is, $\lambda = \hbar/p$. What are the other lengths in this problem? Well, there are none! At least, the way we phrased it: because in fact, expressing the potential by means of the Heaviside function is rather *cavalier*. In reality, the potential must be continuous. We can replace it by an infinitely differentiable potential such as $V(x) = V_0/(1+\mathrm{e}^{-x/a})$, which increases, roughly speaking, on an interval of size $a > 0$ (see Figure 1.4). In the limit where $a \to 0$, the discontinuous Heaviside potential reappears.

Computing the reflection coefficient with this potential is done similarly, but of course the computations are more involved. We refer the reader to [58, chapter 25]. At the end of the day, the reflection coefficient is found to depend not only on $\hbar$, but also on $a$, and is given by

$$R(\hbar, a) = \left(\frac{\sinh a\pi(k - k')}{\sinh a\pi(k + k')}\right)^2.$$

($\hbar$ appears in the definition of $k = \sqrt{2mE}/\hbar$ and $k' = \sqrt{2m(E - V_0)}/\hbar$.) We then see clearly that *for fixed nonzero* $a$, the de Broglie wavelength of the particle may become infinitely small compared to $a$, and this defines the correct classical limit. Mathematically, we have

$$\forall a \neq 0 \qquad \lim_{\substack{\hbar\to 0\\ \hbar\neq 0}} R(\hbar, a) = 0 \qquad\qquad \text{classical limit}$$

On the other hand, if we keep $\hbar$ fixed and let $a$ to to 0, we are converging to the Heaviside potential and we find that

$$\forall \hbar \neq 0 \qquad \lim_{\substack{a\to 0\\ a\neq 0}} R(\hbar, a) = \left(\frac{k - k'}{k + k'}\right)^2 = R(\hbar, 0).$$

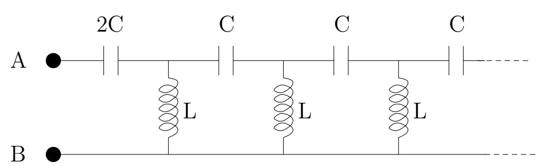
So the two limits $[\hbar \to 0]$ and $[a \to 0]$ cannot be taken in an arbitrary order:

$$\lim_{\substack{\hbar\to 0\\ \hbar\neq 0}} \lim_{\substack{a\to 0\\ a\neq 0}} R(\hbar, a) = \left(\frac{k - k'}{k + k'}\right)^2 \qquad \text{but} \qquad \lim_{\substack{a\to 0\\ a\neq 0}} \lim_{\substack{\hbar\to 0\\ \hbar\neq 0}} R(\hbar, a) = 0.$$

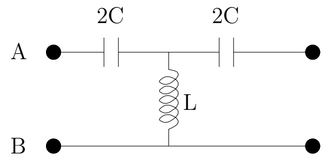
To speak of $R(0, 0)$ has *a priori* no physical sense.

### 1.1.d Semi-infinite filter behaving as waveguide

We consider the cicuit $AB$,




made up of a cascading sequence of "T" cells ($2C$, $L$, $2C$),




(the capacitors of two successive cells in series are equivalent with one capacitor with capacitance $C$). We want to know the total impedance of this circuit.

First, consider instead a circuit made with a finite sequence of $n$ elementary cells, and let $Z_n$ denote its impedance. Kirchhoff's laws imply the following recurrence relation between the values of $Z_n$:

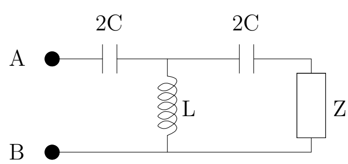
$$Z_{n+1} = \frac{1}{2\mathrm{i}C\omega} + \frac{\mathrm{i}L\omega\left(\dfrac{1}{2\mathrm{i}C\omega} + Z_n\right)}{\mathrm{i}L\omega + \dfrac{1}{2\mathrm{i}C\omega} + Z_n}, \tag{1.2}$$

where $\omega$ is the angular frequency. In particular, note that if $Z_n$ is purely imaginary, then so is $Z_{n+1}$. Since $Z_1$ is purely imaginary, it follows that

$$Z_n \in \mathrm{i}\mathbb{R} \qquad \text{for all } n \in \mathbb{N}.$$

We don't know if the sequence $(Z_n)_{n\in\mathbb{N}}$ converges. But one thing is certain: if this sequence $(Z_n)_{n\in\mathbb{N}}$ converges to some limit, this must be purely imaginary (the only possible real limit is zero).

Now, we compute the impedance of the infinite circuit, noting that this circuit $AB$ is *strictly* equivalent to the following:




Hence we obtain an equation involving $Z$:

$$Z = \frac{1}{2\mathrm{i}C\omega} + \frac{\mathrm{i}L\omega\left(\dfrac{1}{2\mathrm{i}C\omega} + Z\right)}{\mathrm{i}L\omega + \dfrac{1}{2\mathrm{i}C\omega} + Z}. \tag{1.3}$$

Some computations yield a second-degree equation, with solutions given by

$$Z^2 = \frac{1}{C}\cdot\left(L - \frac{1}{4C\omega^2}\right).$$

We must therefore distinguish two cases:

• **If $\omega < \omega_c = \dfrac{1}{2\sqrt{LC}}$**, we have $Z^2 < 0$ and hence $Z$ is purely imaginary of the form

$$Z = \pm \mathrm{i}\sqrt{\frac{1}{4C^2\omega^2} - \frac{L}{C}}.$$

**Remark 1.2** Mathematically, there is nothing more that can be said, and in particular there remains an uncertainty concerning the "sign" of $Z$.

However, this can also be determined by a physical argument: let $\omega$ tend to 0 (continuous regime). Then we have

$$Z(\omega) \underset{\omega\to 0^+}{\sim} \pm\frac{\mathrm{i}}{2C\omega},$$

the modulus of which tends to infinity. This was to be expected: the equivalent circuit is open, and the first capacitor "cuts" the circuit. Physically, it is then natural to expect that, the first coil acting as a plain wire, the first capacitor will be dominant. Then

$$Z(\omega) \underset{\omega\to 0^+}{\sim} -\frac{\mathrm{i}}{2C\omega}$$

(corresponding to the behavior of a single capacitor).

Thus the physically acceptable solution of the equation (1.3) is

$$Z = -\mathrm{i}\sqrt{\frac{1}{4C^2\omega^2} - \frac{L}{C}}.$$

• **If $\omega > \omega_c = \dfrac{1}{2\sqrt{LC}}$**, then $Z^2 > 0$ and $Z$ is therefore *real*:

$$Z = \pm\sqrt{\frac{L}{C} - \frac{1}{4C^2\omega^2}}.$$

**Remark 1.3** Here also the sign of $Z$ can be determined by physical arguments. The real part of an impedance (the "resistive part") is always *non-negative* in the case of a passive component, since it accounts for the dissipation of energy by the Joule effect. Only active components (such as operational amplifiers) can have negative resistance. Thus, the physically acceptable solution of equation (1.3) is

$$Z = +\sqrt{\frac{L}{C} - \frac{1}{4C^2\omega^2}}.$$

In this last case, there seems to be a paradox since $Z$ cannot be the limit as $n \to +\infty$ of $(Z_n)_{n\in\mathbb{N}}$. Let's look at this more closely.

*From the mathematical point of view*, Equation (1.3) expresses nothing but the fact that $Z$ is a "fixed point" for the induction relation (1.2). In other words, this is the equation we would have obtained from (1.2), by continuity, *if* we had known that the sequence $(Z_n)_{n\in\mathbb{N}}$ converges to a limit $Z$. However, there is no reason for the sequence $(Z_n)_{n\in\mathbb{N}}$ to converge.

**Remark 1.4** *From the physical point of view*, the behavior of this infinite chain is rather surprising. How does resistive behavior arise from purely inductive or capacitative components? Where does energy dissipate? And where does it go?

In fact, there is no dissipation of energy in the sense of the Joule effect, but energy does disappear from the point of view of an operator "holding points $A$ and $B$." More precisely, one can show that there is a flow of energy propagating from cell to cell. So at the beginning of the circuit, it looks like there is an "energy well" with fixed power consumption. Still, no energy disappears: an infinite chain can consume energy without accumulating it anywhere.[3] In the regime considered, this infinite chain corresponds to a waveguide.

We conclude this first section with a list of other physical situations where the problem of noncommuting limits arises:

- taking the "classical" (nonquantum) limit, as we have seen, is by no means a trivial matter; in addition, it may be in conflict with a "nonrelativistic" limit (see, e.g., [6]), or with a "low temperature" limit;
- in plasma thermodynamics, the limit of infinite volume ($V \to \infty$) and the nonrelativistic limit ($c \to \infty$) are incompatible with the thermodynamic limit, since a characteristic time of return to equilibrium is $V^{1/3}/c$;
- in the classical theory of the electron, it is often reproached that such a classical electron, with zero naked mass, rotates too fast at the level of the equator (200 times the speed of light) for its magnetic moment and renormalized mass to conform to experimental data. A more careful calculation by Lorentz himself[4] gave about 10 times the speed of light at the equateur. But in fact, the limit $[m \to 0^+]$ requires care, and if done correctly, it *imposes* a limit $[v/c \to 1^-]$ to maintain a constant renormalized mass [7];
- another interesting example is an "infinite universe" limit and a "diluted universe" limit [56].

## 1.2 Sequences

### 1.2.a Sequences in a normed vector space

We consider in this section a normed vector space $(E, \|\cdot\|)$ and sequences of elements of $E$.[5]

[3] This is the principle of Hilbert's infinite hotel.
[4] Pointed out by Sin-Itiro Tomonaga [91].
[5] We recall the basic definitions concerning normed vector spaces in Appendix A.

**DEFINITION 1.5 (Convergence of a sequence)** Let $(E, \|\cdot\|)$ be a normed vector space and $(u_n)_{n\in\mathbb{N}}$ a sequence of elements of $E$, and let $\ell \in E$. The sequence $(u_n)_{n\in\mathbb{N}}$ **converges to $\ell$** if, for any $\varepsilon > 0$, there exists an index starting from which $u_n$ is at most at distance $\varepsilon$ from $\ell$:

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \qquad n \geqslant N \Longrightarrow \|u_n - \ell\| < \varepsilon.$$

Then $\ell$ is called the **limit of the sequence $(u_n)_{n\in\mathbb{N}}$**, and this is denoted

$$\ell = \lim_{n\to\infty} u_n \qquad \text{or} \qquad u_n \xrightarrow[n\to\infty]{} \ell.$$

**DEFINITION 1.6** A sequence $(u_n)_{n\in\mathbb{N}}$ of real numbers **converges to $+\infty$** (resp. **to $-\infty$**) if, for any $M \in \mathbb{R}$, there exists an index $N$, starting from which all elements of the sequence are larger than $M$:

$$\forall M \in \mathbb{R} \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \qquad n \geqslant N \Longrightarrow u_n > M \qquad (\text{resp. } u_n < M).$$

In the case of a complex-valued sequence, a type of convergence to infinity, in modulus, still exists:

**DEFINITION 1.7** A sequence $(z_n)_{n\in\mathbb{N}}$ of complex numbers **converges to infinity** if, for any $M \in \mathbb{R}$, there exists an index $N$, starting from which all elements of the sequence have modulus larger than $M$:

$$\forall M \in \mathbb{R} \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \qquad n \geqslant N \Longrightarrow |z_n| > M.$$

**Remark 1.8** There is only "one direction to infinity" in $\mathbb{C}$. We will see a geometric interpretation of this fact in Section 5.4 on page 146.

**Remark 1.9** The strict inequalities $\|u_n - \ell\| < \varepsilon$ (or $|u_n| > M$) in the definitions above (which, in more abstract language, amount to an emphasis on *open* subsets) may be replaced by $\|u_n - \ell\| \leqslant \varepsilon$, which are sometimes easier to handle. Because $\varepsilon > 0$ is arbitrary, this gives an equivalent definition of convergence.

## 1.2.b Cauchy sequences

It is often important to show that a sequence converges, without explicitly knowing the limit. Since the definition of convergence depends on the limit $\ell$, it is not conveninent for this purpose.[6] In this case, the most common tool is the *Cauchy criterion*, which depends on the convergence "of elements of the sequence with respect to each other":

---

[6] As an exemple: how should one prove that the sequence $(u_n)_{n\in\mathbb{N}}$, with

$$u_n = \sum_{p=1}^{n} \frac{(-1)^{p+1}}{p^4},$$

converges? Probably not by guessing that the limit is $7\pi^4/720$.

**DEFINITION 1.10 (Cauchy criterion)** A sequence $(u_n)_{n\in\mathbb{N}}$ in a normed vector space **is a Cauchy sequence**, or **satisfies the Cauchy criterion**, if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall p, q \in \mathbb{N} \qquad q > p \geqslant N \Longrightarrow \left\| u_p - u_q \right\| < \varepsilon$$

or, equivalently, if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall p, k \in \mathbb{N} \qquad p \geqslant N \Longrightarrow \left\| u_{p+k} - u_p \right\| < \varepsilon.$$

A common technique used to prove that a sequence $(u_n)_{n\in\mathbb{N}}$ is a Cauchy sequence is therefore to find a sequence $(\alpha_p)_{p\in\mathbb{N}}$ of real numbers such that

$$\lim_{p\to\infty} \alpha_p = 0 \qquad \text{and} \qquad \forall p, k \in \mathbb{N} \quad \left\| u_{p+k} - u_p \right\| \leqslant \alpha_p.$$

**PROPOSITION 1.11** *Any convergent sequence is a Cauchy sequence.*

This is a trivial consequence of the definitions. But we are of course interested in the converse. Starting from the Caucy criterion, we want to be able to conclude that a sequence converges – without, in particular, requiring the limit to be known beforehand. However, that is not always possible: there exist normed vector spaces $E$ and Cauchy sequences in $E$ which do not converge.

**Example 1.12** Consider the set of rational numbers $\mathbb{Q}$. With the absolute value, it is a normed $\mathbb{Q}$-vector space. Consider then the sequence

$$u_0 = 3 \quad u_1 = 3.1 \quad u_2 = 3.14 \quad u_3 = 3.141 \quad u_4 = 3.1415 \quad u_5 = 3.14159\cdots$$

(you can guess the rest[7]...). This is a sequence of rationals, which is a Cauchy sequence (the distance between $u_p$ and $u_{p+k}$ is at most $10^{-p}$). However, it does not converge in $\mathbb{Q}$, since its limit (in $\mathbb{R}$!) is $\pi$, which is a notoriously irrational number.

The space $\mathbb{Q}$ is not "nice" in the sense that it leaves a lot of room for Cauchy sequences to exist without converging *in* $\mathbb{Q}$. The mathematical terminology is that $\mathbb{Q}$ *is not complete*.

**DEFINITION 1.13 (Complete vector space)** A normed vector space $(E, \|\cdot\|)$ is **complete** if all Cauchy sequences in $E$ are convergent.

**THEOREM 1.14** *The spaces $\mathbb{R}$ and $\mathbb{C}$, and more generally all finite-dimensional real or complete normed vector spaces, are complete.*

Proof

**First case:** It is first very simple to show that a Cauchy sequence $(u_n)_{n\in\mathbb{N}}$ of real numbers is bounded. Hence, according to the Bolzano-Weierstrass theorem (Theorem A.41, page 581), it has a convergent subsequence. But any Cauchy sequence which has a convergent subsequence is itself convergent (its limit being that of the subsequence), see Exercise 1.6 on page 43. Hence any Cauchy sequence in $\mathbb{R}$ is convergent.

**Second case:** Considering $\mathbb{C}$ as a normed real vector space of dimension 2, we can suppose that the base field is $\mathbb{R}$.

[7] This is simply $u_n = 10^{-n} \cdot \lfloor 10^n \pi \rfloor$ where $\lfloor \cdot \rfloor$ is the integral part function.

Consider a basis $\mathcal{B} = (b_1, \ldots, b_d)$ of the vector space $E$. Then we deduce that $E$ is complete from the case of the real numbers and the following two facts: (1) a sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ of vectors, with coordinates $(x_n^1, \ldots, x_n^d)$ in $\mathcal{B}$, converges in $E$ if and only if each coordinate sequence $(x^k{}_n)_{n\in\mathbb{N}}$; and (2), if a sequence is a Cauchy sequence, then each coordinate is a Cauchy sequence.

Both facts can be checked immediately when the norm of $E$ is defined by $\|\boldsymbol{x}\| = \max|x^k|$, and other norms reduce to this case since all norms are equivalent on $E$.

**Example 1.15** The space $L^2$ of square integrable functions (in the sense of Lebesgue), with the norm $\|f\|_{L^2}^2 \stackrel{\text{def}}{=} \int_{\mathbb{R}} |f|^2$, is complete (see Chapter 9). This infinite-dimensional space is used very frequently in quantum mechanics.

**Counterexample 1.16** Let $E = \mathbb{K}[X]$ be the space of polynomials with coefficients in $K$ (and arbitrary degree). Let $P \in E$ be a polynomial, written as $P = \sum \alpha_n X^n$, and define its norm by $\|P\| \stackrel{\text{def}}{=} \max_{i\in\mathbb{N}} |\alpha_i|$. Then the normed vector space $(E, \|\cdot\|)$ is not complete (see Exercise 1.7 on page 43).

Here is an important example of the use of the Cauchy criterion: the fixed point theorem.

## 1.2.c The fixed point theorem

We are looking for solutions to an equation of the type

$$f(x) = x,$$

where $f : E \to E$ is an application defined on a normed vector space $E$, with values in $E$. Any element of $E$ that satisfies this equation is called a **fixed point** of $f$.

**DEFINITION 1.17 (Contraction)** Let $E$ be a normed vector space, $U$ a subset of $E$. A map $f : U \to E$ is a **contraction** if there exists a real number $\rho \in [0, 1[$ such that $\left\|f(y) - f(x)\right\| \leqslant \rho \|y - x\|$ for all $x, y \in U$. In particular, $f$ is continuous on $U$.

**THEOREM 1.18 (Banach fixed point theorem)** *Let $E$ be a complete normed vector space, $U$ a non-empty closed subset of $E$, and $f : U \to U$ a contraction. Then $f$ has a unique fixed point.*

Proof. Chose an arbitrary $u_0 \in U$, and define a sequence $(u_n)_{n\in\mathbb{N}}$ for $n \geqslant 0$ by induction by $u_{n+1} = f(u_n)$. Using the definition of contraction, an easy induction shows that we have

$$\left\|u_{p+1} - u_p\right\| \leqslant \rho^p \|u_1 - u_0\|$$

for any $p \geqslant 0$. Then a second induction on $k \geqslant 0$ shows that for all $p, k \in \mathbb{N}$, we have

$$\left\|u_{p+k} - u_p\right\| \leqslant (\rho^p + \cdots + \rho^{p+k-1}) \cdot \|u_1 - u_0\| \leqslant \frac{\rho^p}{1-\rho} \cdot \|u_1 - u_0\|,$$

and this proves that the sequence $(u_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Since the space $E$ is complete, this sequences has a limit $a \in E$. Since $U$ is closed, we have $a \in U$.

Now from the continuity of $f$ and the relation $u_{n+1} = f(u_n)$, we deduce that $a = f(a)$. So this $a$ is a fixed point of $f$. If $b$ is an arbitrary fixed point, the inequality $\|a-b\| = \|f(a) - f(b)\| \leqslant \rho \|a-b\|$ proves that $\|a-b\| = 0$ and thus $a = b$, showing that the fixed point is unique.

**Remark 1.19** Here is one reason why Banach's theorem is very important. Suppose we have a normed vector space $E$ and a map $g : E \to E$, and we would like to solve an equation $g(x) = b$. This amounts to finding the fixed points of $f(x) = g(x) + x - b$, and we can hope that $f$ may be a contraction, at least locally. This happens, for instance, in the case of the Newton method, if the function used is nice enough, and if a suitable (rough) approximation of a zero is known.

This is an extremely fruitful idea: one can prove this way the Cauchy-Lipschitz theorem concerning existence and unicity of solutions to a large class of differential equations; one can also study the existence of certain fractal sets (the von Koch snowflake, for instance), certain stability problems in dynamical systems, etc.

Not only does it follow from Banach's theorem that certain equations have solutions (and even better, unique solutions!), but the proof provides an *effective* way to find this solution by a successive approximations: it suffices to fix $u_0$ arbitrarily, and to define the sequence $(u_n)_{n\in\mathbb{N}}$ by means of the recurrence formula $u_{n+1} = f(u_n)$; then we know that this sequence converges to the fixed point $a$ of $f$. Moreover, the convergence of the sequence of approximations is exponentially fast: the distance from the approximate solution $u_n$ to the (unknown) solution $a$ decays as fast as $\rho^n$. An example (the *Picard iteration*) is given in detail in Problem 1 on page 46.

### 1.2.d Double sequences

Let $(x_{n,k})_{(n,k)\in\mathbb{N}^2}$ be a double-indexed sequence of elements in a normed vector space $E$. We assume that the sequences made up of each row and each column converge, with limits as follows:

$$\begin{array}{ccccccc}
x_{11} & x_{12} & x_{13} & \cdots & \longrightarrow & A_1 \\
x_{21} & x_{22} & x_{23} & \cdots & \longrightarrow & A_2 \\
x_{31} & x_{32} & x_{33} & \cdots & \longrightarrow & A_3 \\
\vdots & \vdots & \vdots & & & \\
\downarrow & \downarrow & \downarrow & & & \\
B_1 & B_2 & B_3 & & &
\end{array}$$

The question is now whether the sequences $(A_n)_{n\in\mathbb{N}}$ and $(B_k)_{k\in\mathbb{N}}$ themselves converge, and if that is the case, whether their limits are equal. In general, it turns out that the answer is "No." However, under certain conditions, if one sequence (say $(A_n)$) converges, then so does the other, and the limits are the same.

**DEFINITION 1.20** A double sequence $(x_{n,k})_{n,k}$ **converges uniformly with respect to $k$ to a sequence $(B_k)_{k\in\mathbb{N}}$ as $n \to \infty$** if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \forall k \in \mathbb{N} \qquad n \geqslant N \Longrightarrow \left| x_{n,k} - B_k \right| < \varepsilon.$$

In other words, there is convergence with respect to $n$ for fixed $k$, but in such a way that the speed of convergence is *independent* of $k$; or one might say that "all values of $k$ are similarly behaved."

Uniform convergence with respect to $n$ toward the sequence $(A_n)_{n\in\mathbb{N}}$ is similarly defined.

**THEOREM 1.21 (Double limit)** *With notation as above, if the following three conditions hold:*

① *each row converges, and* $A_n = \lim\limits_{k\to\infty} x_{n,k}$ *for all* $n \in \mathbb{N}$,

② *each column converges, and* $B_k = \lim\limits_{n\to\infty} x_{n,k}$ *for all* $k \in \mathbb{N}$,

③ *the convergence is* ***uniform*** *either with respect to n or with respect to k;*

*then the sequences* $(A_n)_{n\in\mathbb{N}}$ *and* $(B_k)_{k\in\mathbb{N}}$ *converge,* $\lim\limits_{n\to\infty} A_n = \lim\limits_{k\to\infty} B_k = \ell$. *One says that* ***the double sequence*** $(x_{n,k})_{n,k}$ ***converges to the limit*** $\ell$.

Be aware that the uniform convergence condition ③ is very important. The following examples gives an illustration: here both limits exist, but they are different.

$$\begin{array}{ccccccc}
1 & 0 & 0 & 0 & \cdots & \to & 0\\
1 & 1 & 0 & 0 & \cdots & \to & 0\\
1 & 1 & 1 & 0 & \cdots & \to & 0\\
1 & 1 & 1 & 1 & \cdots & \to & 0\\
\vdots & \vdots & \vdots & \vdots & & \vdots & \\
\downarrow & \downarrow & \downarrow & \downarrow & & & \\
1 & 1 & 1 & 1 & \cdots & &
\end{array}$$

### 1.2.e Sequential definition of the limit of a function

**DEFINITION 1.22** Let $f : \mathbb{K} \to \mathbb{K}'$ (where $\mathbb{K}, \mathbb{K}' = \mathbb{R}$ or $\mathbb{C}$ or any normed vector space), let $a \in \mathbb{K}$, and let $\ell \in \mathbb{K}'$. Then $f$ **has the limit** $\ell$**, or tends to** $\ell$**, at the point** $a$ if we have

$$\forall \varepsilon > 0 \quad \exists \eta > 0 \quad \forall z \in \mathbb{K} \qquad |z-a| < \eta \Longrightarrow \left|f(z)-\ell\right| < \varepsilon.$$

There are also limits at infinity and infinite limits, defined similarly:

**DEFINITION 1.23** Let $f : \mathbb{K} \to \mathbb{K}'$ (where $\mathbb{K}, \mathbb{K}' = \mathbb{R}$ or $\mathbb{C}$). Let $\ell \in \mathbb{K}'$. Then $f$ **tends to** $\ell$ **at** $+\infty$, resp. **at** $-\infty$ (in the case $\mathbb{K} = \mathbb{R}$), resp. **at infinity** (in the case $\mathbb{K} = \mathbb{C}$), if we have

$$\begin{array}{llllr}
 & \forall \varepsilon > 0 & \exists A \in \mathbb{R} & \forall x \in \mathbb{R} & x > A \Longrightarrow \left|f(x)-\ell\right| < \varepsilon,\\
\big[\text{resp.} & \forall \varepsilon > 0 & \exists A' \in \mathbb{R} & \forall x \in \mathbb{R} & x < A' \Longrightarrow \left|f(x)-\ell\right| < \varepsilon\\
\text{resp.} & \forall \varepsilon > 0 & \exists A \in \mathbb{R}^+ & \forall z \in \mathbb{C} & |z| > A \Longrightarrow \left|f(z)-\ell\right| < \varepsilon\big].
\end{array}$$

Similarly, a function $f : \mathbb{R} \to \mathbb{R}$ **tends to $+\infty$ at $+\infty$** if

$$\forall M > 0 \quad \exists A \in \mathbb{R} \quad \forall x \in \mathbb{R} \qquad x > A \Longrightarrow f(x) > M,$$

and finally a function $f : \mathbb{C} \to \mathbb{C}$ **tends to infinity at infinity** if

$$\forall M > 0 \quad \exists A \in \mathbb{R} \quad \forall z \in \mathbb{C} \qquad |z| > A \Longrightarrow \big|f(z)\big| > M.$$

In some cases, the definition of limit is refined by introducing a *punctured neighborhood*, i.e., looking at the values at points other than the point where the limit is considered:

**DEFINITION 1.24** Let $f : \mathbb{K} - \{a\} \to \mathbb{K}'$ (with $\mathbb{K}, \mathbb{K}' = \mathbb{R}$ or $\mathbb{C}$ and $a \in \mathbb{K}$) and let $\ell \in \mathbb{K}'$. Then **$f(x)$ converges to $\ell$ in punctured neighborhoods of $a$** if

$$\forall \varepsilon > 0 \quad \exists \eta \in \mathbb{R} \quad \forall z \in \mathbb{K}$$
$$(z \neq a \text{ and } |z - a| < \eta) \Longrightarrow \big(\big|f(z) - \ell\big| < \varepsilon\big).$$

This is denoted

$$\ell = \lim_{\substack{z \to a \\ z \neq a}} f(z).$$

This definition has the advantage of being practically identical to the definition of convergence at infinity. It is often better adapted to the physical description of a problem, as seen in Examples 1.1.b and 1.1.c on page 5 and the following pages. A complication is that it reduces the applicability of the theorem of composition of limits.

**THEOREM 1.25 (Sequential characterization of limits)** *Let $f : \mathbb{K} \to \mathbb{K}'$ be a function, and let $a \in \overline{\mathbb{K}}$ and $\ell \in \overline{\mathbb{K}'}$. Then $f(x)$ converges to $\ell$ as $x$ tends to $a$ if and only if, for any convergent sequence $(x_n)_{n\in\mathbb{N}}$ with limit equal to $a$, the sequence $\big(f(x_n)\big)_{n\in\mathbb{N}}$ converges to $\ell$.*

## 1.2.f Sequences of functions

Consider now the case of a sequence of functions $(f_n)_{n\in\mathbb{N}}$ each defined on a same subset $X$ of $\mathbb{R}$ or $\mathbb{C}$ and taking values in $\mathbb{R}$ or $\mathbb{C}$. Denote by $\|\cdot\|_\infty$ the "supremum norm"[8]:

$$\|f\|_\infty = \sup_{x \in X} \big|f(x)\big|.$$

[8] In fact, it is not a norm on the space of all functions, but only on the subspace of bounded functions. We disregard this subtlety and consider here that $\|\cdot\|$ takes values in $\overline{\mathbb{R}^+} = \mathbb{R} \cup \{+\infty\}$.

The definition of convergence of real or complex sequences may be extended to functions in two ways: one is a local notion, called *simple*, or *pointwise, convergence*, and the other, more global, is *uniform convergence*.[9]

**DEFINITION 1.26 (Simple convergence)** Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions, all defined on the same set $X$, which may be arbitrary. Then **the sequence $(f_n)_{n\in\mathbb{N}}$ converges simply** (or: **pointwise on $X$**) **to a function $f$** defined on $X$ if, for any $x$ in $X$, the sequence $\big(f_n(x)\big)_{n\in\mathbb{N}}$ converges to $f(x)$. This is denoted

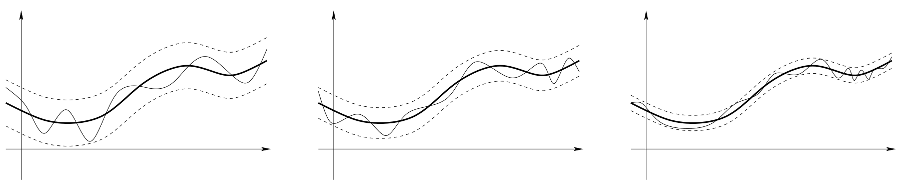
$$f_n \xrightarrow{\text{cv.s}} f.$$

**DEFINITION 1.27 (Uniform convergence)** Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions, all defined on the same set $X$, which may be arbitrary. Then **the sequence $(f_n)_{n\in\mathbb{N}}$ converges uniformly to the function $f$** if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \qquad n \geqslant N \Longrightarrow \|f_n - f\|_\infty < \varepsilon.$$

This is denoted $$f_n \xrightarrow{\text{cv.u}} f.$$

In other words, in the case where $X$ is a subset of $\mathbb{R}$, the graph of the function $f$ is located inside a smaller and smaller band of constant width in which all the graphs of $f_n$ must also be contained if $n$ is large enough:



If we have functions $f_n : X \to E$, where $(E, \|\cdot\|)$ is a normed vector space, we similarly define pointwise and uniform convergence using convergence in $E$; for instance, $(f_n)_{n\in\mathbb{N}}$ converges uniformly to $f$ if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \qquad n \geqslant N \Longrightarrow \sup_{x\in X} \big\|f_n(x) - f(x)\big\| < \varepsilon.$$

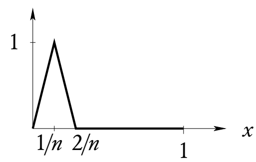
**Remark 1.28** Uniform convergence is an important theoretical mathematical notion. If one wishes to compute numerically a function $f$ using successive approximations $f_n$ (for instance, partial sums of a series expansion), then to get an error of size at most $\varepsilon$ for the value $f(x_1)$ of $f$ at some given point $x_1$, it suffices to find $N_1$ such that $\big|f_n(x_1) - f(x_1)\big| \leqslant \varepsilon$ for any $n \geqslant N_1$. Similarly, if the value of $f$ at other points $x_2, x_3, \ldots, x_p$ is needed, it will be enough to find corresponding integers $N_2, N_3, \ldots, N_p$. However, if it is not known beforehand at which points the function will evaluated, it will be necessary to know an integer $N$ such that $\big|f_n(x) - f(x)\big| \leqslant \varepsilon$ for all $n \geqslant N$ *and* for all $x \in \mathbb{R}$. Uniform convergence is then desirable.

[9] The concept of uniform convergence is due to George Stokes (see page 472) and Philipp Seidel (1821–1896), independently.

It is clear that uniform convergence implies pointwise convergence, but the converse is not true.

**Example 1.29** Define a sequence $(f_n)_{n\geqslant 1}$ a functions on $\mathbb{R}$ by

$$f_n(x) = \begin{cases} nx & \text{if } x \in [0, 1/n], \\ 2 - nx & \text{if } x \in [1/n, 2/n], \\ 0 & \text{if } x \in [2/n, 1]. \end{cases}$$




The reader will have no trouble proving that $(f_n)_{n\geqslant 1}$ converges pointwise to the zero function. However, the convergence is not uniform, since we have $\|f_n - f\|_\infty = 1$ for all $n \geqslant 1$.

**Example 1.30** The sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ defined for $n \geqslant 1$ by

$$f_n : x \longmapsto \sin\left(x + \frac{x}{n}\right)$$

converges uniformly to $f : x \mapsto \sin x$ on the interval $[0, 2\pi]$, and in particular it converges pointwise on this interval. However, although the sequence converges pointwise to the sine function on $\mathbb{R}$, *the convergence is not uniform* on all of $\mathbb{R}$. Indeed, for $n \geqslant 1$, we have

$$f_n\left(\frac{n\pi}{2}\right) = \sin\left(\frac{n\pi}{2} + \frac{\pi}{2}\right) \qquad \text{and} \qquad f\left(\frac{n\pi}{2}\right) = \sin\left(\frac{n\pi}{2}\right),$$

and those two values differ by 1 in absolute value. However, one can check that the convergence is uniform on any bounded segment in $\mathbb{R}$.

♦ **Exercise 1.1** Let $g(x) = \mathrm{e}^{-x^2}$ and $f_n(x) = g(x - n)$. Does the sequence $(f_n)_{n\in\mathbb{N}}$ converge pointwise on $\mathbb{R}$? Does it converge uniformly?

**Remark 1.31** In the case where the functions $f_n$ are defined on a subset of $\mathbb{R}$ with finite measure (for instance, a finite segment), a theorem of Egorov shows that pointwise convergence implies uniform convergence except on a set of arbitrarily small measure (for the definitions, see Chapter 2).

**Remark 1.32** There are other ways of defining the convergence of a sequence of functions. In particular, when some norm is defined on a function space containing the functions $f_n$, it is possible do discuss convergence in the sense of this norm. Uniform convergence corresponds to the case of the $\|\cdot\|_\infty$ norm. In Chapter 9, we will also discuss the notion of *convergence in quadratic mean*, or convergence in $L^2$ norm, and *convergence in mean* or convergence in $L^1$ norm. In pre-Hilbert spaces, there also exists a *weak convergence*, or convergence in the sense of scalar product (which is not defined by a norm if the space is infinite-dimensional).

A major weakness of pointwise convergence is that it does not preserve continuity (see Exercise 1.10 on page 44), or limits in general (Exercise 1.12). Uniform convergence, on the other hand, does preserve those notions.

**THEOREM 1.33 (Continuity of a limit)** *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions defined on a subset $D$ in $\mathbb{K}$ (or in a normed vector space), with values in an arbitrary normed vector space. Assume that the sequence $(f_n)_{n\in\mathbb{N}}$ converges uniformly to a function $f$.*

*i) Let $a \in D$ be such that all functions $f_n$ are continuous at the point $a$. Then $f$ is also continuous at the point $a$.*

*ii) In particular, if each $f_n$ is continuous on $D$, then the limit function $f$ is also continuous on $D$.*

This property extends to the case where $a$ is not in $D$, but is a limit point of $D$. However, it is then necessary to reinforce the hypothesis to assume that the functions $f_n$ have values in a complete normed vector space.

**THEOREM 1.34 (Double limit)** *Let $D$ be a subset of $\mathbb{R}$ (or of a normed vector space) and let $x_0 \in \overline{D}$ be a limit point*[10] *of $D$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions defined on $D$ with values in a* complete *normed vector space $E$. Assume that, for all $n$, the function $f_n$ has a limit as $x$ tends to $x_0$. Denote $\ell_n = \lim_{x\to x_0} f_n(x)$.*

*If $(f_n)_{n\in\mathbb{N}}$ converges uniformly to a function $f$, then*

*i) $f(x)$ has a limit as $x \to x_0$;*

*ii) $(\ell_n)_{n\in\mathbb{N}}$ has a limit as $n \to \infty$;*

*iii) the two limits are equal: $\lim\limits_{x\to x_0} f(x) = \lim\limits_{n\to\infty} \ell_n$.*

*In other words:*
$$\lim_{x\to x_0} \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \lim_{x\to x_0} f_n(x).$$

If we want a limit of differentiable functions to remain differentiable, stronger assumptions are needed:

**THEOREM 1.35 (Differentiation of a sequence of functions)** *Let $I$ be an interval of $\mathbb{R}$ with non-empty interior, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions defined on $I$ with values in $\mathbb{R}$, $\mathbb{C}$, or a normed vector space. Assume that the functions $f_n$ are differentiable on $I$, and moreover that:*

*i) the sequence $(f_n)_{n\in\mathbb{N}}$ converges pointwise to a function $f$;*

*ii) the sequence $(f_n')_{n\in\mathbb{N}}$ converges uniformly to a function $g$.*

*Then $f$ is differentiable on $I$ and $f' = g$.*

**Remark 1.36** If the functions take values in $\mathbb{R}$ or $\mathbb{C}$ or more generally any *complete* normed vector space, it is possible to weaken the assumptions by asking, instead of (i), that the sequence $(f_n(x_0))$ converges at a single point $x_0 \in I$. Assumption (ii) remains identical, and the conclusion is the same: $(f_n)_{n\in\mathbb{N}}$ converges uniformly to a differentiable function with derivative equal to $g$.

**Counterexample 1.37** The sequence of functions given by
$$f_n : x \longmapsto \frac{8}{\pi} \sum_{k=1}^{n} \frac{\sin^2(nx)}{4n^2 - 1}$$

[10] See Definition 4.52 on page 106; the simplest example is $D = ]a, b]$ and $x_0 = a$.

converges uniformly to the function $f : x \mapsto |\sin x|$ (see Exercise 9.3 on page 270), but the sequence of derivatives does not converge uniformly. The previous theorem does not apply, and indeed, although each $f_n$ is differentiable at 0, the limit $f$ is not.

**Remark 1.38** It happens naturally in some physical situations that a limit of a sequence of functions is not differentiable. In particular, in statistical thermodynamics, the state functions of a finite system are smooth. However, as the number of particles grows to infinity, discontinuities in the state functions or their derivatives may appear, leading to phase transitions.

Uniform convergence is also useful in another situation: when trying to exchange a limit (or a sum) and an integration process. However, in that situation, pointwise convergence is often sufficient, using the powerful tools of Lebesgue integration (see Chapter 2).

**THEOREM 1.39 (Integration on a finite interval)** *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of integrable functions (for instance, continuous functions), which converges uniformly to a function $f$ on a finite closed interval $[a,b] \subset \mathbb{R}$. Then $f$ is integrable on $[a,b]$ and we have*

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

**Example 1.40** This theorem is very useful, for instance, when dealing with a power series expansion which is known to converge uniformly on the open disc of convergence (see Theorem 1.66 on page 34). So, if we have $f(x) = \sum_{n=0}^{\infty} a_n x^n$ for $|x| < R$, then for any $x$ such that $|x| < R$, we deduce that

$$\int_0^x f(s)\,\mathrm{d}s = \sum_{n=0}^{\infty} \frac{a_n}{n+1}\, x^{n+1}.$$

To establish that a sequence converges uniformly, in practice, it is necessary to compute $\|f_n - f\|_\infty$, or rather to bound this expression by a quantity which itself converges to 0. This is sometimes quite tricky, and it is therefore useful to know the following two results of Dini:[11]

**THEOREM 1.41 (Dini)** *Let $K$ be a compact subset of $\mathbb{R}^k$, for instance, a closed ball. Let $(f_n)_{n\in\mathbb{N}}$ be an increasing sequence of continuous functions converging pointwise on $K$ to a continuous function $f$. Then the sequence $(f_n)_{n\in\mathbb{N}}$ converges uniformly to $f$ on $K$.*

**THEOREM 1.42 (Dini)** *Let $I = [a,b]$ be a compact interval in $\mathbb{R}$, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of increasing functions from $I$ to $\mathbb{R}$ that converges pointwise on $I$ to a continous function $f$. Then $(f_n)_{n\in\mathbb{N}}$ converges uniformly on $I$.*

---

[11] Ulisse DINI (1845–1918) studied in Pisa and Paris before taking a position in Pisa. His work concerned the theory of functions of a real variable, and he contributed to the early development of functional analysis.

**Remark 1.43** Be careful to distinguish between an *increasing sequence* of functions and a sequence of *increasing functions*. The former is a sequence $(f_n)_{n\in\mathbb{N}}$ of real-valued functions such that $f_{n+1}(x) \geqslant f_n(x)$ for any $x \in \mathbb{R}$ and $n \in \mathbb{N}$. The latter is a sequence of real-valued functions defined on $K \subset \mathbb{R}$ such that for any $x$, $y \in K$: $x \leqslant y \Longrightarrow f_n(x) \leqslant f_n(y)$.

As mentioned briefly already, it is possible with Lebesgue's dominated convergence theorem to avoid requiring uniform convergence to exchange an integral and a limit, as in Theorem 1.39. See Chapters 2 and 3 for details on this theory.

## 1.3 Series

### 1.3.a Series in a normed vector space

We first recall the definition of convergence and absolute convergence of a series in a normed vector space.

**DEFINITION 1.44 (Convergence of a series)** Let $(a_n)_n$ be a sequence with values in a normed vector space. Let $(S_n)_{n\in\mathbb{N}}$ denote the **sequence of partial sums**

$$S_n \stackrel{\text{def}}{=} \sum_{k=0}^{n} a_k.$$

- The **series $\sum a_n$ converges, and its sum is equal to $A$** if the sequence $(S_n)_{n\in\mathbb{N}}$ converges to $A$. This is denoted

$$\sum_{n=0}^{\infty} a_n = A.$$

- The **series $\sum a_n$ converges absolutely** if the series $\sum \|a_n\|$ converges in $\mathbb{R}$.
- In particular, a series $\sum a_n$ of real or complex numbers converges absolutely if the series $\sum |a_n|$ converges in $\mathbb{R}$.

As in the case of sequences, there exists a Cauchy criterion for convergence of series[12]:

[12] This criterion was stated by Bernhard Bolzano (see page 581) in 1817. But Bolzano was isolated in Prague and little read. Cauchy presented this criterion, without proof and as an obvious fact, in his analysis course in 1821.

**THEOREM 1.45 (Cauchy criterion)** *If a series $\sum u_n$ with values in a normed vector space $E$ converges, then it satisfies the Cauchy criterion:*

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall p, q \in \mathbb{N} \qquad (q > p \geqslant N) \Longrightarrow \left\| \sum_{n=p}^{q} u_n \right\| < \varepsilon,$$

*or in other words:*

$$\lim_{p,q\to\infty} \sum_{p\leqslant n\leqslant q} u_n = 0.$$

*Conversely, any series which satisfies the Cauchy criterion and takes values in $\mathbb{R}$, $\mathbb{C}$, any finite-dimensional normed vector space, or more generally, any complete normed vector space, converges.*

From this the following fundamental theorem is easily deduced:

**THEOREM 1.46** *Any absolute convergent series $\sum a_n$ with values in a* complete *normed vector space is convergent.*

*In particular, any absolutely convergent series of real or complex numbers is convergent.*

PROOF. Let $\sum u_n$ be an absolutely convergent series. Although we can write

$$\left\| \sum_{n=0}^{k} u_n \right\| \leqslant \sum_{n=0}^{k} \|u_n\|,$$

nothing can be deduced from this, because the right-hand side does not tend to zero. But we can use the Cauchy critetion: for all $p$, $q \in \mathbb{N}$, we have of course

$$\left\| \sum_{n=p}^{q} u_n \right\| \leqslant \sum_{n=p}^{q} \|u_n\|,$$

and since $\sum \|u_n\|$ satisfies the Cauchy criterion, so does $\sum u_n$. Since $u_n$ lies in a complete space by assumption, this means that the series $\sum u_n$ is indeed convergent.

## 1.3.b Doubly infinite series

In the theory of Fourier series, we will have to deal with formulas of the type

$$\int_0^1 |f(t)|^2 \, \mathrm{d}t = \sum_{n=-\infty}^{+\infty} |c_n|^2.$$

To give a precise meaning to the right-hand side, we must clarify the meaning of the convergence of a series indexed by integers in $\mathbb{Z}$ instead of $\mathbb{N}$.

**DEFINITION 1.47** A **doubly infinite series $\sum_{n\in\mathbb{Z}} a_n$**, with $a_n$ is a normed vector space, **converges** if $\sum a_n$ and $\sum a_{-n}$ are both convergent, the index ranging over $\mathbb{N}$ in each case. Then we denote

$$\sum_{n=-\infty}^{+\infty} a_n \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} a_n + \sum_{n=1}^{\infty} a_{-n}$$

and say that this is the sum of $\sum_{n\in\mathbb{Z}} a_n$.

In other words, a series of complex numbers $\sum_{n\in\mathbb{Z}} a_n$ converges to $\ell$ if and only if, for any $\varepsilon > 0$, there exists $N > 0$ such that

$$\text{for any } i \geqslant N \text{ and } j \geqslant N, \qquad \left| \left( \sum_{n=-i}^{j} a_n \right) - \ell \right| < \varepsilon.$$

**Remark 1.48** It is crucial to allow the upper and lower bounds $i$ and $j$ to be independent. In particular, if the limit of $\sum_{n=-k}^{k} a_n$ exists, it does not follow that the doubly infinite series $\sum_{n\in\mathbb{Z}} a_n$ converges.

For instance, take $a_n = 1/n$ for $n \neq 0$ and $a_0 = 0$. Then we have $\sum_{n=-k}^{k} a_n = 0$ for all $k$ (and so this sequence does converge as $k$ tends to infinity), but the series $\sum_{n\in\mathbb{Z}} a_n$ *diverges* according to the definition, because each of the series $\sum a_n$ and $\sum a_{-n}$ (over $n \geqslant 0$) is divergent.

## 1.3.c Convergence of a double series

As in Section 1.2.d, let $(a_{ij})_{i,j\in\mathbb{N}}$ be a family of real numbers indexed by two integers. For any $p, q \in \mathbb{N}$, we have

$$\sum_{i=1}^{p}\sum_{j=1}^{q} a_{ij} = \sum_{j=1}^{q}\sum_{i=1}^{p} a_{ij},$$

since each sum is finite. On the other hand, even if all series involved are convergent, it is not always the case that

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} \qquad \text{and} \qquad \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij},$$

as the following example shows:

$$(a_{ij}) = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & -1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & -1 & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \end{pmatrix}$$

where we have (note that $i$ is the row index and $j$ is the column index):

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = 0 \qquad \text{but} \qquad \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij} = 1.$$

We can find an even more striking example by putting

$$(a_{ij}) = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 2 & -2 & 0 & 0 & 0 & \dots \\ 0 & 0 & 3 & -3 & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \end{pmatrix}$$

in which case

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = \sum_{i=1}^{\infty} 0 = 0 \qquad \text{but} \qquad \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} 1 = +\infty.$$

### 1.3.d Conditionally convergent series, absolutely convergent series

**DEFINITION 1.49** A series $\sum a_n$ with $a_n$ in a normed vector space is **conditionnally convergent** if it is convergent but not absolutely convergent.

**DEFINITION 1.50** We denote by $\mathfrak{S}$ the group of **permutations**, that is, the group of bijections from $\mathbb{N}$ to $\mathbb{N}$, and we denote by $\mathfrak{S}_n$ the finite group of permutations of the set $\{1, \ldots, n\}$.

**DEFINITION 1.51** A **series $\sum x_n$ is commutatively convergent** if it is convergent with sum $X$, and for any permutation $\varphi \in \mathfrak{S}$, the rearranged series $\sum x_{\varphi(n)}$ converges to $X$.

Is a convergent series necessarily *commutatively* convergent? In other words, *is it legitimate to change arbitrarily the order of the terms of a convergent series?*

At first sight, it is very tempting to say "Yes," almost without thinking, since permuting terms in a finite sum has no effect on the result. The problem is that we have here an *infinite sum*, not a finite sum in the algebraic sense. So there is a limiting process involved, and we will see that this brings a very different picture: only *absolutely convergent* series will be *commutatively convergent*. So, if a series is conditionally convergent (convergent but not absolutely so), changing the order of the terms may alter the value of the sum – or even turn it into a divergent series.

**THEOREM 1.52** *Let $\sum a_n$ be a conditionnally convergent series, with $a_n \in \mathbb{R}$. Then for any $\ell \in \overline{\mathbb{R}}$, there exists a permutation $\varphi \in \mathfrak{S}$ such that the rearranged series $\sum a_{\varphi(n)}$ converges to $\ell$.*

*In fact, for any $a, b \in \overline{\mathbb{R}}$, with $a \leqslant b$, there exists a permutation $\psi \in \mathfrak{S}$ such that the set of limit points of the sequence of rearranged partial sums $\left(\sum_{k=0}^{n} a_{\psi(k)}\right)_{n\in\mathbb{N}}$ is the interval $[a, b]$.*

PROOF. We assume that $\sum a_n$ is conditionnally convergent.
► First remark: *there are infinitely many positive values* and *infinitely many negative values* of the terms $a_n$ of the series. Let $\alpha_n$ denote the sequence of non-negative terms, in the order they occur, and let $\beta_n$ denote the sequence of negative terms.

Here is an illustration:

$$\begin{array}{cccccccccccc} u_n : & 1 & 3 & 2 & -4 & -1 & 2 & -1 & 0 & 2 & \cdots \\ & \alpha_1 & \alpha_2 & \alpha_3 & \beta_1 & \beta_2 & \alpha_4 & \beta_3 & \alpha_5 & \alpha_6 & \cdots \end{array}$$

► Second remark: both series $\sum \alpha_n$ and $\sum \beta_n$ are divergent, their partial sums converging, respectively, to $+\infty$ and $-\infty$. Indeed, if both series were to converge, the

series $\sum a_n$ would be absolutely convergent, and if only one were to converge, then $\sum a_n$ would be divergent (as follows from considering the sequence of partial sums).
► Third remark: both sequences $(\alpha_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ tend to 0 (since $(a_n)_{n\in\mathbb{N}}$ tends to 0, as a consequence of the convergence of $\sum a_n$).

Now consider $\ell \in \mathbb{R}$. Let $S_n$ denote the sequence of sums of values of $\alpha$ and $\beta$ which is constructed as follows. First, sum all consecutive values of $\alpha_n$ until their sum is larger than $\ell$; call this sum $S_1$. Now add to $S_1$ all consecutive values of $\beta_n$ until the resulting sum $S_1 + \beta_1 + \cdots$ is *smaller* than $\ell$; call this sum $S_2$. Then start again adding from the remaining values of $\alpha_n$ until getting a value larger than $\ell$, called $S_3$, and continue in this manner until the end of time.

Now notice that:

- Since at each step we add at least one value of $\alpha$ or one of $\beta$, it is clear that *all* values of $\alpha$ will be used sooner or later, as well as all values of $\beta$, that is, when all is said and done, all values of $a_n$ will have been involved in one of the sums $S_n$.
- Since, at each step, the distance $|\ell - S_n|$ is at most equal to the absolute value of the last value of $\alpha$ or $\beta$ considered, the distance from $S_n$ to $\ell$ tends to 0 as $n$ tends to infinity.

From this we deduce that the sequence $(S_n)$ is a sequence of partial sums of a rearrangement of the series $\sum a_n$, and that it converges to $\ell$. Hence this proves that *by simply changing the order of the terms, one may cause the series to converge to an arbitrary sum.*

Let now $a, b \in \overline{\mathbb{R}}$ with $a < b$ (the case $a = b$ being the one already considered).
• **If $a$ and $b$ are both finite**, we can play the same game of summation as before, but this time, at each step, we either sum values of $\alpha_n$ until we reach a value larger than $b$, or we sum values of $\beta_n$ until the value is less than $a$.
• **If $b = +\infty$ and $a$ is finite**, we sum from $a$ to above $a+1$, then come back to below $a$, then sum until we are above $a + 2$, come back below $a$, etc. Similarly if $a = -\infty$ and $b$ is finite.
• **If $a = -\infty$ and $b = +\infty$**, start from 0 to go above 1, then go down until reaching below $-2$, then go back up until reaching above 3, etc.

**Example 1.53** Consider the sequence $(a_n)_{n\in\mathbb{N}^*}$ with general term $a_n = (-1)^{n+1}/n$. It follows from the theory of power series (Taylor expansion of $\log(1+x)$) that the series $\sum a_n$ converges and has sum equal to $\log 2$. If we sum the same values $a_n$ by taking one positive term followed by two negative terms, then the resulting series converges to $\frac{1}{2}\log 2$. Indeed, if $(S_n)_{n\in\mathbb{N}}$ and $(S'_n)_{n\in\mathbb{N}}$ denote the sequence of partial sums of the original and modified series, respectively, then for $n \in \mathbb{N}$ we have

$$S_{2n} = 1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \cdots + \tfrac{1}{2n-1} - \tfrac{1}{2n}$$

and

$$\begin{aligned} S'_{3n} &= \underbrace{1 - \tfrac{1}{2}} - \tfrac{1}{4} + \underbrace{\tfrac{1}{3} - \tfrac{1}{6}} - \tfrac{1}{8} + \cdots + \underbrace{\tfrac{1}{2n-1} - \tfrac{1}{4n-2}} - \tfrac{1}{4n} \\ &= \tfrac{1}{2} - \tfrac{1}{4} + \tfrac{1}{6} - \tfrac{1}{8} + \cdots + \tfrac{1}{4n-2} - \tfrac{1}{4n} = \tfrac{1}{2} S_{2n}. \end{aligned}$$

As an exercise, the reader can check that if one takes instead two positive terms followed by one negative terms, the resulting series converges with a value equal to $\frac{3}{2}\log 2$.

The following result shows that, on the other hand, one can rearrange at will the order of the terms of an absolutely convergent series.

**THEOREM 1.54** *A series of complex numbers is commutatively convergent if and only if it is absolutely convergent.*

PROOF. Assume first that the terms of the series are real numbers. The theorem above shows that if $\sum a_n$ is commutatively convergent, it must be absolutely convergent. Conversely, assume the series is absolutely convergent and let $\ell$ denote its sum. Let $\psi$ be any permutation. By convergence of the series, there exists $N \geqslant 1$ such that

$$\left|\sum_{k=1}^{n} a_k - \ell\right| < \varepsilon$$

for $n \geqslant N$. For each such $n \geqslant N$, there exists $N'$ such that the set $\{\psi(1), \dots, \psi(N')\}$ contains $\{1, \dots, n\}$ (it suffices that $N'$ be larger than the maximum of the images of 1, ..., $N$ by the inverse permutation $\psi^{-1}$). Then for any $m \geqslant N'$, we have

$$\left|\sum_{k=1}^{m} a_{\psi(k)} - \ell\right| \leqslant \left|\sum_{k=1}^{n} a_k - \ell\right| + \sum_{k>n} |a_k| \leqslant \varepsilon + \sum_{k>n} |a_k|,$$

since the set $\{\psi(1), \dots, \psi(m)\}$ contains $\{1, \dots, n\}$, and possibly additional values which are all larger than $n$. The absolute convergence makes its appearance now: the last sum on the right is the remainder for the convergent series $\sum |a_k|$, and for $n \geqslant N''$ it is therefore itself smaller than $\varepsilon$. Since, given $\varepsilon$, we can take $n = N''$ and find the value $N'$ from it, such that

$$\left|\sum_{k=1}^{m} a_{\psi(k)} - \ell\right| \leqslant 2\varepsilon$$

for $m \geqslant N'$, and so we have proved that the rearranged series converges with sum equal to $\ell$.

If the terms of the series are complex numbers, it suffices to apply the result for real series to the series of real and imaginary parts.

The possibility of rearranging at will the order of summation explains the importance of absolutely convergent series[13].

**Remark 1.55** In statistical mechanics, there are so-called *diagrammatic* methods to compute the values of certains quantities, such as pressure or mean-energy, at equilibrium. Those methods are based on rearrangements of the terms of certain series, summing "by packets" in particular. Those methods are particularly useful when the original series is not absolutely convergent. This means that all results obtained in this manner must be treated carefully, if not suspiciously. They belong to the gray area of *exact* (at least, this is what everyone believes!) results, but which are not *rigorous*. (It is of course much more difficult to obtain results which can be judged with mathematical standards of rigor; the reader is invited to read the beautiful papers [63, 64] for convincing illustrations.)

[13] Peter Gustav LEJEUNE-DIRICHLET showed in 1837 that a convergent series with non-negative terms is commutatively convergent. In 1854, Bernhard RIEMANN wrote three papers in order to obtain a position at the university of Göttingen. In one of them, he describes commutatively convergent series in the general case. However, another paper was selected, concerning the foundations of geometry.

### 1.3.e Series of functions

We can define pointwise and uniform convergence of series of functions just as was done for sequences of functions.

**DEFINITION 1.56 (Pointwise convergence)** Let $X$ be an arbitrary set, $(E, \|\cdot\|)$ a normed vector space. A series $\sum f_n$ of functions $f_n : X \to E$ **converges pointwise to a function $F : X \to E$** if, for any $x \in X$, the series $\sum f_n(x)$ converges to $F(x)$ in $E$, that is, if

$$\forall \varepsilon > 0 \quad \forall x \in X \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \qquad n \geqslant N \Longrightarrow \left\| \sum_{k=1}^{n} f_k(x) - F(x) \right\| < \varepsilon.$$

The function $F$ is called the **pointwise**, or **simple**, **limit** of the series $\sum f_n$, and this is denoted $\sum f_n \xrightarrow{\text{cv.s.}} F$.

**DEFINITION 1.57 (Uniform convergence)** Let $X$ be an arbitrary set, $(E, \|\cdot\|)$ a normed vector space. A series $\sum f_n$ of functions $f_n : X \to E$ **converges uniformly to a function $F : X \to E$** if the sequence of partial sums of the series converges uniformly to $F$, that is, if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall x \in X \quad \forall n \in \mathbb{N} \qquad n \geqslant N \Longrightarrow \left\| \sum_{k=1}^{n} f_k(x) - F(x) \right\| < \varepsilon.$$

This is denoted $\sum f_n \xrightarrow{\text{cv.u}} F$. This amounts to

$$\lim_{n\to\infty} \left\| \sum_{k=1}^{n} f_k - F \right\|_\infty = 0 \qquad \text{where} \qquad \|g\|_\infty = \sup_{x \in X} \left\| g(x) \right\|.$$

**DEFINITION 1.58 (Absolute convergence)** Let $X$ be an arbitrary set, $(E, \|\cdot\|)$ a normed vector space. A series $\sum f_n$ of functions $f_n : X \to E$ **converges absolutely** if the series $\sum \|f_n\|_\infty$ converges, where

$$\|f_n\|_\infty = \sup_{x \in X} \left\| f_n(x) \right\|.$$

The following theorem is the most commonly used to prove uniform convergence of a series of functions:

**THEOREM 1.59** *Any absolutely convergent series with values in a* complete *normed vector space is uniformly convergent, and hence pointwise convergent.*

Corresponding to the continuity and differentiability results for sequences of functions, we have:

**THEOREM 1.60 (Continuity and differentiability of a series of functions)** *Let $D$ be a subset of $\mathbb{R}$ or of a normed vector space. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions $f_n : D \to E$, where $\big(E, \|\cdot\|\big)$ is some normed vector space, for instance, $\mathbb{R}$ or $\mathbb{C}$. Assume that the series $\sum f_n$ converges pointwise to a function $F$.*

*i) If each $f_n$ is continuous on $D$, and if the series $\sum f_n$ converges uniformly on $D$, then $F$ is continuous on $D$.*

*ii) If $D$ is an interval of $\mathbb{R}$, each $f_n$ is differentiable on $D$, and the series $\sum f_n'$ converges uniformly, then $F$ is differentiable and we have*

$$F' = \sum_{n=0}^{\infty} f_n'.$$

## 1.4 Power series, analytic functions

Quite often, physicists encounter series expansions of some function. These expansions may have different origins:

- the superposition of many phenomena (as in the Fabry-Perot interferometer);
- perturbative expansions, when exact computations are too difficult to perform (e.g., hydrodynamics, semiclassical expansions, weakly relativistic expansions, series in astronomy, quantum electrodynamics, etc.);
- sometimes the exact evalution of a function which expresses some physical quantity is impossible; a *numerical* evaluation may then be performed using Taylor series expansions, Fourier series, infinite product expansions, or asymptotic expansions.

We first recall various forms of the Taylor formula. The general idea is that there is an approximate expression

$$f(x) \approx f(a) + (x-a)\,f'(a) + \frac{(x-a)^2}{2!}\,f''(a) + \cdots + \frac{(x-a)^k}{k!}\,f^{(k)}(a)$$

for a function $f$ which is at least $k$ times differentiable on an interval $J$, with values in some normed vector space $\big(E, \|\cdot\|\big)$, and for a given point $a \in J$, where $x$ lies is some neighborhood of $a$.

The question is to make precise the meaning of the symbol "$\approx$"!

Define $R_k(x)$ to be the difference between $f(x)$ and the sum on the right-hand side of the above expression; in other words, we have

$$f(x) = \sum_{n=0}^{k} \frac{(x-a)^n}{n!}\,f^{(n)}(a) + R_k(x) = T_k(x) + R_k(x)$$

Brook TAYLOR (1685–1731), English mathematician, was a student at Cambridge, then member and secretary of the prestigious Royal Society, a venerable institution dedicated to the advancement of Science. He wrote the famous formula

$$f(a+\varepsilon) = f(a) + \varepsilon f'(a) + \tfrac{\varepsilon^2}{2} f''(a) + \cdots$$

without considering the issue of convergence. He was also interested in the physical and mathematical aspects of vibrating strings.

by definition. Of course, we hope that the **Taylor remainder $R_k(x)$** is a "small quantity," so that we may approximate the value of $f(x)$ by the value of the **Taylor polynomial of order $k$ at $x$**, that is, by $T_k(x)$. There are different ways in which this remainder may become small:

- one may let $x$ tend to $a$ (for a fixed value of $k$);
- or let $k$ tend to infinity (for a fixed value of $x$).

The Taylor-Lagrange and Taylor-Young formulas are relevant for the first case, while the second belongs to the theory of power series.

## 1.4.a Taylor formulas

**THEOREM 1.61 (Taylor formula with integral remainder)** *Let $J$ be an interval of $\mathbb{R}$, and $(E, \|\cdot\|)$ a normed vector space. Let $f : J \to E$ be a function of $\mathscr{C}^k$ class on $J$, which is piecewise of $\mathscr{C}^{k+1}$ class on $J$. For any $a$ and $x \in J$, we have*

$$f(x) = \sum_{n=0}^{k} \frac{(x-a)^n}{n!} f^{(n)}(a) + \int_a^x \frac{(x-t)^k}{k!} f^{(k+1)}(t)\,\mathrm{d}t.$$

**THEOREM 1.62 (Taylor-Lagrange formula)** *Let $f : J \to \mathbb{R}$ be a* real-valued *function of $\mathscr{C}^k$ class on an interval $J$ of $\mathbb{R}$, which is $k+1$ times differentiable in the interior of $J$. Let $a \in J$. Then, for any $x \in J$, there exists $\theta \in \,]0,1[$ such that*

$$f(x) = \sum_{n=0}^{k} \frac{(x-a)^n}{n!} f^{(n)}(a) + \frac{(x-a)^{k+1}}{(k+1)!} f^{(k+1)}\big(a + \theta(x-a)\big).$$

**Remark 1.63** This formula is only valid for real-valued functions. However, the following corollary is also true for functions with complex values, or functions with values in a normed vector space.

**COROLLARY 1.64 (Taylor-Lagrange inequality)** *Let $f : J \to E$ be a function of $\mathscr{C}^k$ class on an interval $J$ of $\mathbb{R}$, with values in a normed vector space $E$. Assume $f$ is $k+1$ times differentiable in the interior of $J$. Let $a \in J$. Then for any $x \in J$ we have*

$$\left\| f(x) - \sum_{n=0}^{k} \frac{(x-a)^n}{n!} f^{(n)}(a) \right\|_E \leqslant \frac{|x-a|^{k+1}}{(k+1)!} \sup_{t \in J} \left\| f^{(k+1)}(t) \right\|_E .$$

**THEOREM 1.65 (Taylor-Young formula)** *Let $f$ be a function which is $k$ times differentiable on an interval $J$ of $\mathbb{R}$, with values in a normed vector space $E$. Let $a \in J$. Then we have*

$$f(x) - \sum_{n=0}^{k} \frac{(x-a)^n}{n!} f^{(n)}(a) = \underset{x \to a}{\mathrm{o}} \big( (x-a)^k \big).$$

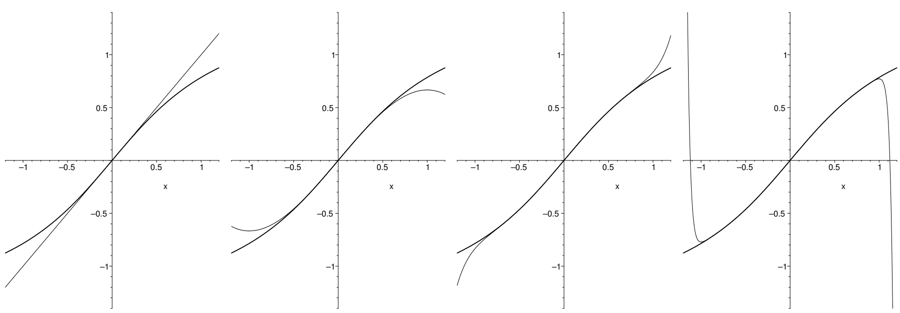
## 1.4.b Some numerical illustrations

Suppose we want to compute numerically some values of the inverse tangent function arctan, which is of course infinitely differentiable on $\mathbb{R}$. It is easy to compute the values of the successive derivatives of this function at 0, and we can write down explicitly the Taylor polynomial at 0 of arbitrary order: this gives the expression

$$\arctan x = \sum_{n=0}^{k} \frac{(-1)^n}{2n+1} x^{2n+1} + R_k(x),$$

for the Taylor formula of order $2n+1$ (notice that only odd powers of $x$ appear, because the inverse tangent function is odd).

If we represent graphically those polynomials with $k = 0, 1, 4, 36$ (i.e., of order 1, 5, 9, and 18, respectively), with the graph of the function itself for comparison, we obtain the following:



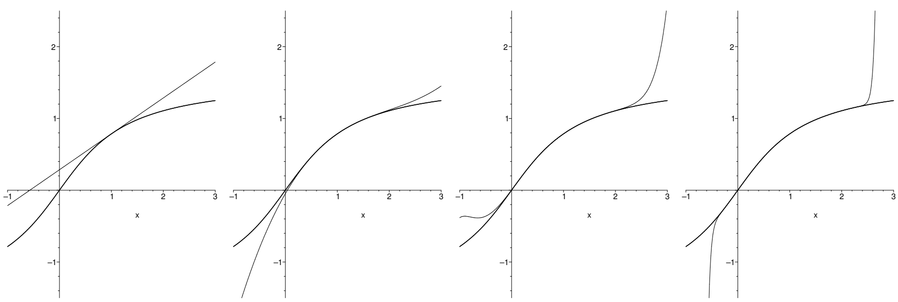
order 1 order 5 order 9 order 37

The following facts appear:

- on each graph (i.e., for fixed $k$), the Taylor polynomial and the inverse tangent functions get closer and closer together as $x$ approaches 0;

- for a fixed real number $x \in [-1, 1]$ (for instance 0.8), the values at $x$ of the Taylor polynomial of increasing degree get closer and closer to the value of the function as $k$ increases;
- on the other hand, for a real number $x$ such that $|x| > 1$, disaster strikes: the larger $k$ is, the further away to $\arctan x$ is the value of the Taylor polynomial!

The first observation is simply a consequence of the Taylor-Young formula. The other two deserve more attention. It seems that the sequence $(T_k)_{k\in\mathbb{N}}$ of the Taylor polynomials converges on $[-1, 1]$ and diverges outside.[14] However, *the function* arctan *is perfectly well-defined, and very regular, at the point* $x = 1$; it does not seem that anything special should happen there. In fact, it is possible to write down the Taylor expansion centered at $a = 1$ instead of $a = 0$ (this is a somewhat tedious computation[15]), and (using approximations of the same order as before), we obtain the following graphs:



We can see the same three basic facts, except that convergence seems to be restricted now to the interval $\left[1-\sqrt{2}, 1+\sqrt{2}\right]$.

In order to understand why such intervals occur, it is necessary to dwell further on the theory of power series (see below) and especially on holomorphic functions of a complex variable (in particular, Theorem 4.40 on page 101). We will only state here that the function arctan can be continued naturally to a function on the complex plane ($\arctan z_0$ is defined as the value of the integral of the function $1/(1+z^2)$ on a certain path[16] joining the origin to $z_0$. The function thus obtained is well-defined, independently of the chosen path, up to an integral multiple of $\pi$[17] and is a well-defined function on $\mathbb{C}$ minus the two single points where $1 + z^2$ vanishes, namely i and $-$i. Then, one shows that for such a function, the sequence of Taylor polynomials centered

[14] To be honest, it is difficult to ascertain from the graphs above if the interval to consider is $[-1, 1]$ or $]-1, 1[$, for instance. The general theory of series shows that $(T_n(x))_{n\in\mathbb{N}}$ converges quickly if $|x| < 1$ and very slowly if $|x| = 1$.

[15] The $n$-th coefficient of the polynomial is $(-1)^{n+1}\sin(n\pi/4)2^{-n/2}/n$ and the constant term is $\pi/4$.

[16] This notion of integral on a path is defined by the formula (4.2) page 94.

[17] This is a consequence of the residue theorem 4.81 on page 115.

at $a$ converges on the open disc centered at $a$ with radius equal to the distance from $a$ to the closest singulirity (hence the radius is $|1-0|=1$ in the first case of Taylor expansions at $a=0$, and is $|1-\mathrm{i}|=\sqrt{2}$ in the second case).

### 1.4.c Radius of convergence of a power series

A **power series centered at $z_0$** is any series of functions of the type

$$z \longmapsto \sum_{n=0}^{\infty} a_n (z-z_0)^n,$$

where $(a_n)_{n\in\mathbb{N}}$ is a given sequence of real or complex numbers, which are sometimes called the *coefficients* of the power series.

**THEOREM 1.66 (Radius of convergence)** *Let $\sum a_n(z-z_0)^n$ be a power series centered at $z_0$. The **radius of convergence** is the element in $\overline{\mathbb{R}}^+$ defined by*

$$R \stackrel{\text{def}}{=} \sup\left\{t \in \mathbb{R}^+ \,;\, (a_n t^n)_{n\in\mathbb{N}} \text{ is bounded}\right\}.$$

*The power series converges absolutely and uniformly on any compact subset in the disc $\mathcal{B}(z_0\,;R) \stackrel{\text{def}}{=} \left\{z\in\mathbb{C}\,;\, |z-z_0|<R\right\}$, in the complex plane $\mathbb{C}$, and it diverges for any $z\in\mathbb{C}$ such that $|z|>r$. For $|z|=r$, the series may be convergent, conditionally convergent, or divergent at $z$.*

Note that "absolute convergence" here refers to absolute convergence as a series of functions, which is stronger than absolute convergence for every $z$ involved: in other words, for any compact set $D\subset\mathcal{B}(z_0\,;R)$, we have

$$\sum_{n=0}^{\infty} \sup_{z\in D} |a_n z^n| < +\infty.$$

**Example 1.67** The power series $-\log(1-z)=\sum_{n=1}^{\infty} z^n/n$ converges for any $z\in\mathbb{C}$ such that $|z|<1$ and diverges if $|z|>1$ (the radius of convergence is $R=1$). Moreover, this series is divergent at $z=1$, but conditionnally convergent at $z=-1$ (by the alternate series test), and more generally, it is conditionnally convergent at $z=\mathrm{e}^{\mathrm{i}\theta}$ for any $\theta\notin 2\pi\mathbb{Z}$ (this can be shown using the Abel transformation, also known as "summation by parts").

**DEFINITION 1.68 (Power series expansion)** Let $\Omega$ be an open subset in $\mathbb{R}$ or $\mathbb{C}$. A function $f:\Omega\to\mathbb{C}$ defined on $\Omega$ **has a power series expansion centered at some $z_0\in\Omega$** if there exist an open subset $V\subset\Omega$ containing $z_0$ and a sequence $(a_n)_{n\in\mathbb{N}}$ of complex numbers such that

$$\forall z\in V \qquad f(z)=\sum_{n=0}^{\infty} a_n(z-z_0)^n.$$

The radius of convergence of a power series depends only weakly on the precise values of the coefficients, so, for instance, if $F=P/Q$ is a rational function with no pole in $\mathbb{N}$, the power series $\sum F(n)\, a_n\, z^n$ and $\sum a_n\, z^n$ have the same radius of convergence. From this and Theorem 1.60, it follows in particular that a power series can be differentiated term by term inside the disc of convergence:

**THEOREM 1.69 (Derivative of a power series)** *Let $J$ be an open subset of $\mathbb{R}$, $x_0 \in J$, and $f : J \to \mathbb{C}$ a function which has a power series expansion centered at $x_0$:*

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n.$$

*Let $R > 0$ be the radius of convergence of this power series. Then $f$ is infinitely differentiable on the open interval $]x_0 - R, x_0 + R[$, and each derivative has a power series expansion on this interval, which is obtained by repeated term by term differentiation, that is, we have*

$$f'(x) = \sum_{n=1}^{\infty} n a_n\, (x - x_0)^{n-1} \qquad \text{and} \qquad f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!}\, a_n (x - x_0)^{n-k}$$

*for any $k \in \mathbb{N}$. Hence the $n$-th coefficient of the power series $f(x)$ can be expressed as*

$$a_n = \frac{f^{(n)}(x_0)}{n!}.$$

**Remark 1.70** The power series $\sum \left(f^{(n)}(x_0)/n!\right) \cdot (x - x_0)^n$ is the **Taylor series of $f$ at $x_0$.** On any compact subset inside the open interval of convergence, it is the uniform limit of the sequence of Taylor polynomials.

**Remark 1.71** In Chapter 4, this result will be extended to power series of one complex variable (Theorem 4.40 on page 101).

## 1.4.d Analytic functions

Consider a function that may be expended into a power series in a neighborhood $V$ of a point $z_0$, so that for $z \in V$, we have

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

Given such a $z \in V$, a natural question is the following: *may $f$ also be expended into a power series centered at $z$?*

Indeed, it might seem possible *a priori* that $f$ can be expanded in power series only around $z_0$, and around no other point. However, this is not the case:

**DEFINITION 1.72 (Analytic function)** A function $f : \Omega \to \mathbb{C}$ defined on an open subset $\Omega$ of $\mathbb{C}$ or $\mathbb{R}$ is **analytic** on $\Omega$ if, for any $z_0 \in \Omega$, $f$ has a power series expansion centered at $z_0$.

Note that the radius of convergence of the power series may (and often does!) vary with the point $z_0$.

**THEOREM 1.73** *Let $\sum a_n z^n$ be a power series with positive radius of convergence $R > 0$, and let $f$ denote the sum of this power series on $\mathcal{B}(0\,;R)$. Then the function $f$ is analytic on $\mathcal{B}(0\,;R)$.*

**Example 1.74** The function $f : x \mapsto 1/(1-x)$ has the power series expansion

$$f(x) = \sum_{n=0}^{\infty} x^n$$

around 0, with radius of convergence equal to 1. Hence, for any $x_0 \in\, ]-1,1[$, there exists a power series expansion centered at $x_0$ (obviously with different coefficients). This can be made explicit: let $h \in \mathcal{B}(0\,;|1-x_0|)$, then with $x = x_0 + h$, we have

$$f(x) = f(x_0+h) = \frac{1}{1-(x_0+h)} = \frac{1}{1-x_0}\cdot\frac{1}{1-h/(1-x_0)} = \sum_{n=0}^{\infty}\frac{(x-x_0)^n}{(1-x_0)^{n+1}}.$$

**Remark 1.75 (Convergence of Taylor expansions)** Let $f : U \to \mathbb{C}$ be a function defined on an open subset $U$ of $\mathbb{R}$. Under what conditions is $f$ analytic?[18] There are two obvious necessary conditions:

- $f$ is infinitely differentiable on $U$;
- for any $x_0 \in U$, there exists an open disc $\mathcal{B}(x_0, r)$ such that the series $\sum \frac{1}{n!} f^{(n)}(x_0)(x - x_0)^n$ converges for any $x \in \mathcal{B}(x_0, r)$.

However, those two conditions are *not* sufficient. The following classical counter-example shows this: let

$$f(x) \stackrel{\text{def}}{=} \exp\left(-1/x^2\right) \quad \text{if } x \neq 0, \qquad f(0) = 0.$$

It may be shown[19] that $f$ is indeed of $\mathscr{C}^\infty$ and that each derivative of $f$ at 0 vanishes, which ensures (!) the convergence of the Taylor series everywhere. But since the function vanishes only at $x = 0$, it is clear that the Taylor series does not converge to $f$ on any open subset, hence $f$ is not analytic.

It is therefore important not to use the terminology "analytic" where "infinitely differentiable" is intended. This is a confusion that it still quite frequent in scientific literature.

The Taylor formulas may be used to prove that a function is analytic. If the sequence $(R_n)_{n\in\mathbb{N}}$ of remainders for a function $f$ converges uniformly to 0 on a neighborhood of $a \in \mathbb{R}$, then the function $f$ is analytic on this neighborhood. To show this, one may use the integral expression of the remainder terms in the Taylor formula. A slightly different but useful approach is to prove that both the function under consideration and its Taylor series (which must be shown to have positive radius of convergence) satisfy the same differential equation, with the corresponding initial conditions; then $f$ is analytic because of the unicity of solutions to a Cauchy problem.

Also, it is useful to remember that if $f$ and $g$ have power series expansions centered at $z_0$, then so do $f+g$ and $fg$. And if $f(z_0) \neq 0$, the function $1/f$ also has a power series expansion centered at $z_0$.

---

[18] The same question, for a function of a complex variable, turns out to have a completely different, and much simpler, answer: if $f$ is differentiable — in the complex sense — on the open set of definition, then it is always analytic. See Chapter 4.

[19] By induction, proving that $f^{(n)}(x)$ is for $x \neq 0$ of the form $x \mapsto Q_n(x) f(x)$, for some rational function $Q_n$.

# 1.5

# A quick look at asymptotic and divergent series

## 1.5.a Asymptotic series

**DEFINITION 1.76 (Asymptotic expansion)** Let $F$ be a function of a real or complex variable $z$, defined for all $z$ with $|z|$ large enough. The function $F$ **has an asymptotic expansion** if there exists a sequence $(a_n)_{n\in\mathbb{N}}$ of complex numbers such that

$$\lim_{z\to\infty} z^N \left\{ F(z) - \sum_{n=0}^{N} \frac{a_n}{z^n} \right\} = 0$$

for any positive integer $N$. This is denoted

$$F(z) \underset{z\to\infty}{\sim} \sum_{n=0}^{\infty} \frac{a_n}{z^n}. \tag{1.4}$$

The definition means that the expansion (1.4) is a good approximation for large values of $z$. Indeed, if we only consider the first twenty terms of the series, for instance, we see that the sum of those approximates $f(z)$ "to order $1/z^{20}$ at least" when $[z \to \infty]$.

However, it frequently turns out that *for fixed* $z$, the behavior of the series in (1.4) is quite bad as $N \to \infty$. In particular, the series may be divergent. This phenomenon was pointed out and studied in detail by Henri Poincaré in the case of asymptotic series used in astronomy, at the beginning of the twentieth century [70].

How can a divergent asymptotic series still be used? Since $\sum a_n/z^n$ is asymptotic to $F$, if there is some $R$ such that $F$ is continuous for $|z| \geqslant R$, then we see that there exist constants $C_1$, $C_2$,… such that

$$\left| F(z) - \sum_{n=0}^{N} \frac{a_n}{z^n} \right| \leqslant \frac{C_N}{|z|^{N+1}} \qquad \text{for } N \in \mathbb{N} \text{ and } |z| \geqslant R.$$

For fixed $z$, we can look for the value of $N$ such that the right-hand side of this inequality is minimal, and truncate the asymptotic series at this point. Of course, *we do not obtain* $F(z)$ *with infinite precision*. But in many cases the actual precision increases with $|z|$, as described in the next section, and may be pretty good.

It is also possible to speak of **asymptotic expansion** as $z \to 0$, which corresponds to the existence of a sequence $(a_n)_{n\in\mathbb{N}}$ such that

$$\lim_{z\to 0} \frac{1}{z^N} \left\{ f(z) - \sum_{n=0}^{N} a_n z^n \right\} = 0,$$

which is denoted

$$f(z) \underset{z \to 0}{\sim} \sum_{n=0}^{\infty} a_n z^n. \tag{1.5}$$

**Remark 1.77** If it exists, an asymptotic expansion of a function is unique, but there may be two different functions with the same asymptotic expansion! For instance, $e^{-x}$ and $e^{-x^2}$ both have asymptotic expansions with $a_n = 0$ for all $n$ as $x \to +\infty$.

A physical example is given by quantum electrodynamics. This quantum theory of electromagnetic interactions gives physical results in the form of series in powers of the coupling constant $\alpha = e^2/\hbar c \approx 1/137$ (the *Sommerfeld fine structure constant*), which means that a *perturbative expansion in* $\alpha$ is performed:

As shown by Dyson [32], when studying a physical quantity $F$ we can expect to find a perturbative series of the following type (with the normalization $\hbar = c = 1$):

$$F(e^2) = F(\alpha) = \sum_{n=0}^{\infty} f_n \alpha^n.$$

Since the value of $\alpha$ is fixed by Nature, with a value given by experiments, only the truncated series can give a physical result if the series is divergent. The truncation must be performed around the 137-th term, which means that we still expect a very precise result — certainly more precise, by far, than anything the most precise experiment will ever give! However, if $F(e^2)$ is not analytic at $e = 0$, the question is raised whether the asymptotic expansion considered gives access to $F$ uniquely or not.

Studying asymptotic series is in itself a difficult task. Their implications in physics (notably field theory) are at the heart of current research [61].

## 1.5.b Divergent series and asymptotic expansions

Since EULER, CAUCHY (page 88), and especially POINCARÉ (page 475), it has been realized that divergent series may be very useful in physics. As seen in the previous section, they appear naturally in computations of asymptotic series.

As a general rule, convergent series are used to prove numerical or functional identities (between power series, Fourier series, etc). Thus the series may be used instead of the value of their sum at any time in a computation. Another remark is that, from the *computational* viewpoint, some series are more interesting than others, because they converge faster. For example, we have the following two identities for $\log 2$:

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n+1}}{n} + \cdots$$

$$-\log \frac{1}{2} = \log 2 = \frac{1}{2} + \frac{1}{2^2 \cdot 2} + \frac{1}{2^3 \cdot 3} + \cdots + \frac{1}{2^n \cdot n} + \cdots$$

The second of those (which comes from expanding $x \mapsto \log(1 - x)$ in power series at $x = 1/2$) converges much faster than the first (which results from a

Leonhard Euler (1707–1783), a Swiss mathematician, an exceptional teacher, obtained a position at the Academy of Sciences of Saint Petersburg thanks to Nicolas and Daniel Bernoulli when he was only twenty. He also spent some years in Berlin, but came back to Russia toward the end of his life, and died there at seventy-six (while drinking tea). His works are uncountable! We owe him the notations e and i and he imposed the use of $\pi$ that was introduced by Jones in 1706. Other notations due to Euler are sin, cos, tang, cot, sec, and cosec. He also introduced the use of complex exponents, showed that $\mathrm{e}^{\mathrm{i}x} = \cos x + \mathrm{i}\sin x$, and was particularly fond of the formula $\mathrm{e}^{\mathrm{i}\pi} + 1 = 0$. He defined the function $\Gamma$, which extends the factorial function from integers to $\mathbb{C} \setminus (-\mathbb{N})$, and used the Riemann zeta function for real values of the variable. No stone of the mathematical garden of his time was left unturned by Euler; let us only add the Euler angles in mechanics and the Euler equation in fluid mechanics.

similar expansion at $x = -1$, hence on the *boundary* of the disc of convergence[20]).

While studying problems of celestial mechanics, Poincaré realized that the meaning of "convergent series" was not the same for mathematicians, with rigor in mind, or astronomers, interested in efficiency:

> Geometers, preoccupied with rigorousness and often indifferent to the length of the inextricable computations that they conceive, with no idea of implementing them in practice, say that a series is *convergent* when the sum of its terms tends to some well-defined limit, however slowly the first terms might diminish. Astronomers, on the contrary, are used to saying that a series converges when the twenty first terms, for instance, diminish very quickly, even though the next terms may well increase indefinitely. Thus, to take a simple example, consider the two series with general terms $\dfrac{1000^n}{1 \cdot 2 \cdot 3 \cdots n}$ and $\dfrac{1 \cdot 2 \cdot 3 \cdots n}{1000^n}$.
>
> Geometers will say that the first series converges, and even that it converges

[20] The number of terms necessary to approximate $\log 2$ within $10^{-6}$, for instance, can be estimated quite precisely for both series. Using the Leibniz test for alternating sums, the remainder of the first series is seen to satisfy

$$|R_n| \leqslant |u_{n+1}| = \frac{1}{n+1},$$

and this is the right order of magnitude (a pretty good estimate is in fact $R_n \approx 1/2n$). If we want $|R_n|$ to be less than $10^{-6}$, it suffices to take $n = 10^6$ terms. This is a very slow convergence. The remainder of the second series, on the other hand, can be estimated by the remainder of a geometric series:

$$R'_n = \sum_{k=n+1}^{\infty} \frac{1}{n \cdot 2^n} \leqslant \sum_{k=n+1}^{\infty} \frac{1}{2^k} = \frac{1}{2^n}.$$

Hence twenty terms or so are enough to approximate $\log 2$ within $10^{-6}$ using this expansion (since $2^{20} \approx 10^6$).

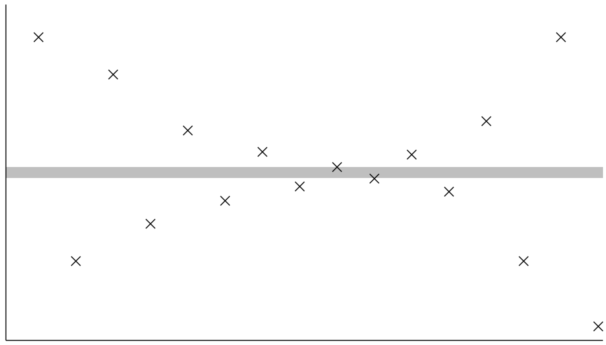


Fig. 1.5 — The precise value of $f(x)$ is always found between two successive values of the partial sums of the serie $\sum f_k(x)$. Hence, it is inside the gray strip.

> rapidly, [...] but they will see the second as divergent.
> Astronomers, on the contrary, will see the first as divergent [...] and the second as convergent. [70]

How is it possible to speak of convergence of a series which is really divergent? We look at this question using a famous example, the **Euler series**. Let

$$\forall x > 0 \qquad f(x) = \int_0^{+\infty} \frac{\mathrm{e}^{-t/x}}{1+t}\,\mathrm{d}t,$$

and say we wish to study the behavior of $f$ for small values of $x$. A first idea is to expand $1/(1+t)$ as $\sum(-1)^k\, t^k$ and exchange the sum and the integral if permitted. Substitute $y = t/x$ and then integrate by parts; a simple induction then shows that

$$\int_0^{+\infty} t^k\,\mathrm{e}^{-t/x}\,\mathrm{d}t = k!\,x^{k+1},$$

and since the $\sum(-1)^k\,k!\,x^{k+1}$ is obviously divergent for any nonzero value of $x$, this first idea is a lamentable failure.

To avoid this problem, it is possible to truncate the power-series expansion of the denominator, and write

$$\frac{1}{1+t} = \sum_{k=0}^{n-1} (-1)^k\, t^k + \frac{(-1)^n\, t^n}{1+t},$$

from which we derive an expression for $f$ of the type $f = f_n + R_n$, where

$$f_n(x) = x - x^2 + 2!\,x^3 - 3!\,x^4 + \cdots + (-1)^{n-1}(n-1)!\,x^n \tag{1.6}$$

and

$$R_n(x) = (-1)^n \int_0^{+\infty} \frac{t^n\,\mathrm{e}^{-t/x}}{1+t}\,\mathrm{d}t.$$

Since $(1+t)^{-1} \leqslant 1$, the remainder satisfies $\left|R_n(x)\right| \leqslant n!\,x^{n+1}$, which means that $R_n(x)$ is of absolute value smaller than the first omitted term; moreover,

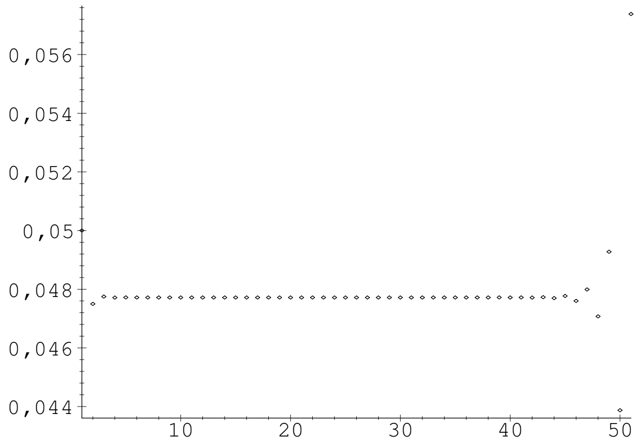



Fig. 1.6 — The first 50 partial sums of the series $\sum(-1)^{k-1}(k-1)!\,x^k$ for $x = 1/20$. Notice that, starting from $k = 44$, the series diverges rapidly. The best precision is obtained for $n = 20$, and gives $f(x)$ with an error roughly of size $2 \cdot 10^{-8}$.

they are of the same sign. It follows (see the proof of the alternating series test) that

$$f_{2n}(x) < f(x) < f_{2n+1}(x), \tag{1.7}$$

although, in contrast with the case of alternating series with terms converging to 0, the general term here $(-1)^n n!\,x^{n+1}$ diverges. Hence it is not possible to deduce from (1.7) an *arbitrarily precise* approximation of $f(x)$. However, *if $x$ is small, we can still get a very good approximation*, as we now explain.

Fix a positive value of $x$. There exists an index $N_0$ such that the distance $\left|f_{2n+1}(x) - f_{2n}(x)\right|$ is smallest (the ratio between consecutive terms is equal to $nx$, so this value of $n$ is in fact $N_0 = \lfloor 1/x \rfloor$). This means that, if we look at the first $N_0$ values, the series "seems to converge," before it starts blowing up. It is interesting to remark that the "convergence" of the first $N_0$ terms is *exponentially fast*, since the minimal distance $\left|f_{N+1}(x) - f_N(x)\right|$ is roughly given by

$$N!\,x^N \approx N!\,N^{-N} \sim \sqrt{2\pi/x}\,\mathrm{e}^{-1/x}$$

(using the Stirling formula, see Exercise 5.4 on page 154.) Thus, if we wish to know the value of $f(x)$ for a "small" value of $x$, and if a precision of the order of $\sqrt{2\pi/x}\,\mathrm{e}^{-1/x}$ suffices, it is possible to use the divergent asymptotic series (1.6), by computing and summing the terms *up to the smallest term* (see Figure 1.5). For instance, we obtain for $x = 1/50$ a precision roughly equal to $6 \cdot 10^{-20}$, which is perfectly sufficient for most physical applications! (see Figure 1.6.)

For a given value of $x$, on the other hand, the asymptotic series does not allow any improvement on the precision.[21] But the convergence is so fast that

[21] For instance, in quantum field theory, the asymptotic series in terms of $\alpha$ has a limited


Sir George Biddel AIRY (1801–1892), English astronomer, is known in particular for discovering the theory of diffraction rings. He determined approximately the solar apex, the direction toward which the sun and the solar system seem to be directed, in the Hercules region. He was also interested in geology. He was director of the Royal Observatory and took part in the controversy concerning priority for the discovery of Neptune (the French pushing the claim of Le Verrier while the English defended Adams).
The picture here represents a fake stamp, painted directly on the envelope, representing a contemporary caricature of Sir Airy; the post office was bluffed and stamped and delivered the letter.

it makes it possible to do some computations which are out of reach of a standard method! And what Poincaré remarked is, in fact, a fairly general rule: *divergent series converge, in general, much more rapidly than convergent series.*

In 1857, George Stokes was studying the Airy integral [3]

$$\mathrm{Ai}(z) \stackrel{\text{def}}{=} \frac{1}{\pi}\int_0^{+\infty} \cos\left(\frac{t^3}{3} + zt\right) \mathrm{d}t,$$

which appears in the computations of caustics. The goal was to find zeros of this function, and compare the with "experimental" zeros (corresponding to dark bands in any optics figure, which had been measured with great precision, at least as far as the first twenty-five). Airy himself, using a convergent series expansion of Ai at 0, managed fairly easily to compute the position of the first band, and with considerable difficulty, found the second one. In fact, his mathematically convergent expansion was "divergent" in the sense of astronomers (all the more so as one gets farther away from the origin). Stokes used instead the "devilish" method of divergent series[22] and, after bypassing some nontrivial difficulties (linked, in particular, to complex integration), obtained *all* the hands[23] with a precision of $10^{-4}$!

**Remark 1.78** There are other well-known techniques to give a sense to the sum of (some) divergent series. The interested reader may read the classic book of Émile Borel [13], the first part of which at least is very readable. Concerning asymptotic expansions, see [72].

precision since $\alpha$ is fixed (equal to $1/137$ approximately) and cannot be made to tend to zero. This suggests that quantum field theory, in its current perturbative form, will one day be replaced by another theory. Of course, as long as a precision to $10^{-100}$ is "enough"...

[22] Niels ABEL wrote in 1826 that divergent series are "the Devil's invention, and it is shameful to base any proof of any kind on them. By using them, one can get from them whatever result is sought: they have done much evil and caused many paradoxes" (letter to his professor Holmboë).

[23] Only the first is less precise, because it is too small and Stokes used an asymptotic expansion at $+\infty$.

# EXERCISES

## Physical "paradoxes"

♦ **Exercise 1.2 (Electrical energy)** Consider an electric circuit consisting of two identical capacitors in series, with capacitance $C$ and resistance $R$. Suppose that for $t \leqslant 0$, the circuit is open, one of the capacitors carries the charge $Q$, and the other has no charge. At $t = 0$, the circuit is closed, and is left to evolve freely. What is the state of equilibrium for this circuit? What is the energy of the system at $t = 0$? What is the energy as $t \to +\infty$? Show that the missing energy depends only on $R$. What happened to this energy?

Now assume that $R = 0$. What is the energy of the system at any arbitrary $t$? What is the limit of this energy as $t \to +\infty$? Do you have any comments?

♦ **Exercise 1.3 (A paradox in optics)** We know that two distinct sources of monochromatic light do not create a clear interference picture in an experiment with Young slits. As the distance between the sources increases, we first see a contrast decrease in the interference picture. This is called a **defect of spatial coherence**.

Hence, a famous experiment gives a measurement of the angular distance between two components of a double star by the observation of the first disappearance of the interference fringes when slowly moving two Young slits apart.

This experiment works very well with monochromatic light. However, if we define two monochromatic sources $S_1$ and $S_2$ mathematically, each emits a signal proportional to $\mathrm{e}^{2\mathrm{i}\pi\nu t}$, and there should be no problem of spatial coherence.

Perform the computation properly. A computation in optics *always* starts with amplitudes (possibly, one may show that the crossed terms cancel out in average, and do the computations with intensity only). Here, the cross terms are fine, and never disappear. In other words, this shows that two different monochromatic light sources are always perfectly coherent.

But experiment shows the opposite: a defect of spatial coherence. How can this be explained?

♦ **Exercise 1.4** In the rubber ball paradox of page 2, give an interpretation of the variation of kinetic energy of the ball, in the moving reference frame, in terms of the work of the force during the rebound. The shock may be modeled by a very large force lasting a very short amount of time, or one can use the formalism of distributions (see Chapter 7).

## Sequences and series

♦ **Exercise 1.5** It is known that $\mathbb{Q}$, and hence also $\mathbb{Q} \cap [0,1]$, is countable. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of rational numbers such that $\mathbb{Q} \cap [0,1] = \{x_n \ ; \ n \in \mathbb{N}\}$. Show that the sequence $(x_n)_{n\in\mathbb{N}}$ diverges.

♦ **Exercise 1.6** In an arbitrary normed vector space, show that a Cauchy sequence which has a convergent subsequence is convergent.

♦ **Exercise 1.7** Show that the space $\mathbb{K}[X]$ of polynomials with coefficients in $\mathbb{K}$ is not complete with the norm given by

$$P = \sum_{i=1}^{n} \alpha_i X^i \longmapsto \|P\| \stackrel{\text{def}}{=} \max_{1\leqslant i\leqslant n} |\alpha_i| .$$

♦ **Exercise 1.8 (Fixed point)** Let $a, b \in \mathbb{R}$ be real numbers with $a < b$, and let $f : [a,b] \to [a,b]$ be a continuous function with a fixed point $\ell$. Assume that there exists a real number

$\lambda$ and an interval $V = [\ell - \varepsilon, \ell + \varepsilon]$ around $\ell$, contained in $[a, b]$ and stable under $f$ (i.e., $f(x) \in V$ if $x \in V$), such that

$$\forall x \in V \qquad \left|f(x) - f(\ell)\right| \leqslant \lambda \left|x - \ell\right|^2.$$

i) Let $\alpha \in V$ be such that $\lambda(\alpha - \ell) < 1$. Let $u$ be the sequence defined by induction by

$$u_0 = \alpha, \qquad u_{n+1} = f(u_n) \quad \text{for all } n \in \mathbb{N}.$$

Show that $(u_n)_{n\in\mathbb{N}}$ converges to $\ell$.

ii) Show in addition that $|u_n - \ell| \leqslant \lambda^{-1} \cdot \big(\lambda(\alpha - \ell)\big)^{2^n}$.

♦ **Exercise 1.9** Let

$$f_n(x) = \begin{cases} 2n^3 x & \text{if } 0 \leqslant x \leqslant 1/2n, \\ n^2 - 2n^3\,(x - 1/2n) & \text{if } 1/2n \leqslant x \leqslant 1/n, \\ 0 & \text{if } 1/n \leqslant x \leqslant 1, \end{cases}$$

for $n \in \mathbb{N}$ and $x \in [0, 1]$. Plot a graph of $f$, and compute

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x \qquad \text{and} \qquad \int_0^1 \lim_{n\to\infty} f_n(x)\,\mathrm{d}x.$$

♦ **Exercise 1.10** Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions converging simply to a function $f$. If each $f_n$ is increasing, show that $f$ is also increasing. Show that the same stability holds for the properties "$f_n$ is convex" and "$f_n$ is $k$-Lipschitz." Show that, on the other hand, it is possible that each $f_n$ is continuous, but $f$ is not (take $f_n(x) = \sin^{2n} x$).

♦ **Exercise 1.11** Let $\varphi_n$ be the function defined on $[-1, 1]$ by

$$\varphi_n(x) = \int_0^x \left(1 - \mathrm{e}^{-1/nt^2}\right) \mathrm{d}t$$

for $n \in \mathbb{N}^*$.

Show that $\varphi_n$ is infinitely differentiable, and that the sequence $(\varphi_n)_{n\in\mathbb{N}^*}$ converges uniformly on $[-1, 1]$. What is its limit?

Let $\varepsilon > 0$ be given. Show that for any $p \in \mathbb{N}$, there exists a map $\Psi_p$ from $[-1, 1]$ into $\mathbb{R}$, infinitely differentiable, such that

i) $\Psi_p^{(k)}(0) = 0$ for $k \neq p$, and $\Psi_p^{(p)}(0) = 1$.

ii) for $k \leqslant p - 1$ and $x \in [-1, 1]$, $\quad \left|\Psi_p^{(k)}(x)\right| \leqslant \varepsilon$.

Now let $(a_n)_{n\in\mathbb{N}}$ be an arbitrary sequence of real numbers. Construct an infinitely differentiable map $f$ from $[-1, 1]$ to $\mathbb{R}$ such that $f^{(n)}(0) = a_n$ for all $n \in \mathbb{N}$.

♦ **Exercise 1.12 (Slightly surprising exercise)** Construct a series of functions $\sum f_n$ defined on $\mathbb{R}$, which converges pointwise to a sum $F(x)$, the convergence being uniform on any finite interval of $\mathbb{R}$, and which, moreover, satisfies:

$$\forall n \in \mathbb{N} \qquad \lim_{x\to+\infty} f_n(x) = +\infty$$

but

$$\lim_{x\to+\infty} F(x) = -\infty.$$

♦ **Exercise 1.13** Consider a power series centered at the point $a \in \mathbb{C}$, given by

$$f(z) \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} c_n (z-a)^n.$$

Let $R$ denote its radius of convergence, and assume $R > 0$.

i) Prove the **Cauchy formula**: for any $n \in \mathbb{N}$ and any $r \in ]0, R[$, we have

$$c_n = \frac{1}{2\pi\, r^n} \int_0^{2\pi} f(a + r\,\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{e}^{-\mathrm{i}n\theta}\,\mathrm{d}\theta.$$

ii) Prove the **Gutzmer formula**: for $r \in ]0, R[$, we have

$$\sum_{n=0}^{\infty} |c_n|^2\, r^{2n} = \frac{1}{2\pi} \int_0^{2\pi} \left| f(a + r\,\mathrm{e}^{\mathrm{i}\theta}) \right|^2 \mathrm{d}\theta.$$

iii) Prove that if $R = +\infty$, in which case the sum $f(z)$ of the power series is said to be an *entire function*, and if moreover $f$ is bounded on $\mathbb{C}$, then $f$ is constant (this is Liouville's theorem, which is due to Cauchy).

iv) Is the sine function a counter-example to the previous result?

♦ **Exercise 1.14** Let $f$ be a function of $\mathscr{C}^\infty$ class defined on an open set $\Omega \subset \mathbb{R}$. Show that $f$ is analytic if and only if, for any $x_0 \in \Omega$, there are a neighborhood $\mathscr{V}$ of $x_0$ and positive real numbers $M$ and $t$ such that

$$\forall x \in \mathscr{V} \quad \forall p \in \mathbb{N} \qquad \left| \frac{f^{(p)}(x)}{p!} \right| \leqslant M\, t^p.$$

## Functions of two variables

♦ **Exercise 1.15** Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function of two real variables. This exercise gives examples showing that the limits

$$\lim_{x\to 0} \lim_{y\to 0} f(x,y) \qquad \lim_{y\to 0} \lim_{x\to 0} f(x,y) \qquad \text{and} \qquad \lim_{(x,y)\to(0,0)} f(x,y)$$

are "independent": each may exist without the other two existing, and they may exist without being equal.

i) Let $f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{if } x^2+y^2 \neq 0, \\ 0 & \text{if } x = y = 0. \end{cases}$

Show that the limits $\lim\limits_{x\to 0} \lim\limits_{y\to 0} f(x,y)$ and $\lim\limits_{y\to 0} \lim\limits_{x\to 0} f(x,y)$ both exist, but that the limit[24] $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ is not defined.

---

[24] The limit of $f$ as the *pair* $(x,y)$ tends to a value $(a,b) \in \mathbb{R}^2$ is defined using any of the natural norms on $\mathbb{R}^2$, for instance the norm $\|(x,y)\|_\infty = \max\big(|x|, |y|\big)$ or the euclidean norm $\|(x,y)\|_2 \stackrel{\text{def}}{=} \sqrt{x^2+y^2}$, which are equivalent. Thus, we have

$$\lim_{\substack{(x,y)\to(a,b)\\(x,y)\neq(a,b)}} f(x,y) = \ell$$

if and only if

$$\forall \varepsilon > 0 \quad \exists \eta > 0 \qquad \big((x,y) \neq (a,b) \text{ and } \big\|(x-a, y-b)\big\|_\infty < \eta\big) \Longrightarrow \big(|f(x,y) - \ell| \leqslant \varepsilon\big).$$

ii) Let $f(x,y) = \begin{cases} y + x\sin(1/y) & \text{if } y \neq 0, \\ 0 & \text{if } y = 0. \end{cases}$

Show that both limits $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ and $\lim\limits_{y\to 0}\lim\limits_{x\to 0} f(x,y)$ exist, but on the other hand $\lim\limits_{x\to 0}\lim\limits_{y\to 0} f(x,y)$ does not exist.

iii) Let $f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} + y\sin\left(\dfrac{1}{x}\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$

Show that $\lim\limits_{\substack{x\to 0\\ x\neq 0}}\lim\limits_{\substack{y\to 0\\ y\neq 0}} f(x,y)$ exists. Show that neither $\lim\limits_{(x,y)\to(0,0)} f(x,y)$, nor $\lim\limits_{\substack{y\to 0\\ y\neq 0}}\lim\limits_{x\to 0} f(x,y)$ exist.

iv) Let $f(x,y) = \begin{cases} \dfrac{x^2-y^2}{x^2+y^2} & \text{if } x^2+y^2 \neq 0, \\ 0 & \text{if } x = y = 0. \end{cases}$

Show that the limits $\lim\limits_{x\to 0}\lim\limits_{y\to 0} f(x,y)$ and $\lim\limits_{y\to 0}\lim\limits_{x\to 0} f(x,y)$ both exist, but are different.

## PROBLEM

♦ **Problem 1 (Solving differential equations)** The goal of this problem is to illustrate, in a special case, the Cauchy-Lipschitz theorem that ensures the existence and unicity of the solution to a differential equation with a given initial condition.

In this problem, $I$ is an interval $[0,a]$ with $a > 0$, and we are interested in the nonlinear differential equation

$$y' = \frac{ty}{1+y^2} \tag{E}$$

with the initial condition

$$y(0) = 1. \tag{CI}$$

The system of two equations (E) + (CI) is called the **Cauchy problem**. In what follows, $E$ denotes the space $\mathscr{C}(I,\mathbb{R})$ of real-valued continuous functions defined on $I$, with the norm $\|f\|_\infty = \sup\limits_{t\in I} |f(t)|$.

i) Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\left(E, \|\cdot\|_\infty\right)$.

(a) Show that for any $x \in I$ the sequence $\left(f_n(x)\right)_{n\in\mathbb{N}}$ converges in $\mathbb{R}$. For $x \in I$, we let $f(x) \stackrel{\text{def}}{=} \lim\limits_{n\to\infty} f_n(x)$.

(b) Show that $(f_n)_{n\in\mathbb{N}}$ converges uniformly $f$ on $I$.

(c) Show that the function $f : I \to \mathbb{R}$ is continuous.

(d) Deduce from this that $\left(E, \|\cdot\|_\infty\right)$ is a complete normed vector space.

ii) For any $f \in E$, define a function $\Phi(f)$ by the formula

$$\Phi(f) : I \longrightarrow \mathbb{R},$$
$$t \longmapsto \Phi(f)(t) = 1 + \int_0^t \frac{u\,f(u)}{1+\left(f(u)\right)^2}\,\mathrm{d}u.$$

Show that the functions $f \in E$ which are solutions of the Cauchy problem (E) + (CI) are exactly the fixed points of $\Phi$.

iii) Show that the function $x \longmapsto \dfrac{x}{1+x^2}$ is 1-Lipschitz on $\mathbb{R}$, i.e.,

$$\left|\frac{y}{1+y^2} - \frac{x}{1+x^2}\right| \leqslant |y-x|\,. \tag{1.8}$$

iv) Show that $\Phi$ is a contracting map if $a$ is sufficiently small.

v) Show that there exists a unique solution to the Cauchy problem. Give an explicit iterative method to solve the system numerically (**Picard iterations**).

**Remark 1.79** In general, all this detailed work need not be done: the Cauchy-Lipschitz theorem states that for any continous function $\psi(x, y)$ which is locally Lipschitz with respect to the second variable, the Cauchy problem

$$y' = \psi(t, y),$$

has a unique maximal solution (i.e., a solution defined on a maximal interval).

## SOLUTIONS

♦ **Solution of exercise 1.2.** The energy of the circuit at the beginning of the experiment is the energy contained in the charged capacitor, namely $E = Q^2/2C$. At equilibrium, when $[t \to \infty]$, no current flows, and the charge of each capacitor is $Q/2$ (it is possible to write down the necessary differential equations and solve them to check this). Thus the final energy is $E' = 2(Q/2)^2/C = E/2$. The energy which is dissipated by the Joule effect (computed by the integral $\int_0^{+\infty} Ri^2(t)\,\mathrm{d}t$, where $t \mapsto i(t)$ is the current flowing through the circuit at time $t$) is of course equal to $E - E'$, and does not depend on $R$.

However, if $R = 0$, one observes oscillations of charge in each capacitor. The total energy of the system is conserved (it is not possible to compute it from relations in a quasi-stationary regime; one must take magnetic fields into account!). In particular, as $[t \to +\infty]$, the initial energy is recovered. The explanation for this apparent contradiction is similar to what happened for Romeo and Juliet: the time to reach equilibrium is of order $2/RC$ and tends to infinity as $[R \to 0]$. This is a typical situation where the limits $[R \to 0]$ and $[t \to +\infty]$ do not commute.

Finally, if we carry the computations even farther, it is possible to take into account the electromagnetic radiation due to the variations of the electric and magnetic fields. There is again some loss of energy, and for $[t \to +\infty]$, the final energy $E - E' = E/2$ is recovered.

♦ **Solution of exercise 1.3.** Light sources are never purely monochromatic; otherwise there would indeed be no spatial coherence problem. What happens is that light is emitted in wave packets, and the spectrum of the source necessarily has a certain width $\Delta\lambda > 0$ (in a typical example, this is order of magnitude $\Delta\nu = 10^{14}\ \mathrm{s}^{-1}$, corresponding to a coherence length of a few microns for a standard light-bulb; the coherence length of a small He-Ne laser is around thirty centimeters, and that of a monomode laser can be several miles). All computations must be done first with $\Delta\lambda \neq 0$ before taking a limit $\Delta\lambda \to 0$. Thus, surprisingly, spatial coherence is *also* a matter of temporal coherence. This is often hidden, with the motto being "since the sources are not coherent, I must work by summing intensities instead of amplitudes."

In fact, when considering an interference figure, one must always sum amplitudes, and then (this may be a memory from your optics course, or an occasion to read Born and Wolf [14]) perform a *time average* over a period $\Delta t$, which may be very small, but not too much (depending on the receptor; the eyes are pretty bad in this respect, an electronic receptor is better, but none can have $\Delta t = 0$).

The delicate issue is to be careful of a product $\Delta t \cdot \Delta\lambda$. If you come to believe (wrongly!) that two purely monochromatic sources interfere without any spatial coherence defect, this means that you have assumed $\Delta t \cdot \Delta\lambda = 0$. To see the spatial coherence issue arise, one must keep $\Delta\lambda$ large enough so that $\Delta t \cdot \Delta\lambda$ cannot be neglected in the computation.

◆ **Solution of exercise 1.7**. Let $P_n = \sum_{k=1}^{n} X^k/k$. It is easy to see that $(P_n)_{n\in\mathbb{N}}$ is a Cauchy sequence: for any integers $k$ and $p$, we have $\left\|P_{p+k} - P_p\right\| = \frac{1}{p+1}$.

However, this sequence does not converge in $\mathbb{K}[X]$; indeed, if it were to converge to a limit $L$, we would have $L \in \mathbb{K}[X]$, and all coefficients of $L$ of degree large enough ($\geqslant N$, say) would be zero, which implies that $\|P_k - L\| \geqslant 1/(\deg L + 1)$ if $k \geqslant N$, contradicting that $(P_n)_{n\in\mathbb{N}}$ converges to $L$.

This example shows that $\mathbb{K}[X]$ is not complete.

◆ **Solution of exercise 1.9**. For any $n \in \mathbb{N}$, we have $\int f_n = n/2$, and for any $x \in [0,1]$, the sequence $\big(f_n(x)\big)_n$ tends to 0, showing that

$$\lim_{n\to\infty} \int f_n(x)\,\mathrm{d}x = +\infty \qquad \text{whereas} \qquad \int \Big(\lim_{n\to\infty} f_n(x)\Big)\,\mathrm{d}x = 0.$$

◆ **Solution of exercise 1.11**. The sequence $(\varphi_n)_{n\in\mathbb{N}}$ converges uniformly to 0.

Notice also the property

$$\varphi_n(0) = 0, \qquad \varphi_n'(0) = 1, \qquad \varphi_n^{(k)}(0) = 0 \qquad \forall k \geqslant 2.$$

For given $\varepsilon > 0$, it suffices to define $\Psi_p$ as the $(p-1)$-st primitive of $\varphi_N$, where $N$ is sufficiently large so that $\sup_{x\in[-1,1]} \left|\psi_p^{(p-1)}(x)\right| \leqslant \varepsilon$. Here, each primitive is selected to be the one vanishing at 0 (i.e., the integral from 0 to $x$ of the previous one). It is easy to see that the successive derivatives of this function satisfy the require condition, and the last property follows from the construction.

Now let $(a_n)_{n\in\mathbb{N}}$ be an arbitrary sequence of real numbers. For all $n \in \mathbb{N}$, one can apply the previous construction to find $\Psi_n$ such that

$$\sup_{x\in[-1,1]} \left|\psi_n^{(n-1)}(x)\right| \leqslant \frac{1}{2^n \cdot \max(1, |a_n|)}.$$

It is then immediate that the series $\sum a_n \Psi_n$ converges uniformly to a function $f$ having all desired properties.

Of course, the function $f$ thus constructed is by no means unique: one may add a term $\alpha(\varphi_n' - 1)$, where $\alpha \in \mathbb{R}$, without changing the values of the derivatives at 0.

◆ **Solution of exercise 1.12**. Let

$$f_n(x) = -\frac{x^{4n-1}}{(4n-1)!} + \frac{x^{4n+1}}{(4n+1)!}$$

for $n \geqslant 1$; the series $\sum f_n$ converges to $F(x) = \sin x - x$.

◆ **Solution of exercise 1.13**

i) The power series for $f(a + r\,\mathrm{e}^{\mathrm{i}\theta})$ may be integrated term by term because of its absolute convergence in the disc centered at $a$ of radius $r < R$. Since we have

$$\int_0^{2\pi} \mathrm{e}^{\mathrm{i}(k-n)\theta}\,\mathrm{d}\theta = \delta_{kn} = \begin{cases} 1 & \text{if } k = n, \\ 0 & \text{otherwise,} \end{cases}$$

the stated formula follows.

ii) Similarly, expand $\left|f(a+r\,e^{i\theta})\right|^2$ as a product of two series and integrate term by term. Most contributions cancel out using the formula above, and only the terms $|c_n|^2\,r^{2n}$ remain.

iii) If $f$ is bounded on $\mathbb{C}$, we have $|c_n\,r^n| \leqslant \|f\|_\infty$. Letting $r\to+\infty$, it follows that $c_n=0$ for $n\geqslant 1$, which means that $f$ is constant.

iv) The function sin is *not* bounded on $\mathbb{C}$! Indeed, we have for instance $\lim\limits_{x\to+\infty}\left|\sin(ix)\right| = +\infty$. So there is no trouble.

### ♦ Solution of problem 1

i) (a) Let $x\in I$. For any $p,q\in\mathbb{N}$, we have

$$\left|f_p(x)-f_q(x)\right| \leqslant \sup_{y\in I}\left|f_p(y)-f_q(y)\right| = \left\|f_p-f_q\right\|_\infty,$$

and this proves that the sequence $\big(f_n(x)\big)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$, so it converges.

(b) Let $\varepsilon>0$ be fixed. There exists $N$ such that $\left\|f_p-f_q\right\|_\infty \leqslant \varepsilon$ for all $p>q>N$. Let $x\in I$. We then have

$$\left|f_p(x)-f_q(x)\right| \leqslant \varepsilon \qquad \text{for any } p>q>N,$$

and since this holds for all $p$, we may fix $q$ and let $p\to\infty$. We obtain

$$\left|f(x)-f_q(x)\right| \leqslant \varepsilon \qquad \text{for all } q>N.$$

This bound holds independently of $x\in I$. Thus we have shown that

$$\left\|f-f_q\right\|_\infty \leqslant \varepsilon \qquad \text{for any } q\geqslant N.$$

Finally, this being true for any $\varepsilon>0$, it follows that the sequence $(f_n)_{n\in\mathbb{N}}$ converges uniformly to $f$.

***Remark:*** *At this point, we haven't proved that there is convergence in the normed vector space $\big(E,\|\cdot\|_\infty\big)$. It remains to show that the limit $f$ is in $E$, that is, that $f$ is continuous. This follows from Theorem 1.33, but we recall the proof.*

(c) Let $x\in I$; we now show that $f$ is continuous at $x$.

Let $\varepsilon>0$ be fixed. From the preceding question, there exists an integer $N$ such that $\|f_n-f\|_\infty \leqslant \varepsilon$ for all $n\geqslant N$, and in particular $\|f_N-f\|_\infty\leqslant\varepsilon$. Since $f_N$ is an element of $E$, it is continuous. So there exists $\eta>0$ such that

$$\forall y\in I \qquad |y-x|\leqslant\eta \Longrightarrow \left|f_N(y)-f_N(x)\right|\leqslant\varepsilon.$$

Using the triangle inequality, we deduce from this that for all $y\in I$ such that $|x-y|\leqslant\eta$, we have

$$\left|f(y)-f(x)\right| \leqslant \left|f(y)-f_N(y)\right| + \left|f_N(y)-f_N(x)\right| + \left|f_N(x)-f(x)\right| \leqslant 3\varepsilon.$$

This proves the continuity of $f$ at $x$, and since $x$ is arbitrary, this proves that $f$ is continuous on $I$, and hence is an element of $E$.

(d) For any Cauchy sequence $(f_n)_{n\in\mathbb{N}}$ in $E$, the previous questions show that $(f_n)_{n\in\mathbb{N}}$ converges in $E$. Hence the space $\big(E,\|\cdot\|_\infty\big)$ is complete.

ii) Let $f$ be a fixed point of $\Phi$. Then we have

$$\forall t\in I \qquad f'(t) = \big(\Phi(f)\big)'(t) = \frac{t\,f(t)}{1+\big(f(t)\big)^2}.$$

Moreover, it is easy to see that $f(0)=\Phi(f)(0)=1$.

Conversely, let $f$ be a solution of the Cauchy problem $(E)+(CI)$. Then $\Phi(f)$ is differentiable and we have

$$\forall t \in I \qquad \Phi(f)'(t) = \frac{t\, f(t)}{1+\big(f(t)\big)^2} = f'(t).$$

The functions $f$ and $\Phi(f)$ have the same derivative on $I$, and moreover satisfy

$$\Phi(f)(0) = 1 \qquad \text{and} \qquad f(0) = 1.$$

It follows that $\Phi(f) = f$.

iii) Looking at the derivative $g' : x \mapsto \dfrac{1-x^2}{(1+x^2)^2}$ we see that $\left|g'(x)\right| \leqslant 1$ for all $x \in \mathbb{R}$. The mean-value theorem then proves that $g$ is 1-Lipschitz, as stated.

iv) Let $f, g \in E$. Then we have

$$\begin{aligned}
\left\|\Phi(f)-\Phi(g)\right\|_\infty &= \sup_{t\in I} \left| \int_0^t \left( \frac{u\, f(u)}{1+\big(f(u)\big)^2} - \frac{u\, g(u)}{1+\big(g(u)\big)^2} \right) \mathrm{d}u \right| \\
&\leqslant \sup_{t\in I} \int_0^t u \left| \frac{f(u)}{1+\big(f(u)\big)^2} - \frac{g(u)}{1+\big(g(u)\big)^2} \right| \mathrm{d}u \\
&\leqslant \int_0^a u \left| \frac{f(u)}{1+\big(f(u)\big)^2} - \frac{g(u)}{1+\big(g(u)\big)^2} \right| \mathrm{d}u
\end{aligned}$$

by positivity. Using the inequality of the previous question, we get

$$\left\|\Phi(f)-\Phi(g)\right\|_\infty \leqslant \int_0^a u \left|f(u)-g(u)\right| \mathrm{d}u \leqslant \|f-g\|_\infty \int_0^a u\, \mathrm{d}u = \frac{a^2}{2}\, \|f-g\|_\infty .$$

This is true for any $f, g \in E$, and hence $\Phi$ is $(a^2/2)$-Lipschitz; if $0 \leqslant a < \sqrt{2}$, this map $\Phi$ is a contraction.

v) According to the fixed-point theorem, the previous results show that $\Phi$ has a unique fixed point in $E$.

According to Question ii), this means that there exists a unique solution of the Cauchy Problem $(E)+(CI)$ on an interval $[0,a]$ for $a < \sqrt{2}$.

To approximate the solution numerically, it is possible to select an arbitrary function $f_0$ (for instance, simply $f_0 = 0$), and construct the sequence $(f_n)_{n\in\mathbb{N}}$ defined by $f_{n+1} = \Phi(f_n)$ for $n \geqslant 0$. This requires computing (numerically) some integrals, which is a fairly straightforward matter (numerical integration is usually numerically stable: errors do not accumulate in general[25]). The speed of convergence of the sequence $(f_n)_{n\in\mathbb{N}}$ to the solution $f$ of the Cauchy problem is exponential: with $I = [0,1]$, the distance (from the norm on $E$) between $f_n$ and $f$ is divided by 2 (at least) after each iterative step. It is therefore possible to expect a good numerical approximation after few iterations (the precision after ten steps is of the order of $\|f_0 - f\|_\infty / 1000$ since $2^{10} = 1024$).

[25] On the other hand, numerical differentiation tends to be much more delicate.