My Horror Chamber

Some unexpected examples in Calculus and Mathematics

Leif Mejlbro

bookboon
The eBook company

LEIF MEJLBRO

MY HORROR CHAMBER

My Horror Chamber
1st edition

ISBN 978-87-403-1563-9

CONTENTS

List of Figures

Preface

Many years ago when I was teaching a class in Functional Analysis at the Technical University of Denmark, I was quite often challenged by my students who would ask: *"If we make the following assumptions is it then possible to conclude some desired result not in the textbooks?"* The answer was usually "no", though sometimes the questions also struck some unexpected positive results. Anyway, I always had to prepare a proof or counter-example for the following lecture. In this way a large collection of strange examples in Mathematics grew up. One of my students, Eske Rahn, called this collection *The Horror Chamber of Mathematics*. This explains the title of this book.

The purpose of this book is to show that Mathematics and even Calculus can be very difficult to handle with lots of unexpected results, so one cannot just rely on some standardized exercises. I still remember my first encounter with such a strange example, which was not in my elementary books in Mathematics. I had bought the newly published book of *Riesz & Sz.–Nagy* [47], in which I for the first time came across a *singular function* in the form of a continuous and strictly increasing function, which is differentiable "almost everywhere", and whenever it is differentiable, its differential quotient is 0. In other words, if we first differentiate this function, and then take the Lebesque integral of the result, we just get 0 and not the original function! This means that differentiation and integration are not always the inverse of each other. Such examples force the reader to think in an alternative way instead of using the standard procedures. Without going into details, it can be mentioned that it is the operation of *differentiation*, which is the villain and not the Lebesque integral. Unfortunately, the traditional teaching gives the impression that differentiation is easy, while integration is an art, to quote a professor from one of his lectures in the early 1960s. It should be added that the singular functions were known long before the 1960s. The subject was not included in the usual elementary courses.

Such weird examples can be used for inspiration of thinking in a new and different way. One example is the clearly convergent integral

$$\int_0^{\frac{\pi}{2}} \frac{\mathrm{d}x}{1+\tan^\alpha x} \qquad \text{for any fixed } \alpha \in \mathbb{R}.$$

It is not difficult to calculate the integral for $\alpha = 0$, 1, 2, and the result is in all cases the same, namely $\pi/4$. A head-on approach in the general case is doomed to fail, and computer programs like MAPLE are also in trouble, though one gets a result as a decimal fraction. We need an idea, and the idea is to change variable to $t = \pi/2 - x$ and then use the symmetry. In fact,

$$\begin{aligned}\int_0^{\pi/2} \frac{\mathrm{d}x}{1+\tan^\alpha x} &= \int_0^{\pi/2} \frac{\mathrm{d}x}{1+\cot^\alpha x} = \int_0^{\pi/2} \frac{\tan^\alpha x}{\tan^\alpha x+1}\,\mathrm{d}x \\ &= \frac{1}{2}\int_0^{\pi/2}\left\{\frac{1}{1+\tan^\alpha x}+\frac{\tan^\alpha x}{1+\tan^\alpha x}\right\}\mathrm{d}x = \frac{1}{2}\int_0^{\pi/2} 1\,\mathrm{d}x = \frac{\pi}{4}.\end{aligned}$$

One may of course add this result to the catalogue of known integrals in MAPLE, but then one just constructs another example which MAPLE does not recognize and which can be solved by some nonstandard mathematical argument. The point is that Kurt Gödel in 1931 proved his famous theorem which implies that there will always exist mathematical problems which the given version of MAPLE cannot solve. The reader should therefore be aware of that sometimes the standard procedures of solving problems are not sufficient for a given problem under consideration.

But even if one is not looking for some example which can be applied in some context, this collection could be of some interest – maybe even fun – for some readers. It is classical Calculus and Mathematics, and yet giving quite unexpected results, like e.g. “a continuous curve which passes through every point of the unit square and which is almost nowhere almost everywhere” (and “almost everywhere almost nowhere”).

It is impossible to give a full list of all such strange examples. The reader is referred to the References which is larger than necessary, so one may search other examples in the university libraries.

November 9, 2016

Leif Mejlbro

1 Elementary theory of functions

1.1 Introduction

In this chapter we give some examples which only involve "simple functions". Some of them are extremely elementary and they only show that one shall always be very meticulous when checking the assumptions.

In the first example I have deliberately "forgotten" an assumption, which many readers will tacitly assume to be fulfilled. We only use that a function can have relative minima everywhere without being continuous. In the following example we also lack a crucial assumption.

The next two examples will probably be more surprising for the reader. In the first example we construct a function which is discontinuous everywhere in the set of rational numbers $\mathbb{Q}$, and continuous everywhere in the complementary set $\mathbb{R} \setminus \mathbb{Q}$ of irrational numbers.

Then we construct a sequence of functions, which in some sense "converges" towards *Dirac's δ-function*", and which clearly does not converge pointwise towards *Dirac's δ-function*". This example gives us a warning that the two natural definitions of convergence in fact are different, and they are useful in different situations which may have nothing to do with each other. This shows that one should always be very precise when we specify some definition, e.g. of convergence.

We then construct a continuous function, which is not differentiable at any point. We shall later in Chapter 2 meet some other bizarre examples of continuous functions, concerning their differentiability properties. However, the present example will for the time being be bad enough.

We continue discussing an hypothesis, which intuitively "must be correct". Then we demonstrate that when we try to prove it, we really get into trouble! Therefore we go to plan B, which is to assume the negation of the original hypothesis . . . and then again end up with another mathematical mess! In this way we may continue with this mathematical ping-pong without getting any further, unless we get a quite different idea. I shall of course for illustration finally solve the problem. Although only Calculus known from high school is applied it is seen that the proof becomes far from obvious.

In the next example we must assume some elementary knowledge of *Complex Functions Theory*. Notice that we in one of the constructions implicitly are using *Banach's fix point theorem*. We learn from this example that Taylor series are not always the best tool for approximation, which also follows from one of the previous examples, where we constructed a continuous function, which is not differentiable at any point. In this case it does not make sense to talk about the corresponding *Taylor series*. It simply does not exist. Instead we may use the classical and important *Weierstraß's approximation theorem*

Many years ago one of my friends asked me if some functional equation did only have one class of solutions, depending on a single arbitrary constant. This would indeed be desirable from at statistical point of view. Unfortunately, this hypothesis was wrong without some additional assumptions, so the task became to find such additional conditions, which assures that the functional equation indeed had only the desired family of solutions. This has been included here as an illustrative example. The project was quite a task, though not as bad as an exercise many years ago during my first year of study. My professor asked us to find the error in some given hypothesis, then to correct the statement and finally prove the right theorem. First year students! One would not challenge students today with exercises of that sort.

Finally, we include a vicious example of a linear partial differential equation of polynomial coefficients and a C^∞-function as its inhomogeneous term. The point is that this equation does not have a solution, not even in the sense of *Schwartz's distributions* (generalized functions). By a small twist of this example we may even obtain linear partial differential operators, which are *injective*, which should violate one's intuition.

1.2 A nonconstant real function with relative minima everywhere

The headline looks like a paradox because "if a function has a relative minimum everywhere it cannot be growing, so it must be a constant". In order to resolve this enigma we go back to analyzing the definition of a relative minimum at a given point $x_0 \in \mathbb{R}$. By definition, there exists an $\varepsilon > 0$, such that

$$f(x) \geq f(x_0) \qquad \text{for all } x \in \,]\, x_0 - \varepsilon \, , \, x_0 + \varepsilon \, [\, .$$

Then note that we in this definition *do not require that the function f is continuous!* This gives us the idea of searching a discontinuous function of this property. One extreme example of a nonconstant function with a relative minimum at every point $x \in \mathbb{R}$ is then given by

$$f(x) = n, \qquad \text{for } x \in]n-1, n] \quad \text{and} \quad n \in \mathbb{Z}.$$

It can be proved (left to the reader) that if $f : \mathbb{R} \to \mathbb{R}$ has a relative minimum at every real point $x \in \mathbb{R}$, then the range

$$\{f(x) \mid x \in \mathbb{R}\}$$

is at most countable, i.e.

$$\{f(x) \mid x \in \mathbb{R}\} = \{y_n \mid n \in \mathbb{N}\}$$

for some sequence (y_n) (or finite set).

1.3 A continuous absolutely integrable real function which does not tend towards 0 at infinity

For some unknown reason, many students believe that if a function is continuous and absolutely integrable, then it must converge towards 0 for $x \to +\infty$ (or $x \to -\infty$, the construction of a counter-example being similar). This is in general not true. One is probably misled by the usual catalogue of nice functions which are dominant in the elementary Calculus courses. This will "almost always" tend towards 0 for $x \to +\infty \, (-\infty)$.

A simple counter-example is given by

$$f(x) = \begin{cases} n \quad \text{for } x = n, & \text{for } n \in \mathbb{N}, \\[2ex] 0 \quad \text{for } x = n \pm 2^{-n}, & \text{for } n \in \mathbb{N}, \\[2ex] \text{piecewise linear.} \end{cases}$$

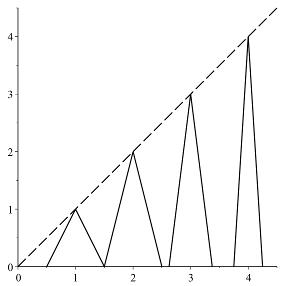

Figure 1.1: *The graph of a continuous and absolutely integrable real function on $\mathbb{R}$ which does not tend towards 0 for $x \to +\infty$.*

It is obvious that f is continuous and nonnegative, and since by taking the restriction of f to $\mathbb{N}$ we get the sequence

$$f(n) = n \to +\infty \qquad \text{for } n \to +\infty,$$

we only have to check that $\int_0^{+\infty} f(x)\,\mathrm{d}x < +\infty$.

Using the interpretation of the integral of f as the area between its graph and the x-axis, we get

$$\int_0^{n+\frac{1}{2}} f(x)\,\mathrm{d}x = \sum_{j=1}^{n} j \cdot 2^{-j} \to \sum_{n=1}^{+\infty} n \cdot 2^{-n} = 2 < +\infty,$$

because when $x \in]-1,1[$,

$$\frac{x}{(1-x)^2} = x\,\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1-x}\right) = x\,\frac{\mathrm{d}}{\mathrm{d}x}\left\{\sum_{n=0}^{+\infty} x^n\right\} = x\sum_{n=1}^{+\infty} n\,x^{n-1} = \sum_{n=1}^{+\infty} n \cdot x^n.$$

We therefore get for $x = \frac{1}{2}$,

$$\sum_{n=1}^{+\infty} n\left(\frac{1}{2}\right)^n = \frac{\frac{1}{2}}{\left(1-\frac{1}{2}\right)^2} = 2.$$

We then add the missing assumption and prove the following theorem

Theorem 1.1 *Every* uniformly *continuous and absolutely integrable function $f : \mathbb{R} \to \mathbb{R}$ tends towards 0 for $x \to +\infty$.*

Proof. Assume that f is uniformly continuous and that $f(x)$ does not tend towards 0 for $x \to +\infty$. We shall prove that f is not absolutely integrable, because then the claim follows by contraposition.

By assumption $f(x)$ does not tend towards 0 for $x \to +\infty$. This means that we can find an $\varepsilon > 0$, such that for all $N > 0$ there is an $x_N \geq N$, such that

$$|f(x_N)| \geq \varepsilon. \tag{1.1}$$

Also, by assumption, f is uniformly continuous. Then, corresponding to the chosen $\varepsilon > 0$ above there is a $\delta > 0$, such that for $x, y \in \mathbb{R}$,

$$(1.2) \qquad \text{if } |x-y| \leq \delta, \text{ then } |f(x) - f(y)| \leq \frac{\varepsilon}{2}.$$

Choose any $x_1 \in \mathbb{R}_+$ satisfying (1.1), i.e. $|f(x_1)| \geq \varepsilon$. Then by (1.2),

$$(1.3) \qquad |f(y)| \geq \frac{\varepsilon}{2} \qquad \text{for all } y \in [x_1 - \delta, x_1 + \delta],$$

so

$$(1.4) \qquad \int_{x_1-\delta}^{x_1+\delta} |f(y)| \, \mathrm{d}y \geq \varepsilon \cdot \delta > 0.$$

Then construct by induction an increasing sequence (x_n), such that

$$x_{n+1} \geq x_n + 2\delta \qquad \text{and} \qquad |f(x_n)| \geq \varepsilon.$$

Since (1.3) and (1.4) hold for every x_n, we trivially get the estimate

$$\int_{-\infty}^{+\infty} |f(y)| \, \mathrm{d}y \geq \sum_{n=1}^{+\infty} \int_{x_n-\delta}^{x_n+\delta} |f(y)| \, \mathrm{d}y \geq \varepsilon \cdot \delta \sum_{n=1}^{+\infty} 1 = +\infty,$$

proving that f is not absolutely integrable. We conclude by contraposition that if f is absolutely integrable, then either f is not uniformly continuous, or $f(x) \to 0$ for $x \to +\infty$, and the theorem is proved. $\square$

1.4 A real function, which is discontinuous at every rational number, and continuous at every irrational number

It is well-known – or easy to prove – that the function

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{for } x \in \mathbb{Q}, \\ \\ 0 & \text{for } x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases}$$

is discontinuous at every point in $\mathbb{R}$.

We modify the construction above by defining

$$f(x) = \begin{cases} 1, & \text{for } x = 0, \\ \\ \dfrac{1}{q}, & \text{for } x = \dfrac{p}{q},\ p \in \mathbb{Z} \setminus \{0\},\ q \in \mathbb{N},\ q \not| \, p, \\ \\ 0, & \text{for } x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases}$$

where the symbol $q \not| \, p$ means that q is not a divisor in p.

We shall prove that f is discontinuous in $\mathbb{Q}$ and continuous in $\mathbb{R} \setminus \mathbb{Q}$.

Assume that $x_0 \in \mathbb{Q}$. Then $f(x_0) = 1/q > 0$ for some $q \in \mathbb{N}$, and we see that for every $0 < \varepsilon < 1/q$ and every $\delta > 0$ there is an $x \in \mathbb{R}$, where $|x_0 - x| < \delta$, such that

$$|f(x_0) - f(x)| = \frac{1}{q} - 0 = \frac{1}{q} > \varepsilon.$$

In fact, just choose x as any irrational number in the interval $]x_0 - \delta, x_0 + \delta[$. This proves that f is discontinuous at every rational number $x_0 \in \mathbb{Q}$.

Then assume that $x_0 \in \mathbb{R} \setminus \mathbb{Q}$. For every given $\varepsilon > 0$ we can find $q_0 \in \mathbb{N}$, such that

$$0 < \frac{1}{q} < \varepsilon \qquad \text{for all } q > q_0.$$

We define for every fixed $j \in \mathbb{N}$,

$$\delta_j := \min \left\{ \left| x_0 - \frac{m}{j} \right| \;\middle|\; m \in \mathbb{Z},\ j \not| \, m \right\}.$$

The points of the set $\{m/j \mid m \in \mathbb{Z}\}$ lie for neighbouring points at the mutual distance $1/j$ in $\mathbb{Q}$. Since $x_0 \notin \mathbb{Q}$, it follows that $\delta_j > 0$ for all $j \in \mathbb{N}$. Define

$$\delta := \min \{\delta_j \mid j = 1, \dots, q_0\} > 0,$$

(a minimum of a *finite* set of positive numbers, hence $\delta > 0$), where q_0 was given above. For every $x \in \mathbb{R}$, for which $|x - x_0| < \delta$, we get

$$|f(x) - f(x_0)| < \varepsilon.$$

In fact, if x is irrational, then the left hand side is 0, and if x is rational, then

$$|f(x) - f(x_0)| = \frac{1}{q} - 0 = \frac{1}{q} < \varepsilon \qquad \text{for all } q > q_0,$$

because then $1/q < \varepsilon$.

Summing up, for every irrational x_0 and every chosen $\varepsilon > 0$ there is a $\delta > 0$, such that

$$|f(x) - f(x_0)| < \varepsilon, \qquad \text{whenever } |x - x_0| < \delta.$$

This argument shows that f is continuous a every irrational point.

Remark 1.1 It follows from the theory of the *Lebesgue integral* that the constructed function f above is *Riemann integrable* of integral 0. It is left to the reader to prove this claim by using the definition of the Riemann integral over the restriction of f to the unit interval $[0, 1]$. $\Diamond$

1.5 A sequence of functions (φ_n), which converges towards Dirac's δ "function" in the "sense of distributions", and at the same time pointwisely towards $-\infty$ at $x = 0$, and towards 0 for $x \neq 0$.

At the first glance this looks strange, because Dirac's δ "function" in some sense is *nonnegative*, and yet the pointwise limit is $-\infty$ at $x = 0$. This shows that the two concepts of convergency are not at all equal.

We shall first define what is meant by *convergency in the sense of distributions*. We say that a sequence of real function (φ_n) converges towards Dirac's δ "function" in the sense of distributions, if

$$\text{(1.5)}\qquad \lim_{n\to+\infty}\int_{-\infty}^{+\infty} f(t)\varphi_n(t)\,\mathrm{d}t = f(0)$$

for every *continuous* function f, which is 0 outside a bounded interval on the real line.

The reason for this definition is that Dirac's δ is not a function in the ordinary sense of the word. The idea is that the functions φ_n represent a "mass" 1, which for $n \to +\infty$ is concentrated in a very small interval containing 0. Since f is assumed to be continuous, it is almost constant in this small interval of the value $f(0)$, so the integral is close to $f(0)$. So instead of this mysterious element δ one may use any φ_n for sufficiently large n as an approximation of this ideal element.

One usually uses the following construction of the possible sequences (φ_n). Choose a frequency φ of some probability function, i.e.

$$\varphi(t) \geq 0 \quad \text{for all } x \in \mathbb{R}, \qquad \text{and} \qquad \int_{-\infty}^{+\infty}\varphi(t)\,\mathrm{d}t = 1.$$

Then define the sequence by

$$\text{(1.6)}\qquad \varphi_n(t) := n\,\varphi(n\,t) \qquad \text{for all } t \in \mathbb{R} \text{ and all } n \in \mathbb{R}.$$

Then by a change of variable,

$$\int_{-\infty}^{+\infty}\varphi_n(t)\,\mathrm{d}t = \int_{-\infty}^{+\infty} n\,\varphi(n\,t)\,\mathrm{d}t = \int_{-\infty}^{+\infty}\varphi(t)\,\mathrm{d}t = 1,$$

where the mass of φ_n is concentrated in a small interval containing 0.

The point in this section is that φ in this construction does not have to be a frequency. We shall only need that φ satisfies $\int_{-\infty}^{+\infty}\varphi(t)\,\mathrm{d}t = 1$. We shall later in Theorem 1.2 below formulate some conditions on φ which assure that the sequence (φ_n) given by (1.6) also satisfies (1.5). However, in order to become familiar with this construction we first give some examples.

The simplest construction uses

$$\varphi(t) = \chi_{\left[-\frac{1}{2},\frac{1}{2}\right]}(t) = \begin{cases} 1 & \text{for } -\dfrac{1}{2} \leq x \leq \dfrac{1}{2}, \\ 0, & \text{otherwise,} \end{cases}$$

which clearly is a frequency. Then

$$\varphi_n(t) = n \cdot \chi_{\left[-\frac{1}{2n},\frac{1}{2n}\right]}(t) = \begin{cases} n & \text{for } -\dfrac{1}{2n} \leq x \leq \dfrac{1}{2n}, \\ 0 & \text{otherwise,} \end{cases}$$

which again is a frequency, and we see that the mass is concentrated in the small interval $\left[-\frac{1}{2n}, \frac{1}{2n}\right]$.

It is obvious that the pointwise limit is

$$(1.7) \qquad \Phi(t) := \lim_{n\to+\infty} \varphi_n(t) = \begin{cases} +\infty & \text{for } t \to +\infty, \\ 0 & \text{otherwise.} \end{cases}$$

This gives us a clue where to search for more strange examples. If we e.g. consider

$$\varphi(t) = \chi_{[-1,-\frac{1}{2}]}(t) + \chi_{[\frac{1}{2},1]}(t),$$

then this φ is again a frequency. The usual argument shows that for fixed $t \neq 0$ the sequence $\varphi_n(t) \to 0$ for $n \to +\infty$, and since $\varphi_n(0) = 0$ for all $n \in \mathbb{N}$, it follows in this case that

$$\Phi(t) := \lim_{n\to+\infty} \varphi_n(t) = 0 \qquad \text{for all } t \in \mathbb{R},$$

so already here we obtain a "strange" behaviour of the chosen sequence of functions.

If we skip the claim that φ should be a frequency, a good candidate would be

$$\tilde{\varphi}(t) = \frac{1}{2}\,\chi_{[-3,-1]}(t) - \frac{1}{2}\,\chi_{[-1,1]}(t) + \frac{1}{2}\,\chi_{[1,3]}(t).$$

In this case

$$\int_{-\infty}^{+\infty} |\tilde{\varphi}(t)|\,\mathrm{d}t = 3, \qquad \text{and} \qquad \int_{-\infty}^{+\infty} \tilde{\varphi}(t)\,\mathrm{d}t = 1,$$

so $\tilde{\varphi}$ satisfies the conditions of Theorem 1.2 below.

Since $\tilde{\varphi}_n(t) = 0$ for $|t| > \frac{3}{n}$, it is obvious that for $t \neq 0$ fixed,

$$\lim_{n\to+\infty} \tilde{\varphi}(t) = 0,$$

and since $\tilde{\varphi}(0) = -n/2$, we clearly get

$$\lim_{n\to+\infty} \tilde{\varphi}(0) = -\infty,$$

and we have constructed our example, provided that Theorem 1.2 holds.

Theorem 1.2 *Let φ be a (measurable) real function satisfying*

$$\int_{-\infty}^{+\infty} |\varphi(t)|\,dt < +\infty \qquad \textit{and} \qquad \int_{-\infty}^{+\infty} \varphi(t)\,dt = 1.$$

We define the sequence of concentrated sequences (φ_n) by

$$\varphi_n(t) = n \cdot \varphi(n \cdot t), \qquad \textit{for } n \in \mathbb{N}.$$

Then for every fixed continuous real function f, which is zero outside a bounded interval,

$$\lim_{n\to+\infty} \int_{-\infty}^{+\infty} f(t)\,\varphi_n(t)\,dt = f(0) \quad [= \delta(f)].$$

PROOF. We shall only prove the claim, when φ is Riemann integrable.

Assume that f is a continuous function which is 0 outside a bounded set. Then f is also bounded (one of the main theorems in Mathematics), so $|f(t)| \leq M$ for some $M > 0$, which is fixed in the following, and all $t \in \mathbb{R}$.

Clearly,

$$\begin{aligned}\int_{-\infty}^{+\infty} f(t)\varphi_n(t)\,\mathrm{d}t &= f(0)\int_{-\infty}^{+\infty}\varphi_n(t)\,\mathrm{d}t + \int_{-\infty}^{+\infty}\{f(t)-f(0)\}\varphi_n(t)\,\mathrm{d}t \\ &= f(0) + \int_{-\infty}^{+\infty}\{f(t)-f(0)\}\varphi_n(t)\,\mathrm{d}t,\end{aligned}$$

so the theorem follows, if we can prove

$$(1.8)\qquad \int_{-\infty}^{+\infty}\{f(t)-f(0)\}\varphi_n(t)\,\mathrm{d}t = \int_{-\infty}^{+\infty}\left\{f\left(\frac{y}{n}\right)-f(0)\right\}\varphi(y)\,\mathrm{d}y \to 0 \qquad \text{for } n \to +\infty,$$

where the first equality follows from the change of variable, $y = nt$.

Let $\varepsilon > 0$ be given. From $\int_{-\infty}^{+\infty}|\varphi(t)|\,\mathrm{d}t < +\infty$ follows that we can find a constant $a > 0$, such that

$$(1.9)\qquad \int_{-\infty}^{-a}|\varphi(t)|\,\mathrm{d}t + \int_{a}^{+\infty}|\varphi(t)|\,\mathrm{d}t < \frac{\varepsilon}{8M}.$$

Then we use that f is continuous, so there exists an $\eta > 0$, such that

$$(1.10) \qquad |f(t)-f(0)| \int_{-\infty}^{+\infty} |\varphi(y)|\,\mathrm{d}y < \frac{\varepsilon}{2}, \qquad \text{whenever } |t| \le \eta.$$

Choose $N \in \mathbb{N}$, such that $N \ge a/\eta$. Then $0 < a/n \le \eta$ for all $n \geqq N$. We therefore get for $n \ge N$,

$$\begin{aligned}
\left|\int_{-\infty}^{+\infty} \{f(t)-f(0)\}\varphi_n(t)\,\mathrm{d}t\right| &= \left|\int_{-\infty}^{+\infty} \left\{f\left(\frac{y}{n}\right)-f(0)\right\}\varphi(y)\,\mathrm{d}y\right| \\
&\le \left|\int_{-\infty}^{-a} \left\{f\left(\frac{y}{n}\right)-f(0)\right\}\varphi(y)\,\mathrm{d}y\right| + \left|\int_{-a}^{a} \left\{f\left(\frac{y}{n}\right)-f(0)\right\}\varphi(y)\,\mathrm{d}y\right| \\
&\qquad + \left|\int_{a}^{+\infty} \left\{f\left(\frac{y}{n}\right)-f(0)\right\}\varphi(y)\,\mathrm{d}y\right| \\
&\le \int_{-\infty}^{-a} \left|f\left(\frac{y}{n}\right)-f(0)\right| \cdot |\varphi(y)|\,\mathrm{d}y + \int_{a}^{+\infty} \left|f\left(\frac{y}{n}\right)-f(0)\right| \cdot |\varphi(y)|\,\mathrm{d}y \\
&\qquad + \int_{-a}^{a} \left|f\left(\frac{y}{n}\right)-f(0)\right| \cdot |\varphi(y)|\,\mathrm{d}y \\
&\le 2M \int_{-\infty}^{-a} |\varphi(y)|\,\mathrm{d}y + 2M \int_{a}^{+\infty} |\varphi(y)|\,\mathrm{d}y + \max_{|t|\le\eta} |f(t)-f(0)| \cdot \int_{-a}^{a} |\varphi(y)|\,\mathrm{d}y \\
&\le 2 \cdot 2M \cdot \frac{\varepsilon}{8M} + \frac{\varepsilon}{2} = \varepsilon,
\end{aligned}$$

where we have applied (1.9) and (1.10). Then it follows from the estimate above that also (1.8) holds, which again implies the theorem. □

Remark 1.2 A closer look at the proof above shows that we only need that the function f is continuous at 0. The tedious technical details ar left to the reader. ◊

It is not difficult to obtain the same phenomenon starting with some properly chosen *continuous* φ. One simple example is given by the graph of Figure 1.2.

One can also construct a real analytic function φ, such that the derived sequence (φ_n) has the same two properties, namely that $(\varphi_n) \to \delta$ in the sense of distributions, while $\varphi_n(0) \to -\infty$ for $n \to +\infty$ and $(\varphi_n(t)) \to 0$ for $n \to +\infty$ for all fixed $x \ne 0$. One such example is given by

$$\varphi^a(t) := \begin{cases} \dfrac{\sin^2 t}{t^2} - \dfrac{\pi-1}{a\pi} \cdot \dfrac{\sin^2(at)}{t^2}, & \text{for } t \ne 0, \\[2ex] 1 - a \cdot \dfrac{\pi-1}{\pi}, & \text{for } t = 0. \end{cases}$$

It is obvious that we must choose the constant $a > \pi/(\pi-1)$ in order to get $\varphi_n^a(0) \to -\infty$ for $n \to +\infty$. The graph of φ^2 is given on Figure 1.3.

Since the proof requires some knowledge of Complex Functions Theory, I shall only sketch the proof and leave it to the interested reader to fill in all the missing details.

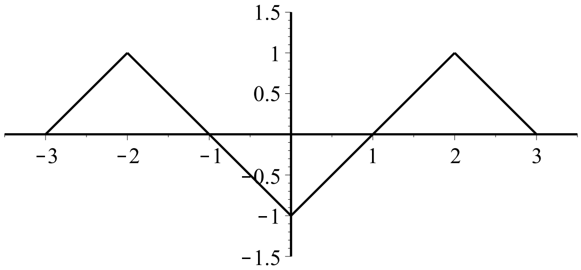

Figure 1.2: *The graph of a continuous function φ, for which $(\varphi_n) \to \delta$ in the sense of distributions, while $(\varphi_n(t))$ converges pointwisely towards $-\infty$ for $t = 0$, and towards 0 otherwise.*

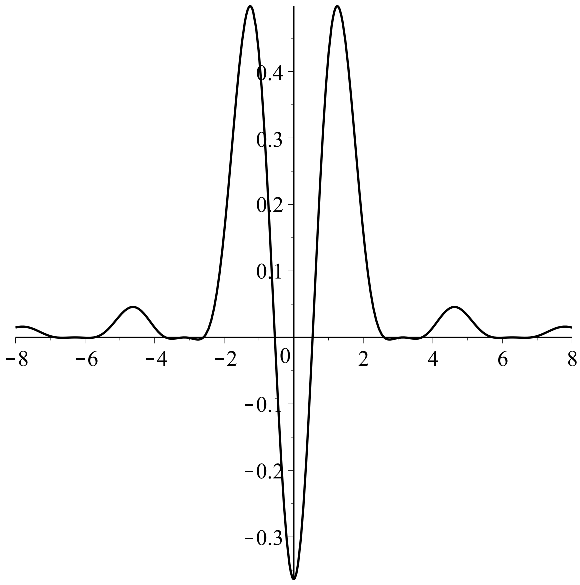

Figure 1.3: *The graph of the real analytic function φ^2, where we have chosen $a = 2 > \pi/(\pi - 1)$, such that $\left(\varphi_n^2\right) \to \delta$ in the sense of distributions, while $\left(\varphi_n^2(t)\right)$ converges pointwisely towards $-\infty$ for $t = 0$, and towards 0 otherwise.*

SKETCH OF PROOF.

1) It is easy to prove, using a Laurent series expansion for $t \neq 0$ and $t_0 = 0$, that $\varphi^a(t)$ is real analytic with a removable singularity at $t = 0$.

2) We then prove that φ^a is absolutely integrable, $\int_{-\infty}^{+\infty} |\varphi^a(t)| \, \mathrm{d}t < +\infty$. This follows from the following estimates, where all the integrands are trivially nonnegative,

$$\begin{aligned}\int_{-\infty}^{+\infty} \frac{\sin^2 t}{t^2} \, \mathrm{d}t &= \int_{-\infty}^{-\pi} + \int_{-\pi}^{\pi} + \int_{\pi}^{+\infty} \frac{\sin^2 t}{t^2} \, \mathrm{d}t \\ &\leq \int_{-\infty}^{-\pi} \frac{1}{t^2} \, \mathrm{d}t + \int_{-\pi}^{\pi} \frac{\sin^2 t}{t^2} \, \mathrm{d}t + \int_{\pi}^{+\infty} \frac{1}{t^2} \, \mathrm{d}t < +\infty,\end{aligned}$$

because $\sin^2 t / t^2$ can be considered as a continuous function on the closed and bounded interval $[-\pi, \pi]$ with the value 1 at the singular point $t = 0$, so all three integrals are positive and bounded.

3) The condition of 2) assures that we can apply some residuum formula to derive (a small trick is used here) that

$$\int_{-\infty}^{+\infty} \frac{\sin^2 t}{t^2} \, \mathrm{d}t = \pi,$$

which implies that

$$\begin{aligned}\int_{-\infty}^{+\infty} \varphi^a(t) \, \mathrm{d}t &= \int_{-\infty}^{+\infty} \frac{\sin^2 t}{t^2} \, \mathrm{d}t - \frac{\pi - 1}{\pi} \int_{-\infty}^{+\infty} a \cdot \frac{\sin^2(at)}{(at)^2} \, \mathrm{d}t \\ &= \left(1 - \frac{\pi - 1}{\pi}\right) \int_{-\infty}^{+\infty} \frac{\sin^2 t}{t^2} \, \mathrm{d}t = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\sin^2 t}{t^2} \, \mathrm{d}t = 1.\end{aligned}$$

4) Since for $a > \pi/(\pi - 1)$

$$\varphi_n^a(t) = \begin{cases} \dfrac{\sin^2 nt}{nt^2} - \dfrac{\pi - 1}{\pi a n} \cdot \dfrac{\sin^2(ant)}{t^2}, & \text{for } t \neq 0, \\[2ex] 1 - a \cdot \dfrac{\pi - 1}{\pi} < 0, & \text{for } t = 0. \end{cases},$$

it follows from 2) and 3) that $\varphi_n^a \to \delta$ in the sense of distributions, while it follows from 4) that

$$\lim_{n \to +\infty} \varphi_n^a(t) = \begin{cases} -\infty & \text{for } t = 0, \\ 0 & \text{for } t \neq 0. \end{cases}$$

The examples above should convince the reader that convergency in the sense of distributions is not the same as the more familiar pointwise convergency. For the same reason one should avoid function notations as

$$\text{“}\delta(t)\text{''} = \begin{cases} +\infty & \text{for } t = 0, \\ 0 & \text{for } t \neq 0, \end{cases}$$

of Dirac's δ "function", because it is *not* a function! Instead one should use either *Lighthill's* definition, which is preferred by engineers, because you start with ordinary functions, or apply the more abstract *Theory of Distributions*, developped by *Laurent Schwartz* [49], which is preferred by mathematicians.

We have above used Lighthill's approach, i.e. we have represented Dirac's δ function by a sequence (φ_n) of functions, where

$$\lim_{n\to+\infty} \int_{-\infty}^{+\infty} f(t)\varphi_n(t)\,\mathrm{d}t = f(0),$$

whenever f is absolutely integrable and continuous at $t = 0$.

Laurent Schwartz is using a more abstract setup, where δ is defined as a *functional* defined on the space of all continuous (absolutely integrable) functions f by

$$\delta(f) := f(0).$$

The difference between a function ψ and a functional, like δ, is that the function is defined on points in e.g. $\mathbb{R}$, while a functional is defined on some set of *functions*, and its *values* are numbers.

1.6 A continuous function which is not differentiable anywhere

The continuous functions, which we usually consider, are also differentiable, with only probably the exception of some points, where the corresponding graphs have bends of cusps. We are guided by the dominance of the real analytic functions in our daily life, and compositions of such functions, to intuitively believe that this is correct in general, and Ampère [3] tried as early as in 1806 to prove this statement, however, in vain.

The problem was solved by *Karl Weierstraß* in 1861. He proved that every function of the form

$$f(x) = \sum_{n=0}^{+\infty} \frac{1}{a^n} \sin\left(b^n x\right), \qquad \text{for } x \in \mathbb{R},$$

where the constants a, b satisfy

$$a,\, b > 1 \qquad \text{and} \qquad b > \left(\frac{3\pi}{2} + 1\right) a,$$

is continuous and not differentiable anywhere.

We have in Figure 1.4 shown what MAPLE does to the finite sum

$$\sum_{n=0}^{10} 2^{-n} \sin\left(11^n x\right),$$

where $11 > (1 + 3\pi/2) \cdot 2$. Clearly, something is wrong, though we have not proved by this graph that the function is not differentiable. The function pictured in Figure 1.4 is of course of class C^∞, because it is a finite sum of C^∞ functions.

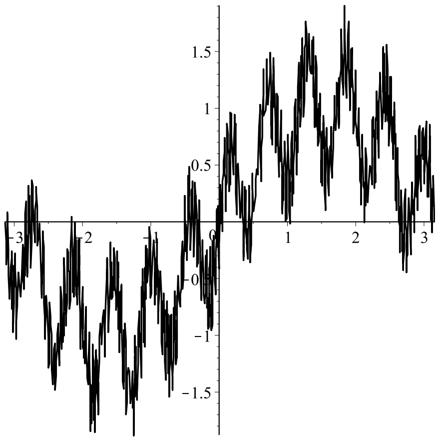

Figure 1.4: *The graph of the function* $\sum_{n=0}^{10} 2^{-n} \sin\left(11^n x\right)$, *where* $x \in [-\pi, \pi]$.

That such a function is continuous follows from $a > 1$, and from that the series is uniformly convergent,

$$\left| \sum_{n=0}^{+\infty} a^{-n} \sin\left(b^n x\right) \right| \leq \sum_{n=0}^{+\infty} a^{-n} = \frac{a}{a-1} < +\infty.$$

It is more difficult to prove that it is nowhere differentiable. We shall here not go into details, since we below choose a simpler construction. For some unknown reason Weierstraß's example was first published by *du Bois-Reymond* in 1875.

The construction chosen in the following was developed by *van der Waerden* [52] in 1930. The same construction can also be found in *Riesz, Sz.-Nagy* [47].

In order to avoid any confusion we shall here emphasize that we only consider differentiation in the classical sense, and not using other concepts of differentiation like "weak differentiation" or "differentiation in the sense of distributions". We define

> A continuous function f is differentiable at a point x, where $f(x)$ is defined and finite, if the difference quotient
>
> $$\frac{f(x+h)-f(x)}{h} \tag{1.11}$$
>
> is convergent for $h \to 0$, when $h \neq 0$ goes through the set, where $f(x+h)$ is defined. In this case we define the *differential quotient* as
>
> $$f'(x) := \lim_{\substack{h \to 0 \\ f(x+h) \text{ defined}}} \frac{f(x+h)-f(x)}{h}.$$

When we construct a continuous function f which is not differentiable anywhere, we shall "only" check (1.11) for every fixed x.

We first define an auxiliary function $\varphi : \mathbb{R} \to \left[0, \frac{1}{2}\right]$ by

$$\varphi(x) = \operatorname{dist}(x, \mathbb{Z}) := \min_{y \in \mathbb{Z}} |x-y|, \qquad x \in \mathbb{R}.$$

This definition means that $\varphi(x)$ denotes the distance from x to the set $\mathbb{Z}$.

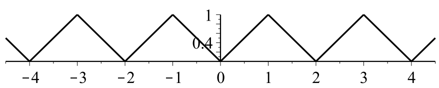

Figure 1.5: *The graph of the saw tooth function $\varphi(x)$.*

Then we define a real function $f : \mathbb{R} \to \mathbb{R}$ by

$$(1.12) \qquad f(x) := \sum_{n=0}^{+\infty} \frac{1}{10^n} \varphi(10^n x) \qquad \text{for } x \in \mathbb{R}.$$

Every single term of this series is continuous, and we have the trivial estimate

$$|f(x)| \leq \sum_{n=0}^{+\infty} 10^{-n} = \frac{10}{9} \qquad \text{for all } x \in \mathbb{R}.$$

Hence, (1.12) is uniformly convergent, so f is indeed a continuous function on $\mathbb{R}$.

We shall then prove that the difference quotient (1.11) does not have a limit for $h \to 0$ for any point $x \in \mathbb{R}$. Since f is defined for all $x \in \mathbb{R}$, it suffices to prove the claim for just one sequence (h_m), where the definition of this sequence depends on the point x under consideration.

Clearly, f is periodic of period 1, so it suffices to consider $x \in [0, 1[$ and afterwards use the periodicity. When $x \in [0, 1[$, we can write x as a uniquely determined decimal fraction

$$(1.13) \qquad x = 0.a_1 a_2 \cdots a_n \cdots , \qquad \text{where } a_n \in \{0, 1, \ldots, 9\},$$

where we shall never allow a representation, where for some e.g. $a_N \neq 9$, while $a_n = 9$ for all $n > N$. In fact, such a representation can be replaced by another one, where a_N is replaced by $a_N + 1 \in \{1, 2, \ldots, 9\}$, and $a_n = 0$ for all $n > N$.

For given $n \in \mathbb{N}_0$ we have two possibilities: Either

$$a) \ \ 0.a_{n+1} a_{n+2} \cdots \leq \frac{1}{2}, \qquad \text{or} \qquad b) \ \ 0.a_{n+1} a_{n+2} \cdots > \frac{1}{2}.$$

We get respectively,

$$a)\ \ \varphi\left(10^m x\right) = 0.a_{n+1}a_{n+2}\cdots \qquad \text{and} \qquad b)\ \ \varphi\left(10^n x\right) = 1 - 0.a_{n+1}a_{n+2}\cdots .$$

Fix $x \in [0,1[$ with the uniquely determined decimal fraction (1.13). We choose (h_m) in the following way,

$$h_m = \begin{cases} -10^{-m}, & \text{if } a_m \in \{4,9\}, \\ +10^{-m}, & \text{if } a_m \in \{0,1,2,3,5,6,7,8\}. \end{cases}$$

The change of sign of h_m, when $a_m \in \{4,9\}$, implies that if $n < m$, then $\varphi\left(10^n\left(x+h_m\right)\right)$ and $\varphi\left(10^n x\right)$ lie under the same "half roof" of the graph of the function $\varphi\left(10^n y\right)$. Hence,

$$\text{(1.14)} \qquad \varphi\left(10^n\left(x+h_m\right)\right) - \varphi\left(10^n x\right) = \pm 10^n\, h_m \qquad \text{for } n < m,$$

where the sign depends on x, m and n.

The property (1.14) is apparently not sufficient to prove that (1.11) does not converge for the function f given by (1.12). However, if we consider the difference quotient (1.11) for f, then it follows that

$$\begin{aligned} \frac{1}{h_m}\left\{f\left(x+h_m\right) - f(x)\right\} &= \frac{1}{h_m}\sum_{n=0}^{+\infty} 10^{-n}\left\{\varphi\left(10^n\left(x+h_m\right)\right) - \varphi\left(10^n x\right)\right\} \\ &= \frac{1}{h_m}\sum_{n=0}^{m-1} 10^{-n}\left\{\varphi\left(10^n\left(x+h_m\right)\right) - \varphi\left(10^n x\right)\right\} \\ &= \frac{1}{h_m}\sum_{n=0}^{m-1} 10^{-n}\left\{\pm 10^n h_m\right\} = \sum_{n=0}^{m-1}(\pm 1). \end{aligned}$$

Here, the second equality follows from the fact that φ is periodic of period 1, so $\varphi\left(10^n\left(x+h_m\right)\right) = \varphi\left(10^n x\right)$ for $n \geq m$. The third equality follows by an insertion of (1.14).

Summing up, we conclude that

$$\frac{1}{h_m}\left\{f\left(x+h_m\right) - f(x)\right\} = \sum_{n=0}^{m-1}(\pm 1) = \begin{cases} \text{an even number for } m \text{ even}, \\ \text{an odd number for } m \text{ odd}. \end{cases}$$

It follows then immediately that the limit does not exist for $m \to +\infty$. Since $x \in \mathbb{R}$ was chosen arbitrarily, we have proved that the function f given by (1.12) is not differentiable anywhere in the classical sense.

Remark 1.3 Since f is a continuous and bounded function, it can be proved that the derivation of f does exist *in the sense of distributions*, i.e. generalized functions. We shall not go further into this theory. $\Diamond$

1.7 Let $f : \mathbb{R}_+ \to \mathbb{R}$ be continuous and assume that $f(a_n) \to 0$ for $n \to +\infty$ for every equidistant sequence (a_n) on $\mathbb{R}_+$. Is it possible to conclude that $f(x) \to 0$ for $x \to +\infty$?

This is an old problem which belongs to the folklore in Mathematics. I have not been able to find the origin of it. In the light of the previously constructed strange functions one would be inclined to intuitively guess that this is not possible. Wrong again! The claim is actually true. We shall here follow *L. Mejlbro* [36].

We shall prove

Theorem 1.3 *Assume that $f : \mathbb{R}_+ \to \mathbb{R}$ is continuous and that*

$$(1.15) \qquad f(na) \to 0 \qquad \text{for } n \to +\infty, \quad n \in \mathbb{N},$$

for every fixed $a > 0$. Then

$$f(x) \to 0 \qquad \text{for } x \to +\infty, \quad x \in \mathbb{R}_+.$$

PROOF. Assume that $f : \mathbb{R}_+ \to \mathbb{R}$ is continuous and fulfils (1.15). We shall prove that whenever $\varepsilon > 0$ is given, we can find an $N \in \mathbb{R}_+$, such that

$$|f(x)| \le \varepsilon \qquad \text{for all } x \ge N.$$

So given $\varepsilon > 0$, we define a sequence of sets (A_n) by

$$A_n = \{x \in \mathbb{R}_+ \mid |f(m \cdot x)| \le \varepsilon \text{ for all } m \ge n\} = \bigcap_{m=n}^{+\infty} \{x \in \mathbb{R}_+ \mid |f(m \cdot x)| \le \varepsilon\}.$$

Since f is continuous, every set A_n is closed in $\mathbb{R}_+$. Furthermore, it is obvious that $A_{n+1} \supseteq A_n$, so the sequence (A_n) is increasing.

Choose any $a > 0$. Then by assumption, $f(m \cdot a) \to 0$ for $m \to +\infty$ through $\mathbb{N}$, and there exists an $n \in \mathbb{N}$, such that

$$|f(m \cdot a)| \le \varepsilon \qquad \text{for all } m \in \mathbb{N}, \text{ such that } m \ge n.$$

This implies that $a \in A_n$. Since $a \in \mathbb{R}_+$ was chosen arbitrarily, we conclude that

$$(1.16) \qquad \mathbb{R}_+ = \bigcup_{n=1}^{+\infty} A_n \qquad \left(= \lim_{n \to +\infty} A_n \right),$$

so we have written $\mathbb{R}_+$ as a *countable* union of the *closed sets* A_n. According to *Baire's theorem*, at least one of them must contain an interior point, i.e. $A_n^\circ \ne \emptyset$. (The proof continues after the following remark.)

Remark 1.4 Since it cannot be assumed that all readers know *Baire's theorem*, we include a proof in this remark.

INDIRECT PROOF. Assume that none of the sets A_n in (1.16) contains interior points. The set A_1 is closed, so there are constants $a_1, b_1 \in \mathbb{R}_+$, $a_1 < b_1$, such that

$$[a_1, b_1] \cap A_1 = \emptyset.$$

By assumption, the next set A_2 does not contain any interior point either, so $[a_1, b_1] \setminus A_2 \neq \emptyset$. Since A_2 is also closed, we can find a closed interval $[a_2, b_2] \subseteq [a_1, b_1]$, where $a_2 < b_2$, such that $[a_2, b_2] \cap A_2 = \emptyset$, hence in particular,

$$[a_2, b_2] \cap (A_1 \cup A_2) = \emptyset.$$

In this way we continue the construction of closed intervals $[a_n, b_n]$, where $0 < a_n < b_n$, such that $[a_n, b_n] \subseteq [a_{n-1}, b_{n-1}]$, and such that

$$[a_n, b_n] \cap \bigcup_{j=1}^{n} A_j = \emptyset. \tag{1.17}$$

Clearly,

$$[a_1, b_1] \supseteq [a_2, b_2] \supseteq \cdots \supseteq [a_n, b_n] \supseteq \cdots$$

is a *decreasing* sequence of *closed* intervals, so their intersection is not the empty set. Choose

$$x \in \bigcap_{n=1}^{+\infty} [a_n, b_n], \qquad x \in \mathbb{R}_+, \tag{1.18}$$

where we have proved above that such an $x \in \mathbb{R}_+$ exists. Then by (1.17),

$$x \notin \bigcup_{n=1}^{+\infty} A_n = \mathbb{R}_+,$$

because we also have that $x \notin A_n$ for any $n \in \mathbb{N}$.

Due to (1.18) this is not possible, so we have come to a contradiction. We therefore conclude that at least one of the sets A_n must contain interior points, and *Baire's theorem* is proved. $\lozenge$

We shall return to the main proof. Let $x \in A_n^\circ$ be an interior point. By definition we can find $0 < a < b$, such that $[a, b] \subseteq A_n^\circ \subseteq A_n$. It then follows from the definition of A_n that

$$|f(m \cdot x)| \leq \varepsilon, \qquad \text{for all } x \in [a, b] \text{ and all } m \geq n, \quad m \in \mathbb{N}. \tag{1.19}$$

Define

$$B_n := \bigcup_{m=n}^{+\infty} [m\,a, m\,b].$$

Then it follows from (1.19) that $|f(y)| \leq \varepsilon$ for all $y \in B_n$.

Put

$$P = \min\left\{ p \in \mathbb{N} \;\middle|\; p \geq n \text{ and } p \geqq \frac{a}{b-a} \right\}.$$

It follows from a small calculation that every point $y \in [P \cdot a, +\infty[$ belongs to at least one interval of the type $[m\,a, m\,b]$, where $m \in \mathbb{N}$ and $m \geq n$. In other words, $[P \cdot a, +\infty[\subseteq B_n$.

Finally, it follows from the above that to every given $\varepsilon > 0$ we can find $N = P \cdot a$, such that

$$|f(y)| \leq \varepsilon \qquad \text{for all } y \geq N,$$

and we have proved the theorem. $\square$

It is left to the reader to prove the following generalization of Theorem 1.3.

Theorem 1.4 *Let (c_n) be a strictly increasing sequence on $\mathbb{R}_+$, such that*

$$c_n \to +\infty \quad \textit{and} \quad \frac{c_{n+1}}{c_n} \to 1 \qquad \textit{for } n \to +\infty.$$

Assume that the continuous function $f : \mathbb{R}_+ \to \mathbb{R}$ for every $a > 0$ satisfies

$$f(a \cdot c_n) \to 0, \qquad \textit{for } n \to +\infty, \quad n \in \mathbb{N}.$$

Then $f(x) \to 0$ for $x \to +\infty$ through $\mathbb{R}_+$.

Example 1.1 When we choose the sequence $c_n = n^p$ for fixed $p > 0$, the conditions of Theorem 1.4 are fulfilled, so $f(a \cdot n^p) \to 0$ for $n \to +\infty$ through $\mathbb{R}_+$ for all $a > 0$ implies that also $f(x) \to 0$ for $x \to +\infty$ through $\mathbb{R}_+$. If, however, the strictly increasing sequence (c_n) does not satisfy the condition that $c_{n+1}/c_n \to 1$ for $n \to +\infty$, it is possible to construct a continuous function f, which does not tend towards 0 for $x \to +\infty$, and yet $f(a \cdot c_n) \to 0$ for $n \to +\infty$ for every $a > 0$. $\Diamond$

1.8 Two examples of functions from $C^\infty(\mathbb{R})$ which are nowhere real analytic

It is often believed that a C^∞ function is also analytic and "that a real C^∞ function in $x \in \mathbb{R}$ is made analytic by substituting $x \in \mathbb{R}$ by $z \in \mathbb{C}$." We shall in this section demonstrate that this intuitive claim in general is wrong.

A real function f is real analytic at a point $x_0 \in \mathbb{R}$, if its Taylor series expanded from x_0 is convergent in a neighbourhood of x_0 with the right sum, namely $f(x)$,

$$f(x) = \sum_{n=0}^{+\infty} \frac{1}{n!} f^{(n)}(x_0) \cdot (x - x_0)^n, \qquad \text{for } |x - x_0| < r.$$

Clearly, every C^∞ function has a Taylor series. So if the function is not analytic, then either its Taylor series is *divergent* for all $x \neq x_0$, or it is convergent,but it does not converge towards $f(x)$.

We shall in the first example follow a standard procedure. The second example, which is due to *T. Bang* [6] from 1946 is probably more exciting.

Example 1.2 We shall use *Cauchy's function* φ, which is defined by

$$\varphi(x) = \begin{cases} \exp\left(-\dfrac{1}{x}\right), & \text{for } x > 0, \\[2ex] 0, & \text{for } x \leq 0. \end{cases}$$

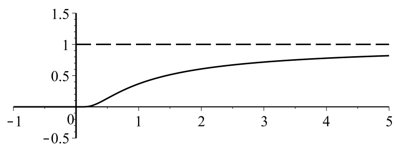

Figure 1.6: *The graph of Cauchy's function $\varphi(x)$.*

This function is obviously of class C^∞ for $x \neq 0$, and for $x = 0$ we only have to convince ourselves that $\varphi(x)$ is more "flat" than any monomial x^n at $x = 0$, so it is easy to prove that

$$\varphi^{(n)}(0) = 0 \qquad \text{for all } n \in \mathbb{N}_0.$$

The Taylor series of φ developed from $x_0 = 0$ is trivially convergent and represents the identical zero function, while it does not converge towards $\varphi(x)$ for $x > 0$. In particular, $\varphi(x)$ is not analytic at $x_0 = 0$. It is, however, analytic at every other point $x \neq 0$, because it is composed of analytic functions.

The idea is to move the point $x_0 = 0$, where $\varphi(x)$ is C^∞, but not (real) analytic, through a dense and countable set in $\mathbb{R}$. Choose e.g. the dense and countable set of rational numbers, $\mathbb{Q}$. Due to the countability, the set can be written as a sequence, $\mathbb{Q} = \{q_n \mid n \in \mathbb{N}\}$.

Then define the function

$$f(x) := \sum_{n=1}^{+\infty} 2^{-n} \varphi\left(x - q_n\right).$$

Since $\varphi^{(k)}(0) = 0$ and $\lim_{x \to +\infty} \varphi^{(k)}(x) = 0$ for every fixed $k \in \mathbb{N}$, we conclude that

$$\sup_{x \in \mathbb{R}} \left|\varphi^{(k)}(x)\right| = \max_{x \in \mathbb{R}} \left|\varphi^{(k)}(x)\right| < +\infty, \qquad \text{for every } k \in \mathbb{N}_0.$$

Then (it is not hard to prove that) each of the termwisely differentiated series

$$\sum_{n=1}^{+\infty} 2^{-n}\, \varphi^{(k)}\left(x - q_n\right), \qquad k \in \mathbb{N}_0,$$

is uniformly convergent, so its sum must be $f^{(k)}(x)$, which is defining a continuous function. Since this is true for all $k \in \mathbb{N}_0$, we conclude that f is of class $C^{\mathbb{R}}$.

Then choose any rational number $x_0 \in \mathbb{Q}$, so $x_0 = q_m$ for some $m \in \mathbb{N}$, and we have the splitting of the function

$$f(x) = \sum_{n \neq m} 2^{-n}\, \varphi\left(x - q_n\right) + 2^{-m}\, \varphi\left(x - q_m\right).$$

The sum is analytic at $x = q_m$, while the latter term is not. We therefore conclude that $f(x)$ is not analytic at $x_0 = q_m$. Since x_0 is any rational number, we conclude that $f(x)$ is not analytic at any rational number, though of course of class C^∞ everywhere.

Finally, choose any $x_0 \in \mathbb{R} \setminus \mathbb{Q}$. Then just note that every open interval containing x_0 will also contain a rational number, where $f(x)$ is not analytic. Hence, we cannot find any open interval containing x_0, in which $f(x)$ is analytic, so f is not analytic in x_0.

Summing up, we conclude that $f(x)$ is of class $C^\infty(\mathbb{R})$ and not analytic at any point of $\mathbb{R}$. $\Diamond$

Example 1.3 *Bang* [6] (1946). This example uses a totally different idea. We consider the nonlinear differential equation

$$\text{(1.20)} \qquad \frac{\mathrm{d}f}{\mathrm{d}x}(x) = 2\, f(2x), \qquad \text{for } x \geq 0,$$

where $f(0) = 0$. We shall prove that it has a C^∞ solution, which satisfies the following two conditions,

a) In the interval $[0, 2]$ the graph of the solution is symmetric with respect to the vertical line $x = 1$,

$$f(1 - x) = f(1 + x) \qquad \text{for } x \in [0, 1].$$

b) The function has the values $f(0) = 0$ and $f(1) = 1$.

When we integrate (1.20) from $x_0 = 0$ we get the equivalent equation

$$(1.21) \qquad f(x) = \int_0^{2x} f(t)\,\mathrm{d}t \qquad \text{for } x \geq 0.$$

Using (1.21), b) and a) we get

$$(1.22) \qquad 1 = f(1) = \int_0^2 f(t)\,\mathrm{d}t = 2\int_0^1 f(t)\,\mathrm{d}t = 2\,f\left(\frac{1}{2}\right),$$

so

$$f\left(\frac{1}{2}\right) = \frac{1}{2}.$$

If $x \in \left[0, \frac{1}{2}\right]$, then it follows from a) and (1.22) that

$$\begin{aligned} f\left(\frac{1}{2} - x\right) + f\left(\frac{1}{2} + x\right) &= \int_0^{1-2x} f(t)\,\mathrm{d}t + \int_0^{1+2x} f(t)\,\mathrm{d}t \\ &= \int_0^{1.2x} f(t)\,\mathrm{d}t + \int_{2-(1+2x)}^{2-0} f(t)\,\mathrm{d}t = \int_0^2 f(t)\,\mathrm{d}t = f(1) = 1, \end{aligned}$$

so we conclude that

c) For $x \in [0,1]$ the graph of f is symmetric with respect to the point $\left(\frac{1}{2}, \frac{1}{2}\right)$.

Using an iteration process we then prove in the following that there exists a uniquely determined function which satisfies the following four conditions,

$$(1.23) \qquad f(x) = \int_0^{2x} f(t)\,\mathrm{d}t \qquad \text{for } x \geq 0.$$

a) $f(1-x) = f(1+x) \qquad \text{for } x \in [0,1].$

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

c) $f\left(\frac{1}{2}-x\right) + f\left(\frac{1}{2}+x\right) = 1, \qquad \text{for } x \in \left[0,\frac{1}{2}\right], \text{ and } f\left(\frac{1}{2}\right) = \frac{1}{2}.$

Let $f_0(x)$ be any continuous function, which fulfils a), b), c) above. We may e.g. choose

$$f_0(x) := 1 - |1-x|.$$

We defined by induction a sequence of functions (f_n), satisfying a), b) and c), $n \in \mathbb{N}_0$, such that the function f_{n+1} following after f_n is given by

$$(1.24) \qquad f_{n+1}(x) = \int_0^{2x} f_n(t)\,\mathrm{d}t \qquad \text{for } x \in [0,1],$$

and we put, cf. a),

$$f_{n+1}(x) := f_{n+1}(2-x) \qquad \text{for } x \in [1,2],$$

where we have used (1.24). Clearly, a) is fulfilled by definition, so we shall prove that is also satisfies b) and c).

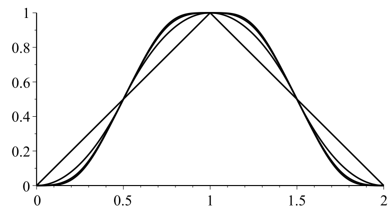

Figure 1.7: *The graph of the first three iteration functions, where we have used f_0 as our initial function. The graphs of f_2 and f_3 can hardly be distinguished.*

It follows from (1.24) that $f_{n+1}(0) = 0$, and since by assumption, f_n satisfies a) and c), we get

$$\begin{aligned} f_{n+1}(1) &= \int_0^2 f_n(t)\,\mathrm{d}t = 2\int_0^1 f_n(t)\,\mathrm{d}t = 2\int_0^{\frac{1}{2}} f_n(t)\,\mathrm{d}t + 2\int_{\frac{1}{2}}^1 f_n(t)\,\mathrm{d}t \\ &= 2\int_0^{\frac{1}{2}} \left\{ f_n\left(\frac{1}{2}-t\right) + f_n\left(\frac{1}{2}+t\right)\right\} \mathrm{d}t = 2\int_0^{\frac{1}{2}} 1\,\mathrm{d}t = 1, \end{aligned}$$

proving b) for f_{n+1}.

Finally, for $x \in \left[0, \frac{1}{2}\right]$,

$$f_{n+1}\left(\frac{1}{2}-x\right)+f_{n+1}\left(\frac{1}{2}+x\right)=\int_0^{1-2x} f_n(t)\,\mathrm{d}t+\int_0^{1+2x} f_n(t)\,\mathrm{d}t$$

$$=\int_0^{1-2x} f_n(t)\,\mathrm{d}t+\int_{2-(1+2x)}^{2-0} f_n(t)\,\mathrm{d}t=\int_0^2 f_n(t)\,\mathrm{d}t=f_n(1)=1,$$

and we have proved c). Hence, the three conditions a), b) and c) are preserved by the iterations.

Summing up, we have found that

$$(1.25)\qquad f_n(0)=0, \quad f_n\left(\frac{1}{2}\right)=\frac{1}{2} \quad \text{and} \quad f_n(1)=1 \qquad \text{for all } n \in \mathbb{N}.$$

Since for every $x \in [0,1]$ there is an $\alpha \in \left\{0, \frac{1}{2}, 1\right\}$, such that $|2x-2\alpha| \le \frac{1}{2}$, we get the estimate

$$\begin{aligned} |f_{n+1}(x)-f_n(x)| &= |\{f_{n+1}(x)-f_{n+1}(\alpha)\}-\{f_n(x)-f_n(\alpha)\}| \\ &= \left|\int_{2\alpha}^{2x}\{f_n(t)-f_{n-1}(t)\}\,\mathrm{d}t\right| \le |2x-2\alpha|\cdot\|f_n-f_{n-1}\|_\infty \\ &\le \frac{1}{2}\,\|f_n-f_{n-1}\|_\infty\,, \end{aligned}$$

where we define

$$\|\varphi\|_\infty := \sup_{x\in[0,2]} |\varphi(x)|.$$

It follows from the estimate above by iteration that

$$\|f_{n+1}-f_n\|_\infty \le \frac{1}{2}\,\|f_n-f_{n-1}\|_\infty \le \cdots \le \frac{1}{2^n}\,\|f_1-f_0\|_\infty\,,$$

so (f_n) is a *uniformly convergent Cauchy sequence.*

The uniformly convergent Cauchy sequence (f_n) of continuous functions has a continuous limit function f. By taking the limit in (1.25) it follows that the limit function f satisfies the integral equation

$$(1.26)\qquad f(x)=\int_0^{2x} f(t)\,\mathrm{d}t,$$

and f satisfies of course also the conditions a), b) and c).

Since f is continuous, the right hand side of (1.26) is of class C^1, hence f itself is of class C^1, which again implies that the right hand side of (1.26) is of class C^2, so f is of class C^2. By using this bootstrap method we conclude that $f \in C^\infty(\mathbb{R})$. Furthermore, we get by differentiating (1.26),

$$(1.27)\qquad \frac{\mathrm{d}f}{\mathrm{d}x}(x)=2\,f(2x), \qquad \text{for } x \in [0,1].$$

Then we extend (1.27) to hold for all $x \geq 0$. Since $f'(x)$ is known for $x \in [0,2]$, we get $f(x)$ for $x \in [2,4]$ from the equation

$$f(x) = \frac{1}{2}\, f'\left(\frac{x}{2}\right) \qquad \text{for } x \in [2,4].$$

Finally, it is seen by another bootstrap method (iteration) that $f(x)$ is uniquely determined for all $x > 0$, and that f belongs to the class $C^\infty(\mathbb{R}_+)$.

When $x = 0$, it follows from (1.27) that $f'(0+) = 0$, and then by induction, that also $f^{(n)}(0+) = 0$ for all $n \in \mathbb{N}$. Hence, 0 is a zero for f of infinite order. This implies that we can extend f to all of $\mathbb{R}$ by the definition $f(-x) = f(x)$, i.e. f becomes an even function, and we have finally constructed our $C^\infty(\mathbb{R})$ function f.

When we iterate (1.27) it follows that

$$(1.28) \qquad f^{(n)}(x) = 2^{\frac{1}{2}\,n(n+1)}\, f(2^n x) \qquad \text{for } x \geq 0 \text{ and } n \in \mathbb{N}.$$

Using a) we see that $f(2) = f(0) = 0$, so $x = 2$ is also a zero of f. Since the graph of f in $[0,2]$ is symmetric with respect to the line $x = 1$, and since $x = 0$ is a zero of infinite order, $x = 2$ must also be a zero of infinite order.

Furthermore, for $x \in [0,2]$,

$$f\left(1 - \frac{x}{2}\right) = \int_0^{2-x} f(t)\,\mathrm{d}t = f\left(1 + \frac{x}{2}\right) = \int_0^{2+x} f(t)\,\mathrm{d}t,$$

hence

$$\int_0^{2-x} f(t)\,\mathrm{d}t = \int_0^{2+x} f(t)\,\mathrm{d}t, \qquad \text{for } x \in [0,2].$$

Then by a differentiation,

$$f(2+x) = -f(2-x), \qquad \text{for } x \in [0,2].$$

In particular, $f(4) = f(2+2) = -f(2-2) = f(0) = 0$.

Repeating this argument it follows that the graph of f is symmetric in the interval $[0.2 \cdot 4^n]$ with respect to the line $x = 4^n$, $n \in \mathbb{N}_0$, and also symmetric with respect to the point $(2 \cdot 4^n, 0)$ in the interval $\left[0, 4^{n+1}\right]$, $n \in \mathbb{N}_0$.

Since all points $\{2n \mid n \in \mathbb{Z}\}$ are zeros of infinite order, we get that also $\varphi(x) := |f(x)|$ is of class $C^\infty(\mathbb{R})$, and even periodic of period 2.

Choose some $x_0 = 2^{-p}\,h$, where $h \in \mathbb{Z}$ and $p \in \mathbb{N}_0$. It follows from (1.28) that

$$f^{(n)}\left(x_0\right) = 0 \qquad \text{for all } n \geq p+1,$$

so the Taylor series, expanded from x_0, degenerates to a polynomial, which clearly does not agree with the given function $f(x)$, which is not a polynomial. In other words, the Taylor series *does* converge, but *not* towards $f(x)$.

Then choose $x_0 \neq 2^{-p}, h$ for all $h \in \mathbb{Z}$ and all $p \in \mathbb{N}_0$. Clearly, x_0 can uniquely be written in the form

$$x_0 = [x_0] + \sum_{n=1}^{+\infty} 2^{-n}\,a_n,$$

where the integer $[x_0] \in \mathbb{Z}$ is the largest one, for which $[x_0] \leq x_0$, and where $a_n \in \{0,1\}$. By assumption, x_0 does not have the form $2^{-p}\,h$ for some $h \in \mathbb{Z}$ and $p \in \mathbb{N}_0$. Therefore, the sequence (a_n) contains infinitely often 0 and 1. Using (1.28) we find a constant $C > 0$ and an infinite index set $I \subseteq \mathbb{N}$, such that

$$\left|f^{(n)}\left(x_0\right)\right| \geq C \cdot 2^{\frac{1}{2}\,n(n+1)}, \qquad \text{for all } n \in I.$$

Therefore, the radius of convergence of the Taylor series, expanded from x_0, is given by

$$\varrho = \left\{\limsup_{n\to+\infty} \sqrt[n]{\frac{1}{n!}\,\left|f^{(n)}\left(x_0\right)\right|}\right\}^{-1} \leq \left\{\limsup_{n\to+\infty}\left\{\frac{1}{n^n} \cdot C \cdot 2^{\frac{1}{2}\,n(n+1)}\right\}^{1/n}\right\}^{-1}$$

$$= \left\{\lim_{n\to+\infty} \sqrt[n]{C} \cdot \frac{1}{n} \cdot 2^{\frac{1}{2}(n+1)}\right\}^{-1} = 0.$$

We conclude that the Taylor series for $f(x)$, expanded from $x_0 \in \mathbb{R} \setminus \{2^{-p}\,h \mid h \in \mathbb{Z}, p \in \mathbb{N}_0\}$, is divergent for all $x \neq x_0$, which shows that f does not have an analytic extension. It then follows by a small argument that the same is true for the function $\varphi(x) = |f(x)|$. $\Diamond$

The example above shows that it is possible to construct a function $\varphi \in C^{\infty}(\mathbb{R})$, such that

1) $\|\varphi\|_{\infty} = 1$, i.e. φ is in particular bounded.
2) φ is periodic.
3) The Taylor series from a point $x_0 \in \mathbb{R}$ is divergent for almost every $x_0 \in \mathbb{R}$. And if the Taylor series, expanded from some $x_0 \in \mathbb{R}$ is convergent, then it does not converge towards φ.

Another consequence is that a Taylor series expansion of a C^{∞} function does not always give us a good approximation of the function under consideration. Fortunately, the *Weierstraß's approximation theorem* states that every *continuous* function (so we do not require that $\varphi \in C^{\infty}$) in a *closed and bounded* (i.e. compact) interval can be even *uniformly* approximated by a sequence of polynomials. If we therefore in our calculations allow a uniform error $< \varepsilon$ in a closed and bounded interval I, then we can replace any continuous function in I by some polynomial.

1.9 The functional equation $xf(x) = \frac{1}{x} f\left(\frac{1}{x}\right)$ for $x \in \mathbb{R}_+$.

Many years ago I was asked by an engineer if the only solutions of the functional equation

$$xf(x) = \frac{1}{x} f\left(\frac{1}{x}\right), \qquad \text{for } x \in \mathbb{R}_+,$$

were of the structure $f(x) = c/x$ for some constant c. The reason for this question was that such a result would be very convenient in Statistics. Unfortunately, this is in general not the case. We first prove the following theorem.

Theorem 1.5 *All solutions of the functional equation*

$$(1.29) \qquad xf(x) = \frac{1}{x} f\left(\frac{1}{x}\right), \qquad \textit{for } x \in \mathbb{R}_+,$$

are given by

$$(1.30) \qquad f(x) = \frac{1}{x}\left\{h(x) + h\left(\frac{1}{x}\right)\right\}, \qquad \textit{for } x \in \mathbb{R}_+,$$

where $h : \mathbb{R}_+ \to \mathbb{R}$ is an arbitrary function.

PROOF a) If $f(x)$ is given by (1.30), then by a small calculation,

$$xf(x) = h(x) + h\left(\frac{1}{x}\right) = h\left(\frac{1}{x}\right) + h\left(\left(\frac{1}{x}\right)^{-1}\right) = \frac{1}{x} f\left(\frac{1}{x}\right),$$

where we e.g. use the substitution $u = 1/x$ and (1.30). It follows that (1.29) holds.

b) Then assume that f is a solution of (1.29) and define

$$h(x) := \frac{1}{2}\, x\, f(x).$$

Then

$$\frac{1}{x}\left\{h(x)+h\left(\frac{1}{x}\right)\right\}=\frac{1}{x}\left\{\frac{1}{2}\,x\,f(x)+\frac{1}{2}\,\frac{1}{x}\,f\left(\frac{1}{x}\right)\right\}=\frac{1}{2}\cdot\frac{1}{x}\{xf(x)+xf(x)\}=f(x). \qquad \square$$

If we e.g. choose

$$h(x)=c\cdot x\left\{\exp\left(\frac{x^2}{\left(x^2+1\right)^2}\right)-1\right\},$$

and use (1.30) to define f, then for some constant $c>0$,

$$f(x)\geq 0 \quad \text{for } x\in\mathbb{R}_+, \qquad \text{and} \qquad \int_0^{+\infty} f(x)\,\mathrm{d}x=1,$$

so f becomes a frequency on $\mathbb{R}_+$.

Recently, in August 2015, onee of my former students, Jeppe Ammitzbøll, sent me a letter in which he improved the theorem above, so we only get the wanted solutions $f(x) = c/x$ for $x \in \mathbb{R}_+$, where $c \in \mathbb{R}$ is an arbitrary constant.

Theorem 1.6 (Jeppe Ammitzbøll, 2015; private communication). *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a function. Then the following three conditions taken together,*

$$(1.31) \qquad x\,f(x) = \frac{1}{x}\,f\left(\frac{1}{x}\right), \qquad \textit{for } x \in \mathbb{R}_+,$$

$$(1.32) \qquad f(x)f\left(\frac{1}{x}\right) = f(1)^2, \qquad \textit{for } x \in \mathbb{R}_+,$$

$$(1.33) \qquad f(x)\,f(1) \geq 0, \qquad \textit{for } x \in \mathbb{R}_+,$$

are equivalent to

$$f(x) = \frac{f(1)}{x}, \qquad \textit{for } x \in \mathbb{R}_+.$$

Proof. The proof is incredibly simple.

a) If $x\,f(x) = f(1)$, then it is easy to check the three conditions, (1.31), (1.32) and (1.33).

b) If instead f satisfies the three conditions, (1.31), (1.32) and (1.33), then

$$\begin{aligned} (x\,f(x))^2 &= x\,f(x) \cdot \frac{1}{x} \cdot \frac{1}{x}\,f\left(\frac{1}{x}\right) && \text{from (1.31)} \\ &= f(1)^2 && \text{from (1.32).} \end{aligned}$$

Since (1.33) only means that $f(x)$ and $f(1)$ have the same sign, and $x > 0$, we get by taking the square root

$$x\,f(x) = +f(1), \qquad \text{for all } x \in \mathbb{R}_+,$$

and the theorem is proved. We note that in this case $c = f(1)$. □

A more abstract approach would be the following. We want that the only solutions are of the form $f(x) = c/x$ for $x \in \mathbb{R}_+$, where $c \neq 0$ is a constant. These desired solutions are all monotonous, so they have an inverse. If $y = f(x) = c/x$, then the inverse is of course $f^{-1}(y) = x = c/y$ of the same structure. It would therefore be natural to add the assumption that the inverse $f^{-1}(y)$ of $f(x)$ exists on $\mathbb{R}_+$. Unfortunately this is not enough. If we choose $h(x) = \sqrt{x}$ in (1.30) of Theorem 1.5, then we get that

$$f(x) = \frac{1}{x}\left\{h(x) + h\left(\frac{1}{x}\right)\right\} = \frac{1}{\sqrt{x}} + \frac{1}{x\sqrt{x}},$$

which satisfies (1.29) without having the right structure. It is easy to see that $f(x)$ is monotonous, so its inverse exists. Thus, f^{-1} should fulfil something more.

By adding some missing assumptions we arrive at the following theorem

Theorem 1.7 *Assume that f is a strictly monotonous C^1 function, and that both f and its inverse f^{-1} satisfy the functional equation (1.29). Then there is a constant $c \in \mathbb{R}$, such that*

$$f(x) = \frac{c}{x} \qquad \text{for } x \in \mathbb{R}_+. \tag{1.34}$$

The proof, which is given here for completeness, is rather technical and probably not of interest for all readers. One may just accept the result of Theorem 1.7.

PROOF. It follows from (1.29) that $f(x)$ and $f\left(\frac{1}{x}\right)$ must have the same sign, so they are both either positive, or both negative. The monotony of f then implies that $f(1) \neq 0$.

We define the auxiliary function

$$g(x) := \frac{f(x)}{f(1)}.$$

Then g also satisfies the assumption (1.29). We may therefore in the following assume that $f(1) = 1$, and accordingly $f^{-1}(1) = 1$.

When $x > 1$, it follows from the functional equation that we have the estimate

$$f(x) = \frac{1}{x^2}\, f\left(\frac{1}{x}\right) < f\left(\frac{1}{x}\right).$$

Since f is monotonous, we conclude that f is decreasing in $\mathbb{R}_+$, i.e. $f'(x) < 0$ for all $x \in \mathbb{R}_+$.

When we differentiate the functional equation (1.29), we get

$$\begin{aligned} x f'(x) + f(x) &= -\frac{1}{x^2}\, f\left(\frac{1}{x}\right) - \frac{1}{x^3}\, f'\left(\frac{1}{x}\right) = -\frac{1}{x}\left\{\frac{1}{x}\, f\left(\frac{1}{x}\right)\right\} - \frac{1}{x^3}\, f'\left(\frac{1}{x}\right) \\ &= -\frac{1}{x}\,\{x f(x)\} - \frac{1}{x^3}\, f'\left(\frac{1}{x}\right) = -f(x) - \frac{1}{x^3}\, f'\left(\frac{1}{x}\right), \end{aligned}$$

hence by a rearrangement,

$$f'(x) = -\frac{2}{x}\, f(x) - \frac{1}{x^4}\, f'\left(\frac{1}{x}\right) = -\frac{2}{x}\, f(x) + \frac{1}{x^4}\left|f'\left(\frac{1}{x}\right)\right| \geq -\frac{2}{x}\, f(x).$$

Using that $f(1) = 1$ it follows that $f(x) > 0$ for all $x > 0$. Therefore, this inequality can also be written

$$\frac{f'(x)}{f(x)} = \frac{\mathrm{d}}{\mathrm{d}x}\ln f(x) \geq -\frac{2}{x}.$$

This inequality is integrated from $x_0 = 1$. Then for $x > 1$

$$\ln f(x) - \ln f(1) = \ln f(x) \geq -\ln\left(x^2\right) = \ln\left(\frac{1}{x^2}\right),$$

hence

$$f(x) \geq \frac{1}{x^2}, \qquad \text{for } x \geq 1,$$

because the exponential is increasing.

If instead $x \in]0,1[$, then

$$\int_x^1 \frac{f'(t)}{f(t)}\,\mathrm{d}t = -\ln f(x) \geq \int_x^1 \left\{-\frac{2}{t}\right\}\mathrm{d}t = \ln\left(x^2\right),$$

so

$$f(x) \leq \frac{1}{x^2}, \qquad \text{for } x \in]0,1[.$$

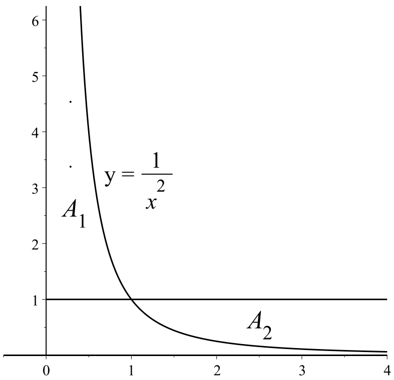

Figure 1.8: *The graph of f lies in $A_1 \cup A_2$ between the graphs of $y = 1/x^2$ and $y = 1$.*

Summing up, we conclude that

$$(1.35) \qquad \begin{cases} 1 \leq f(x) \leq \dfrac{1}{x^2}, & \text{for } x \in]0,1], \\[2ex] \dfrac{1}{x^2} \leq f(x) \leq 1, & \text{for } x \in [1,+\infty[. \end{cases}$$

The trick is that because $f'(x) < 0$ everywhere, the inverse function f^{-1} satisfies the same assumptions as f, so the inverse f^{-1} must also fulfil (1.35). Accordingly,

$$\begin{cases} 1 \leq f^{-1}(y) \leq \dfrac{1}{y^2}, & \text{for } y \in]0,1], \\[2ex] \dfrac{1}{y^2} \leq f^{-1}(y) \leq 1, & \text{for } y \in [1,+\infty[. \end{cases}$$

Put $y = f(x)$. Then $x = f^{-1}(y)$. Furthermore, $x \in [1,+\infty[$, when $y \in]0,1]$, and *vice versa.* Hence

$$1 \leq x \leq \frac{1}{f(x)^2}, \quad \text{i.e.} \quad f(x) \leq \frac{1}{\sqrt{x}}, \qquad \text{for } x \geq 1,$$

and

$$\frac{1}{f(x)^2} \leq x \leq 1, \quad \text{i.e.} \quad f(x) \geq \frac{1}{\sqrt{x}}, \qquad \text{for } x \in]0,1].$$

When these improved estimates are inserted into (1.35), then

$$(1.36)\qquad \begin{cases} \dfrac{1}{\sqrt{x}} \le f/x) \le \dfrac{1}{x^2}, & \text{for } x \in \,]0,1], \\[2ex] \dfrac{1}{x^2} \le f(x) \le \dfrac{1}{\sqrt{x}}, & \text{for } x \in [1,+\infty[. \end{cases}$$

Due to the symmetry of f and f^{-1} we also have

$$(1.37)\qquad \begin{cases} \dfrac{1}{\sqrt{y}} \le f^{-1}(y) \le \dfrac{1}{y^2}, & \text{for } y \in \,]0,1], \\[2ex] \dfrac{1}{y^2} \le f^{-1}(y) \le \dfrac{1}{\sqrt{y}}, & \text{for } y \in [1,+\infty[, \end{cases}$$

Then we note that (1.36) is the special case $n = 2$ of the following inequalities for $n \in \mathbb{N} \setminus \{1\}$,

$$(1.38)\qquad \begin{cases} x^{-\frac{n-1}{n}} \le f(x) \le x^{-\frac{n}{n-1}}, & \text{for } x \in \,]0,1], \\[2ex] x^{-\frac{n}{n-1}} \le f(x) \le x^{-\frac{n-1}{n}}, & \text{for } x \in [1,+\infty[. \end{cases}$$

We prove the following lemma of induction

Lemma 1.1 *Assume that f and f^{-1} satisfy the assumptions of Theorem 1.7 and that $f(1)=1$. If f, hence also f^{-1}, satisfies (1.38) for some $n \geq 2$, then f and f^{-1} also satisfy (1.38), when n is replaced $n+1$.*

Proof. When $x \in]0,1]$, it follows from (1.38) that

$$x^{\frac{1}{n}} \leq x f(x) = \frac{1}{x}\, f\left(\frac{1}{x}\right),$$

hence

$$f\left(\frac{1}{x}\right) \geq x^{\frac{n+1}{n}}, \qquad \text{for } x \in]0,1].$$

When x is replaced by $1/x$, then we get

$$f(x) \geq x^{-\frac{n+1}{n}}, \qquad \text{for } x \in [1,+\infty[,$$

which is the first inequality of the second line of (1.38), when n is replaced by $n+1$.

Similarly,

$$f^{-1}(y) \geq y^{-\frac{n+1}{n}}, \qquad \text{for } y \in [1,+\infty[.$$

Then put $y = f(x)$ for $x \in]0,1]$ (when $y \in [1,+\infty[$). Then

$$x \geq f(x)^{-\frac{n+1}{n}}, \qquad \text{for } x \in]0,1],$$

and therefore

$$f(x) \geq x^{-\frac{n}{n+1}}, \qquad \text{for } x \in]0,1],$$

which is the first inequality of the first line of (1.38), when n is replaced by $n+1$.

Similarly, we get for $x \in [1,+\infty[$ from the second line of (1.38) that

$$x^{\frac{1}{n}} \geq x f(x) = \frac{1}{x}\, f\left(\frac{1}{x}\right),$$

hence

$$f\left(\frac{1}{x}\right) \leq x^{\frac{n+1}{n}}, \qquad \text{for } x \in [1,+\infty[.$$

Replacing x by $1/x$ we get

$$f(x) \leq x^{-\frac{n+1}{n}}, \qquad \text{for } x \in]0,1],$$

which is the second inequality of the first line of (1.38), when n is replaced by $n+1$.

Similarly,

$$f^{-1}(y) \le y^{-\frac{n+1}{n}}, \qquad \text{for } y \in]0,1].$$

Put $y = f(x)$. Then $x \in [1,+\infty[$ for $y \in]0,1]$, so

$$x \le f(x)^{-\frac{n+1}{n}}, \qquad \text{for } x \in [1,+\infty[,$$

and therefore

$$f(x) \le x^{-\frac{n}{n+1}}, \qquad \text{for } x \in [1,+\infty[,$$

which is the second inequality of the second line of (1.38), where n has been replaced by $n+1$. The lemma follows by the symmetry of f and f^{-1}. $\square$

Return to the proof of Theorem 1.7. We know already that (1.38) holds for $n = 2$. It then follows from Lemma 1.1 that for every fixed $x \in]0,1]$,

$$\frac{1}{x} = \lim_{n\to+\infty} x^{-\frac{n-1}{n}} \le f(x) \le \lim_{n\to+\infty} x^{-\frac{n}{n-1}} = \frac{1}{x},$$

i.e. we must have equality everywhere. We have proved that

$$f(x) = \frac{1}{x}, \qquad \text{for } x \in]0,1].$$

If instead $x \in [1,+\infty[$, then

$$\frac{1}{x} = \lim_{n\to+\infty} x^{-\frac{n}{n-1}} \le f(x) \le \lim_{n\to+\infty} x^{-\frac{n-1}{n}} = \frac{1}{x},$$

and we have equality. Hence also

$$f(x) = \frac{1}{x}, \qquad \text{for } x \in [1,+\infty[.$$

We have proved the theorem in the special case, when $f(1) = 1 = c$. However, due to the introductory remarks of the proof this is sufficient, so the theorem is proved. $\square$

1.10 A linear partial differential equation without solution

Lewy [33] constructed in 1957 an extremely simple inhomogeneous partial differential equation with polynomial coefficients and where the right hand side is a C^∞ function, such that this equation did not have any solutions, not even in the sense of distributions, i.e. in the sense of generalized functions. Later on, *Lars Hörmander* [27] found some sufficient conditions for linear partial differential operations being of this strange type.

A simple twist of the present example gives us an *injective* differential operator, which is hard to believe, because one usually loses some information by differentiation. We shall here in the large follow *Lewy* [33]. His idea was to create a linear partial differential operator, which turns every function, or distribution, into an analytic function, so the range is a subset of the analytic functions. If we then choose a C^∞ function f, which is not analytic anywhere, cf. e.g. Section 1.8, as the right hand side of the equation, then f does not lie in the range of the linear partial differential operator, so the equation does not have a solution.

We consider the space $\mathbb{R}^3$ with the coordinates (x,y,s), and let $\psi \in C^1(\mathbb{R})$. first prove

Theorem 1.8 *Assume that* $u \in C^1$ *is a solution of the linear partial differential equation*

$$(1.39) \quad -\frac{\partial u}{\partial x} - i\,\frac{\partial u}{\partial y} + 2i(x+iy)\,\frac{\partial u}{\partial s} = \psi'(s),$$

in a neighbourhood Ω *of the point* $(0,0,s_0)$. *Then* $\psi(s)$ *is an analytic function in a neighbourhood of* s_0.

PROOF. We may assume that Ω is an open ball of radius $\delta > 0$. Let Γ be a circle contained in Ω and lying in a plane parallel with the XY plane, i.e.

$$\Gamma = \left\{(x, y, s) \in \mathbb{R}^3 \mid s = s_1 \text{ and } x^2 + y^2 = t\right\},$$

where s_1 and t are chosen constants, which satisfy

$$|s_0 - s_1| < \delta \qquad \text{and} \qquad 0 < t < \delta^2 - (s_1 - s_0)^2 .$$

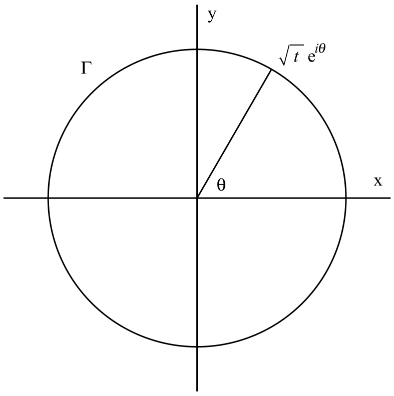

Figure 1.9: *The circle* Γ *in the proof of Theorem 1.8.*

The radius of Γ is $r = \sqrt{t}$, and it has the parametric description

$$x = e^{\ln \sqrt{t}} \cdot \cos\theta = \sqrt{t} \cdot \cos\theta \qquad \text{and} \qquad y = e^{\ln \sqrt{t}} \cdot \sin\theta = \sqrt{t} \cdot \sin\theta,$$

so

$$x + iy = \sqrt{t} \cdot e^{i\theta} = e^{\ln \sqrt{t} + i\theta}.$$

We shall change variables to

$$\varrho = \ln \sqrt{t} \qquad \text{and} \qquad \theta.$$

Then by the chain rule,

$$\frac{\partial}{\partial \varrho} = \frac{\partial x}{\partial \varrho}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \varrho}\frac{\partial}{\partial y} = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y},$$

and

$$\frac{\partial}{\partial \theta} = \frac{\partial x}{\partial \theta}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta}\frac{\partial}{\partial y} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}.$$

Using that $t = x^2 + y^2 = (x + iy)(x - iy)$ we get from the above,

$$\frac{\partial}{\partial \varrho} + i\frac{\partial}{\partial \theta} = (x - iy)\frac{\partial}{\partial x} + (y + ix)\frac{\partial}{\partial y} = (x - iy)\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) = \frac{t}{x + iy}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right),$$

hence,

$$\frac{\partial}{\partial x}+i\,\frac{\partial}{\partial y}=\frac{x+iy}{t}\left(\frac{\partial}{\partial \varrho}+i\,\frac{\partial}{\partial \theta}\right)=\frac{e^{i\theta}}{\sqrt{t}}\left\{\frac{\partial}{\partial \varrho}+i\,\frac{\partial}{\partial \theta}\right\}.$$

The first two terms of the left hand side of (1.39) are then given by

$$-\frac{\partial u}{\partial x}-i\,\frac{\partial u}{\partial y}=-\left\{\frac{\partial}{\partial x}+i\,\frac{\partial}{\partial y}\right\}u=-\frac{e^{i\theta}}{\sqrt{t}}\left\{\frac{\partial}{\partial \varrho}+i\,\frac{\partial}{\delta \theta}\right\}u.$$

We integrate (1.39) along Γ, where $s=s_1$ is constant. Then also $\psi'(s)$ is a constant, so only $\theta\in[0,2\pi]$ varies. Hence,

$$\begin{aligned}
2\pi\psi'(s) &= \int_0^{2\pi}\left\{-\left(\frac{\partial}{\partial x}+i\,\frac{\partial}{\partial y}\right)u+2i(x+iy)\frac{\partial u}{\partial s}\right\}\mathrm{d}\theta\\
&= -\int_0^{2\pi}\frac{e^{i\theta}}{\sqrt{t}}\left\{\frac{\partial u}{\partial \varrho}+i\,\frac{\partial u}{\partial \theta}\right\}\mathrm{d}\theta+2i\int_0^{2\pi}\sqrt{t}\cdot e^{i\theta}\,\frac{\partial u}{\partial s}\,\mathrm{d}\theta\\
&= -\int_0^{2\pi}\frac{e^{i\theta}}{\sqrt{t}}\,\frac{\partial u}{\partial \varrho}\,\mathrm{d}\theta-i\int_0^{2\pi}\frac{e^{i\theta}}{\sqrt{t}}\,\frac{\partial u}{\partial \theta}\,\mathrm{d}\theta+2i\int_0^{2\pi}\sqrt{t}\cdot e^{i\theta}\,\frac{\partial u}{\partial s}\,\mathrm{d}\theta.
\end{aligned}$$

By a partial integration,

$$\int_0^{2\pi}e^{i\theta}\,\frac{\partial u}{\partial \theta}\,\mathrm{d}\theta=\left[e^{i\theta}\,u\right]_{\theta=0}^{2\pi}-i\int_0^{2\pi}e^{i\theta}\,u\,\mathrm{d}\theta=-i\int_0^{2\pi}e^{i\theta}\,u\,\mathrm{d}\theta,$$

and then by insertion,

$$(1.40)\qquad 2\pi\psi'(s)=-\int_0^{2\pi}\frac{e^{i\theta}}{\sqrt{t}}\left\{\frac{\partial u}{\partial \varrho}+u\right\}\mathrm{d}\theta+2i\int_0^{2\pi}\sqrt{t}\cdot e^{i\theta}\,\frac{\partial u}{\partial s}\,\mathrm{d}\theta.$$

Using that $\varrho=\ln\sqrt{t}$, we get

$$\begin{aligned}
\frac{\partial}{\partial t}\left\{2e^{i\theta}\sqrt{t}\cdot u\right\} &= 2e^{i\theta}\cdot\frac{1}{2\sqrt{t}}\cdot u+2e^{i\theta}\sqrt{t}\,\frac{\partial u}{\partial t}=\frac{e^{i\theta}}{\sqrt{t}}\,u+2e^{i\theta}\sqrt{t}\cdot\frac{\mathrm{d}\ln\sqrt{t}}{\mathrm{d}t}\,\frac{\partial u}{\partial \varrho}\\
&= \frac{e^{i\theta}}{\sqrt{t}}\,u+2e^{i\theta}\,\frac{\sqrt{t}}{2t}\,\frac{\partial u}{\partial \varrho}=\frac{e^{i\theta}}{\sqrt{t}}\left\{\frac{\partial u}{\partial \varrho}+u\right\},
\end{aligned}$$

and we can reduce (1.40) to

$$\pi\psi'(s)=\frac{\partial}{\partial s}\left\{i\int_0^{2\pi}e^{i\theta}\sqrt{t}\,u\,\mathrm{d}\theta\right\}+i\,\frac{\partial}{\partial t}\left\{i\int_0^{2\pi}e^{i\theta}\sqrt{t}\,u\,\mathrm{d}\theta\right\},$$

because θ is independent of t, and because s is a constant. This shows that it is quite natural to introduce the function

$$(1.41)\qquad U(s,t):=i\int_0^{2\pi}e^{i\theta}\sqrt{t}\,u(s,t,\theta)\,\mathrm{d}\theta.$$

In particular, $U(s,0)=0$, so the equation above is reduced to

$$(1.42)\qquad \frac{\partial U}{\partial s}+i\,\frac{\partial U}{\partial t}=\pi\,\psi'(s).$$

We then put

$$V(s,t) = V(z) := U(s,t) - \pi\pi(s), \qquad \text{where } z = s + it.$$

Then V is of class C^1, and according to (1.42) this function V satisfies the *complex Cauchy-Riemann equation*

$$\frac{\partial V}{\partial s} + i\,\frac{\partial V}{\partial t} = 0.$$

In fact, if we split V into its real and imaginary parts, $V = V_1 + iV_2$, where V_1 and V_2 are C^1 functions, then

$$\frac{\partial V}{\partial s} + i\,\frac{\partial V}{\partial t} = \frac{\partial V_1}{\partial s} + i\,\frac{\partial V_2}{\partial s} + i\,\frac{\partial V_1}{\partial t} - \frac{\partial V_2}{\partial t} = 0.$$

Then by taking the real and imaginary parts and rearranging,

$$\frac{\partial V_1}{\partial s} = \frac{\partial V_2}{\partial t} \qquad \text{and} \qquad \frac{\partial V_1}{\partial t} = -\frac{\partial V_2}{\partial s},$$

and we have proved the usual real Cauchy-Riemann equations.

We conclude that $V(z)$ is analytic in a least a domain which contains the set

$$\omega = \left\{ z = s + it \in \mathbb{C} \mid 0 < t < \delta^2 - (s - s_0)^2 \right\}.$$

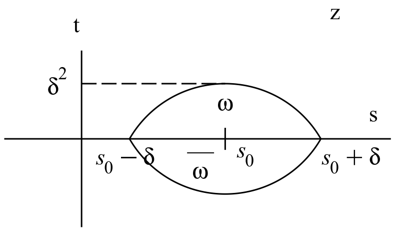

Figure 1.10: *The Schwarz reflection.*

It follows from $U(s,0)=0$ that $V=-\pi\psi$ for $t=0$, where ψ by assumption is real. Then using the *Schwarz reflection* V can be extended analytically across the line $t=0$, so V is analytic in a set, which at least contains the sets

$$\omega \qquad \text{and} \qquad \overline{\omega}=\{z\in\mathbb{C}\mid \overline{z}\in\omega\}\,,$$

and the line segment $]\,s_0+\delta,s_0+\delta\,[$ on the s-axis between ω and $\overline{\omega}$. In particular, V is analytic on the s-interval $]\,s_0+\delta,s_0+\delta\,[$, so we conclude that $-V/\pi$ is an analytic extension of ψ. Then for every chosen $a_1 \in\ ]\,s_0+\delta,s_0+\delta\,[$, the Taylor series of ψ is equal to ψ in a neighbourhood of a_1, i.e. ψ is analytic. The theorem is proved. □

It follows from Theorem 1.8 that if $u\in C^1\left(\mathbb{R}^3\right)$ and the expression

$$-\frac{\partial u}{\partial x}-i\,\frac{\partial u}{\partial y}+2i(x+iy)\,\frac{\partial u}{\partial s}$$

can be written as the derivative of a C^1 function, $\psi'(s)$, in the single variable s, then $\psi(s)$ – and hence also $\pi'(s)$ – is an analytic function.

Corollary 1.1 *There exists a function $f\in C^\infty\left(\mathbb{R}^3\right)$, such that the equation*

$$(1.43)\qquad -\frac{\partial u}{\partial x}-i\,\frac{\partial u}{\partial y}+2i(x+iy)\,\frac{\partial u}{\partial s}=f,$$

is never fulfilled for any $u\in C^1\left(\mathbb{R}^3\right)$, i.e. the set of solutions is the empty set.

Proof. Choose $f(x,y,s)=f(s)$ as a real C^∞ function, such that f is not analytic anywhere, cf. Section 1.8. The corollary follows by contraposition of Theorem 1.8. □

Corollary 1.1 still holds if the operation of differentiation is replaced by the more general *differentiation in the sense of distributions*. The reader may consult e.g. *Hörmander* [27] for further details.

It is also possible to prove that (1.43) is the simplest possible example of a linear partial differential equation with C^∞ coefficients and without solutions.

Corollary 1.2 (Hörmander [27]) *Given $N \geq 3$, $N \in \mathbb{N}$. There exists an* injective *linear partial differential operator of first order with C^∞ coefficients.*

Proof. Without loss of generality we may assume that $N = 3$. Let f be any function, such that (1.43) does not have a solution. We claim that the only solution of the linear partial differential equation

$$\frac{\partial u}{\partial x} + i\,\frac{\partial u}{\partial y} - 2i(x + iy = \frac{\partial u}{\partial s} + f \cdot u = 0, \tag{1.44}$$

is the zero solution $u = 0$.

We shall assume that this is not the case, so we assume that there exists a solution $u \neq 0$ of (1.44). There exists a branch of the complex logarithm and an open set ω, such that $v = \ln u$ of class $C^1(\omega)$.

When we put $u = e^v$ into (1.44) and divide the result by $e^v \neq 0$ in ω, we get (1.43) with u replaced by v, which contradicts that (1.43) does not have a solution. We conclude that the only solution of (1.44) is $u = 0$, and since the operator is linear, it is also injective. □

The nasty consequence of this example is that if we e.g. try so solve the equation

$$f'(x) = 2f(2x)$$

from Section 1.8 on a computer, then the approximation will always be an analytic function, i.e. in the possible range of the differential operator. If we therefore try to let the computer solve the problem, it will most likely find a "solution". So the computer is trying to approximate the empty set! This is an extreme example showing that one cannot always trust the results obtained on a computer.

2 Integrals

2.1 Some necessary measure theory

The present chapter is containing examples from *Measure Theory*, so we shall start with introducing the most necessary concepts from this theory. Most of them are intuitively clear, and yet their consequences may sometimes seem very strange. Only one example requires more than just the knowledge of *nullsets*. The basics given here are also used in the Chapters 3, 6 and 8.

Apart from finite pointsets the simplest sets in $\mathbb{R}$ we can think of are the *intervals*, of which we have four types,

$$\begin{array}{lll} [a,b] = \{x \in \mathbb{R} \mid a \le x \le b\} & \text{for } -\infty < a \le b < +\infty, & \text{(closed interval)} \\ ]a,b] = \{x \in \mathbb{R} \mid a < x \le b\}, & \text{for } -\infty \le a < b < +\infty, & \text{(halfopen interval)} \\ [a,b[= \{x \in \mathbb{R} \mid a \le x < b\}, & \text{for } -\infty < a < b \le +\infty, & \text{(halfopen interval)} \\ ]a,b[= \{x \in \mathbb{R} \mid a < x < b\}, & \text{for } -\infty \le a < b \le +\infty, & \text{(open interval).} \end{array}$$

Whenever I is such an interval with finite a and b, the *measure* of I is defined as $|I| = b - a$, i.e. the measure is the *length* of the interval, no matter whether the endpoints are included in I or not.

Based on the 1-dimensional intervals above we introduce in the N-dimensional real space $\mathbb{R}^N$,

Definition 2.1 *An* interval *in* $\mathbb{R}^N$*,* $N \in \mathbb{N}$*, is a set of the form*

$$I = I_1 \times \cdots \times I_N,$$

where each I_j*,* $j = 1, \dots, N$*, is an interval in* $\mathbb{R}$*.*

If each I_j*,* $j = 1, \dots, N$*, is bounded with the measure* $|I_j| < +\infty$*, we say that* $I = I_1 \times \cdots \times I_N$ *is a bounded interval of the measure*

$$|I| = |I_1 \times \cdots \times I_N| = |I_1| \cdots |I_N| .$$

Clearly, the bounded intervals in the plane $\mathbb{R}^2$ are rectangles, and the bounded intervals of the 3-space $\mathbb{R}^3$ are axiparallel boxes. It follows immediately from Definition 2.1 that there are four different types of intervals in $\mathbb{R}^2$, and in general 2^N different intervals in $\mathbb{R}^N$, $N \in \mathbb{N}$. Whenever they are bounded, the 2^N different intervals based on the same box have all the same measure. Since the emphasis is put on the *measures* here, we shall in general not need to distinguish between the formally 2^N different interval based on the same box.

Intuitively, if we in some way may give a set $A \subseteq \mathbb{R}^N$ a measure, denoted by $|A|$, and if (I_n) is a sequence of intervals, such that $A \subseteq \bigcup_n I_n$ (we call the sequence (I_n) a *covering* of the set A), then we get the inequality

$$|A| \le \sum_n |I_n| .$$

Note that we do not require that the intervals I_n are mutually disjoint. We shall be content, if only A is contained in the union of the I_n. This relaxation is quite convenient, because it will often be difficult explicitly to describe a *disjoint* covering. Furthermore, it opens up for the following apparently simple and "obvious", and yet crucial definition

Definition 2.2 *A set $A \subseteq \mathbb{R}^N$ is called a* nullset $\left(\text{in } \mathbb{R}^N\right)$, *if to every $\varepsilon > 0$ we can find a sequence of intervals (I_n) (a covering), such that*

$$A \subseteq \bigcup_n I_n \qquad \text{and} \qquad \sum_n |I_n| < \varepsilon.$$

This "innocent" definition implies latently the difference between the *Riemann integral* and the *Lebesgue integral*. So this "trivial" definition turns up to be far from trivial.

We shall first prove the following theorem.

Theorem 2.1 *Every* countable *subset of $\mathbb{R}^N$ is a nullset.*

Proof. Since the set A is countable, we can write it as a sequence, $A = \left\{x^{(n)} \mid n \in \mathbb{N}\right\}$. When $\varepsilon > 0$ is given, we choose

$$\varepsilon := \frac{1}{2}\sqrt[N]{2^{n-1}\varepsilon}, \qquad \text{for } n \in \mathbb{N}.$$

We write in coordinates in $\mathbb{R}^N$,

$$x^{(n)} = \left(x_1^{(n)}, \dots, x_N^{(n)}\right),$$

and define the intervals $I_n \subset \mathbb{R}^N$, $n \in \mathbb{N}$, by

$$I_n := \left[x_1^{(n)} - \varepsilon_n, x_1^{(n)} + \varepsilon_n\right] \times \cdots \times \left[x_N^{(n)} - \varepsilon_n, x_N^{(n)} + \varepsilon_n\right],$$

so each I_n is an N-dimensional cube of centre $x^{(n)}$ and edge length $2\varepsilon_n$, i.e. of measure

$$|I_n| = (2\varepsilon_n)^N = 2^{-n-1}\,\varepsilon.$$

Since each point of $A = \left\{x^{(n)} \mid n \in \mathbb{N}\right\}$ belongs to at least one of the I_j, we clearly have a covering of A, so

$$A = \left\{x^{(n)} \mid n \in \mathbb{N}\right\} \subseteq \bigcup_{n=1}^{+\infty} I_n.$$

Furthermore,

$$\sum_{n=1}^{+\infty} |I_n| = \sum_{n=1}^{+\infty} 2^{-n-1}\,\varepsilon = \frac{\varepsilon}{2} < \varepsilon,$$

and it follows that A is a nullset, □

The rational numbers form a countable set $\mathbb{Q}$, so $\mathbb{Q}$ is a nullset, i.e. of measure 0. This may seem strange, because every $x \in \mathbb{R}$ can be approximated by a sequence (q_n) of rational numbers $q_n \in \mathbb{Q}$.

We shall later also give examples of nullsets which are *not* countable, so the family of nullsets in $\mathbb{R}^N$ is larger and not equal to the family of countable sets in $\mathbb{R}^N$.

We mention that a proper linear subspace of $\mathbb{R}^N$, hence of dimension $< N$, is always a nullset in $\mathbb{R}^N$.

In connection with the nullsets it is also natural to introduce *nullfunctions*.

Definition 2.3 *A function* $f : \mathbb{R} \to \mathbb{R} \cup [-\infty, +\infty\}$ *is called a* nullfunction,, *if there exists a nullset* $A \subset \mathbb{R}$*, such that*

$$f(x) = 0, \qquad \textit{whenever } x \notin A.$$

If f *is a nullfunction, we* define *its* Lebesgue integral *by*

$$\int_{-\infty}^{+\infty} f(x)\, dx := 0.$$

This formulation may again seem quite reasonable, until we consider the standard example of a function, which is not Riemann integrable, namely

$$\chi_{Q \cap [0,1]}(x) = \begin{cases} 1 & \text{for } x \in \mathbb{Q} \cap [0,1], \\ 0 & \text{otherwise.} \end{cases}$$

Since $\mathbb{Q} \cap [0,1]$ is countable, it is a nullset, so $\chi_{\mathbb{Q} \cap [0,1]}$ is a nullfunction, and its *Lebesgue integral* is

$$\int_{-\infty}^{+\infty} \chi_{\mathbb{Q} \cap [0,1]}(x)\, \mathrm{d}x = 0,$$

while its Riemann integral is not defined.

Definition 2.4 *We say that two functions f and g are equal* almost everywhere, *and we write*

$$f = g \qquad a.e.,$$

if the set

$$A := \{x \mid f(x) \neq g(x)\}$$

is a nullset.
If f and g are equal almost everywhere, and f is Lebesgue integrable, then so is g, and

$$\int f\, dx = \int g\, dx.$$

This means that the Lebesgue integral cannot distinguish between function which are equal almost everywhere. It can be proved that "= a.e." is an equivalence relation, and if an element of one of its equivalence classes is Lebesgue integrable, then all functions in this equivalence class are Lebesgue integrable with the same value of the their Lebesgue integral.

Roughly speaking, one may say that we get the Lebesgue integrable functions by taking all Riemann integrable functions and add to them all nullfunctions. This statement is not quite correct, "but almost". This inaccuracy will, however, not be of any importance in the following, so in order to just give a hint we mention that the standard example of a Riemann integrable function which by definition is not Lebesgue integrable, is the function $f(x) = \sin(x)/x$ for $x > 0$.

For convenience and later use we here also introduce the L^p spaces for $p \in [1, +\infty[$. This is easiest in the 1-dimensional case $L^p(\mathbb{R})$, so we shall only give the definitions in this 1-dimensional case. The trivial and tedious task of extending the definitions to $L^p\left(\mathbb{R}^N\right)$ is left to the reader. Finally we briefly introduce the spaces $L^p(\Omega)$ for subsets of $\mathbb{R}^N$.

2.1.1 The space $L^1(\mathbb{R})$

A slightly imprecise definition, which however is sufficient for us, is the following.

Given a function $f : \mathbb{R} \to \mathbb{R}$, or $\mathbb{C}$. Then f belongs to the space $L^1(\mathbb{R})$, if there exists a Riemann integrable function $g : \mathbb{R} \to \mathbb{R}$, or $\mathbb{C}$, such that

$$f = g \quad a.e., \qquad \text{and} \qquad \int_{-\infty}^{+\infty} |g(x)|\, dx < +\infty.$$

When this is the case, we define the Lebesgue integral of f by

$$\int_{-\infty}^{+\infty} f(x)\, dx := \int_{-\infty}^{+\infty} g(x)\, dx.$$

We define the 1-norm, $\|\cdot\|_1$, on $L^1(\mathbb{R})$ by

$$\|f\|_1 := \int_{-\infty}^{+\infty} |f(x)|\, dx = \int_{-\infty}^{+\infty} |g(x)|\, dx.$$

The elements of $L^1(\mathbb{R})$ are more precisely equivalence classes of functions, where two functions belong to the same class, if their difference is a nullfunction. For the same reason, $\|\cdot\|_1$ is strictly speaking not a norm, but a so-called semi-norm. In practice this will not bother us much, so instead of "an equivalence class of functions" we just say "a function", and we call $\|\cdot\|_1$ a norm in the following.

The elements of $L^1(\mathbb{R})$ are as mentioned above *equivalence classes* of functions with respect to the equivalence relation "$=$ a.e.". It is then possible to prove that $\|\cdot\|_1$ becomes a (semi-)norm on the space of these functions, and a norm on the space of equivalence classes, and that $\left(L^1(\mathbb{R}), \|\cdot\|_1\right)$ becomes a complete normed vector space. The latter means that every *Cauchy sequence* of elements (f_n) from $L^1(\mathbb{R})$ with respect to this (semi-)norm converges towards some limit function $f \in L^1(\mathbb{R})$. More precisely, if $(f_n) \subset L^1(\mathbb{R})$ satisfies the condition that to every $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}$, such that

$$\|f_m - f_n\|_1 < \varepsilon, \qquad \text{for all } m,\, n \geq n_0,$$

then there is a limit function $f \in L^1(\mathbb{R})$, i.e. for every $\varepsilon > 0$ there exists an $n_1 \in \mathbb{N}$, such that also

$$\|f - f_n\|_1 < \varepsilon, \qquad \text{for all } n \geq n_1.$$

A complete normed vector space is called a *Banach space*, a notation we shall use in the following.

2.1.2 The space $L^p(\mathbb{R})$ for $1 < p < +\infty$.

We use $L^1(\mathbb{R})$ to define the space $L^p(\mathbb{R})$ for $1 < p < +\infty$.

A function f belongs to $L^p(\mathbb{R})$, if there exists a locally Riemann integrable function g, i.e. g is Riemann integrable over every bounded open set, such that

$$f = g \quad a.e., \qquad and \qquad |f|^p \in L^1(\mathbb{R}).$$

When $f \in L^p(\mathbb{R})$, we define its (semi-)norm – called a norm in the following – by

$$\|f\|_p := \left\{ \int_{-\infty}^{+\infty} |f(x)|^p \, dx \right\}^{1/p}.$$

It should be mentioned as a warning that the condition $|f|^p \in L^1(\mathbb{R})$ is not enough. There exist functions f, such that $|f|^p \in L^1(\mathbb{R})$, and such that no locally Riemann integrable function g exists, such that also $f = g$ a.e.. By definition, such functions are *not* elements of $L^p(\mathbb{R})$.

It can be proved that $(L^p(\mathbb{R}), \|\cdot\|_p)$ is a *Banach space*.

Two positive numbers, p, $q > 1$, are called *conjugated*, if

$$\frac{1}{p} + \frac{1}{q} = 1.$$

We mention without proof the important

Hölder's inequality. *Let p, $q > 1$ be conjugated numbers. If $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$, then the pointwise product $f \cdot g \in L^1(\mathbb{R})$, and we have the estimate*

$$\|f \cdot g\|_1 \le \|f\|_p \cdot \|g\|_q.$$

The case $p = q = 2$ is of particular interest. In this case Hölder's inequality is called

Cauchy-Schwarz's inequality. *If f, $g \in L^2(\mathbb{R})$, then*

$$\left| \int_{-\infty}^{+\infty} f(x) \cdot \overline{g(x)} \, \mathrm{d}x \right| \le \|f\|_2 \cdot \|g\|_2.$$

We have here added a complex conjugation to the second factor in the integrand. The reason is that then the function $\langle \cdot, \cdot \rangle : L^2(\mathbb{R}) \times L^2(\mathbb{R}) \to \mathbb{C}$, given by

$$\langle f, g \rangle := \int_{-\infty}^{+\infty} f(x) \cdot \overline{g(x)} \, \mathrm{d}x, \qquad \text{for } f,\, g \in L^2(\mathbb{R}),$$

has properties, which are similar to the usual *inner product* in an *Euclidean vector space* or a finite dimensional complex vector space with an inner product. We therefore call the present function $\langle \cdot, \cdot \rangle$ (in the space $L^2(\mathbb{R})$) an inner product.

Then *Cauchy-Schwarz's inequality* is written

$$|\langle f,g\rangle| \le \|f\|_2 \cdot \|g\|_2, \qquad \text{for } f,\, g \in L^2(\mathbb{R}),$$

which has the same structure as the Cauchy-Schwarz inequality in Euclidean spaces.

We mentioned above that all L^p spaces are complete, i.e. *Banach spaces.* In particular, $L^2(\mathbb{R})$ is complete and equipped with an inner product, where the norm $\|\cdot\|_2$ is given by the inner product,

$$\|f\|_2^2 := \langle f,f\rangle, \qquad \text{for } f \in L^2(\mathbb{R}).$$

A complete vector space with an inner product is called a *Hilbert space.* The real finite dimensional Hilbert spaces are of course as before also called *Euclidean spaces.*

2.1.3 The space $L^\infty(\mathbb{R})$.

A function f belongs to $L^\infty(\mathbb{R})$, if there exists a locally Riemann integrable function g on $\mathbb{R}$, such that $f = g$ a.e., and if furthermore there is a constant $\lambda > 0$, such that the set

$$\{x \in \mathbb{R} \mid |f(x)| > \lambda\}$$

is a nullset. When this is the case we define the (semi-)norm $\|\cdot\|_\infty$ by

$$\|f\|_\infty := \inf\{\lambda > 0 \mid \text{the set } \{x \in \mathbb{R} \mid |f(x)| \lambda\} \text{ is a nullset}\}.$$

We obtain with respect to the equivalence relation "= a.e." that $(L^\infty(\mathbb{R}), \|\cdot\|_\infty)$ becomes a *Banach space.* Furthermore, *Hölder's inequality* is in this case written

$$\|f\cdot g\|_1 \le \|f\|_1 \cdot \|g\|_\infty, \qquad \text{for } f \in L^1(\mathbb{R}) \text{ and } g \in L^\infty(\mathbb{R}),$$

or, in a more understandable form,

$$\int_{-\infty}^{+\infty} |f(x)g(x)|\,\mathrm{d}x \le \|g\|_\infty \int_{-\infty}^{+\infty} |f(x)|\,\mathrm{d}x \qquad \text{for } f \in L^1(\mathbb{R}) \text{ and } g \in L^\infty(\mathbb{R}).$$

The extensions to the spaces $L^p\left(\mathbb{R}^N\right)$ of the spaces above are left to the reader.

2.1.4 The space $L^p(\Omega)$, where $\Omega \subseteq \mathbb{R}^N$ is measurable.

Let $\Omega \subseteq \mathbb{R}^N$ be any subset. We define the *indicator function* χ_Ω of Ω by

$$\chi_\Omega(x) = \begin{cases} 1 & \text{for } x \in \Omega, \\ 0 & \text{for } x \notin \Omega. \end{cases}$$

The set Ω is called a *measurable set*, if the indicator function for the set $\Omega \cap [-n,n]^N$ is Lebesgue integrable for every $n \in \mathbb{N}$. When this is the case, the *measure* of Ω is defined in the following way

$$|\Omega| := \lim_{n\to+\infty} \int_{\mathbb{R}^N} \chi_{\Omega\cap[-n,n]^N}(x)\,\mathrm{d}x = \lim_{n\to+\infty} \int_{\Omega\cap[-n,n]^N} \mathrm{d}x.$$

We notice that the measure $|\Omega|$ can be both finite and infinite.

A function f belongs to the space $L^p(\Omega)$, $\Omega \subset \mathbb{R}^N$, if the related function F, defined by

$$F(x) := \begin{cases} f(x) & \text{for } x \in \Omega, \\ 0 & \text{for } x \notin \Omega, \end{cases}$$

belongs to the space $L^p\left(\mathbb{R}^N\right)$, defined previously. In this case we define the p-norm over Ω by

$$\|f\|_{\Omega,p} := \|F\|_p.$$

When there is no risk of misunderstanding we just write $\|f\|_p$, when we should have written $\|f\|_{\Omega,p}$.

If Ω is a measurable set, then $L^p(\Omega)$ is a *Banach space* for all $p \in [1, +\infty]$, and a *Hilbert space* for $p = 2$.

2.2 Short overview of the contents of this chapter.

The first three examples require only some knowledge of the Riemann integral. In Section 2.3 we demonstrate that integration of a continuous function over a bounded, though not closed domain in the plane may cause some problems. For the chosen function the two possible axiparallel double integrals both exists, but they are not equal to each other. We shall not go too deep into the explanation, because that would involve the Lebesgue integral, and this is not the main subject here.

In Section 2.4 we construct a function, which is differentiable everywhere in $\mathbb{R}$, but its derivative is not absolutely integrable over $[0, 1]$. This gives us the hint that the classical concept of differentiation may have some nasty properties, which at best should be avoided. Later on we shall see even worse examples, both in the present chapter, as well as in Chapter 5.

In Section 2.5 we play a little with integration. We shall not yet apply the Lebesgue integral, although the example is better explained in Lebesgue theory. However, the basic calculations are easy to understand even with only the Riemann integral at hand. The starting point is the following nonsense, "one half indefinite integration from 0". Once the meaning of this phrase has been understood (it is called *fractional integration*), it is easy also to introduce "indefinite integration α times from 0", where α is any real, positive number, like e.g. $\sqrt{\pi}$, just to mention one of the more bizarre possible numbers. We shall not go through all details in the general formula, because this would require some knowledge of the Γ and the B functions.

In spite of the above, fractional integration is not as exotic as one would think. It actually occurs quite natural in some practical cases, including situations from the technical sciences. The earliest implicit application of this theory can even be traced back to *Abel* [1], [2], in connection with the problem of the tautochron.

We continue in Section 2.6 and Section 2.7 demonstrating that convergence in L^p is not in harmony with the usual well-known pointwise convergence. This is a hidden warning! The inexperienced student would always prefer to apply the pointwise convergence instead of the more useful convergence in L^p. In practice, when one has made some measurements, the natural concept will mostly be L^2, or L^1, or L^∞. These are used in the maximum likelihood method and in convergence in energy.

In Sections 2.8 and 2.9 we show that if $p \neq q$, then there are L^p functions which do not belong to L^q, and *vice versa*.

In Section 2.10 we construct a function $f : [0,1] \to \mathbb{R}$, which is differentiable everywhere with a bounded derivative f', and yet f' is not Riemann integrable. This can be proved by using the Riemann theory alone, but in order not to be too elaborated we shall here allow ourselves without proof to apply a theorem from the Lebesgue theory.

Finally, we construct in Section 3.3 an "universal indefinite integral" (in a sense which shall be explained later on) of every measurable function.

2.3 A continuous function f on $]0,1[^2$, where both iterated integrals exist without having the same value.

The simplest possible function of this type is left as an exercise to the reader at the end of this section. However, due to a discontinuity of this example at some points, we shall here show the same phenomena, where the function under consideration is continuous.

For simplicity we first consider the function $g : [1,+\infty[\times [0,1[\to \mathbb{R}$ given by

$$g(x,y) = e^{-xy} - 2\,e^{-2xy}, \qquad \text{for } (x,y) \in I_1 \times I_2 = [1,+\infty[\times [0,1].$$

We shall later use a coordinate transformation to construct a function which is continuous and bounded on a bounded set. Obviously g is continuous and bounded in $I_1 \times I_2$.

Then we calculate the iterated integral

$$\int_{I_1}\left\{\int_{I_2} g(x,y)\,\mathrm{d}y\right\}\mathrm{d}x=\int_1^{+\infty}\left\{\int_0^1\left(e^{-xy}-2e^{-2xy}\right)\,\mathrm{d}y\right\}\mathrm{d}x.$$

Fix any $x\geq 1$. Then we get for the inner integral

$$\int_0^1\left\{e^{-xy}-2e^{-2xy}\right\}\,\mathrm{d}y=\left[-\frac{1}{x}\,e^{-xy}+\frac{2}{2x}\,e^{-2xy}\right]_{y=0}^{1}=\frac{1}{x}\left\{e^{-2x}-e^{-x}\right\}.$$

Using for $x\geq 1$ the inequality

$$\left|\frac{1}{x}\left\{e^{-2x}-e^{-x}\right\}\right|\leq e^{-x}-e^{-2x}\leq e^{-x},\qquad\text{for } x\geq 1,$$

and the result $\int_1^{+\infty}e^{-x}\,\mathrm{d}x=e^{-1}<+\infty$, we conclude that

$$\int_1^{+\infty}\left\{\int_0^1\left(e^{-xy}-2e^{-2xy}\right)\,\mathrm{d}y\right\}\mathrm{d}x=\int_1^{+\infty}\frac{1}{x}\left(e^{-2x}-e^{-x}\right)\,\mathrm{d}x$$

exists. Accordingly, the inequality $\left(e^{-2x}-e^{-x}\right)/x<0$ for $x\geq 1$ implies that

$$\int_1^{+\infty}\left\{\int_0^1\left(e^{-xy}-2e^{-2xy}\right)\,\mathrm{d}y\right\}\mathrm{d}x<0.$$

Then we consider the other iterated integral, where the order of integration is reversed,

$$\int_{I_2}\left\{\int_{I_1} f(x,y)\,\mathrm{d}x\right\}\mathrm{d}y=\int_0^1\left\{\int_1^{+\infty}\left(e^{-xy}-2e^{-2xy}\right)\,\mathrm{d}x\right\}\mathrm{d}y.$$

Let $y>0$ be fixed. Then we get for the inner integral

$$\int_1^{-\infty}\left(e^{-xy}-2e^{-2xy}\right)\,\mathrm{d}x=\left[-\frac{1}{y}\,e^{-xy}+\frac{2}{2y}\,e^{-2xy}\right]_{x=1}^{+\infty}=\frac{1}{y}\left(e^{-y}-e^{-2y}\right).$$

Since

$$\frac{1}{y}\left(e^{-y}-e^{-2y}\right)=e^{-y}\cdot\frac{1-e^{-y}}{y}\to 1,\qquad\text{for } y\to 0+,$$

we define $\left(e^{-y}-e^{-2y}\right)/y$ as a continuous function on the closed interval $[0,1]$, when we add the value 1 for $y=0$. Then

$$\int_0^1\left\{\int_1^{+\infty}\left(e^{-xy}-2e^{-2xy}\right)\,\mathrm{d}x\right\}\mathrm{d}y=\int_0^1\frac{1}{y}\left(e^{-y}-e^{-2y}\right)\,\mathrm{d}y$$

clearly exists. It follows from

$$\frac{1}{y}\left(e^{-y}-e^{-2y}\right)=\frac{1}{y}\,e^{-y}\left(1-e^{-y}\right)>0,\qquad\text{for } y\in\,]0,1],$$

that

$$\int_0^1\left\{\int_1^{+\infty}\left(e^{-xy}-2e^{-2xy}\right)\,\mathrm{d}x\right\}\mathrm{d}y>0,$$

and we have proved that

$$\int_1^{+\infty}\left\{\int_0^1\left(e^{-xy}-2e^{-2xy}\right)\,\mathrm{d}y\right\}\mathrm{d}x<0 \qquad\text{and}\qquad \int_0^1\left\{\int_1^{+\infty}\frac{1}{x^2}\left(e^{-xy}-2e^{-2xy}\right)\,\mathrm{d}x\right\}\mathrm{d}y>0.$$

Using the coordinate transformation $x = 1/t$ these iterated integrals are transformed into

$$\int_0^1\left\{\int_0^1\frac{1}{x^2}\left\{\exp\left(-\frac{y}{x}\right)-2\,\exp\left(-\frac{2y}{x}\right)\right\}\mathrm{d}y\right\}\mathrm{d}x<0,$$

and

$$\int_0^1\left\{\int_0^1\frac{1}{x^2}\left\{\exp\left(-\frac{y}{x}\right)-2\,\exp\left(-\frac{2y}{x}\right)\right\}\mathrm{d}x\right\}\mathrm{d}y>0.$$

To finish the example we just note that the integrand

$$f(x,y)=\frac{1}{x^2}\left\{\exp\left(-\frac{y}{x}\right)-2\,\exp\left(-\frac{2y}{x}\right)\right\},\qquad\text{for } (x,y)\in\,]0,1],$$

is continuous.

It is of course inconvenient that the two iterated integrals do not agree. We therefore quote the correct theorem for Lebesgue theory,

Fubini's theorem *Let $f = f(x,y)$ be a Lebesgue integrable function in $\mathbb{R}^2$. For almost every fixed y the function $f(x,y)$ a Lebesgue integrable function of x. Its integral $\int f(x,y)\,dx$ defines (almost everywhere) a Lebesgue integrable funktion of y, and we get*

$$\int f(z)\mathrm{d}z=\int\left(\int f(x,y)\mathrm{d}x\right)\mathrm{d}y=\int\left(\int f(x,y)\mathrm{d}y\right)\mathrm{d}x.$$

where $\mathrm{d}z$ denotes the area element of $\mathbb{R}^2$.

Alternatively, it can be proved that if f is continuous – or piecewise continuous – on $\mathbb{R}^2$, then

$$\text{(2.1)}\qquad \int\left(\int|f(x,y)|\,\mathrm{d}y\right)\mathrm{d}x<+\infty,$$

implies that f is a Lebesgue integrable fnction, and we can then interchange the order of integration.

In the present case, the function

$$\text{(2.2)}\qquad g(x,y)=e^{-xy}-2e^{-2xy}=e^{-2xy}\left(e^{xy}-2\right)$$

is continuous for $(x,y)\in A=[1,+\infty[\,\times[0,1]$, so we must show that (2.1) does not hold for g.

It follows from (2.2) that

$$g(x,y)>0\quad\text{for } xy>\ln 2,\qquad\text{and}\qquad g(x,y)<0\quad\text{for } xy<\ln 2.$$

Hence,

$$\int_A|g(x,y)|\,\mathrm{d}x\,\mathrm{d}y=\int_{\{(x,y)\in A|xy>\ln 2\}}g(x,y)\,\mathrm{d}x\,\mathrm{d}y+\int_{\{(x,y)\in A|xy\le\ln 2\}}(-g(x,y))\,\mathrm{d}x\,\mathrm{d}y,$$

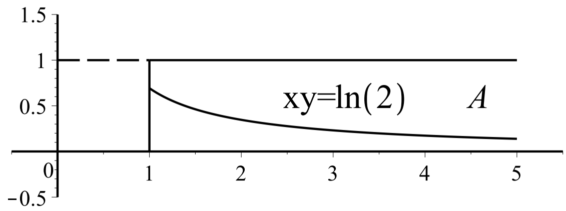

Figure 2.1: *The domain A of integration in Section 2.3.*

where both integrals are either positive or even $+\infty$. In order to prove (2.1) it therefore suffices to show that at least one of the two integrals on the right hand side is $+\infty$. Let us just for the exercise calculate the second integral,

$$\int_{\{(x,y)\in A \mid xy \leq \ln 2\}} (-g(x,y))\,\mathrm{d}x\,\mathrm{d}y = \int_1^{+\infty} \left\{ \int_0^{\ln(2)/x} \left\{ 2\,e^{-2xy} - e^{-xy} \right\} \mathrm{d}y \right\} \mathrm{d}x$$

$$= \int_1^{+\infty} \left[\frac{1}{x} \left(e^{-xy} - e^{-2xy} \right) \right]_{y=0}^{\ln(2)/x} = \int_1^{+\infty} \frac{1}{x} \left\{ e^{-\ln 2} - e^{-2\ln 2} \right\} \mathrm{d}x = \frac{1}{4} \int_1^{+\infty} \frac{1}{x}\,\mathrm{d}x = +\infty,$$

and (2.1) is not fulfilled.

It is left to the reader to go through the similar, though simpler analysis for the function

$$f(x,y) = \begin{cases} \dfrac{x^2 - y^2}{\left(x^2+y^2\right)^2} & \text{for } (x,y) \in [0,1]^2 \setminus \{(0,0)\}, \\[2ex] 0 & \text{for } (x,y) = (0,0). \end{cases}$$

Show that $f(x,y)$ is not continuous at $(0,0)$, and that

$$\int_0^1 \left\{ \int_0^1 f(x,y)\,\mathrm{d}x \right\} \mathrm{d}y = \frac{\pi}{4}, \qquad \text{and} \qquad \int_0^1 \left\{ \int_0^1 f(x,y)\,\mathrm{d}y \right\} \mathrm{d}x = -\frac{\pi}{4}.$$

The function is simpler in structure than the previous one considered. It is, however, discontinuous at the point $(0,0)$.

2.4 A function $f : \mathbb{R} \to \mathbb{R}$, which is differentiable everywhere, though the derivative f' is not absolutely integrable over $[0,1]$.

Let us consider the function

$$f(x) = \begin{cases} x^2 \sin\left(\dfrac{1}{x^2}\right), & \text{for } x \in \mathbb{R} \setminus \{0\}, \\ 0 & \text{for } x = 0. \end{cases}$$

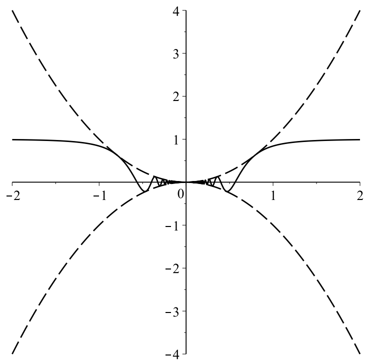

Figure 2.2: *The graph of the function $f(x)$ of Section 2.4. (Not to scale)*

It is obvious that $f : \mathbb{R} \to \mathbb{R}$ is continuous and that $f \in C^\infty(\mathbb{R} \setminus \{0\})$. The derivative of f is for $x \neq 0$ given by

$$f'(x) = 2x \sin\left(\frac{1}{x^2}\right) - \frac{2}{x} \cos\left(\frac{1}{x^2}\right), \qquad \text{for } x \in \mathbb{R} \setminus \{0\}.$$

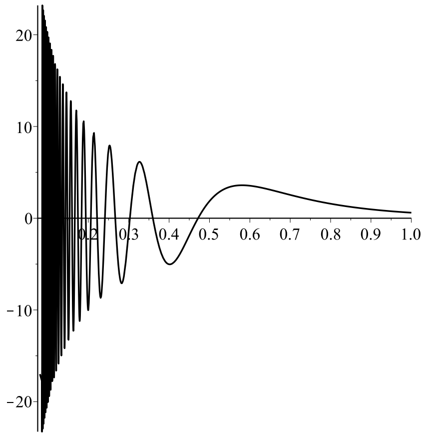

Figure 2.3: *The graph of the derivative* $f(x)$ *of Section 2.4 for* $0.1 < x < 1$. *(Not to scale)*

It is seen on Figure 2.3 that something violent is taking place close to $x = 0$. It is amazing that the derivative exists at $x = 0$. Here we apply the classical definition of the derivation,

$$f'(0) = \lim_{x \to 0} \frac{f(x) - f(0)}{x} = \lim_{x \to 0} x \cdot \sin\left(\frac{1}{x^2}\right) = 0,$$

because $\left|x \sin\left(x^{-2}\right)\right| \leq |x|$ for all $x \neq 0$. Therefore, the derivative $f'(x)$ is given by

$$f'(x) = \begin{cases} 2x \sin\left(\frac{1}{x^2}\right) - \frac{2}{x} \cos\left(\frac{1}{x^2}\right), & \text{for } x \neq 0, \\ 0, & \text{for } x = 0. \end{cases}$$

The derivative f' is clearly discontinuous at $x = 0$.

Since the function

$$\varphi(x) = \begin{cases} 2x \sin\left(\frac{1}{x^2}\right), & \text{for } x \in]0,1], \\ 0, & \text{for } x = 0, \end{cases}$$

is continuous, and the interval $[0, 1]$ is closed and bounded, the absolute value $|\varphi(x)|$ is also continuous, and the integral $\int_0^1 |\varphi(x)| \, \mathrm{d}x$ is finite. Therefore, we shall only prove that

$$\frac{2}{x} \cos\left(\frac{1}{x^2}\right)$$

is not absolutely integrable over $]0,1]$ to conclude that f' is not absolutely integrable.

Choose any $\varepsilon \in]0,1]$. Then by the change of variable $y = 1/x^2$,

$$\int_{\varepsilon}^{1} \left| \frac{2}{x} \cos\left(\frac{1}{x^2}\right) \right| \mathrm{d}x = \int_{1}^{1/\varepsilon^2} |2\sqrt{y}\,\cos y| \cdot \frac{1}{2}\,\frac{1}{y\sqrt{y}}\,\mathrm{d}y = \int_{1}^{1/\varepsilon^2} \frac{1}{y}\,|\cos y|\,\mathrm{d}y.$$

Hence, by taking the limit $\varepsilon \to 0+$, followed by a trivial estimate

$$\int_{0}^{1} \left| \frac{2}{x} \cos\left(\frac{1}{x^2}\right) \right| \mathrm{d}x = \int_{1}^{+\infty} \frac{1}{y}\,|\cos y|\,\mathrm{d}y \geq \sum_{p=0}^{+\infty} \int_{\frac{3\pi}{4}+p\pi}^{\frac{5\pi}{4}+p\pi} \frac{1}{y}\,|\cos y|\,\mathrm{d}y.$$

Notice that

$$|\cos y| \geq \frac{1}{\sqrt{2}}, \qquad \text{for } y \in \left[\frac{3\pi}{4}+p\pi, \frac{5\pi}{4}+p\pi\right], \qquad p \in \mathbb{N}_0.$$

We get the estimates

$$\begin{aligned}\int_{0}^{1} \left| \frac{2}{x} \cos\left(\frac{1}{x^2}\right) \right| \mathrm{d}x &\geq \sum_{p=0}^{+\infty} \frac{1}{\sqrt{2}} \int_{\frac{3\pi}{4}+p\pi}^{\frac{5\pi}{4}+p\pi} \frac{1}{y}\,\mathrm{d}y \geq \sum_{p=0}^{+\infty} \frac{1}{\sqrt{2}} \int_{\frac{3\pi}{4}+p\pi}^{\frac{5\pi}{4}+p\pi} \frac{1}{\frac{5\pi}{4}+p\pi}\,\mathrm{d}y \\ &= \sum_{p=0}^{+\infty} \frac{1}{\sqrt{2}} \cdot \frac{\pi}{2} \cdot \frac{4}{5\pi+4p\pi} = \sqrt{2} \sum_{p=0}^{+\infty} \frac{1}{5+4p} \geq \frac{\sqrt{2}}{4} \sum_{p=2}^{+\infty} \frac{1}{p} = +\infty,\end{aligned}$$

and the claim is proved.

2.5 Fractional integration

We shall in this section define what it means to *"integrate a function α times from 0"* for all $\alpha \in \mathbb{R}_+$, and not just for $\alpha \in \mathbb{N}$. We shall use the Γ function, defined by

$$\Gamma(\alpha) := \int_0^{+\infty} t^{\alpha-1}\, e^{-t}\,\mathrm{d}t, \qquad \text{for } \alpha > 0,$$

and without proof take for granted that

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \qquad \text{and} \qquad \Gamma(\alpha+1) = \alpha\Gamma(\alpha) \quad \text{for } \alpha > 0,$$

and of course that

$$\Gamma(n) = (n-1)! \qquad \text{for } n \in \mathbb{N}.$$

We shall also need the B function (the Beta function), defined by

$$B(\alpha,\beta) := \int_0^1 (1-t)^{\alpha-1}\, t^{\beta-1}\,\mathrm{d}t = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Let us first consider the linear operator A, given on $L^1([0,1])$ by

$$(2.3) \qquad Af(x) := \frac{1}{\sqrt{\pi}} \int_0^x \frac{f(x)}{\sqrt{x-t}}\,\mathrm{d}t, \qquad \text{for } x \in [0,1] \text{ and } f \in L^1([0,1]).$$

If $f \geq 0$, then clearly $Af \geq 0$, and it follows from *Fubini's theorem*, cf. page 59, that

$$\begin{aligned}\|Af\|_1 &= \int_0^1 Af(x)\,\mathrm{d}x = \frac{1}{\sqrt{\pi}}\int_0^1\left\{\int_0^x \frac{f(x)}{\sqrt{x-t}}\,\mathrm{d}t\right\}\mathrm{d}x = \frac{1}{\sqrt{\pi}}\int_0^1\left\{\int_t^1 \frac{f(t)}{\sqrt{x-t}}\,\mathrm{d}x\right\}\mathrm{d}t \\ &= \frac{2}{\sqrt{\pi}}\int_0^1 \sqrt{1-t}\,f(t)\,\mathrm{d}t \leq \frac{2}{\sqrt{\pi}}\int_0^1 f(t)\,\mathrm{d}t = \frac{1}{\sqrt{\pi}}\,\|f\|_1.\end{aligned}$$

Suppose that f is real. Then there exist functions f^+ and $f^- \geq 0$, such that $f = f^? - f^-$ and $|f| = f^+ + f^-$. In fact, the unique splitting is given by

$$f^+ = \frac{1}{2}(|f| + f) \qquad \text{and} \qquad f^- = \frac{1}{2}(|f| - f).$$

By the estimate above,

$$\begin{aligned}\|Af\|_1 &= \int_0^1 |Af(x)|\,\mathrm{d}x \leq \int_0^1 \left\{Af^+(x) + Af^-(x)\right\}\mathrm{d}x = \int_0^1 A|f|(x)\,\mathrm{d}x \\ &= \frac{2}{\sqrt{\pi}}\int_0^1 \sqrt{1-t}\cdot|f(t)|\,\mathrm{d}t \leq \frac{2}{\sqrt{\pi}}\,\|f\|_1.\end{aligned}$$

Finally, if f is a complex function we get by copying the previous method that also $|Af(x)| \leq A|f|(x)$, and we have proved that in general,

$$\|Af\|_1 \leq \frac{2}{\sqrt{\pi}}\,\|f\|_1, \tag{2.4}$$

and also that $Af \in L^1([0,1])$, when $f \in L^1([0,1])$, so A is a bounded linear operator,

$$A : L^1([0,1]) \to L^1([0,1]),$$

where

$$\|A\| := \sup\left\{\|Af\|_1 \mid f \in L^1([0,1]) \text{ and } \|f\|_1 = 1\right\} \leq \frac{2}{\sqrt{\pi}}.$$

Remark 2.1 If we in particular choose

$$f_n(x) := n \cdot \chi_{[0,\frac{1}{n}]}(x) = \begin{cases} n, & \text{for } n \in \left[0, \dfrac{1}{n}\right], \\ 0, & \text{otherwise,}\end{cases}$$

then it is easy to prove that $\|A\| = 2/\sqrt{\pi}$. We shall not use this in the following, so the details are left to the reader.

Furthermore, it is also left to the reader to prove that $Af \in L^\infty([0,1])$, when $f \in L^1([0,1])$. The reason is that the estimate (2.4) is sufficient here, so we shall not prove this claim either. $\Diamond$

We shall use the monotonous substitution

$$t := (x-u)\sin^2\theta + u \qquad \text{for } \theta \in \left[0, \frac{\pi}{2}\right].$$

Let u and x be fixed numbers, where $u < x$. Then

$$(2.5) \quad \int_u^x \frac{\mathrm{d}t}{\sqrt{(x-t)(t-u)}} = \int_0^{\frac{\pi}{2}} \frac{(x-u)\cdot 2\sin\theta\cdot\cos\theta}{\sqrt{\left((x-u)-(x-u)\sin^2\theta\right)(x-u)\cdot\sin^2\theta}}\,\mathrm{d}\theta$$

$$= 2\int_0^{\frac{\pi}{2}} \frac{\cos\theta\cdot\sin\theta}{\sqrt{\cos^2\theta\cdot\sin^2\theta}}\,\mathrm{d}\theta = 2\int_0^{\frac{\pi}{2}} \mathrm{d}\theta = \pi.$$

If $f \in L^1([0,1])$, then it follows from the above that also $Af \in L^1([0,1])$, so $A^2f(x) = A(Af)(x)$ makes sense. By iterating (2.3) it follows from *Fubini's theorem*, cf. page 59, and (2.5) that

$$A^2f(x) = \frac{1}{\sqrt{\pi}}\int_0^x \frac{1}{\sqrt{x-t}}\,Af(t)\,\mathrm{d}t = \frac{1}{\sqrt{\pi}}\int_0^x \frac{1}{\sqrt{x-t}}\left\{\frac{1}{\sqrt{\pi}}\int_0^t \frac{f(u)}{\sqrt{t-u}}\,\mathrm{d}u\right\}\mathrm{d}t$$

$$= \frac{1}{\pi}\int_0^x \left\{\int_u^x \frac{\mathrm{d}t}{\sqrt{(x-t)(t-u)}}\right\} f(u)\,\mathrm{d}u = \frac{1}{\pi}\int_0^x \pi f(u)\,\mathrm{d}u = \int_0^x f(t)\,\mathrm{d}t.$$

Define the integration operator $B : L^1([0,1]) \to L^1([0,1])$ by

$$Bf(x) := \int_0^x f(t)\,\mathrm{d}t, \qquad \text{for } x \in [0,1] \text{ and } f \in L^1([0,1]).$$

We see that Bf becomes the function we get by integrating f from $x_0 = 0$. Then it follows from the above that $B = A^2 = A \circ A$, so roughly speaking, the operator A may be interpreted as *"integration $\frac{1}{2}$ times from 0"*.

We then prove by induction

$$(2.6) \qquad B^n(x) = \frac{1}{(n-1)!}\int_0^x (x-t)^{n-1} f(t)\,\mathrm{d}t, \qquad \text{for } f \in L^1([0,1]).$$

Clearly, this formula is according to the above correct for $n = 1$. Then assume that it holds for some $n \in \mathbb{N}$. We get by *Fubini's theorem*

$$\begin{aligned} B^{n+1}f(x) &= B^n(Bf)(x) = \frac{1}{(n-1)!}\int_0^x (x-t)^{n-1}\left\{\int_0^t f(u)\,\mathrm{d}u\right\}\mathrm{d}t \\ &= \frac{1}{(n-1)!}\int_0^x f(u)\left\{\int_u^x (x-t)^{n-1}\,\mathrm{d}t\right\}\mathrm{d}u = \frac{1}{(n-1)!}\int_0^x f(u)\,\frac{1}{n}(x-u)^n\,\mathrm{d}u \\ &= \frac{1}{n!}\int_0^x (x-t)^n f(t)\,\mathrm{d}t, \end{aligned}$$

and we get (2.6), when n is replaced by $n+1$. The formula then follows by induction.

In particular, n successive integrations from 0 can be replaced by one single integration from 0 by multiplying the integrand by the kernel

$$\frac{1}{(n-1)!}\,(x-t)^{n-1}, \qquad \text{for } t \in [0,x],$$

and where the kernel is put $= 0$ outside this interval.

Using that $(n-1)! = \Gamma(n)$ and $\sqrt{\pi} = \Gamma\left(\frac{1}{2}\right)$, it follows that (2.6) and (2.3) can be written in thee general form

$$(2.7) \qquad B^\alpha f(x) = \frac{1}{\Gamma(\alpha)}\int_0^x (x-t)^{\alpha-1} f(t)\,\mathrm{d}t, \qquad \text{for } f \in L^1([0,1]),$$

where either $\alpha = n \in \mathbb{N}$ or $\alpha = \frac{1}{2}$. Let us consider (2.7) for a general $\alpha > 0$.

If $f \geq 0$ and $f \in L^1([0,1])$, it follows from the definition that $B^\alpha f \geq 0$, thus by *Fubini's theorem*,

$$\begin{aligned} \|B^\alpha f\|_1 &= \int_0^1 B^\alpha f(x)\,\mathrm{d}x = \frac{1}{\Gamma(\alpha)}\int_0^1\left\{\int_0^x (x-t)^{\alpha-1} f(t)\,\mathrm{d}t\right\}\mathrm{d}x \\ &= \frac{1}{\Gamma(\alpha)}\int_0^1 f(t)\left\{\int_t^1 (x-t)^{\alpha-1}\,\mathrm{d}x\right\}\mathrm{d}t = \frac{1}{\alpha\Gamma(\alpha)}\int_0^1 (1-t)^\alpha f(t)\,\mathrm{d}t \\ &\leq \frac{1}{\Gamma(\alpha+1)}\,\|f\|_1, \end{aligned}$$

which proves that $B^\alpha f \in L^1([0,1])$ for $f \in L^1([0,1])$ and $f \geq 0$.

Using once more the previous tedious arguments (left to the reader) we extend this estimate to all $f \in L^1([0,1])$, so we have proved for all $\alpha > 0$ that B^α is linear,

$$(2.8) \qquad B^\alpha : L^1([0,1]) \to L^1([0,1]), \qquad \text{and} \qquad \|B^\alpha\| \leq \frac{1}{\Gamma(\alpha+1)}.$$

So far we have strictly speaking only considered α in B^α as an index. However, if we also can prove that in general $B^\alpha B^\beta = B^{\alpha+\beta}$ for all α and $\beta > 0$, then α may also be considered as an exponent, and that we can interpret B^α as an *"integration α times from 0"*, because this is true for $\alpha = 1$.

So, let us assume that α, $\beta > 0$ are given. Then we get by composition of the operators,

$$
\begin{aligned}
B^\alpha B^\beta f(x) &= \frac{1}{\Gamma(\alpha)} \iint_0^x (x-t)^{\alpha-1} B^\beta f(t)\,\mathrm{d}t \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x (x-t)^{\alpha-1} \left\{ \int_0^t (t-u)^{\beta-1} f(u)\,\mathrm{d}u \right\} \mathrm{d}t \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x \left\{ \int_u^x (x-t)^{\alpha-1}(t-u)^{\beta-1}\,\mathrm{d}t \right\} f(u)\,\mathrm{d}u.
\end{aligned}
$$

We calculate the inner integral,

$$
\begin{aligned}
\int_u^x (x-t)^{\alpha-1}(t-u)^{\beta-1}\,\mathrm{d}t &= \int_0^{x-u} (x-u-t)^{\alpha-1} t^{\beta-1}\,\mathrm{d}t \\
&= (x-u)^{\alpha-1}(x-u)^{\beta-1}(x-u) \int_0^1 (1-\tau)^{\alpha-1}\tau^{\beta-1}\,\mathrm{d}\tau \\
&= (x-u)^{\alpha+\beta-1} B(\alpha,\beta) = (x-u)^{\alpha+\beta-1} \cdot \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)},
\end{aligned}
$$

where we have used the definition of the Beta function. Assuming that α and $\beta > 0$ we get by insertion,

$$
\begin{aligned}
B^\alpha B^\beta f(x) &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \int_0^x (x-u)^{\alpha+\beta-1} f(u)\,\mathrm{d}u \\
&= \frac{1}{(\alpha+\beta)} \int_0^x (x-t)^{\alpha+\beta-1} f(t)\,\mathrm{d}t = B^{\alpha+\beta} f(x),
\end{aligned}
$$

and the claim is proved.

2.6 A sequence of functions (f_n), which converges in $L^p(\mathbb{R})$ for all $p \geq 1$, and which does not converge pointwise at any point $x \in \mathbb{R}$.

Given any $p \geq 1$. We remind the reader of that according to Section 2.1 a measurable function f belongs to $L^p(\mathbb{R})$, if

$$\|f\|_p := \left\{ \int_{-\infty}^{+\infty} |f(x)|^p \,\mathrm{d}x \right\}^{\frac{1}{p}} < +\infty.$$

A sequence (f_n) from $L^p(\mathbb{R})$ converges in $L^p(\mathbb{R})$ towards a function $f \in L^p(\mathbb{R})$, if

$$\|f - f_n\|_p \to 0, \qquad \text{for } n \to +\infty.$$

A sequence (f_n) of functions on $\mathbb{R}$ converges pointwise towards a function f at a fixed point $x \in \mathbb{R}$, if

$$f_n(x) \to f(x), \qquad \text{for } n \to +\infty.$$

After these reminders we construct our sequence of functions in the following way,

$$
\begin{array}{ccccc}
f_0 = \chi_{]-1,1]}, & * & & * & * \\
f_1 = \chi_{]-2,-1]}, & f_2 = \chi_{[-1,0]} & & f_3 = \chi_{]0,1]}, & f_4 = \chi_{]1,2]}, \\
f_5 = \chi_{]-3,-\frac{5}{2}]}, & f_6 = \chi_{]-\frac{5}{2},-2]}, & \cdots & f_{15} = \chi_{]2,\frac{5}{2}]}, & f_{16} = \chi_{]\frac{5}{2},3]} \\
* & * & \vdots & * & * \\
f_{(n-2)2^{n-1}+1} = \chi_{]-n,-n+2^{2-n}]} & * & \cdots & * & f_{(n-1)2^n} = \chi_{]n-2^{n-2},n]} \\
* & * & \vdots & * & *
\end{array}
$$

The construction goes as follows. In the n-th row we define $n \cdot 2^{n-1}$ functions of the type $\chi_{]a,b]}$, where $b - a = 2^{2-n}$, such that every point from the interval $]-n, n]$ lies in precisely one of these intervals. Then for every fixed $x \in \mathbb{R}$, the sequence $(f_m(x))$ takes the values 0 and 1, both taken infinitely often. In particular it cannot be convergent at any $x \in \mathbb{R}$, so (f_m) is not pointwise convergent anywhere on the set of real numbers, $\mathbb{R}$.

On the other hand, if $m \in \left\{(n-2) \cdot 2^{n-1} + 1, \dots, (n-1) \cdot 2^n\right\}$, then clearly

$$\|f_m - 0\|_p = \|f_m\|_p = 2^{\frac{2-n}{p}} \to 0, \text{for } n \to +\infty,$$

because then also $m \to +\infty$. We therefore conclude that (f_n) converges in $L^p(\mathbb{R})$ for every $p \in [1, +\infty[$ towards the zero function. Notice, however, that (f_n) does not converge in $L^\infty(\mathbb{R})$! (Why not?)

2.7 A sequence from $L^p(\mathbb{R})$, which does not converge in $L^p(\mathbb{R})$, and which is pointwise convergent for all $x \in \mathbb{R}$.

If we define

$$f_n(x) := \begin{cases} n^2, & \text{for } x \in \left[n, n + \dfrac{1}{n}\right], \\[2ex] 0, & \text{otherwise,} \end{cases}$$

then clearly

$$f_n(x) \to 0, \qquad \text{for } n \to +\infty, \qquad \text{all } x \in \mathbb{R},$$

so the sequence of function (f_n) converges pointwise everywhere towards the zero function.

On the other hand, for every $p \in [1, +\infty[$,

$$\|f_n\|_p = \left\{\int_n^{n+\frac{1}{n}} n^{2p}\,\mathrm{d}x\right\}^{\frac{1}{p}} = n^{\frac{2p-1}{p}} \to +\infty, \qquad \text{for } n \to +\infty.$$

This implies that (f_n) does not converge in $L^p(\mathbb{R})$ for any $p \in [1, +\infty[$. Notice that it is trivial that it does not converge in the space $L^\infty(\mathbb{R})$ either.

2.8 A function from $L^1([0,1])$, which does not belong to any $L^p([0,1])$ for $p > 1$.

It follows from *Hölder's inequality*, cf. page 54, that

$$L^p([0,1]) \subseteq L^1([0,1]), \qquad \text{for all } p > 1.$$

In fact, $\chi_{[0,1]} \in L^q([0,1])$ trivially, where (p,q) are conjugate numbers, and thus

$$\|f\|_1 = \int_0^1 |f(x)| \cdot \chi_{[0,1]}(x)\,\mathrm{d}x \le \|f\|_p \left\{ \int_0^1 \chi_{[0,1]}{}^q(x)\,\mathrm{d}x \right\}^{\frac{1}{q}} = \|f\|_p,$$

so if $\|f\|_p < +\infty$, then also $\|f\|_1 < +\infty$, and we have proved that $L^p([0,1]) \subseteq L^1([0,1])$.

Fix $p > 1$. We define a function $f_p \in L^1([0,1])$, which does not belong to $L^p([0,1])$, by

$$f_p(x) := \begin{cases} x^{-\frac{1}{p}}, & \text{for } x \in]0,1], \\ 0, & \text{for } x = 0. \end{cases}$$

In fact, $f_p(x) \geq 0$ for all $x \in]0,1]$, and

$$\int_0^1 \{f_p(x)\}^p = \int_0^1 \frac{1}{x}\,\mathrm{d}x = +\infty,$$

so $f_p \notin L^p([0,1])$, and as it is not bounded, also $f_p \notin L^\infty([0,1])$.

On the other hand,

$$\|f\|_1 = \int_0^1 x^{-\frac{1}{p}}\,\mathrm{d}x = \frac{p}{p-1} < +\infty. \tag{2.9}$$

Notice also that f_p is only discontinuous at the nullset $\{0\}$.

The above is the standard example of a function, which lies in $L^1([0,1])$, but not in $L^p([0,1])$ for $p > 1$.

Then we use the construction above to define a function $f \in L^1([0,1])$, which does not belong to any $L^p([0,1])$, $p > 1$. Using a well-known trick we define

$$f(x) = \sum_{n=1}^{+\infty} \frac{1}{2^n}\, f_{1+\frac{1}{n}}(x), \qquad x \in [0,1].$$

We shall show that $f \notin L^p([0,1])$ for all $p > 1$, and that $f \in L^1([0,1])$.

First notice that for every $p > 1$ there is an $n \in \mathbb{N}$, such that $1 + \frac{1}{n} < p$. Then because all $f_p \geq 0$,

$$f(x) \geq 2^{-n}\, f_{1+\frac{1}{n}}(x) \geq 2^{-n}\, f_p(x).$$

Since f_p does not belong to $L^p([0,1])$, it follows that also $f \notin L^p([0,1])$. This holds for all $p > 1$, so we conclude that

$$f \notin \bigcup_{p>1} L^p([0,1]).$$

On the other hand, we get from (2.9) that

$$\|f\|_1 = \sum_{n=1}^{+\infty} \frac{1}{2^n} \left\| f_{1+\frac{1}{n}} \right\|_1 = \sum_{n=1}^{+\infty} (n+1)2^{-n} = 3 < +\infty,$$

so $f \in L^1([0,1])$. We have proved that

$$L^1([0,1]) \setminus \bigcup_{p>1} L^p([0,1]) \neq \emptyset.$$

2.9 A function $f \in L^p(\mathbb{R})$, which does not belong to any $L^q(\mathbb{R})$ for $q \neq p$ and p, $q \in [1, +\infty]$.

For $p = 1$ this was already proved in Section 2.8, because by an obvious embedding we trivially have $L^1([0,1]) \subset L^1(\mathbb{R})$. The present task is to prove this statement in general.

If $p = +\infty$, then $f = \chi_{\mathbb{R}}$ is bounded, so it belongs to L^∞. Since $\mathbb{R}$ has infinite measure, it is obvious that $\chi_{\mathbb{R}} \notin L^q(\mathbb{R})$ for any $q \in [1, +\infty[$.

Finally, we consider the remaining case, where $1 < p < +\infty$. We may somehow use the construction given in Section 2.8, which we shall do here. We shall, however, at the end of this section shall describe an alternative function.

We define as in Section 2.8,

$$g_q(x) := \begin{cases} x^{-\frac{1}{q}}, & \text{for } x \in]0,1], \\ \\ 0, & \text{otherwise,} \end{cases} \qquad\qquad \text{for } q > p.$$

Then $g_q \geq 0$, and (g_q) is decreasing in q, and $g_q \notin L^q(\mathbb{R})$, cf. Section 2.8, while

$$\|g_q\|_p = \left\{ \int_0^1 x^{-\frac{p}{q}}\,\mathrm{d}x \right\}^{\frac{1}{q}} = \left\{ \frac{q}{q-p} \right\}^{\frac{1}{p}}, \qquad \text{for } q > p,$$

so $g_q \in L^p(\mathbb{R})$.

If instead $q \in [1, p[$, we define

$$h_q(x) = \begin{cases} x^{-\frac{1}{q}}, & \text{for } x \in [1, +\infty[, \\ \\ 0, & \text{otherwise,} \end{cases} \qquad\qquad \text{for } q \in [1, p[.$$

Then $h_q \geq 0$, and $h_q \leq h_r$ for $q \leq r$, and h_q does not belong to $L^q(\mathbb{R})$, while

$$\|h_q\|_p = \left\{ \int_1^{+\infty} x^{-\frac{p}{q}}\,\mathrm{d}x \right\}^{\frac{1}{p}} = \left\{ \frac{q}{p-q} \right\}^{\frac{1}{p}} < +\infty,$$

because $q < p$ by assumption, so $h_q \in L^p(\mathbb{R})$.

Finally, we define

$$f(x) := \sum_{n=1}^{+\infty} \frac{1}{2^n} \cdot g_{p+\frac{1}{n}}(x) + \sum_{n \geq \frac{1}{p-1}} \frac{1}{2^n} \cdot h_{p-\frac{1}{n}}(x).$$

Then we get the estimates

$$\begin{aligned} \|f\|_p \;&\leq\; \sum_{n=1}^{+\infty} \frac{1}{2^n} \left\| g_{p+\frac{1}{n}} \right\|_p + \sum_{n \geq \frac{1}{p-1}} \frac{1}{2^n} \left\| h_{p-\frac{1}{n}} \right\|_p \\ &=\; \sum_{n=1}^{+\infty} 2^{-n} \{np+1\}^{\frac{1}{p}} + \sum_{n \geq \frac{1}{p-1}} 2^{-n} \{np-1\}^{\frac{1}{p}} \leq 2 \sum_{n=1}^{+\infty} 2^{-n} \{np+1\} < +\infty, \end{aligned}$$

and it follows that $f \in L^p(\mathbb{R})$.

If $q > p$, then we can find $n \in \mathbb{N}$, such that $p + \frac{1}{n} < q$. Hence,

$$f(x) \geq 2^{-n}\, g_{p+\frac{1}{n}}(x) \geq 2^{-n}\, g_q(x).$$

From $g_q \notin L^q(\mathbb{R})$ follows that $f \notin L^q(\mathbb{R})$ for all $q > p$.

If instead $q \in [1, p[$, then there is an $n \in \mathbb{N}$, such that $q < p - \frac{1}{n}$, so $n > \frac{1}{p-q} \geq \frac{1}{p-1}$. Hence,

$$f(x) \geq 2^{-n}\, h_{p-\frac{1}{n}}(x) \geq 2^{-n}\, h_q(x).$$

It follows from $h_q \notin L^q(\mathbb{R})$ that also $f \notin L^q(\mathbb{R})$.

Summarizing, we have for every $p \in [1, +\infty]$ constructed a function f, such that

$$f \in L^p(\mathbb{R}) \setminus \bigcup_{q \in [1,+\infty] \setminus \{p\}} L^q(\mathbb{R}).$$

It is left to the reader to check that the function

$$\tilde{f}_p(x) := \begin{cases} x^{-\frac{1}{p}} \cdot (1 + |\ln x|)^{-\frac{2}{p}}, & \text{for } x > 0, \\ 0, & \text{for } x \leq 0, \end{cases}$$

belongs to $L^p(\mathbb{R})$ for $p \in [1, +\infty[$, and it does not belong to any other $L^q(\mathbb{R})$, where $q \neq p$.

2.10 A function $f : [0,1] \to \mathbb{R}$, which everywhere has a bounded derivative f', and yet f' is not Riemann integrable.

The following example was constructed by *V. Volterra* in 1881. Notice that the Lebesgue integral was first introduced by *Lebesgue* in 1902, from which we conclude that the integral involved must be the Riemann integral. However, in order to shorten the presentation, we shall (without proof) apply the following theorem from Lebesque theory.

Theorem 2.2 *Let f be a bounded function, defined on a bounded interval $[a, b]$. Then f is Riemann integrable, if and only if f is continuous almost everywhere, i.e. the set of discontinuity points for f is a nullset (of Lebesgue measure 0),*

Remark 2.2 Theorem 2.2 explains why the function $\chi_{\mathbb{Q}}$ is not Riemann integrable. The set of discontinuity points in every bounded interval of the form $[-n, n]$ is $[-n, n]$ itself, which is not a nullset. $\Diamond$

Let us first consider the function

$$g(x) = \begin{cases} x^2 \sin \dfrac{1}{x}, & \text{for } x > 0, \\ 0, & \text{for } x \leq 0. \end{cases}$$

From

$$\lim_{x \to 0+} \frac{g(x) - g(0)}{x - 0} = \lim_{x \to 0+} x \cdot \sin \frac{1}{x} = 0 = \lim_{x \to 0-} \frac{g(x) - g(0)}{x - 0},$$

follows that g is differentiable everywhere with the derivative

$$g'(x) = \begin{cases} -\cos \dfrac{1}{x} + 2x \sin \dfrac{1}{x}, & \text{for } x > 0, \\ 0, & \text{for } x \leq 0. \end{cases}$$

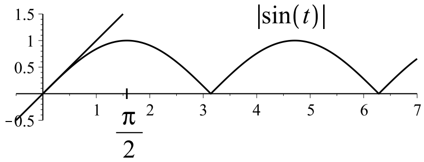

Figure 2.4: *The graphs of* t *and* $|\sin t|$ *for* $t \geq 0$.

When we consider the graphs of t and $|\sin t|$, cf. Figure 2.4, we see that $|\sin t| < t$ for all $t > 0$. Therefore, $\left|\sin \frac{1}{x}\right| < \frac{1}{x}$ for $x > 0$, and it follows that

$$|g'(x)| < \left|\cos \frac{1}{x}\right| + 2 \cdot x \cdot \frac{1}{x} \leq 3, \qquad \text{for } x \geq 0,$$

so g' is bounded.

Then notice that

$$g'\left(\frac{1}{2p\pi}\right) = -1 \quad \text{and} \quad g'\left(\frac{1}{(2p+1)\pi}\right) = +1, \qquad \text{for } p \in \mathbb{N}.$$

Thus, in any (small) neighbourhood of 0 there are points x where $g'(x) = 1$, and points x where $g'(x) = -1$; cf. Figure 2.5. In particular, g' is discontinuous at $x = 0$, which should not be surprising.

Furthermore, the continuous graph of g itself oscillates infinitely often between the two parabolæ of equations $y = x^2$ and $y = -x^2$, cf. Figure 2.6.

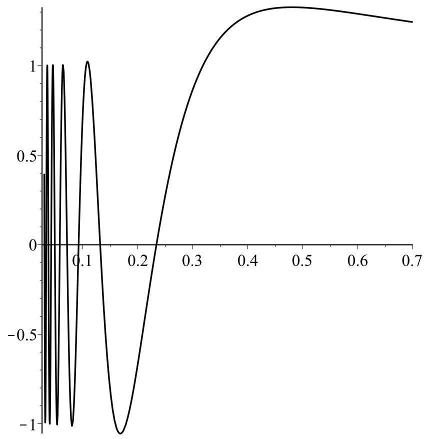

Figure 2.5: *The graph of* $g'(x) = -\cos\dfrac{1}{x} + 2x\,\sin\dfrac{1}{x}$ *for* $x \geq 0$.

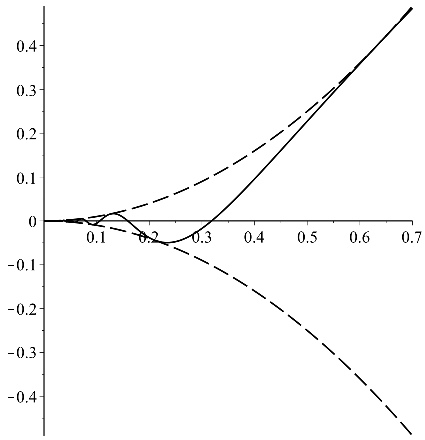

Figure 2.6: *The graph of* $g(x) = x^2\,\sin\dfrac{1}{x}$ *for* $x \geq 0$.

For any given open and bounded interval $]a,b[$, where $a<b$, we define a function $g_{a,b} :]a,b[\to \mathbb{R}$ by

$$g_{a,b}(x) := \begin{cases} (x-a)^2 \sin \dfrac{1}{x-a}, & \text{for } x \in]a, ak], \\[2ex] k^2 \sin \dfrac{1}{k}, & \text{for } x \in [a+k, b-k], \\[2ex] (b-x)^2 \sin \dfrac{1}{b-x}, & \text{for } x \in [b-k, b[, \end{cases}$$

where $k \in \left]a, \frac{a+b}{2}\right]$ is defined as the largest number in this interval, for which the function $g(x-a)$ has an extremum. Thus,

$$k := \max\left\{ x \in \left]a, \frac{a+b}{2}\right] \mid g'(x-a) = 0 \right\}.$$

We notice that it follows from the discussion of g itself above that for given a, $b \in \mathbb{R}$, where $a<b$, there exists such a k. It is of course obvious that k depends on $b-a$, but we shall not need to know this relationship.

The functions $g_{a,b}$ are derived from g, so it follows immediately that $g_{a,b}$ is differentiable, and that the derivative $g'_{a,b}$ is continuous in $]a,b[$, so in every neighbourhood of a (or b) we can find an $x \in]a,b[$, such that $g'_{a,b}(x) = 1$, as well as an $x \in]a,b[$, such that $g'_{a,b}(x) = -1$. Also, $\left|g'_{a,b}(x)\right| < 3$ for all $x \in]a.b[$.

We extend $g_{a,b}(x)$ to all of $\mathbb{R}$ by putting $g_{a,b}(x) = 0$, whenever $x \notin]a,b[$. This extension, also denoted by $g_{a,b}$, is differentiable in all of $\mathbb{R}$, and $g'_{a,b}$ is only discontinuous at a and b.

The idea is then to sum a sequence of functions of the type $g_{a,b}$, such that the set of discontinuity points of the sumfunction f has a positive measure, because the claim then follows from Theorem 2.2. We shall do this by using a Cantor-like construction, cf. e.g. Section 3.2.

We choose an interval $]a_1, b_1[$ of length $\frac{1}{4}$ and symmetric with respect to the midpoint of the interval $[0,1]$. Clearly, $]a_1, b_1[\; = \; \left]\frac{3}{8}, \frac{5}{8}\right[$, but we have chosen this vague description, because the explicit description of the following intervals very soon becomes really messy. Therefore, it is better only to use this loose geometric description.

When we remove $]a_1, b_1[$ from $[a_0, b_0] := [0,1]$, we obtain two closed intervals, $[a_0, a_1]$ and $[b_q, b_0]$, each of length $\frac{1}{2}\left(1-\frac{1}{4}\right) = \frac{3}{8}$. Then choose two open intervals $]a_2, b_2[$ and $]a_3, b_3[$, each of length 4^{-2}, such that $]a_2, b_2[$ lies symmetric with respect to the midpoint of $[a_0, a_1]$, and such that $]a_3, b_3[$ lies symmetric with respect to the midpoint of $[b_1, b_0]$, cf. Figure 2.7. When we remove these intervals, the remainder set is the union of four closed intervals of the same length,

$$2^{-2}\left(1 - \frac{1}{4} - 2 \cdot \frac{1}{4^2}\right).$$

We proceed in this way. At step, or level, n the remainder set consists of 2^n closed intervals of equal length, namely,

$$2^{-n}\left\{1 - \frac{1}{4} - 2 \cdot \frac{1}{4^2} - \cdots - 2^{n-1} \cdot \frac{1}{4^n}\right\} = 2^{-n-1}\left(1 + 2^{-n-1}\right).$$

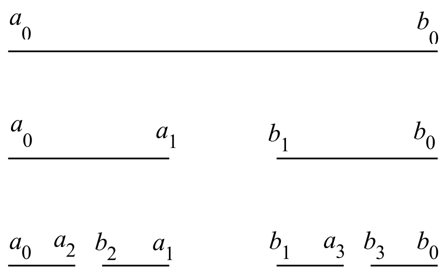

Figure 2.7: *The Cantor-like construction of Section 2.10.*

We choose in each of these intervals an open interval of length 4^{-n}, which lies symmetric with respect to the midpoint of the interval under consideration. These are then removed and the remainder set is the union of 2^{n+1} closed intervals, each of length $2^{-n-2}\left(1+2^{-n-1}\right)$. After n steps the measure of the remainder set is $\frac{1}{2}\left(1+2^{-n-1}\right)$.

When this process is continued indefinitely, we finally obtain a remainder set A, given by

$$A = [0,1] \setminus \bigcup_{n=1}^{+\infty}]a_n, b_n[\,,$$

where the measure of A is

$$|A| = \lim_{n\to+\infty} \frac{1}{2}\left(1+2^{-n-1}\right) = \frac{1}{2} > 0.$$

Then define $f : [0,1] \to \mathbb{R}$, by

$$f(x) = \sum_{n=1}^{+\infty} g_{a_n,b_n}(x) = \begin{cases} g_{a_n,b_n}(x), & \text{for } x \in\ ]a_n, b_n[\,, \quad n \in \mathbb{N}, \\ 0, & \text{for } x \in A. \end{cases}$$

The family of intervals $\{]a_n, b_n[\ \mid n \in \mathbb{N}\}$ are mutually disjoint. Thus, f is a C^1-function in each of the intervals $]a_n, b_n[$, $n \in \mathbb{N}$.

We shall prove that f is differentiable at every point $x \in A$, and that $f'(x) = 0$ for $x \in A$. The claim follows, if for every $x \in A$ and every $h \neq 0$, where $x + h \in [0,1]$, we can prove that

$$\left|\frac{f(x+h)-f(x)}{h}\right| = \left|\frac{f(x+h)}{h}\right| \le |h|. \tag{2.10}$$

Let us for convenience assume that $h > 0$. The proof is analogous for $h < 0$.

a) If $x + h \in A$, then $f(x+h) = 0$, and (2.10) is trivial.

b) If $x+h \notin A$, then there is an n, such that $x+h \in \,]a_n, b_n[$. Now, $x \in A$ and $h > 0$, so $x \le a_n \le a + hzb_n$, hence

$$|f(x+h)| = |g_{a_n,b_n}(x+h)| \le (x+h-a_n)^2 \le h^2,$$

and the estimate (2.10) follows.

We have proved that f is differentiable in all of $[0,1]$. The intervals $]a_n, b_n[$, $n \in \mathbb{N}$, are mutually disjoint, so we conclude from the discussion of $g_{a,b}$ that f' is bounded, $|f'(x)| < 3$ for every $x \in [0,1]$.

Let $x \in A$, and let $\delta > 0$. We can find an endpoint, a_n or b_n, of distance $< \delta$ from x. In fact, the distance from x to the nearest endpoint a_j, b_j, from level n is smaller than 2^{-n}. Hence, choose $n \in \mathbb{N}$, such that $2^{-n} < \frac{\delta}{2}$, and let c denote the closest of the points a_j, b_j from level n. The properties of $g_{a,b}$ imply that we can always in a distance $< \frac{\delta}{2}$ from c find points t , such that $f'(1) = -1$ and $f'(t) = +1$. This shows that f' is discontinuous at every point of A, because $f'(x) = 0$, and x was a point from A. We have proved that A is the set of discontinuity points of the derivative f'.

The measure of A is $|A| = \frac{1}{2} > 0$. It follows from Theorem 2.2 (which Volterra did not know, because it was first proved after his result) that f' is not Riemann integrable.

It is possible to show that the constructed function f above is indeed Lebesgue integrable with the expected value,

$$\int_0^1 f'(x)\,\mathrm{d}x = f(1) - f(0) = 0.$$

2.11 Marcinkiewicz's function

The following strange example concerning the so-called *universal primitives* was given by *Marcinkiewicz* [34] in 1935. The core is the following: To a given sequence (h_n) of real numbers $h_n \neq 0$, where $h_n \to 0$ for $n \to +\infty$, there exists a *continuous* function Φ on $[0,1]$, such that to *every* measurable function φ, defined on $[0,1]$ we can construct an infinite suhsequence $(h_{n_k}) \subseteq (h_n)$, such that φ can be found as the limit of the difference quotients

$$\lim_{k\to+\infty} \frac{1}{h_{n_k}} \left\{\Phi\left(x+h_{n_k}\right)-\Phi(x)\right\} = \varphi(x) \qquad \text{for almost every } x \in [0,1].$$

In other words, there exist some continuous functions Φ which really have a weird behaviour, and which is not easy to comprehend.

Since

$$\left(\frac{1}{h_{n_k,\varphi}} \left\{\Phi(x+h_{n_k,\varphi})-\Phi(x)\right\}_{k\in\mathbb{N}}\right)$$

is a sequence of difference quotients, which converges towards the chosen function $\varphi(x)$ almost everywhere, and since φ could be any measurable function, Marcinkiewicz called such a Φ a *universal primitive.*

Remark 2.3 Marcinkiewicz furthermore proved that the set of all universal primitives Φ with respect to a given sequence (h_n), $h_n \neq 0$, and $h_n \to 0$ for $n \to +\infty$, is of *second category.* I have chosen here not to prove this result. The interested reader is referred to *Marcinkiewicz* [34]. $\Diamond$

The construction of the function Φ relies on the existence of a *singular function*, i.e. a continuous bijective and strictly monotonous function

(2.11) $\quad H : [\alpha,\beta] \to [a,b],$

such that $H'(x) = 0$ for almost every $x \in [\alpha,\beta]$.

We have already proved the existence of such a singular function in Section 3.3 in the special case, when $[\alpha,\beta] = [a,b] = [0,1]$. It is a simple exercise using an affine map (left to the reader) to generalize this construction to (2.11).

We shall use the singular functions in the proof of the following lemma.

Lemma 2.1 *Given $\varepsilon > 0$. Let $F_1(x)$ and $F_2(x)$ be two continuous functions on $[0,1]$. Assume furthermore that $F_2(x)$ is differentiable almost everywhere in $[0,1]$.*
There exists a continuous function $G(x)$, which is differentiable almost everywhere in $[0,1]$, such that

a) $\qquad G'(x) = F_2'(x)$ *for almost every* $x \in [0,1]$.

b) $\qquad |F_1(x) - G(x)| < \varepsilon$ *for all* $x \in [0,1]$.

PROOF. The difference $F_1 - F_2$ is continuous on the closed and bounded (i.e. compact) interval $[0,1]$, hence also *uniformly continuous.* Thus, to the given $\varepsilon > 0$ above, the interval $[0,1]$ can be divided into a finite number of subintervals

$$I_n = [a_n, a_{n+1}], \quad n = 0, 1, \dots, n_0 - 1, \qquad \text{where } 0 = a_0 < a_1 < \cdots < a_{n_0} = 1,$$

such that

$$(2.12) \qquad |\{F_1(x) - F_2(x)\} - \{F_1(y) - F_2(y)\}| < \varepsilon \qquad \text{for all } x,\, y \in I_n.$$

Using the methods of Section 3.3 we construct a continuous function on $[0,1]$, such that

$$H'(x) = 0 \qquad \text{for almost every } x \in [0,1],$$

and such that H in each of the intervals $I_n = [a_n, a_{n+1}]$ is monotonous of type (2.11), and where furthermore,

$$H(a_n) = F_1(a_n) - F_2(a_n)\,.$$

It follows from (2.12) that

$$|F_1(x) - F_2(x) - H(x)| < \varepsilon \qquad \text{for all } x \in [0,1],$$

and we conclude that the function

$$G(x) := F_2(x) + H(x) \qquad \text{for } x \in [0,1],$$

fulfils the two conditions of the lemma. □

We shall then prove

Theorem 2.3 Marcinkiewicz [34], 1935. *Let (h_n) be a sequence of real numbers $h_n \neq 0$, converging towards 0 for $n \to +\infty$. There exists a* continuous *function Φ, defined on $[0,1]$, such that*

(2.13) *to* every *measurable function φ on $[0,1]$ there exists an infinite subsequence $(h_{n_k,\varphi})$ of (h_n), such that*

$$\lim_{k \to +\infty} \frac{1}{h_{n_k,\varphi}} \{\Phi(x + h_{n_k,\varphi}) - \Phi(x)\} = \varphi(x) \quad \textit{for almost every } x \in [0,1].$$

Proof. Without loss of generality we assume that all $h_n > 0$. Then notice that the set of all polynomials of fixed degree $n \in \mathbb{N}_0$ and of rational coefficients is countable. A countable union of countable sets is again countable, from which we conclude that the set of all polynomials of rational coefficients is countable. We can therefore structure the set of all such rational polynomials in a *sequence* $\{P_n(x) \mid n \in \mathbb{N}\}$.

We define by induction a sequence $\{\Phi_k(x)\}$ of continuous functions, which are differentiable almost everywhere, as well as a subsequence (t_k) of the given sequence (h_n), and a sequence of sets (E_k) of measure $|E_k| \leq 2^{-k}$, such that

$$(2.14) \qquad |\Phi_k(x) - \Phi_{k-1}(x)| \leq \frac{1}{k-1}\, t_k, \qquad \text{for all } x \in [0,1] \text{ and } k \in \mathbb{N} \setminus \{1\},$$

$$(2.15) \qquad \left| \frac{1}{t_k} \{\Phi_k(x + t_k) - \Phi_k(x)\} - P_k(x) \right| < \frac{1}{k}, \qquad \text{for all } x \in [0,1] \setminus E_k \text{ for } k \in \mathbb{N},$$

$$(2.16) \qquad t_k < \frac{1}{2}\, t_{k-1}, \qquad \text{for } k \in \mathbb{N} \setminus \{1\}.$$

The induction goes as follows.

Let $\Phi_1(x)$ be any primitive of the first polynomial $P_1(x)$ from the sequence, and choose t_1, such that (2.15) holds. The other two conditions, (2.14) and (2.16), are satisfied of formal reasons, because $k = 1 \notin \mathbb{N} \setminus \{1\}$.

Assume that we have found the first $k-1$ terms, $k \in \mathbb{N} \setminus \{1\}$, of $\{\Phi_j(x)\}$, $\{E_j\}$ and $\{t_j\}$. Choose $\varepsilon := \dfrac{1}{k-1}\, t_{k-1}$. By Lemma 2.1 we then construct a continuous function Φ_k, such that (2.14) holds, and such that $\Phi_k' = P_k(x)$ almost everywhere.

Once Φ_k has been fixed, it is easy to find $t_k \in \{h_n \mid n \in \mathbb{N}\}$, such that (2.15) and (2.16) are satisfied, because we can always remove a set E_k of measure $\leq 2^{-k}$, such that (2.15) is fulfilled for some $t_k > 0$.

It follows from (2.14) and (2.16) that $\{\Phi_k\}$ is a uniform Cauchy sequence, thus it is convergent with a continuous limit function

$$\Phi(x) = \lim_{n \to +\infty} \Phi_n(x).$$

Then notice that for all $x \in [0,1]$,

$$|\Phi(x) - \Phi_k(x)| \leq \sum_{n=k}^{+\infty} |\Phi_{n+1}(x) - \Phi_n(x)| \leq \sum_{n=k}^{+\infty} \frac{1}{n}\, t_n \leq \frac{1}{k}\, t_k \sum_{n=0}^{+\infty} 2^{-n} = \frac{1}{k} \cdot 2t_k.$$

If $x \notin E_k$, then it follows from the triangle inequality and (2.15) that

$$\begin{aligned} &\left| \frac{1}{t_k} \left\{ \Phi\left(x + t_k\right) - \Phi(x) \right\} - P_k(x) \right| \\ &\qquad \leq \left| \frac{1}{t_k} \left\{ \Phi_k(x + t_k) - \Phi_k(x) \right\} - P_k(x) \right| + \left| \frac{1}{t_k} \left\{ \Phi_k(x + t_k) - \Phi_k(x + t_k) \right\} \right| + \left| \frac{1}{t_k} \left\{ \Phi(x) - \Phi_k(x) \right\} \right| \\ &\qquad \leq \frac{1}{k} + \frac{1}{t_k} \cdot \frac{1}{k} \cdot 2t_k + \frac{1}{t_k} \cdot \frac{1}{k} \cdot 2t_k = \frac{5}{k}. \end{aligned}$$

Hence,

$$\left| \frac{1}{t_k} \left\{ \Phi(x + t_k) - \Phi(x) \right\} - P_k(x) \right| \leq \frac{5}{k} \qquad \text{for all } k \in \mathbb{N}. \tag{2.17}$$

Let f be any measurable function on $[0,1]$. Using *Weierstraß's approximation theorem* we find a subsequence $\{P_{k_m}\}$ of the sequence $\{P_k\}$, such that

$$\lim_{m \to +\infty} P_{k_m}(x) = f(x) \qquad \text{for almost every } x \in [0,1].$$

Considering if necessary a subsequence we may furthermore assume that outside a set H_m of measure $\leq 2^{-m}$,

$$|P_{k_m}(x) - f(x)| < \frac{1}{m}, \qquad \text{for all } x \notin H_m.$$

Clearly, $k_m \geq m$, so when $x \notin E_{k_m} \cup H_m$, it follows from (2.17) that

$$\begin{aligned} &\left| \frac{1}{t_{k_m}} \left\{ \Phi(x + t_{k_m}) - \Phi(x) \right\} - f(x) \right| \qquad (2.18) \\ &\qquad \leq \left| \frac{1}{t_{k_n}} \left\{ \Phi(x + t_{k_m}) \right\} - P_{k_m}(x) \right| + |P_{k_m}(x) - f(x)| < \frac{5}{k_m} + \frac{1}{m} \leq \frac{6}{m}. \end{aligned}$$

Finally, $\sum_m |E_{k_m} \cup H_m| < +\infty$, so we can find m_0, such that

$$\sum_{m \geq m_0} |E_{k_m} \cup H_m| < \varepsilon,$$

and we conclude from (2.18) that for almost every $x \in [0,1]$,

$$\lim_{m \to +\infty} \frac{1}{t_{k_n}} \{\Phi(x + t_{k_m}) - \Phi(x)\} = f(x).$$

Since the measurable function $f(x)$ was anarbitrarily chosen on $[0,1]$, the theorem is proved. $\square$

3 Measurable sets

3.1 Introduction

In Section 3.2 and Section 3.3 we show some classical examples, which should be included in every collection of "strange" examples. Cf. also *Gelbaum and Olmsted* [22] and *Mejlbo* [35] (not to be confused with the present author). These examples are also connected with the theory of fractals, though they are much older.

The example of Section 3.4 was the first one which many years ago aroused my attention to these examples. I have not been able to find the right reference, but it may probably be found in *Riesz-Sz. Nagy* [47]. We construct a continuous and strictly increasing function, which is differentiable almost everywhere, and its derivative is 0, when it exists,

The *Kakeya problem* is the subject of Section 3.5. The task is to find the Lebesgue measure of the smallest subset Ω of $\mathbb{R}^2$, such that an infinitely thin needle of length 1 can be moved inside Ω and eventually come back to its original position, only with the needle pointing in the opposite direction. The *Nikodym set* of Section 3.6 is closely connected with the Kakeya problem.

We show in Section 3.7 a counterexample from *Complex Functions Theory.* The example is put here, because the methods used are more in line with the already mentioned examples than with analytic functions.

Finally, we construct in Section 3.8 a metric space and two different probability measures, which nevertheless are identical on every ball in this metric space.

3.2 The Cantor set C.

(G. Cantor, 1845–1918). This classical set, which should be known by all mathematicians, is an example of a perfect nowhere dense set. It is also an example of a non-countable nullset in $\mathbb{R}$.

We first introduce

Definition 3.1 *A closed set A is* perfect*, if every point $x \in A$ is an accumulation point of A, i.e. for every neighbourhood U,*

$$A \cap (U \setminus \{x\}) \neq \emptyset.$$

Remark 3.1 It can be proved that if A is a perfect set in a complete separable metric space, then A is not countable. For completeness we shall nevertheless give a proof for that the Cantor set C is not countable. $\Diamond$

Definition 3.2 *A set A in a topological space is called* nowhere dense*, if its closure $\overline{A}$ has empty interior, $(\overline{A})^\circ = \emptyset$.*

Construction of the Cantor set.
Let $C_0 = I = [0,1]$ be the unit interval. We divide it into three subintervals of equal length and remove the middle open interval. We get a smaller subset

$$C_1 = I_0 \cup I_2 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right],$$

consisting of two closed intervals of equal length $\frac{1}{3}$.

Then iterate this process by removing the open middle interval in the next trisection of each interval,

$$C_2 = I_{00} \cup I_{02} \cup I_{22} = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right],$$

where the trisections with an obvious notation are

$$I_0 = I_{00} \cup I_{01} \cup I_{0,2} \qquad \text{and} \qquad I_2 = I_{20} \cup I_{21} \cup I_{22}.$$

Proceed in this way. Cf. Figure 3.1.

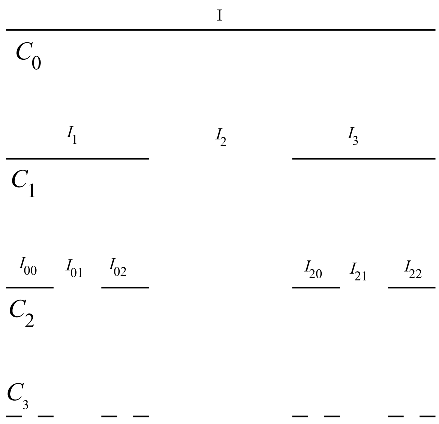

Figure 3.1: *The Cantor-like construction of Section 2.10.*

Then notice that if $x \in I_{a_1 a_2 \cdots a_n}$, where all $a_j \in \{0,1,2\}$, then x has the following triadic description

$$x = \sum_{k=1}^{+\infty} x_k \cdot 3^{-k} = 0_3 x_1 x_2 x_3 \cdots, \qquad \text{where } x_k \in \{0,1,2\},$$

and where also $x_j = a_j$ for $j = 1, 2, \dots, n$.

Clearly, $C_0 \supseteq C_1 \supseteq C_2 \supseteq \cdots \supseteq C_n \supseteq \cdots$, and all C_n are closed. Hence,

$$C := \bigcap_{n=1}^{+\infty} C_n = \lim_{n \to +\infty} C_n \neq \emptyset.$$

If in particular $a_j \in \{0,2\}$, then the endpoints of the subinterval $I_{a_1 a_2 \cdots a_n}$ will always belong to C.

We shall prove that C is a perfect set. Given $x \in C$ and $\varepsilon > 0$. Choose $n \in \mathbb{N}$, such that $3^{-n} < \varepsilon$. Then choose the subinterval I of C_n, which contains x. Since I has two endpoints, at least one of them, y, is $\neq x$. So $y \in C$, and since $|x-y| \leq 3^{-n} < \varepsilon$, it follows that C is perfect, cf. Definition 3.1.

Then we prove that C is nowhere dense. We shall prove that the closed set $C = \bigcup_{n=1}^{+\infty} C_n$ does not contain interior points. Let $]a,b[$ be any open interval, and choose $n \in \mathbb{N}$, such that $3^{-n} < b-a$. Then $]a,b[$ cannot be a subset of C_n, because the lengths of the subintervals of C_n is 2^{-n}. Hence $]a,b[$ is not a subset of C. This holds for every open interval $]a,b[$, so we conclude that C does not contain any open interval, and C is nowhere dense.

We then prove that C is not countable. This is done by a nice direct proof instead of relying on the unproved claim that every perfect set is not countable. This proof will contain some elements, which previously were often used in proofs of this type.

It follows from Figure 3.1 that every element of C has a uniquely determined triadic representation

$$x = 0_3 x_1 x_2 x_3 \cdots, \qquad \text{where } x_n \in \{0,2\},$$

i.e. without the number 1. Notice that e.g. $1 = 0_3 222 \cdots$, and similarly for every other right endpoint.

Therefore, we have a unique correspondence between the Cantor set C and the interval $[0,1]$, given by the map

$$x=\sum_{n=1}^{+\infty} x_n \cdot 3^{-n} \leftrightarrow y=\sum_{n=1}^{+\infty}\left(\frac{1}{2}\cdot x_n\right)\cdot 2^{-n},$$

because $\frac{1}{2}\,x_n \in \{0,1\}$, when $x_n \in \{0,2\}$. Notice that every $y \in [0,1]$ similarly is expressed as a dyadic fraction,

$$y = 0_2 y_1 y_2 y_3 \cdots \qquad \text{or} \qquad y=\sum_{j=1}^{+\infty} y_j \cdot 2^{-j}, \quad \text{where } y_j \in \{0,1\}.$$

This correspondence $x \leftrightarrow y$ is clearly bijective, hence, the two sets C and $[0,1]$ must contain "equally many" elements.

It is well-known that $[0,1]$ has a non-countable number of elements which is e.g. proved by Cantor's diagonalization. An alternative simple proof is the following. We have already seen that every countable subset of $\mathbb{R}$ is a Lebesgue nullset. Since $[0,1]$ is *not* a nullset, (its measure is 1), it follows that $[0,1]$ is not countable.

Finally, it follows from the iterative construction of C that

$$|C_{n+1}| = \left(1-\frac{1}{3}\right)|C_n| = \frac{2}{3}\,|C_n|\,,$$

where $|C_n|$ denotes the measure of C_n. Hence, by recursion

$$|C| = \lim_{n\to+\infty}|C_n| = \lim_{n\to+\infty}\left(\frac{2}{3}\right)^n = 0,$$

and we have proved that C is a nullset, and all the claims have been proved.

Remark 3.2 We may slightly alter the construction above in the following way. In the n-th step we let the middle interval A_n which should be removed from the subinterval I_n, before we proceed to step $n+1$, have the measure

$$|A_n| = \varepsilon_n\,|I_n|\,,$$

where the fraction $\varepsilon_n \in]0,1[$. Repeating the proof above with obvious modifications we again obtain that the limit set C becomes a nowhere dense and perfect set. All such sets C are called *Cantor sets.* The measure is then

$$|C| = \prod_{n=1}^{+\infty}(1-\varepsilon_n)\,.$$

It is well-known, and can also easily be proved by taking the logarithm of the sequence

$$\left(\prod_{j=1}^{n}(1-\varepsilon_j)\right),$$

that the infinite product is positive, if and only if

$$\sum_{n=1}^{+\infty} \varepsilon_n < +\infty.$$

It is therefore possible to choose the sequence (ε_n), where $0 < \varepsilon_n < 1$ for all $n \in \mathbb{N}$, such that the corresponding Cantor set C becomes a perfect and nowhere dense set of positive measure. It is left to the reader to prove that given any $\delta \in]0,1[$ it is possible to construct C, such that $|C| = 1 - \delta$. $\Diamond$

3.3 The Cantor function

Using the *Cantor set* of Section 3.2 as a step stone we shall here construct the *Cantor function* f. This function is a continuous and (weakly) monotonous function which is differentiable almost everywhere with the derivative 0, and yet $f(0) < f(1)$.

We shall later in Section 3.4 construct a Cantor function, which is even strictly monotonous. We shall, however, as a start in this section prove the weaker claim, which is easier to construct.

In the first step we put

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

Let I_1 denote the middle open subinterval by the trisection, i.2. $I_1 = \left]\frac{1}{3}, \frac{2}{3}\right[$. We define

$$f(x) := \frac{1}{2} = \frac{1}{2}\{f(0) + f(1)\} \qquad \text{for all } x \in I_1.$$

Then repeat this construction.

If I_{01} is the middle open subinterval of the interval I_0 from C_1, and I_{21} is the middle open interval of I_2, cf. Section 3.2, then we put

$$f(x) = \begin{cases} \dfrac{1}{4} = \dfrac{1}{2}\left\{f(0) + f\left(I_1\right)\right\}, & \text{for } x \in I_{01}, \\[2ex] \dfrac{1}{2} + \dfrac{1}{4} = \dfrac{1}{2}\left\{f\left(I_1\right) + f(1)\right\}, & \text{for } x \in I_{21}. \end{cases}$$

All middle subintervals must be of the form $I_{x_1 \cdots x_n 1}$, where $x_1, \dots, x_n \in \{0, 2\}$. If $x \in I_{x_1 \cdots x_n 1}$, then we define

$$f(x) = \left(\frac{1}{2}\, x_1\right) \cdot 2^{-1} + \cdots + \left(\frac{1}{2}\, x_n\right) \cdot 2^{-n} + 2^{-n-1},$$

and it is easy to prove that $f(x)$ is the mean between the values on the two closest intervals from an earlier step, where f already has been defined. We consider here for convenience $\{0\}$ and $\{1\}$ as degenerated intervals of values $f(0) = 0$ and $f(1) = 1$, respectively.

In this way we define f in the set $([0,1] \setminus C) \cup \{0,1\}$. Obviously, f is (weakly) increasing, and the range of f is dense in the interval $[0,1]$.

If $x \in C \setminus \{0,1\}$, then we define

$$f(x) = \sup\{f(y) \mid y \notin C,\, y < x\} = \inf\{f(y) \mid y \notin C,\, y > x\},$$

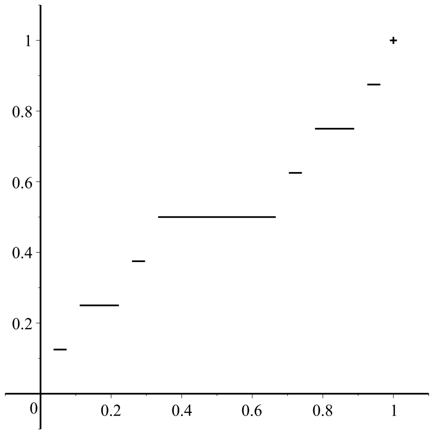

Figure 3.2: *The first steps of the construction of the Cantor function.*

where the latter inequality follows, because the range is dense in $[0,1]$, and because f is increasing. In particular, f must be continuous everywhere in $[0,1]$.

When $x \in [0,1] \setminus C$, then $f'(x) = 0$, because x belongs to some open interval, where f is constant.

It is left to the reader to prove that if $x \in C$, then the corresponding difference quotients tend towards $+\infty$ for the increment tending towards $0+$.

3.4 A strictly increasing function f, where $f'(0) = 0$ almost everywhere.

We shall in this section follow the construction given by *Riesz, Sz.-Nagy* [47]. The authors do not mention who actually created this example. A vague hint given by *Marcinkiewicz* is that it was probably constructed by *Lusin*.

We shall construct a continuous function f, defined on $[0,1]$, with the following properties,

a) The function f is *strictly increasing* in the interval $[0,1]$.

b) The function f is in the classical sense as a limit of difference quotients *differentiable almost everywhere* in $[0,1]$.

c) Whenever f is differentiable at a point $x \in [0,1]$, then $f'(x) = 0$.

A continuous function f, which satisfies the three conditions above, is called a *singular function.* We notice that for all x_1, $x_2 \in [0,1]$, for which $x_1 < x_2$,

$$f(x_2) - f(x_1) \neq \int_{x_1}^{x_2} f'(x)\,\mathrm{d}x.$$

The conditions above are intuitively contradicting each others, and since we have claimed that it nevertheless is possible to construct such a function, we may expect a fairly long theoretical detour. We shall start with the following technical lemma.

Lemma 3.1 (F. Riesz, 1932). *Given a real continuous function $g \in C^0([a,b])$ on the closed interval $[a,b]$. Define an open set $E = E_g$ by*

$$E = \{x \in]a,b[\mid g(\xi) > g(x) \text{ for some } \xi \in]x,b[\}.$$

If $E \neq \emptyset$, then we can write E as a disjoint union

$$E = \bigcup_k]a_k, b_k[, \qquad \text{where } g(a_k) \leq g(b_k).$$

Proof. We assume that $E \neq \emptyset$, because otherwise there is nothing to prove. Let $x_0 \in E$. Then we can find $\xi \in]x_0, b]$, such that $g(\xi) > g(x_0)$.

We put $\varepsilon = g(\xi) - g(x_0)$. It follows from the continuity of g that there is a $\delta > 0$, such that

$$|g(x) - g(x_0)| < g(\xi) - g(x_0), \qquad \text{for all } x \in]a,b[, \text{ for which } |x - x_0| < \delta.$$

Hence, if $x \in]a,b[$ satisfies the estimate $|x - x_0| < \delta$, then

$$g(x) \leq g(x_0) + |g(x) - g(x_0)| < g(x_0) + g(\xi) - g(x_0) = g(\xi),$$

and we have proved that $x \in E$ by the definition of E. We therefore conclude that for every $x_0 \in E$ there is a $\delta > 0$, such that

$$]a,b[\cap]x_0 - \delta, x_0 + \delta[\subseteq E,$$

which proves that E is an open set.

There is only one way, in which we can write

$$E = \bigcup_k \,]a_k, b_k[\,,$$

where k runs through a finite or infinite index set, and where the intervals $]a_k, b_k[$ are all mutually disjoint. So if $b_k \neq b$ (i.e. the right endpoint of the interval), then

$$g(\xi) \leq g(b_k) \qquad \text{for all } \xi \in \,]b_k, b]\,, \tag{3.1}$$

cf. Figure 3.3.

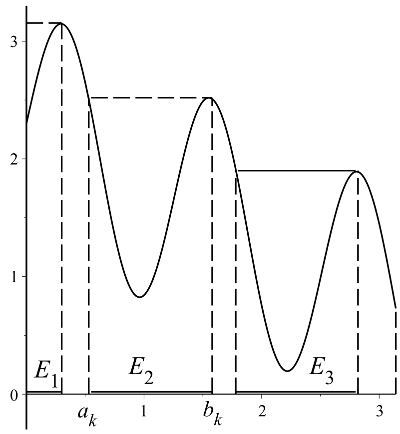

Figure 3.3: *Illustration of the set E in Section 3.4.*

Assume that $x \in \,]a_k, b_k[$. We shall prove that $g(x) \leq g(b_k)$. We put

$$x_1 := \sup_\xi \{\xi \in \,]x, b_k] \mid g(\xi) \geq g(x_k)\}\,.$$

Then $x_1 \in \,]x, b_k]$ is the closest point with respect to b_k in this interval, for which also $g(x_1) \geq g(x)$. We shall prove that $x_1 = b_k$.

We first prove that we can find $\xi \in \,]x, b_k]$, such that $g(\xi) \geq g(x)$.

Since $x \in E$, there is a $\xi \in]x, b]$, such that $g(\xi) \geq g(x)$.

If $\xi \in \,]b_k, b]$, then (3.1) implies that $g(b_k) \geq g(x)$, and we can even choose $\xi \in \,]x, b_k]$. It follows that

$$\{\xi \in \,]x, b_k] \mid g(\xi) \geq g(x)\}$$

is a nonempty set, which is bounded from above, so the supremum x_1 exists.

Finally, it follows from the continuity of g that $g(x) \leq g(x_1)$.

Then we assume that $x_1 < b_k$, and hope for a contradiction, because this will prove the lemma. By assumption, $x_1 \in \,]a_k, b_k[\, \subseteq E$, so there exists a $\xi \in \,]x_1, b]$, such that (cf. the definition of x_1),

$$g(x) \leq g(x_1) \leq g(\xi).$$

If $\xi \in \left]x_1, b_k\right[$, this contradicts the definition of x_1.

If instead $\xi \in \left]b_k, b\right]$, then $g\left(b_k\right) \leq g\left(x_1\right) z g(\xi)$, which proves that $b_k \in E$, contradicting the assumption that $b_k \notin E$.

We conclude that $\xi = b_k$, which implies that $g(x) \leq g\left(b_k\right)$, hence also $x_1 = b_k$, which contradicts the assumption above that $x_1 < b_k$.

Summing up we conclude that $x_1 = b_k$, and we have proved that

$$g(x) \leq g\left(b_k\right), \qquad \text{for all } x \in \left]a_k, b_k\right[.$$

Finally, it follows from the continuity of g that

$$g\left(a_k\right) \leq g\left(b_k\right), \tag{3.2}$$

and the lemma is proved. □

Theorem 3.1 (Lebesgue, 1902) *Every continuous and monotonous function f is differentiable almost everywhere in the classical sense.*

Proof. We may without loss of generality assume that f is increasing. Put

$$\lambda_r(x) := \liminf_{\Delta x \to 0+} \frac{f(x+\Delta x) - f(x)}{\Delta x}, \qquad \Lambda_r(x) = \limsup_{\Delta x \to 0+} \frac{f(x+\Delta x) - f(x)}{\Delta x},$$

(limits from the right), and

$$\lambda_l(x) := \liminf_{\Delta x \to 0-} \frac{f(x+\Delta x) - f(x)}{\Delta x}, \qquad \Lambda_l(x) = \limsup_{\Delta x \to 0-} \frac{f(x+\Delta x) - f(x)}{\Delta x},$$

(limits from the left). Then for every $x \in]a, b[$,

$$\lambda_r(x) \leq \Lambda_r(x) \qquad \text{and} \qquad \lambda_l(x) \leq \Lambda_l(x).$$

If we can prove that

$$a) \quad \Lambda_r(x) < +\infty, \qquad \text{and} \qquad b) \quad \Lambda_r(x) \leq \lambda_l(x),$$

for almost every $x \in]a, b[$, we get by applying b) on the increasing function $g(x) := -f(-x)$ for $x \in [-b, -a]$ that

$$\Lambda_l(x) \leq \lambda_r(x) \qquad \text{for almost every } x \in]a, b[,$$

from which we infer that

$$\Lambda_r(x) \leq \lambda_l(x) \leq \Lambda_l(x) \leq \lambda_r(x) \leq \Lambda_r(x), \qquad \text{for almost every } x \in]a, b[.$$

Therefore, we must have equality almost everywhere, i.e.

$$\limsup_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \liminf_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = f'(x).$$

In other words, the theorem follows, if we can prove the conditions a) and b).

We shall first prove that

$$a) \quad \Lambda_r(x) < +\infty \qquad \text{almost everywhere.}$$

For every constant $C > 0$ define the set E_C by

$$E_C := \left\{ x \;\middle|\; \Lambda_r(x) = \limsup_{\Delta x \to 0+} \frac{f(x+\Delta x) - f(x)}{\Delta x} > \right\}.$$

Hence, if $x \in E_C$, then there is a $\xi > x$, such that

$$\frac{f(\xi) - f(x)}{\xi - x} > C.$$

Defining $g(x) := f(x) - C \cdot x$ we infer that $g(\xi) > g(x)$ for all $x \in E_C$, so $E_C = E$ is the set corresponding to the continuous function g defined in Lemma 3.1. Hence,

$$E_C = \bigcup_k \left]a_k, b_k\right[,$$

where $g(a_k) \le g(b_k)$, thus $f(a_k) - C \cdot a_k \le f(b_k) - C \cdot b_k$, and then by a rearrangement,

$$(3.3) \qquad C(b_k - a_k) \le f(b_k) - f(a_k).$$

The function f is increasing, hence by summing over k,

$$C \cdot \sum_k (b_k - a_k) \le \sum_k \{f(b_k) - f(a_k)\} \le f(b) - f(a) = \text{ constant.}$$

The measure of the set E_C is $|E_C| = \sum_k (b_k - a_k)$. Using (3.3) we get

$$|E_C| \le \frac{1}{C} \cdot \{f(b) - f(a)\}, \qquad \text{for every } C > 0.$$

Now, for every $C > 0$,

$$E_\infty := \left\{ x \;\middle|\; \Lambda_r(x) = \limsup_{\Delta x \to 0+} \frac{f(x+\Delta x) - f(x)}{\Delta x} = +\infty \right\} \subseteq E_C.$$

We therefore conclude that the measure

$$|E_\infty| \le \frac{1}{C} \cdot \{f(b) - f(a)\}, \qquad \text{for every } C > 0.$$

It follows for $C \to +\infty$ that $|E_\infty| = 0$, so E_∞ is a nullset. We conclude that $\Lambda_r(x) < +\infty$ almost everywhere.

Then we prove that

$$b) \quad \Lambda_r(x) \le \lambda_l(x), \qquad \text{almost everywhere.}$$

Choose any two constants c and C, such that $0 < c < C$. Then consider the continuous function

$$g(x) = f(-x) + c \cdot x,$$

and let

$$E_1 = \bigcup_k \left]a_k^1, b_k^1\right[$$

be the system of intervals corresponding to the function $g(x)$ given by Lemma 3.1.

First notice that if $\lambda_l(x) < c$, then $x \in E_1$. This is why we in the definition of $g(x)$ use $f(-x)$ instead of $f(x)$.

Consider in each of the intervals $\left]a_k^1, b_k^1\right[$ the function

$$G(x) := f(x) - C \cdot x.$$

Apply Lemma 3.1 once more, whereby we obtain a system of intervals $]a_{k\ell}, b_{k\ell}[$, such that

$$A_k := \bigcup_\ell \,]a_{k\ell}, b_{k\ell}[\; \subseteq \; \left]a_k^1, b_k^1\right[,$$

and

$$C\left(a_{k\ell} - a_{k\ell}\right) \leq f\left(b_{k\ell}\right) - f\left(a_{k\ell}\right).$$

The function f is increasing, so when we sum over ℓ we get

$$|A_k| \leq \frac{1}{C} \sum_\ell \left\{f\left(b_{k\ell}\right) - f\left(a_{k\ell}\right)\right\} \leq \frac{1}{C}\left\{f\left(b_k^1\right) - f\left(a_k^1\right)\right\}. \tag{3.4}$$

On the other hand, it follows from the definition of E_1 that

$$f\left(b_k^1\right) - f\left(a_k^1\right) \le c\left(b_k^1 - a_k^1\right),$$

hence, by summing over k,

$$\sum_k \left\{f\left(b_k^1\right) - f\left(a_k^1\right)\right\} \le c\sum_k \left\{b_k^1 - a_k^1\right\} = c\,|E_1|\,. \tag{3.5}$$

Write $E_2 = \bigcup_k A_k$ (disjoint union). When we sum (3.4) with respect to k, it follows from (3.5) that

$$|E_2| = \sum_k |A_k| \le \frac{1}{C}\sum_k \left\{f\left(b_k^1\right) - f\left(a_k^1\right)\right\} \le \frac{c}{C}\sum_k \left\{b_k^1 - a_k^1\right\} = \frac{c}{C}\cdot|E_1|\,.$$

Assume that x satisfies the estimates $\lambda_l(x) < c$ and $\Lambda_r(x) > C$. Then $x \in E_2$.

Consider the function

$$g(x) := f(-x) + c\cdot x$$

defined on the set E_2. Then E_2 can be written as a union of disjoint intervals. When we apply Lemma 3.1 on each of these, we get the system

$$E_3 := \bigcup_k \left]a_k^3, b_k^3\right[\subseteq E_2,$$

on which we apply Lemma 3.1 with respect to the continuous function

$$G(x) = f(x) - C\cdot x.$$

Using Lemma 3.1 we obtain an open set $E_4 \subseteq E_3$, such that

$$|E_4| \le \frac{c}{C}\,|E_3| \le \frac{c}{C}\,|E_2| \le \left(\frac{c}{C}\right)^2|E_1|\,.$$

We conclude that if $\lambda_l(x) < c$ and $\Lambda_r(x) > C$ for some x, then $x \in E_4$.

Proceeding in this way we construct a sequence (E_n) of open sets, such that $E_{n+1} \subseteq E_n$ and

$$|E_{2n}| \le \frac{c}{C}\,|E_{2n-1}| \le \frac{c}{C}\,\left|E_{2(n-1)}\right| \le \cdots \le \left(\frac{c}{C}\right)^n|E_1|\,,$$

where the two inequalities $\lambda_l(x) < c$ and $\Lambda_r(x) > C$ imply that $x \in E_{2n}$.

Since

$$\Omega_{cC} := \{x \mid \lambda_l(x) < c,\, \Lambda_r(x) > C\} \subseteq E_{2n}, \qquad \text{for all } n \in \mathbb{N},$$

we conclude that

$$|\Omega_{cC}| \le \lim_{x\to+\infty}\left(\frac{c}{C}\right)^n|E_1| = 0.$$

Thus, if $0 < c < C$, then Ω_{cC} is a nullset.

Let x be a point, for which $\lambda_l(x) < \Lambda_r(x)$. Using that f is increasing we can find *rational numbers* c, $C \in \mathbb{Q}_+$, such that

$$\lambda_l(x) < c < C < \Lambda_l(x),$$

so $x \in \Omega_{cC}$, and

$$\{x \mid \lambda_l(x) < \Lambda_r(x)\} \subseteq \bigcup_{c\in\mathbb{Q}_+} \bigcup_{C\in\mathbb{Q}_+, C>c} \Omega_{cC} = \Omega.$$

This shows Ω as a countable union of nullsets, hence, Ω itself is a nullset, and so is the subset

$$\{x \mid \Lambda_l(x) < \Lambda_r(x)\} \subseteq \Omega.$$

This implies that $\lambda_r(x) \le \lambda_l(x)$ almost everywhere, and the theorem is proved. $\square$

Theorem 3.2 *There exists a* continuous *function* $f : [0,1] \to \mathbb{R}$*, which fulfils the following properties,*

a) *The function* f *is strictly increasing on* $[0,1]$.

b) *The function* f *is differentiable almost everywhere.*

c) *Whenever* $f'(x)$ *exists, then* $f'(x) = 0$.

Proof. Fix any $\alpha \in]0,1[$. One could, alternatively, in the following assume that $\alpha \in]-1,0[$ instead, but it will not provide us with anything new. The choice $\alpha \in]0,1[$ will be sufficient for the time being.

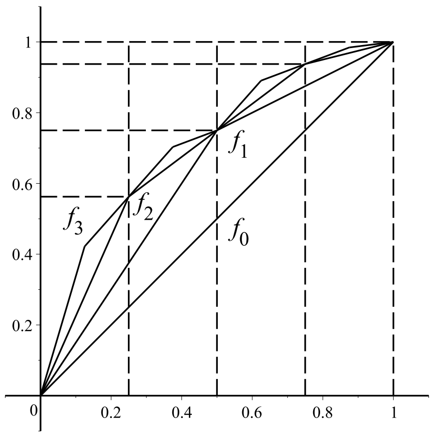

Figure 3.4: *The first iterations* f_0, f_1, f_2, f_3, *when* $\alpha = \frac{1}{2}$ *in the proof of Theorem 3.2.*

Choose $f_0(x) = x$ for $x \in [0,1]$. Then define by induction the continuous functions f_n, such that f_n is piecewise linear in each of the intervals $[k \cdot 2^{-n}, (k+1) \cdot 2^{-n}]$ for $k = 0, 1, \dots, 2^n - 1$. At the coordinate $x = k \cdot 2^{-n}$ we define

$$(3.6) \qquad f_{n+1}\left(k \cdot 2^{-n}\right) = f_n\left(k \cdot 2^{-n}\right),$$

hence, by induction,

$$(3.7) \qquad f_{n+m}\big(k \cdot 2^{-n}\big) = f_n\big(k \cdot 2^{-n}\big)\,, \qquad \text{for all } m \in \mathbb{N}.$$

This means that (3.6) fixes f_{n+1} at all points of the form $x = k \cdot 2^{-n} = (2k) \cdot 2^{-(n+1)}$. If we furthermore fix all the values

$$f_{n+1}\Big((2k+1) \cdot 2^{-(n+1)}\Big)\,, \qquad \text{for } k = 0,\, 1,\, \dots,\, 2^{n-1},$$

then the sequence (f_n) is uniquely determined, because it is also required to be piecewise linear.

We use the chosen $\alpha \in]0,1[$ in the beginning of the proof to define

$$(3.8) \qquad f_{n+1}\left(\frac{2k+1}{2^{n+1}}\right) := \frac{1-\alpha}{2} \cdot f_n\left(\frac{k}{2^n}\right) + \frac{1+\alpha}{2} \cdot f_n\left(\frac{k+1}{2^n}\right), \qquad \text{for } k = 0,\, 1,\, \dots,\, 2^n - 1.$$

A small calculation, or a consideration of the graph, shows that $f_n(x)$ for fixed $n \in \mathbb{N}$ is increasing in x. Furthermore,

$$0 \le f_n(x) \le f_{n+1}(x) \le 1, \qquad \text{for fixed } x \in [0,1],$$

so $\{f_n(x)\}_{n\in\mathbb{N}}$ is a bounded and increasing sequence of numbers for every $x \in [0,1]$. It is therefore convergent, and we have proved that (f_n) converges pointwise towards a function, which we denote

$$f(x) := \lim_{n\to+\infty} f_n(x).$$

For every $x \in [0,1]$ we choose a sequence of intervals

$$I_n := \big[k_n \cdot 2^{-n}, (k_n+1) \cdot 2^{-n}\big]\,, \qquad n \in \mathbb{N},$$

such that $x \in I_n$ for all $n \in \mathbb{N}$, where $k_n \in \{0,\, 1,\, \dots .\, 2^n - 1\}$.

Then it follows from (3.6) and (3.8) that there exists a sequence $\varepsilon = \pm 1$, such that

$$(3.9) \qquad \begin{aligned} &f_{n+1}\Big((k_{n+1}+1)\,2^{-(n+1)}\Big) - f_{n+1}\Big(k_{n+1}2^{-(n+1)}\Big) \\ &= \frac{1}{2}\{1+\varepsilon_{n+1}\alpha\} \cdot \big\{f_n\big((k_n+1)\,2^{-n}\big) - f_n\big(k_n 2^{-n}\big)\big\}\,. \end{aligned}$$

When we let $m \to +\infty$ in (3.7), we get

$$f\left(k \cdot 2^{-n}\right) = f_n\left(k \cdot 2^{-n}\right)\,, \qquad \text{for all } n \in \mathbb{N},$$

and we get by insertion into (3.9),

$$(3.10) \quad f\Big((k_{n+1}+1)\,2^{-(n+1)}\Big) - f\Big(k_{n+1}2^{-(n+1)}\Big) = \frac{1}{2}(1+\varepsilon_{n+1}\alpha)\,\big\{f\big((k_n+1)\,2^{-n}\big) - f\big(k_n 2^{-n}\big)\big\}\,.$$

By recursion in (3.10) and a change of index,

$$(3.11) \quad \begin{aligned} &f\big((k_n+1)\,2^{-n}\big) - f_n\left(k_n 2^{-n}\right) \\ &= \frac{1}{2}(1+\varepsilon_n\alpha)\left\{f\left((k_{n-1}+1)\,2^{-(n-1)}\right) - f\Big(k_{n-1}2^{-(n-1)}\Big)\right\} \\ &= \prod_{k=1}^{n} \frac{1}{2}(1+\varepsilon_k\alpha)\,\{f(1) - f(0)\} = \prod_{k=1}^{n} \frac{1}{2}(1+\varepsilon_k\alpha)\,, \end{aligned}$$

where $\varepsilon_k = \pm 1$.

Since f is increasing and $\alpha \in]0,1[$, we obtain the estimate

$$0 < \frac{1}{2}\left(1+\varepsilon_n \alpha\right) \le \frac{1}{2}(1+\alpha) < 1.$$

We conclude from (3.11) that

$$0 < f_n\left(\left(k_n+1\right)2^{-n}\right) - f_n\left(k_n\, 2^{-n}\right) = f\left(\left(k_n+1\right)2^{-n}\right) - f\left(k_n\, 2^{-n}\right) \le \left(\frac{1+\alpha}{2}\right)^n \to 0 \qquad \text{for } n \to +\infty.$$

It follows that f is continuous and strictly increasing on $[0,1]$.

If the derivative $f'(x)$ exists in the classical sense, then $f'(x)$ is given by (cf. also (3.11))

$$\text{(3.12)} \qquad f'(x) = \lim_{n \to +\infty} \frac{f\left(\dfrac{k_n+1}{2^n}\right) - f\left(\dfrac{k_n}{2^n}\right)}{\dfrac{k_n+1}{2^n} - \dfrac{k_n}{2^n}} = \lim_{n \to +\infty} 2^n \prod_{k=1}^n \frac{1+\varepsilon_k \alpha}{2} = \lim_{n \to +\infty} \prod_{k=1}^n \left(1+\varepsilon_k \alpha\right).$$

Notice that $0 < 1-\alpha < 1 < 1+\alpha < 2$, so none of the factors $1+\varepsilon_k\alpha$ is equal to 0, 1 or $+\infty$. We can therefore write

$$\prod_{k=1}^n \left(1+\varepsilon_k \alpha\right) = \exp\left(\sum_{k=1}^n \ln\left(1+\varepsilon_k \alpha\right)\right),$$

and it suffices to consider the series

$$\text{(3.13)} \qquad \sum_{k=1}^{+\infty} \ln\left(1+\varepsilon_k \alpha\right).$$

This series is divergent, because the sequence of its terms, $(\ln(1+\varepsilon_k\alpha))$ does not tend towards 0 for $k \to +\infty$. In fact, $1+\varepsilon_k\alpha \in \{1-\alpha, 1+\alpha\}$ for all $k \in \mathbb{N}$. We notice that the series (3.13) can only diverge in one of the following three ways,

i)	$\sum_{k=1}^{+\infty} \ln(1+\varepsilon_k\alpha) = -\infty$,	i.e.	$\prod_{k=1}^{+\infty}(1+\varepsilon_k\alpha) = 0$,
ii)	$\sum_{k=1}^{+\infty} \ln(1+\varepsilon_k\alpha) = +\infty$,	i.e.	$\prod_{k=1}^{+\infty}(1+\varepsilon_k\alpha) = +\infty$,
iii)	$\sum_{k=1}^{+\infty} \ln(1+\varepsilon_k\alpha)$ indetermined,	i.e.	$\prod_{k=1}^{+\infty}(1+\varepsilon_k\alpha)$ indetermined.

Returning to (3.12), where we formally found that

$$f'(x) = \lim_{n\to+\infty} \prod_{k=1}^{n}(1+\varepsilon_k\alpha) = \prod_{k=1}^{+\infty}(1+\varepsilon_k\alpha)\,,$$

it follows that if $f'(x)$ *exists* at a point x, then only i) above is possible, so we conclude that

$$f'(x) = \prod_{k=1}^{+\infty}(1+\varepsilon_k\alpha) = 0,$$

whenever f is differentiable at the point x.

The limit function f is monotonous an continuous, so it follows from *Lebesgue's theorem*,, i.e. Theorem 3.1, page 90, that $f'(x) = 0$, whenever f is differentiable at x. □

3.5 The Kakeya problem

The *Kakeya problem* is also called the *needle problem*. The problem is to find the infimum of the area $|A|$ of all sets $A \subseteq \mathbb{R}^2$, which have the following property. An infinitely thin needle of length 1 is moved around inside A in such a way that it can return to its original position, but in its reverse sense, where its two endpoints are interchanged.

The problem was originally posed by *Kakeya* [30] in 1917. It was solved by *Besicovitch* [7] in 1928. In the same year his construction was improved by *Perron* [44]. Later improvements were given in 1962 by *Rademacher* [46] and *Schoenberg* [48]. The presentation in the following is in broad outline following *Rademacher* [46].

Theorem 3.3 (The Perron tree). *Let A_k, $k = 1, 2, \dots, 2^n$, denote the triangle of the corners*

$$((k-1)\ell, 0),\ (k\ell, 0) \ \textit{and}\ (a, b),$$

where ℓ, $h > 0$ and $a \in \mathbb{R}$. Then the union

$$\bigcup_{k=1}^{2^n} A_k$$

is the triangle of the corners $(0,0)$, $(2^n\ell, 0)$ and (a, b).
For given $\alpha \in \left]\frac{1}{2}, 1\right[$ each triangle A_k can be translated parallel to the X axis to a congruent triangle $\overline{A}_k$, such that we have the following estimate of the measures,

$$\left|\bigcup_{k=1}^{2^n} \overline{A}_k\right| \le \left\{\alpha^{2n} + 2(1-\alpha)\right\}\left|\bigcup_{k=1}^{2^n} A_k\right|.$$

Proof. *Basic construction.* We shall first consider, what we here call the *basic construction.* Let T_1 and T_2 be any two neighbouring triangles with their bases lying on the X axis and of the same height h and of the same length b of their basis, cf. Figure 3.5.

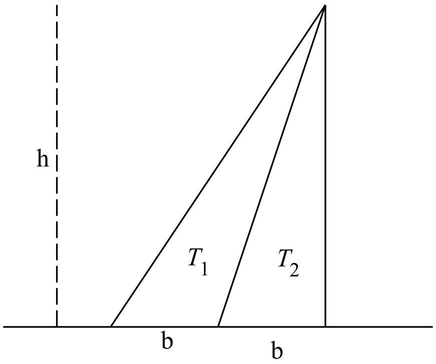

Figure 3.5: *Two neighbouring triangles in the basic construction.*

We translate T_2 along the X axis to the triangle T_2^*, such that the two non-parallel edges of the triangles intersect each other at the height αh, cf. Figure 3.6.

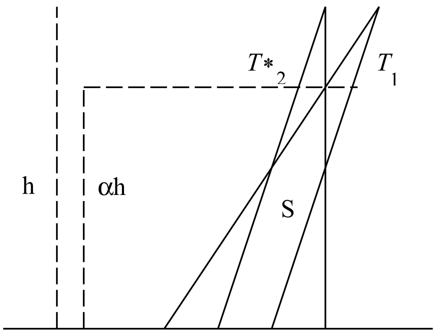

Figure 3.6: *The set $T_1 \cup T_2^*$, where T_2 has been translated along the X axis to T_2^*.*

The union $T_1 \cup T_2^*$ consists of a triangle S and two extra triangles. The triangles S and $T_1 \cup T_2$ are similar, where $T_1 \cup T_2$ is of height h, while S is of height αh. We therefore conclude that

$$|S| = \alpha^2 \, |T_1 \cup T_2| \, .$$

When we divide the two extra triangles along the line af height αh, cf. Figure 3.6, and use that they are pairwise congruent, and similar to either T_1 or T_2, we conclude that the two surplus triangles have the total area

$$2(1-\alpha)^2\left|T_1 \cup T_2\right|,$$

because the height of each of the four small triangles is $(1-\alpha)h$.

Summing up we get

$$|T_1 \cup T_2^*| = \left\{\alpha^2 + 2(1-\alpha)^2\right\} \cdot |T_1 \cup T_2|,$$

and we have finished the description of the *basic construction*.

Then return to the 2^n subtriangles A_k of the triangle given by its corners $(0,0)$, $(2^n\ell, 0)$ and (a,h). We organize these subtriangles in 2^{n-1} pairs, (A_1, A_2), (A_3, A_4), $\dots$, (A_{2^n-1}, A_{2^n}).

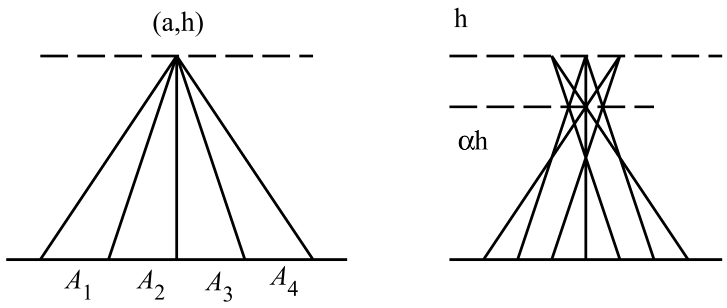

Figure 3.7: *The splitting of a triangle followed by the basic construction on each pair.*

We apply the basic construction on each pair. We get 2^{n-1} triangles S_1, S_2, $\dots$, $S_{2^{n-1}}$ and some extra triangles. The extra triangles have at most the total area

$$2(1-\alpha)^2\left|\bigcup_{k=1}^{2^n} A_k\right|. \tag{3.14}$$

Translate each pair, such that the triangles S_k, $k = 1, \ldots, 2^{n-1}$, form a splitting of a triangle of height αh,

$$\bigcup_{k=1}^{2^{n-1}} S_k, \qquad \text{which is similar to} \qquad \bigcup_{k=1}^{2^n} A_k \quad \text{of height } h.$$

Thus, concerning the areas,

$$\left|\bigcup_{k=1}^{2^{n-1}} S_k\right| = \alpha^2 \left|\bigcup_{k=1}^{2^n} A_k\right|.$$

In most cases the extra triangles mutually intersect each other, so (3.14) is an *upper limit* of their total area. It is, however, no need to find the exact area, because we shall move these triangles many times in the following, and it will be almost impossible to calculate the exact area in the general case.

In the next step we repeat the process on the triangles $S_1, S_2, \dots, S^{2^{n-1}}$. These are paired in the same way as for the A_k, and we apply the basic construction on each pair. This means that we in total translate four of the original triangles A_k for each pair (S_{2j-1}, S_{2j}). The result is 2^{n-2} new triangles $T_1, T_2, \dots, T_{2^{n-2}}$, where the area is

$$\left|\bigcup_{k=1}^{2^{n-2}} T_k\right| = \alpha^2 \left|\bigcup_{k=1}^{2^{n-1}} S_k\right| = \left(\alpha^2\right)^2 \left|\bigcup_{k=1}^{2^n} A_k\right|,$$

and some surplus triangles, the union of which has at most the area

$$2(1-\alpha)^2 \left|\bigcup_{k=1}^{2^{n-1}} S_k\right| = 2(1-\alpha)^2\alpha^2 \left|\bigcup_{k=1}^{2^n} A_:k\right|. \tag{3.15}$$

Proceed in this way. At step number k we have 2^{n-k} triangles of the total area

$$\left(\alpha^2\right)^k \left|\bigcup_{k=1}^{2^n} A_k\right|,$$

and some extra triangles of at most the area

$$2(1-\alpha)^2 \left(\alpha^2\right)^{k-1} \left|\bigcup_{k=1}^{2^n} A_k\right|.$$

After n steps the 2^n triangles A_k, $k = 1, \dots, 2^n$, have been translated along the X axis in such a way that the final set resembles a tree of a small trunk (the diminished triangle, similar to the original one), and a lot of branches (all the extra small triangles). We call this figure a *Perron tree*.

Using the results above we estimate the area of the *Perron tree* $\bigcup_{k=1}^{2^n} \overline{A}_k$ in the following way

$$\begin{aligned}
\left|\bigcup_{k=1}^{2^n} \overline{A}_k\right| &\le \left(\alpha^2\right)^n \left|\bigcup_{k=1}^{2^n} A_k\right| + 2(1-\alpha)^2 \sum_{k=1}^{n} \left(\alpha^2\right)^{k-1} \left|\bigcup_{k=1}^{2^n} A_k\right| \\
&= \left\{\alpha^{2n} + 2(1-\alpha)^2 \cdot \frac{1-\alpha^{2n}}{1-\alpha^2}\right\} \cdot \left|\bigcup_{k=1}^{2^n} A_k\right| = \left\{\alpha^{2n} + 2(1-\alpha) \cdot \frac{1-\alpha^{2n}}{1+\alpha}\right\} \cdot \left|\bigcup_{k=1}^{2^n} A_k\right| \\
&< \left\{\alpha^{2n} + 2(1-\alpha)\right\} \cdot \left|\bigcup_{k=1}^{2^n} A_k\right|,
\end{aligned}$$

and the theorem is proved. □

For given $\varepsilon > 0$, we choose $\alpha \in \left]\frac{1}{2}, 1\right[$, such that $2(1-\alpha) < \frac{\varepsilon}{2}$. Then choose $n \in \mathbb{N}$, such that $\alpha^{2n} < \frac{\varepsilon}{2}$. If we carry out the construction of the *Perron tree* above on any triangle with its base line on the X axis and of area A (and using the parameters α an n above) we obtain a Perron tree of area $< \varepsilon A$. Once this has been realized, it is easy to solve the Kakeya problem.

Theorem 3.4 *For every $\varepsilon > 0$ and every straight line segment AB of length 1 in $\mathbb{R}^2$ we can* construct *a figure $F \subset \mathbb{R}^2$, such that $|F| < \varepsilon$, and such that the line segment AB can be moved continuously inside F, such that it finally returns to its original position with the endpoints A and B reversed.*

Proof. A simple geometric analysis, cf. Figure 3.8, shows that the needle can be moved from a line to a parallel line by adding a set of area $2 \cdot 1 \cdot v = 2v$, where v is measured in radians, and a straight line segment, which is a nullset.

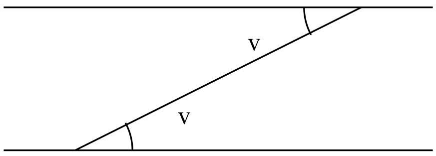

Figure 3.8: *Translation of the needle from one straight line to a parallel straight line.*

We can of course choose $v > 0$ as small as needed, so moving the needle from one line to a parallel one will only require a set of a measure which can be made as small as we want, as long as we only use a finite number of movements of the needle from a line segment to a parallel one.

Then notice that it suffices to prove that the needle AB can be moved inside a set of area $\frac{\varepsilon}{3}$ to some straight line segment which forms the angle $\frac{\pi}{3}$ with its original position. In fact, if we can do this, then we just repeat this construction three times followed by a translation from one line segment to the original one, and the needle AB has been turned around the angle of π, i.e. A and B have been reversed.

Let A be an equilateral triangle of height > 1 and its base on the X axis. Let the initial position of the needle be one of the edges not lying on the X axis. We shall move the needle to the other edge not lying on the X axis.

Choose $\eta > 0$, such that

$$|A|\eta < \frac{\varepsilon}{6}, \qquad \text{i.e.} \qquad \eta < \frac{\varepsilon}{6|A|},$$

and then use the construction of the Perron tree with an $\alpha \in \left]\frac{1}{2}, 1\right[$, such that $2(1-\alpha) < \frac{\eta}{2}$, and with an $n \in \mathbb{N}$, such that $\alpha^{2n} < \frac{\eta}{2}$.

The height of A is > 1 by assumption, hence the needle can be turned inside each $\overline{A}_k$ from one of its longer edges to the other one, if only the needle can enter $\overline{A}_k$.

The needle starts at the utmost edge (to the left) in $\overline{A}_1$, and is then turned inside $\overline{A}_1$ to the other edge of $\overline{A}_1$. This is parallel to one of the edges of $\overline{A}_2$, and by switching from the final edge of $\overline{A}_1$ to the initial edge (to the left) of $\overline{A}_2$ we shall only require a set of measure $2f$, where we choose

$$(0 <)\, v < \frac{\varepsilon}{12} \cdot \frac{1}{2^n}.$$

This operation requires a set of area

$$2v < \frac{\varepsilon}{6 \cdot 2^n},$$

which is possible according to the analysis above. Once the needle lies on the initial edge of $\overline{A}_2$, which also is of height > 1, it is easy to move the needle inside $\overline{A}_2$ to the final edge of $\overline{A}_2$.

Proceed in this way, until the needle has been turned the angle $\frac{\pi}{3}$ from is initial position. Notice that we have only added a set of area

$$2v \cdot 2^n < \frac{\varepsilon}{6}.$$

In this way the needle is turned the angle $\frac{\pi}{3}$ within a set of at most the area

$$|A|\eta + \frac{\varepsilon}{6} < \frac{\varepsilon}{6} + \frac{\varepsilon}{6} = \frac{\varepsilon}{3}.$$

When we repeat this construction three times, we have turned the needle π radians, i.e. the needle has been reversed inside a set F of area $< 3 \cdot \frac{\varepsilon}{3} = \varepsilon$. $\square$

3.6 The Nikodym set

A *Nikodym set* is a set $N \subset [0,1]^2$ with the following two properties,

1) Its measure is $|N| = 1$.

2) To every $x \in N$ we can find an infinite straight line $\ell(x)$, such that $N \cap \ell(x) = \{x\}$.

.

Condition 2) means that every $x \in N$ can be reached by a full straight line lying in the complementary set of N, with the exception of the point $\{x\}$ itself. And yet condition 1) tells us that $|N| = 1$, so

$$\left|\bigcup_{x \in N} \ell(x) \setminus N\right| = 0.$$

The existence of such a set was first shown by *Nikodym* [41] in 1927. His elementary, though extremely complicated proof was later simplified by *Davies* [17] and *Cunninngham* [15], [16]. The titles of Cunningham's two papers show that this problem is closely connected with the *Kakeya problem* of Section 3.5.

The construction here is based on the following

Theorem 3.5 *There exists a nullset $K \subset \mathbb{R}^2$, such that to every $x \in \mathbb{R}^2$ there exists an infinite straight line $\ell(x)$, such that*

$$x \in \ell(x) \qquad \text{and} \qquad \ell(x) \subset K \cup \{x\}.$$

So every point $x \in \mathbb{R}^2$ lies on a straight line $\ell(x)$, which is contained in the union of $\{x\}$ and the nullset K.

Once Theorem 3.5 has been proved, and K is a nullset with the properties of Theorem 3.5, we get a *Nikodym set* by $N = [0,1]^2 \setminus K$. In fact, $|N| = 1$ is obvious, and if $x \in N$, then it follows from Theorem 3.5 that we can find a straight line $\ell(x)$ in the plane, such that

$$\ell(x) \cap B = \{x\}.$$

The proof of Theorem 3.5 relies on two lemmata, of which the first one below is the hard one. The proof is a purely classical geometric proof of a type which today is almost forgotten.

Lemma 3.2 *Let R_1 and R_2 be two closed parallelograms in $\mathbb{R}^2$, where $R_1 \subset R_2$. Choose any $\varepsilon > 0$, and denote by ω one of the two closed strips, which is determined by two parallel edges of R_1.*
There exists a finite family $\Omega = \{\omega_1, \omega_2, \dots, \omega_k\}$ of closed strips, such that

$$\omega_i \cap R_2 \subset \omega \cap R_2 \qquad \text{for every } i \in \{1,2,\dots,k\}, \tag{3.16}$$

$$R_1 \subset \bigcup_{i=1}^k \omega_i, \tag{3.17}$$

$$\left|\left(\bigcup_{i=1}^k \omega_i\right) \cap (R_2 \setminus R_1)\right| \leq \varepsilon. \tag{3.18}$$

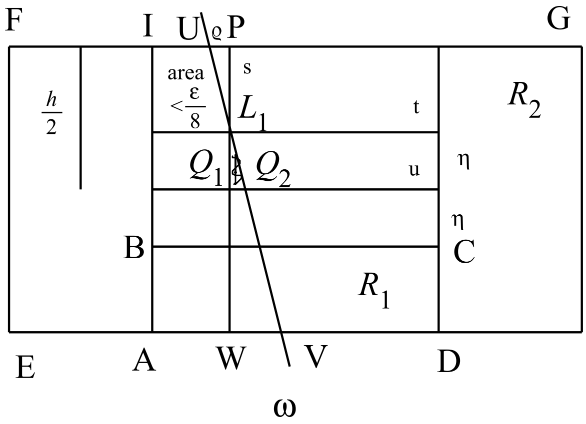

Figure 3.9: *The first case of the proof of Lemma 3.2.*

PROOF. We shall first prove the lemma in the special case, where R_1 and R_2 are axiparallel rectangles with their base lines lying on the same line.

Let R_1 be the rectangle $ABCD$ on Figure 3.9, and let R_2 be the larger rectangle $EFGH$ containing R_1, where the line segment AD (the base line of R_1) is contained in the base line EH of R_2. Let ω be the closed parallel strip determined by the line segments AB and CD. Finally, choose any $\varepsilon > 0$.

Choose s as the vertical line PW in ω, which is parallel with AB, such that the rectangle determined by the vertical line segments BI and s, and IP and BC, has area $< \frac{\varepsilon}{8}$. For clarity this has been exaggerated on Figure 3.9.

Then choose a horizontal line segment t, parallel with BC, such that the subset in ω between t and the line BC has area $< \frac{\varepsilon}{8}$.

The distance between t and BC is put equal to 2η. Let u be the line segment in ω in parallel with t and BC and of distance η from both of them,

For a line segment like e.g. AB we denote its length by $|AB|$. The intersection point of the two perpendicular line segments t ans s is denoted by L_1, and Q_1 denotes the intersection point between the two perpendicular line segments u and s. We then choose a point U on the line segment IP at the distance ϱ from P, such that also the straight line through U and L intersects the base line AD of R_1, and such that if Q_2 denotes the intersection point of this oblique line and the line u, then

$$|IU| > |Q_1Q_2| \qquad \text{and} \qquad \angle UL_1P < \frac{\pi}{4},$$

and such that the intersection point V between the line through UL_2 and AD has a distance from A, which satisfies

$$|AV| < \frac{1}{4} AD|.$$

By choosing $\eta > 0$ sufficiently small it is obvious that U can be chosen, such that the above is fulfilled. In this way we have constructed a triangle $\triangle L_1Q_1Q_2$.

Let $k \in \mathbb{N}$ denote the smallest natural number, for which

$$2k\,|Q_1Q_2| \geq |WD|.$$

Let $\frac{h}{2}$ denote the distance between u and the upper line segment FG. Let A be the triangle of height h with two of its edges lying on the perpendicular lines s and u, similar to $\triangle L_1Q_1Q_2$. (This triangle is for clarity not sketched on Figure 3.9.) Its notation is due to the previous notation of the Perron tree. There should be from now on only a small danger of confusing it with the point A on the base line of R_1.

It follows from $\eta < \frac{h}{2} < h$, that we can choose $n \in \mathbb{N}$, such that

$$\left\{1 - \sqrt[n]{\frac{\eta}{j}}\right\} h^2 < \frac{\varepsilon}{8k}.$$

Once n has been fixed, find $\alpha \in]0,1[$, such that $\alpha^n h = \eta$. Then apply the construction of the Perron tree of Section 3.5 on the triangle A with respect to these parameters. This means that we split A into 2^2 triangles $A_1, A_2, \dots, A_{2^n}$, and then translate them along the line u to thee triangles $\overline{A}_1, \overline{A}_2, \dots, \overline{A}_{2^n}$, such that the "trunk" of the Perron tree is precisely the triangle $\triangle L_1Q_1Q_2$. Here we use the choice of α, such that $\alpha^n h = \eta$. Furthermore, the "branches" of the Perron tree has a total area, which is at most

$$2(1-\alpha)|A| = 2\left\{1 - \sqrt[n]{\frac{\eta}{h}}\right\} \cdot \frac{1}{2}\,h^2 \tan\left(\angle Q_1L_1Q_2\right) < \left\{1 - \sqrt[n]{\frac{\eta}{h}}\right\} h^2 < \frac{\varepsilon}{8k}. \tag{3.19}$$

We consider one of the translated triangles $\overline{A}_i$, cf. Figure 3.10, which is a part of Figure 3.9, where we have removed all irrelevant objects,

Let α and β denote the intersection points o the longer edges of $\overline{A}_i$. Let γ and δ be the intersection points of the line segment AD and the same edges of $\overline{A}_i$. Then cover the line segment $\gamma\delta$ with a finite number of closed line segments, all contained in the line segment $\gamma\delta$, and all of length $|\alpha\beta|$. Finally, consider all strips, which are obtained by joining α and β by line segments with the endpoints of each of these minor segments. We have shown two of them on Figure 3.10.

The construction guarantees that these strips cover $\triangle Q_1WV$ on the previous Figure 3.9.

It follows from $|Q_1Q_2| < |UI|$ that the intersection of each of these strips with the rectangle R_2 is contained in the strip ω, defined by R_1. In fact, all translated triangles $\overline{A}_1, \overline{A}_2, \dots, \overline{A}_{2^n}$ have their base line lying within the line segment Q_1Q_2, and their edges from an angle with the base line which is at most $\angle Q_1L_1Q_2$. This implies that the left hand side of each of the strips, which we have defined above, intersects the line segment FG at a point to the left of U and at a distance from U, which is smaller that $|Q_1Q_2|$.

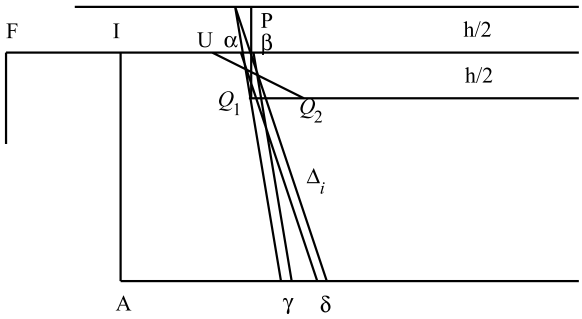

Figure 3.10: *The construction of Perron tree in the proof of Lemma 3.2.*

Figure 3.9 and Figure 3.10 show that the part of the union of these strips stemming from $\overline{A}_1$, and which also lie between the line t and the line segment FG, is contained in $\overline{A}_i$. Hence, the union of them all for $i = 1, , 2, \ldots, 2^n$, lying between these lines, must be contained in the "branches" of the Perron tree. Using (3.19) we conclude that their total area is $< \varepsilon/(8k)$.

We have above described the procedure on the first triangle $\triangle L_1Q_1Q_2$. We then define k triangles $\triangle L_jQ_jQ - j+1$, $j = 1, 2, \dots, k$, by simply translating $\triangle L_1Q_1Q_2$, along the lines t and u, cf. Figure 3.11.

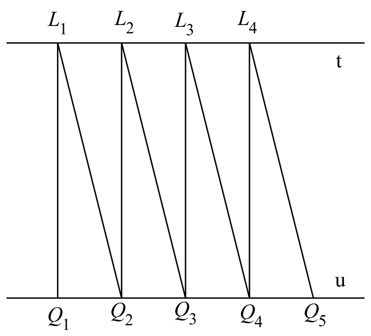

Figure 3.11: *The translated triangles $\triangle L_jQ_jQ_{j+1}$, $j = 1, 2, \dots, k$ in the proof of Lemma 3.2.*

We have already chosen the smallest possible k, for which $k\,|Q_1Q_2| \geq \frac{1}{2}\,|WD|$. This definition implies that Q_{k+1} lies in the right hand side of the strip ω.

It should be no surprise to the reader that we then apply the same procedure as above on all the triangles $\triangle L_jQ_jQ_{j+1}$, $j = 1, 2, \dots, k$. In this way we obtain a finite number of strips, which together with the strip defined by the parallel line segments IA and PW cover the left half of the rectangle R_1. Each of these strips have an intersection with R_2, which is contained in ω. Finally, the part of the union of all strips derived from one of the triangles between the lines t and FG has an area, which according to the construction of the Perron tree is smaller than $k \cdot \frac{\varepsilon}{7}(8k) = \varepsilon/8$.

The area of the domain in ω between t and BC is smaller than $\frac{\varepsilon}{8}$. Hence, the union of all these strips from $\overline{A}_i$ has an intersection with $R_2 \setminus R_1$ of an area, which is $< \frac{\varepsilon}{8} + \frac{\varepsilon}{8} = \frac{\varepsilon}{4}$. We add the vertical strip between IA and PW, and the horizontal strip between t and BC, thus covering the left half of the rectangle R_1 with strips, where the union of all these inside $R_2 \setminus R_1$ has an area, which is smaller than $\frac{\varepsilon}{4} + \frac{\varepsilon}{8} + \frac{\varepsilon}{8} = \frac{\varepsilon}{2}$.

Then repeat the same construction on the right half of R_1, considering all the mirror images instead. It follows from $2 \cdot \frac{\varepsilon}{2} = \varepsilon$ that the lemma is proved in this special case.

Then we start on removing all the restrictions given on R_1 and R_2. The simplest case is when R_1 and R_2 are parallelograms lying as shown on Figure 3.12.

Just use that an affine map will transform this situation back to the previous one. It is well-known that an affine map preserves area, so the case follows.

Assume that R_1 and R_2 are the parallelograms $ABCD$ and $EFGH$ on Figure 3.13, where R_1 and R_2 have their base lines lying on the same line, while the transversal edges of R_1 and R_2 are not parallel. We replace R_2 with the larger R_2^*, where we add two triangles, such that R_1 and R_2^* have

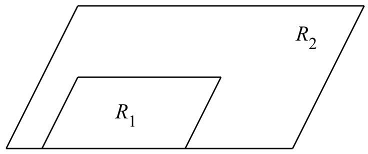

Figure 3.12: *Affine transformation of the rectangles in the proof of Lemma 3.2.*

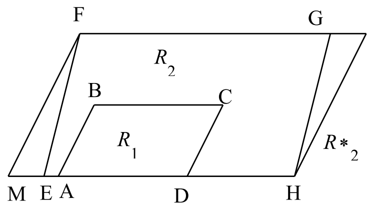

Figure 3.13: *Reduction of an even more general situation in the proof of Lemma 3.2.*

parallel transversal edges. Then R_1 and R_2^* are lying as in Figure 3.12. Therefore, Lemma 3.2 holds for the pair (R_1, R_2^*), and because $R_2 \subset R_2^*$, the lemma also holds for the pair (R_1, R_2) on Figure 3.13, because R_2 is a smaller set than R_2^*.

We assume that R_1 and R_2 have parallel edges and that R_1 lies entirely in the interior of R_2, cf. Figure 3.14. We have already proved the lemma for $R_1^* = MBCN$ and R_2 with $\frac{\varepsilon}{2}$. Each of the strips $\omega_j^* \cap R_2$, $j = 1, \dots, k$, lies totally inside the parallelogram $MPQN$. Then apply the lemma on each of the parallelograms $\omega_j^* \cap APQD$ and R_2 with the constant $\frac{\varepsilon}{2k}$. We get for each $j = 1, 2, \dots, k$ the strips $(\omega_j)_{i=1,2,\dots,r_j}$. The family Ω of all such strips ω_j satisfy the conditions (3.15), (3.16) and (3.17) of Lemma 3.2.

Finally, we consider the general situation as in Figure 3.15. Then replace R_2 by the larger parallelogram $R_1^* \supset R_2$, the edges of which are parallel with the edges of R_1. Using the result above, the lemma holds for R_1 and R_2^*, cf. Figure 3.15. The set R_2 is contained in R_2^*, so the lemma also holds for R_1 and R_2. This proves the lemma in general. $\square$

We immediately conclude from Lemma 3.2

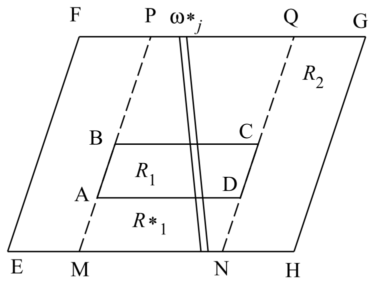

Figure 3.14: *The case in the proof of Lemma 3.2, where* R_1 *lies in the interior of* R_2*, and with parallel edges.*

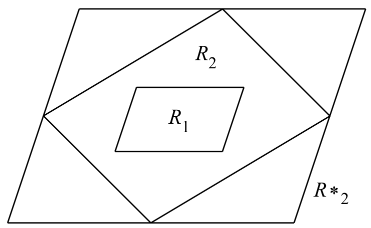

Figure 3.15: *The general case in the proof of Lemma 3.2, where* R_1 *lies in the interior of* R_2*.*

Lemma 3.3 *Let* R_1 *and* R_2 *be two closed parallelograms, where* $R_1 \subset R_2$*. Let* $\Omega = \{\omega_1, \omega_2, \dots, \omega_k\}$ *be a finite family of closed strips, the union of which covers* R_1*. Let* $\varepsilon > 0$*. For every strip* ω_i*,* $i = 1, 2, \dots, k$*, there exists another finite family of closed strips* $\omega_i^1, \omega_i^2, \dots, \omega_i^{j_i}$*, such that if we collect them all,*

$$\Omega^* := \left\{ \omega_i^j \mid i = 1, 2, \dots, k \text{ and } j = 1, 2, \dots, j_i \right\},$$

then

$$\bigcup_{\omega \in \Omega^*} \omega \subset R_1, \tag{3.20}$$

$$\omega_i^j \cap R_2 \subset \omega_i \qquad \text{for all } i \text{ and } j, \tag{3.21}$$

$$\left| \left(\bigcup_{\omega \in \Omega^*} \omega \right) \cap (R_2 \setminus R_1) \right| \leq \varepsilon. \tag{3.22}$$

The simple proof is left to the reader.

Finally, we are able to prove the theorem.

Proof of Theorem 3.5. Let $Q(H)$, $H > 0$, denote the closed axiparallel square of centre $(0,0)$ and edge length $2H$, i.e. $Q(H) = [-H, H] \times [-H, H]$. We write for short $Q(1) = Q$.

Given any sequence (ε_n), $0 < \varepsilon_n < 1$, which converges towards 0. Apply Lemma 3.2 on $R_1 = Q$ and $R_2 = Q(2)$ with $\frac{1}{4}\varepsilon_1$. We obtain a family Ω_1 of strips.

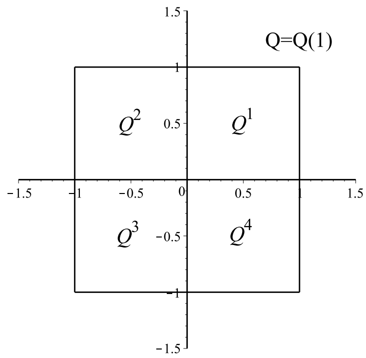

Figure 3.16: *The division of Q into four subsquares.*

We divide Q into four closed squares of equal size, $Q^i = Q_1^i$, $i = 1,\ 2,\ 3,\ 4$, as shown on Figure 3.16. Clearly, the edge length of each Q_1^i is 1. For each fixed $i \in \{1,2,3,4\}$ we apply Lemma 3.3 on $R_1 = Q_1^i$ and $R_2 = Q(3)$ with $\Omega = \Omega_1$ and $\varepsilon := 4^{-2}\,\varepsilon_2$. We get four families Ω_2^i, $i = 1,2,3,4$, of closed strips. These are all collected in

$$\Omega_2 := \bigcup_{i=1}^{4} \Omega_2^i.$$

Then divide each Q_1^i into four closed squares of equal size, i.e. of edge length $\frac{1}{2}$. We get 4^2 squares Q_2^i, $i = 1,2,\dots,4^2$. Fix i and apply Lemma 3.3 on $R_1 = Q_2^i$ and $R_2 = Q(4)$ with $\Omega = \Omega_2$ and $\varepsilon := 4^{-3}\varepsilon_3$. We obtain 4^2 families Ω_3^i, $i = 1,2,\dots,4^2$ of strips. We collect these in

$$\Omega_3 := \bigcup_{i=1}^{4^2} \Omega_3^i.$$

Proceed in this way.

Notice that for every fixed k the square Q_{k-1}^i is covered by the union of all strips in Ω_k^i.

For given $\omega \in \Omega_k^i$ we define

$$\hat{\omega} := \overline{\omega \setminus \Omega_{k-1}^i},$$

and

$$K_k := \bigcup_{\omega \in \Omega_k} \hat{\omega}.$$

The system Ω_k^i was constructed by an application of Lemma 3.3, thus

$$\left| \left(\bigcup_{\omega \in \Omega_k^i} \hat{\omega} \right) \cap Q(k+1) \right| \le 4^{-k}\varepsilon_k,$$

so

$$|K_k \cap Q(k)| \le \varepsilon_k, , \qquad \text{for every } k.$$

We finally define

$$K^* := \liminf_{k \to +\infty} K_k = \bigcup_{h=1}^{+\infty} \bigcap_{k=h}^{+\infty} K_k. \tag{3.23}$$

Choose for fixed N and h a $j > \max\{h,N\}$, such that

$$\left| \left(\bigcap_{k=h}^{+\infty} K_k \right) \cap Q(N) \right| \le |K_j \cap Q(j)| \le \varepsilon_j.$$

This estimate holds for all $j > \max\{h,N\}$, so it follows from the assumption $\varepsilon_j \to 0$ that

$$\left| \left(\bigcap_{k=h}^{+\infty} K_k \right) \cap Q(N) \right| = 0, \qquad \text{for all } N.$$

We finally get for $N \to +\infty$,

$$\left|\bigcap_{k=h}^{+\infty} K_k\right| = 0 \qquad \text{for every } k,$$

and it follows from (6.2) that K^* is a nullset, $|K^*| = 0$.

In the next step we prove that to every $x \in Q$ there exists a straight line $\ell(x)$ through x and contained in the nullset $K^* \cup \{x\}$.

Let $x \in Q$ be given, and let $Q_n^{j(x,n)}$, $n \in \mathbb{N}$, be a subsequence of the sequence of squares we chose previously, such that

$$x \in Q_n^{j(x,n)} \quad \text{for all } n, \qquad \text{and} \qquad \left|Q_n^{j(x,n)}\right| \to 0 \quad \text{for } n \to +\infty.$$

For $n = 1$ we choose a strip $\omega_1 \in \Omega_1^{j(x,1)}$, such that $x \in \omega_1$.

Then choose for $n = 2$ a strip $\omega_2 \in Q_2^{j(x,2)}$, such that $x \in \omega_2$, and such that $\omega_2 \cap Q(2) \subset \omega_1$.

Proceed in this way.

For $n = k$ we can choose a strip $\omega_k \in \Omega_k$, such that $x \in \omega_k$ and such that $\omega_k \cap Q(k) \subset \omega_{k-1}$. Obviously, there also exists a sequence of straight lines $\{\ell_k(x)\}$, all passing through x, such that $\ell_k(x) \subset \omega_k$. Then it follows from $\varepsilon_k \to 0$ for $k \to +\infty$ that the width of the strips ω_k tends towards 0 for $k \to +\infty$. We furthermore conclude from $\omega_k \cap Q(k) \subset \omega_{k-1}$ that the directions of the lines $\{\ell_k(x)\}$ converges towards the direction of a well-defined line $\ell(x)$ through x.

Finally, we shall prove that $\ell(x) \subseteq K^* \cup \{x\}$. So let $y \in \ell(x) \setminus \{x\}$. Then we can find $N \in \mathbb{N}$, such that for all $n \geq N$,

$$y \in Q_n^{j(x,n)} \qquad \text{and} \qquad y \in \{Q(n)\}^\circ, \qquad \text{(the interior of } Q(n)).$$

Choose a sequence of points $\{y_k\}$, such that $y_k \in \ell_k(x)$ and $y_k \to y$ for $k \to +\infty$. Then we can find another (larger) $N \in \mathbb{N}$, such that for all $k \geq N$,

$$y_k \in Q(N) \setminus Q_N^{j(x,N)}.$$

If $i > n \geq \max\{M, N\}$, then

$$y_i \in \ell_i(x) \cap Q(N) \subset \omega_i \cap Q(n) \subset \omega_n.$$

We notice that $y_i \in Q_N^{j(x,N)}$ implies that also $y_i \in Q_n^{j(x,n)}$, hence $y \in \hat{\omega}$.

Summing up, we have proved that for every $n \geq \max\{M, N\}$ we have $y_i \in \hat{\omega}$ for all $i \geq n$. The set $\hat{\omega}_n$ is closed, and $y_i \to y$, so we also have $y \in \hat{\omega}_n$, and therefore $y \in K^*$. Since y was any point of $\ell(x) \setminus \{x\}$, we therefore conclude that $\ell(x) \subset K^* \cup \{x\}$.

Notice finally that the procedure above can be applied on every square Q. In other words, for every given $x \in Q$ there exists a straight line $\ell(x)$ through x, such that

$$x \in \ell(x) \subset K^* \cup \{x\}.$$

When we apply the above on the (countable) sequence $\{Q(n)\}_{n\in\mathbb{N}}$ of squares, we obtain a sequence of nullsets $\{K_n^*\}$.

If we put

$$K := \bigcup_{n=1}^{+\infty} K_n^*,$$

then K is a countable union of nullsets, thus itself a nullset, and we see that the set K has all the other required properties, so the theorem is proved. $\square$

3.7 A set which is filling almost all of $\mathbb{R}$ without containing any homogeneous subset.

Some problems in *Complex Functions Theory* require that a given subset of $\mathbb{R}$ contains an homogeneous subset in the sense of Carleson (to be defined below). It was e.g. proved in *Kahanpää and Mejlbro* [29] the already well-known fact that the Hilbert transform of a function f is totally controlled (in the L^p norm) over a set $A \subseteq \mathbb{R}$, if and only if A is homogeneous. It was in the same paper proved that to every $\varepsilon > 0$ there exists a closed set A, such that $|\mathbb{R} \setminus A| < \varepsilon$, and such that no subset $B \subseteq A$ is homogeneous. We include here the construction of the set A of these properties, because it may have interest in itself, once we have defined, what is meant by an homogeneous set.

Let us start with the definition of Carleson's homogeneous set.

Definition 3.3 *A nonempty measurable subset $A \subseteq \mathbb{R}$ is called* homogeneous*, if there exists a constant $c > 0$, such that*

(3.24) $\quad |A \cap [x-r, x+r]| > 2cr = c|[x-r, x+r]| \qquad$ *for all $x \in A$ and all $r > 0$.*

In other words, a nonempty measurable set A is *homogeneous*, if every bounded interval I *with its midpoint lying in A*, contains some portion of A, at least the fraction c of the length of the interval I. For some obscure reason, which is not clearly understood, it is the geometrical assumption that the *midpoint* of I should lie in A, which makes this construction of the example below possible.

The definition means intuitively that no point in A is "sparse" in A. If we let $r \to 0+$ in (3.24), then roughly speaking the homogeneous set A is more or less uniformly distributed.

For convenience later on we get by a negation of Definition 3.3 that a nonempty subset A of $\mathbb{R}$ is *not* homogeneous, if for every constant $c > 0$ we can find $x \in A$ and $r > 0$, such that

(3.25) $\quad |A \cap [x-r, x+r]| < 2cr.$

In order to become familiar with homogeneous sets we include a couple of simpler examples.

Example 3.1 Every halfline is homogeneous. It suffices to prove this for the halfline $A = [0, +\infty[$. We see that the measure $|A \cap [x-r, x+r]|$ is smallest for $x \in A$, if $x = 0$, in which case

$$|A \cap [-r, r] = |[0, r]| = r = \frac{1}{2} \cdot 2r,$$

proving that the homogeneity constant is $c = \frac{1}{2}$. $\Diamond$

Example 3.2 The set

$$A = \bigcup_{n=-\infty}^{+\infty} [2n, 2n+1]$$

is homogeneous with the homogeneity constant $c = \frac{1}{3}$. The simple proof is left to the reader. $\Diamond$

We notice that every homogeneous set must have an infinite measure. In fact, assume that $|A| < +\infty$, and let $x \in A$. Then

$$\frac{|A \cap [x-r, x+r]|}{|[x-r, x+r]|} \leq \frac{|A|}{2r} \to 0 \qquad \text{for } r \to +\infty,$$

so (3.25) holds, and A is not homogeneous. In particular, *no bounded set is homogeneous!* We shall later use this result.

Obviously, if A contains isolated points, then (3.25) holds at each of these, so A itself is not homogeneous. However, if we remove all the isolated points from A, the result A_1 intuitively is intrinsically more packed, so A_1 might be homogeneous, while the larger set A is not. One simple example is $A = \{-1\} \cup [0, +\infty[$, where $A_1 = [0, +\infty[$ according to Example 3.1 is homogeneous. This simple observation may unfortunately lure the reader to jump to the conclusion that *"if $|A| = +\infty$, then A contains an homogeneous subset"*. The tempting idea is that we first remove all points $x \in A$, for which (3.25) holds. Call the remainder set A_1. It is obvious that if (3.25) is only fulfilled in a nullset of A, then A_1 is indeed homogeneous. Unfortunately, (3.25) may be satisfied in all points of a subset of A of positive measure, so by their removal we get $|A \setminus A_1| > 0$. In that case we must repeat the procedure above on A_1, so we remove all points from A_1, in which (3.25) holds (this time with respect to the smaller set A_1), in order to get the subset A_2. In this way we may proceed, though it is by no means obvious that this process will ever stop. We shall prove in the following that even this idea may seem plausible, it is not always successful. In fact, we shall prove

Theorem 3.6 *To every* $\varepsilon > 0$ *there exists a* closed set $A \subset \mathbb{R}$, *such that* $|\mathbb{R} \setminus A| < \varepsilon$, *and such that* A *does not contain any homogeneous subset.*

We shall leave to the reader the simple proof that we in the proof may assume that

$$A = \bigcup_{n=1}^{+\infty} [p_n, q_n],$$

where p_n, $q_n \in \mathbb{Q}$ are rational numbers, such that $p_n < q_n$, and such that the closed intervals $[p_n, q_n]$ are mutually disjoint.

The proof of Theorem 3.6 relies on the following lemma, which proves the existence of sets, which do not contain "locally" homogeneous subsets, a terminology which will be made precise in the following.

Lemma 3.4 *Let the constants* μ, $c \in]0,1[$ *be given. There exists a closed set* $A \subset [0,1]$, *such that* $|A| = \mu$, *and such that for every nonempty and measurable subset* E *of* A *there exist a point* $x \in E$ *and an* $r \in]0,1[$, *for which*

$$(3.26) \qquad [x-r, x+r] \subseteq [0,1] \qquad \textit{and} \qquad |E \cap [x-r, x+r]| < cr.$$

Condition (3.26) means that if $x \in E$, then the interval $I = [x-r, x+r]$ can be chosen, such that it is contained in $[0,1]$. This is the reason for saying that A is not "locally" homogeneous. Typically, μ lies very close to 1, and c lies very close to 0.

Proof of Lemma 3.4. It follows from $\mu < 1$ and

$$\frac{4\mu\left(1+2^{-n}\right)}{(n+1)c} \to 0, \qquad \text{for } n \to +\infty,$$

that we can choose $n \in \mathbb{N}$, such that

$$(3.27) \qquad \mu + \frac{4\mu\left(1+2^{-n}\right)}{(n+1)c} < 1.$$

Once n has been fixed, such that (3.27) is fulfilled, we define a *standard length* ℓ by

$$\ell := \frac{\mu}{2^n(n+1)},$$

and a *gap length* r_0 by

$$r_0 := \frac{2\mu}{2^n(n+1)c}.$$

We construct the set A as the union of $2^{n+1}-1$ closed intervals I_j, $j = 1, 2, \dots, 2^{n+1}-1$. In particular, A is closed.

We spread these intervals over $[0,1]$ in their natural order, separated by open gaps of length r_0, such that I_1 has its left endpoint lying at the coordinate $r_0 \in]0,1[$. We shall only specify the length of each I_j, $j = 1, 2, \dots, 2^{n+1}-1$. For every j there are uniquely determined constants $m \in \{0, 1, \dots, n\}$ and $k \in \left\{0, 1, \dots, 2^{n+1}-1\right\}$, such that $j = 2^m(2k+1)$ We define I_j as a closed interval of length

$$|I_j| = \left|I_{2^m(2k+1)}\right| = 2^m \ell.$$

In particular, $|I_{2k+1}| = \ell$.

The $2^{m+1} - 1$ intervals I_j are separated from each other by $2^{n+1} - 2$ open intervals, all of length r_0. Furthermore, we have a gap of length r_0 from 0 to the first interval I_1, and another gap from the last interval to the coordinate 1. We shall in the following show that the length of this gap is $\geq r_0$. When we also consider these two intervals as gaps, we have in total 2^{n+1} open gaps.

Finally, we define

$$A = A_0 := \bigcup_{k=1}^{2^{n+1}-1} I_j,$$

where it will be convenient in the following to use the index 0.

The lemma will be proved, if we can prove

a) $|A| = \mu$,

b) $A \subseteq [r_0, 1 - r_0]$,

c) For every nonempty subset $E \subseteq A$ we can find $x \in E$ and $r > 0$, such that

$$[x - r, x + r] \subset [0,] \qquad \text{and} \qquad |E \cap [x - r, x + r]| < cr.$$

The proofs go as follows.

a) By a simple calculation,

$$|A| = \sum_{k=1}^{2^{n+1}-1} |I_j| = \sum_{m=0}^{n} \sum_{k=0}^{2^{n-m}-1} \left|I_{2^m(2k+1)}\right| = \sum_{m=0}^{n} 2^{n-m} 2^m \ell = (n+1)2^n \ell = \mu.$$

b) Define B as the union of the set A and all the 2^{n+1} gaps, all of length r_0. Then by (3.27),

$$|B| = |A| + 2^{n+1} \cdot \frac{2\mu}{2^n(n+1)} = \mu + \frac{4\mu}{(n+1)c} < 1,$$

so we conclude that $A \subseteq [r_0, 1 - r_0 0]$.

c) Assume that $E \subseteq A$ satisfies the condition that for every $x \in E$ and every $r > 0$, for which $[x - r, x + r] \subseteq [0, 1]$, we have $|E \cap [x - r, x + r]| \geq c \cdot r$.

We shall prove that this is not possible, so $E = \emptyset$. This will prove the lemma.

We first prove that

$$E \cap \bigcup_k I_{2k+1} = \emptyset.$$

Let $x \in \bigcup_k I_{2k+1}$, and choose $r = r_0 = \frac{2\ell}{c}$. Then by b) the interval $[x - r, x + r]$ is a subset of $[0, 1]$, hence

$$|E \cap [x - r, x + r]| \leq |A_0 \cap [x - r, x + r]| \leq |I_{2k+1}| = \ell = \frac{c}{2} r_0 < c\, r.$$

It follows from the assumptions on E that $x \notin E$. We therefore conclude that

$$E \cap \bigcup_k I_{2k+1} = \emptyset,$$

and thus

$$E \subseteq A_1 := \bigcup_k I_{2k}.$$

We remove $\bigcup_k I_{2k+1}$ from A_1. Then the gaps of A_1 are all of length $2r_0 + \ell > 2r_0$, including the two "gaps" at the endpoints.

Let $x \in \bigcup_k I_{2(2k+1)}$. We choose $r = 2r_0$, which implies that $[x-r, x+r] \subseteq [0,1]$, and that

$$|E \cap [x-r, x+r]| \leq |A_1 \cap [x-r, x+r]| \leq \left|I_{2(2k+1)}\right| = 2\ell = \frac{c}{2} \cdot 2r_0 < c \cdot r.$$

It follows again from the definition of E that $x \notin E$. Hence,

$$E \subseteq A_2 := \bigcup_k I_{2^2(2k+1)}.$$

The gaps of A_2 are all of length $> 2^2 r_0$, so when we repeat the argument above with $r = 2^2 r_0$ we get

$$E \subseteq A_3 := \bigcup_k I_{2^3(2k+1)},$$

etc.. Finally, after a finite number of steps

$$E \subseteq A_{n+1} := \bigcup_k I_{2^{n+1}(2k+1)} = \emptyset,$$

and the lemma is proved. □

Proof of Theorem 3.6. For given $\varepsilon > 0$ we define

$$\varepsilon_n := 2^{-n-2}\varepsilon, \qquad \mu_n := 1 - \varepsilon_n \quad \text{and} \quad c_n := 2^{-n}, \qquad n \in \mathbb{N}_0.$$

Construct A_n as described in Lemma 3.4 with respect to the constants μ_n and c_n. Define

$$A := \bigcup_{n=0}^{+\infty} (n + A_n) \cup (-n - A_n),$$

where e.g. $n + A_n := \{n + x \mid x \in A_n\}$, and similarly. Then

$$\begin{aligned} |\mathbb{R} \setminus A| &= 2\sum_{n=0}^{+\infty} |[n, n+1] \setminus (n + A_n)| = 2\sum_{n=0}^{+\infty} |[0,1] \setminus A_n| \\ &= 2\sum_{n=0}^{+\infty} (1 - \mu_n) = 2\sum_{n=0}^{+\infty} \varepsilon_n = 2\sum_{n=0}^{+\infty} 2^{-n-2}\varepsilon = \varepsilon, \end{aligned}$$

and we have proved that $|\mathbb{R} \setminus A| = \varepsilon$.

Assume that $E \subseteq A$ satisfies (3.24). Using that $c_n = 2^{-n} \to 0$ for $n \to +\infty$, it follows from Lemma 3.4 that whenever $2^{-N} < c$ for some $N \in \mathbb{N}$, then $x \notin E$ for every x, for which $|x| > N$, and we therefore conclude that $E \subseteq [-N, N]$. We noticed previously that no bounded set is homogeneous, so it follows that A does not contain any homogeneous subset, and the theorem is proved. □

3.7.1 A brief review of the Hilbert transform

We sketch for completeness (without proof) some part of the theory of the Hilbert transform, showing why Theorem 3.6 is of interest.

The *Hilbert transform* $\mathfrak{H}$ is defined by

$$\mathfrak{H}f(x) := \frac{1}{\pi} \lim_{\varepsilon \to 0+} \left\{ \int_{-\infty}^{x-\varepsilon} + \int_{x+\varepsilon}^{+\infty} \right\} \frac{f(t)}{x-t}\,\mathrm{d}t = \frac{1}{\pi}\,\mathrm{vp.} \int_{-\infty}^{+\infty} \frac{f(t)}{x-t}\,\mathrm{d}t = \text{“}\frac{1}{\pi}\left(f * \frac{1}{x}\right)(x)\text{”},$$

for all functions f, for which $\mathfrak{H}f(x)$ is defined almost everywhere in $\mathbb{R}$. Here, “vp.” means *“valeur principale”* or in English, “principal value”. It is possible to prove that $\mathfrak{H}$ can be extended to a continuous linear map

$$\mathfrak{H} : L^2(\mathbb{R}) \to L^2(\mathbb{R}),$$

such that

$$\|\mathfrak{H}f\|_2 = \|f\|_2, \qquad \text{i.e.} \qquad \|\mathfrak{H}\| = 1,$$

and

$$\text{(3.28)} \qquad \mathfrak{H}^2 f(x) = -f(x), \qquad \text{i.e.} \qquad \mathfrak{H}^2 = -I,$$

where I denotes the identical mapping, $If(x) := f(x)$ for all f and x.

For comparison we notice the similarity with the *Fourier transform*

$$\mathfrak{F}f(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ixt} f(t)\,\mathrm{d}t, \qquad \text{for } x \in \mathbb{R},$$

which also can be extended to a continuous linear map

$$\mathfrak{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R}),$$

for which

$$\|\mathfrak{F}f\|_2 = \|f\|_2, \qquad \text{i.e.} \qquad \|\mathfrak{F}\| = 1,$$

and

$$\mathfrak{F}^2 f(x) = f(-x), \qquad \text{thus } \mathfrak{F}^4 = I.$$

Both operators on $L^2(\mathbb{R})$ satisfy the operator equation $T^4 = I$. Obviously, $\mathfrak{F}^2 \neq \mathfrak{H}^2$.

We finally define the *Hardy-Littlewood maximum operator* $\mathfrak{M}$ by

$$\mathfrak{M}f(x) := \sup_{r>0} \frac{1}{2r} \int_{x-r}^{x+r} |f(t)|\,\mathrm{d}t,$$

for every locally integrable function f, defined on $\mathbb{R}$. We see that $\mathfrak{M}f(x)$ is the supremum of all means of $|f|$ over all symmetric intervals with x as its middle point. In some sense, the operator $\mathfrak{M}$ "smoothes $|f|$", so $\mathfrak{M}f(x)$ gets more "nice" properties than the function $f(x)$ itself. In particular, if f is not identical 0, then $\mathfrak{M}f(x) \geq 0$ for all $x \in \mathbb{R}$.

In spite of the similarities above the Hilbert transform of a function may still have some "wild" geometric properties. This is one of the reasons why one becomes interested in the homogeneous sets. We shall shortly explain why.

The main result of *Kahanpää and Mejlbro* [29] is that a set A is homogeneous, if and only if for every $p > 1$ there exists a constant $C_p > 0$, such that for every $f \in L^p(\mathbb{R})$,

$$\int_A |\mathfrak{H}f|^p \,\mathrm{d}x \leq C_p \int_A |\mathfrak{M}f|^p \,\mathrm{d}x.$$

This result tells us that we can control the Hilbert transform of a function $f \in L^p(\mathbb{R})$ by means of the maximum operator over every *homogeneous* set. Furthermore, this is not possible for all f, if A is not homogeneous.

The surprising consequence of Theorem 3.6 is therefore, that for every $\varepsilon > 0$ we may remove a set of measure ε from $\mathbb{R}$ (to get A, so $|\mathbb{R} \setminus A| = \varepsilon$), such that the Hilbert transform cannot be controlled by the maximum operator over the remainder set A, even if this is "almost all" of $\mathbb{R}$.

Notice the similarity between (3.28), i.e. $\mathfrak{H}^2 = -I$, and the familiar complex equation $i^2 = -1$. It is, in fact, possible to consider $\mathfrak{H}$ as an "imaginary unit" in the vector space of operators generated by I and $\mathfrak{H}$. It is routine to transfer the theorems from *Complex Functions Theory* to this linear space of operators, generated by I and $\mathfrak{H}$. We shall here not go further into this extension.

3.8 A metric space (M,d) and two different probability measures μ and ν on M, which agree on every ball in M

At a conference in 1968 *Flemming Topsøe* posed the following question:

> *"Is a measure in a metric space uniquely determined by its values on all balls in the space?"*

The surprising negative answer was given three years later by *R. O. Davies* [18] in 1971. In 1982 *R. O. Davies* [19] gave a less complicated proof, which we shall follow here.

It suffices to prove

Lemma 3.5 *There exist a metric space (M,d) of diameter 1 and two positive measures μ and ν on M, such that*

$$\mu(M) = \frac{2}{3} \qquad \textit{and} \qquad \nu(M) = \frac{1}{3}, \tag{3.29}$$

and

$$\mu(B) = \nu(B) \tag{3.30}$$

for every ball B in M of radius < 1.

Let us for the time being assume that this lemma holds. Then we construct in the same way another metric space $\left(\hat{M},\hat{d}\right)$ of diameter 1 and two positive measures $\hat{\mu}$ and $\hat{\nu}$ on $\hat{M}$, such that

$$\hat{\mu}\left(\hat{M}\right) = \frac{1}{3} \qquad \text{and} \qquad \hat{\nu}\left(\hat{M}\right) = \frac{2}{3}, \tag{3.31}$$

and

$$\hat{\mu}\left(\hat{B}\right) = \hat{\nu}\left(\hat{B}\right) \tag{3.32}$$

for every ball $\hat{B}$ in $\hat{M}$ of radius < 1.

We define a metric space (M_0,d_0) by putting $M_0 := M \cup \hat{M}$, and

$$d_0(x,y) := \begin{cases} d(x,y), & \text{for } x,\, y \in M, \\[2mm] \hat{d}(x,y), & \text{for } x,\, y \in \hat{M}, \\[2mm] 1, & \text{for } x \in M \text{ and } y \in \hat{M}, \quad \text{or} \quad x \in \hat{M} \text{ and } y \in M. \end{cases}$$

Then $\{x \in M_0 \mid d_0(x,y) = 1\} = M_0$ for every $y \in M_0$, and every ball of radius < 1 is either a ball in M, or in $\hat{M}$, cf. the previous definitions.

Then define the measures

$$\mu_0(A) := \mu(A \cap M) + \hat{\mu}(A \cap \hat{M}) \quad \text{and} \quad \nu_0(A) := \nu(A \cap M) + \hat{\nu}(A \cap \hat{M}).$$

It follows from (3.30) and (3.32) that $\mu_0(B) = \nu_0(B)$ for every ball of radius < 1. When the radius is 1, then M_0 is the only ball, and it follows from (3.29) and (3.31) that

$$\mu_0(M_0) = \mu(M) + \hat{\mu}(\hat{M}) = 1, \quad \text{and} \quad \nu_0(M_0) = \nu(M) + \hat{\nu}(\hat{M}) = 1,$$

and we have proved that μ_0 and ν_0 agree on every ball in M_0. Notice that the measures are probability measures, because they are nonnegative, and each of their total masses is 1.

On the other hand, it follows from (3.29) that

$$\mu_0(M) = \mu(M) = \frac{2}{3} \qquad \text{and} \qquad \nu_0(M) = \nu(M) = \frac{1}{3},$$

so the two measures μ_0 and ν_0 are indeed different. Notice also that by construction M is not a ball in M_0. □

Proof of Lemma 3.5. We shall construct M as a union of k_1 closed "islands", each of diameter 1/2. For the time being we shall only be interested in the distance between points from two different "islands".

Let I_1 and I_2 be two different islands. We define

$$d(x,y) = d(y,x) := d(I_1, I_2) \in \left\{1, \frac{1}{2}\right\} \qquad \text{for all } x \in I_1 \text{ and all } y \in I_2.$$

In other words, the distance between any two point from different islands is put equal to the distance between the two different islands themselves. At this first level this distance can only take the values 1/2 or 1. Another interpretation is that each island is represented by a point in a complete graph $\mathfrak{G}_1$, in which the valence of each edge indicates the distance between the endpoints of the edge between them.

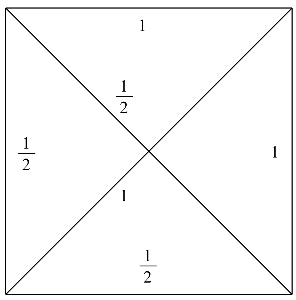

Figure 3.17: *Example of a complete graph $\mathfrak{G}_1$ at level 1, where $k_1 = 4$..*

We let every island (the points of the graph) in the complete graph consist of k_2 smaller closed islands of diameter 1/4. If I_1 and I_2 are two different islands at this level, we put

$$d(x,y) := d(I_1, I_2) \in \left\{\frac{1}{2}, \frac{1}{4}\right\} \qquad \text{for } x \in I_2 \text{ and } y \in I_2.$$

The "internal" distances inside each island at level 1 (i.e. a point in the graph $\mathfrak{G}_1$) is then again represented by a complete graph, consisting og k_2 points, where the valences of each edge can only take the values 1/2 or 1/4, corresponding to the distance between the islands.

We continue in this way by dissolving the islands at a given level into a union of smaller closed islands, the "internal" distances of which are represented by a complete graph. Every island from level $n-1$ consists of k_n smaller closed islands from level n, each of diameter 2^{-n}. If I_1 and I_2 are two different islands from level n, contained in the same island from level $n-1$, then

$$d(x,y) = d(y,x) := d\left(I_1, I_2\right) \in \left\{ \frac{1}{2^{n-1}}, \frac{1}{2^n} \right\} \qquad \text{for } x \in I_1 \text{ and } y \in I_2.$$

We define in this way in total $\prod_{j=1}^{n} k_j$ closed islands at level n.

Every decreasing sequence of islands (from different levels and contained in each other) is assumed to have a non-empty intersection, consisting of just one point. This definition of points implies that the distance between two points is 0, if and only if the points are identical. The other conditions of a metric are trivial, so in this way (M,d) become a metric space. For completeness the triangle inequality follows, when I_1 and I_2 are two different islands from level n contained in the same island from level $n-1$, then $d\left(I_1, I_2\right)$ is either $2 \cdot 2^{-n}$ or 2^{-n}.

The only possible values of the metric are 0 and 2^{-n} for $n \in \mathbb{N}_0$, from which follows in particular that every ball is both open and closed.

The lemma will be proved, if we can specify the sequence (k_n), $k_n \in \mathbb{N}$, and the weight of each island with respect to μ and ν, such that (3.29) and (3.30) are fulfilled.

A ball of radius $1/2$ and centre $x \in M$ is the union of the island I from the first level, for which $x \in I$, and all islands, which have the distance $1/2$ from I. If we consider the graph $\mathfrak{G}_1$ from the first level, then the valences are either $1/2$ or 1. The ball $B[x, 1/2]$ of centre $x \in I$ and radius $1/2$ is therefore I united with every island at this level, for which the connecting edge between them has the valence $1/2$.

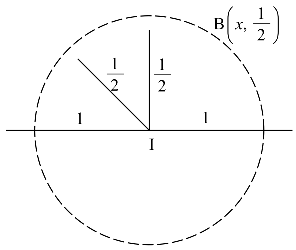

Figure 3.18: *The subgraph $\mathfrak{G}_1$, where we only consider edges connecting I with some other island at level 1. The ball $B(x, \frac{1}{2})$ consists of all points, for which the edge joining it to the island I, have the valence 1/2. If the valence of the edge is 1, then the point is not contained in $B(x, \frac{1}{2})$.*

We can reformulate the problem in the following way: We want to set up a finite graph $\mathfrak{G}_1$ and two different weights at its points (this will define the measures μ and ν on all closed balls of radius $\frac{1}{2}$), such that the total weight of $\mathfrak{G}_1$ with respect to these is $\frac{2}{3}$ and the total weight of $\mathfrak{G}_1$ with respect to the other one is $\frac{1}{3}$, and such that every closed ball of radius $\frac{1}{2}$ gets the same weight, no matter which measure, μ or ν, is chosen.

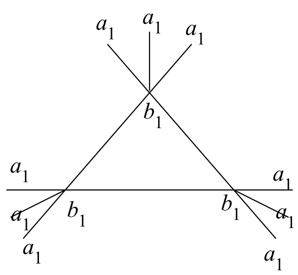

Figure 3.19: *The subgraph of $\mathfrak{G}_1$, where all the edges have the valence 1/2.*

Choose $k_1 = 12$ and fix a subgraph as shown on Figure 3.19 where we only consider edges of valence $1/2$. This means that every other distance is 1.

The triangle of Figure 3.19 is determined by $r_1 = 3$ points, which are called the *interior islands*. Each of them is given the weight b_1. The remaining $r_1^2 = 9$ points are called *exterior islands*. Each of them is given the weight a_1. This determines the μ-measure of the interior and exterior islands at level 1.

In parallel we make the similar construction, where the only difference is that we interchange a_1 and b_1, so the weight a_1 is attached to each of the interior islands, and the weight b_1 is given to each of the exterior islands. In this way we define the μ-measure of all interior and exterior islands at level 1.

Notice that every ball of radius $1/2$ and centre lying in an exterior island has the weight $a_1 + b_1$, no matter which measure, μ or ν, we consider. If instead the centre lies in an interior island, then we obtain the weight $3a_1 + 3b_1$ in both measures. thus, μ and ν agree on every closed ball of radius $1/2$, no matter the choice of a_1 and $b_1 > 0$. The total masses of M in the two measures are

$$\mu(M) = \frac{2}{3} = 9a_1 + 3b_1, \qquad \text{and} \qquad \nu(M) = \frac{1}{3} = 3a_1 + 9b_1,$$

from which follows that we must choose

$$a_1 = \frac{5}{72} \qquad \text{and} \qquad b_1 = \frac{1}{72}.$$

Every closed ball of radius $1/4$ in M is completely contained in one of the islands I from the first level. Here I consists of k_2 islands from the second level. The problem of distributing the masses $\mu(I)$ and $\nu(I)$ among the k_2 islands is completely equivalent to the above with the only modification that the distribution of the mass is now given by either $\frac{5}{72} : \frac{1}{72}$ or $\frac{1}{72} : \frac{5}{72}$, instead of previously by either $\frac{2}{3} : \frac{1}{3}$ or $\frac{1}{3} : \frac{2}{3}$.

We choose a complete graph consisting of r_2 points, i.e. the r_2 interior islands. To every interior island we attach r_2 exterior islands. All the valences in the corresponding subgraph of $\mathfrak{G}_2$ have the valence $1/4$, and the supplementing and remaining edges of $\mathfrak{G}_2$ have all the valence $1/2$. Thus we get $k_2 = r_2\,(r_2 + 1)$ islands from the second level, contained in each island from the first level.

Define the μ-measure of each interior islands as b_2, while every exterior island gets the μ-measure a_2. Concerning the measure ν we swap a_2 and b_2, so the ν-measure of every interior island is a_2, while the ν-measure of every exterior island is b_2. Every ball of radius $1/4$ gets the same weight in both μ and ν, no matter the choice of a_2, $b_2 > 0$. We shall only find the total mass, i.e. we shall find $r_2 \in \mathbb{N}$, such that

$$\text{(3.33)} \qquad r_2^2 a_2 + r_2 b_2 = \frac{5}{12} \qquad \text{and} \qquad r_2 a_2 + r_2^2 b_2 = \frac{1}{72}$$

have positive solutions a_2 and b_2. This is obtained for every $r_2 \geq 6$, $r_2 \in \mathbb{N}$.

In general, when α and β are given, and $r \in \mathbb{N} \setminus \{1\}$ is a parameter, then the system of equations

$$\text{(3.34)} \qquad r^2 a + rb = \alpha \qquad \text{and} \qquad ra + r^2 b = \beta$$

has the solution

$$\text{(3.35)} \qquad a = \frac{r\alpha - \beta}{r\,(r^2 - 1)} \qquad \text{and} \qquad b = \frac{r\beta - \alpha}{r\,(r^2 - 1)}.$$

If we choose

$$r > \max\left\{\frac{\beta}{\alpha}, \frac{\alpha}{\beta}\right\},$$

then $a, b > 0$, in agreement with the requirement $r \geq 6$ in (3.33).

When we continue this process indefinitely, the solution (3.35) of (3.34) guarantees that we can choose r_n and $k_n = r_n(r_n + 1)$, such that the numbers, the positions and the measures of all islands can be successively determined in the desired way, and Lemma 3.5 follows. □

4 Fourier series

4.1 Introduction and basic definitions

4.1.1 Overview of the contents of the chapter

We shall in this chapter give some classical examples from the *Theory of Fourier Series*. In the present section we briefly sketch the most necessary of the theory of Fourier series. In Section 4.2 we discuss the *Gibbs's phenomenon*, which from time to time also occurs in the engineering literature. Then we discuss in Section 4.3 the sequence of functions

$$\left\{\frac{\sin nx}{\pi x}\right\}_{n\in\mathbb{N}},$$

which in some sense to be explicitly explained converges towards *Dirac's delta "function"*.

The next three sections are concerned with pointwise convergence of Fourier series. We construct in Section 4.4 a continuous periodic function, which has a Fourier series which is divergent at some points. We continue in Section 4.5 with *Kolmogorov's* example of a function from $L^1([0,2\pi[)$, the Fourier series of which is pointwise divergent almost everywhere. Finally, we finish the discussion of convergency in Section 4.6 by presenting *Kolmogorov's* example of a function from $L^1([0,2\pi[)$ which has a Fourier series diverging pointwise everywhere.

The most successful application of the theory of Fourier series takes place in $L^2([0,2\pi[)$, which can be considered as a Hilbert space with a countable orthogonal basis. Also, from a physical point of view, a function from $L^2([0,2\pi[)$ represents in some sense an energy, so there is here a strong connection between Physics and Mathematics.

In Section 4.7 we briefly sketch the *Besicovitch space*, which is an example of a *nonseparable Hilbert space*. Such spaces are usually not included in the elementary courses in Mathematics at the universities. The problem with them is that they have "too many" (non-countable) mutually orthonormal unit vectors, However, from a philosophical point of view such spaces might give a far more realistic model of reality than the ordinary Fourier series.

The theory of Fourier series is closely connected with the *Lebesgue integral*. It easily follows from *Hölder's inequality*, cf. Section 2.1, that $L^p([0,2\pi[)$ is contained in $L^1([0,2\pi[)$ for every $p>1$.

4.1.2 Basic definitions

Let $f \in L^1([0,2\pi[)$. Then we define the following constants,

$$(4.1)\qquad \begin{cases} a_n := \dfrac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x, & n\in\mathbb{N}_0,\\[2ex] b_n := \dfrac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)\,\mathrm{d}x, & n\in\mathbb{N},\\[2ex] c_n := \dfrac{1}{2\pi}\int_0^{2\pi} f(x)\,e^{-inx}\,\mathrm{d}x, & n\in\mathbb{Z}, \end{cases}$$

where the condition $f \in L^1([0,2\pi[)$ guarantees that the constants a_n, b_n and c_n all exist. As an example of how to prove this we just mention the following estimate

$$|c_n| \le \frac{1}{2\pi}\int_0^{2\pi} \left|f(x)\,e^{-inx}\right| \mathrm{d}x = \frac{1}{2\pi}\int_0^{2\pi} |f(x)|\,\mathrm{d}x = \frac{1}{2\pi}\|f\|_1 < +\infty.$$

The estimates are similar in the other cases.

The numbers a_n, b_n and c_n are called the *Fourier coefficients* of the function $f \in L^1([0,2\pi[)$, from which they are derived. The series

$$\text{(4.2)} \qquad \begin{cases} \dfrac{1}{2}\,a_0 + \sum_{n=1}^{+\infty}\left\{a_n\cos(nx) + b_n\sin(nx)\right\}, \\ \sum_{n=-\infty}^{+\infty} c_n e^{inx}, \end{cases}$$

are called the (real or complex) *Fourier series* corresponding to the function f. It follows from *Euler's formulæ* that

$$c_n = \frac{1}{2}\left(a_n - ib_n\right), \qquad n \in \mathbb{N},$$

$$c_0 = \frac{1}{2}\,a_0, \qquad n = 0,$$

$$c_{-n} = \frac{1}{2}\left(a_n + ib_n\right), \qquad n \in \mathbb{N},$$

so the two Fourier series of (4.2) are equivalent. When everything is real, one usually prefers the first series, while the second representation is preferred, when we work theoretically in the Hilbert space $L^2([0,2\pi[)$, because then the set $\left\{\frac{1}{\sqrt{2\pi}}\,e^{inx} \mid n \in \mathbb{Z}\right\}$ becomes an orthonormal system with respect to the *complex inner product*

$$\langle f, g\rangle := \int_0^{2\pi} f(x)\,\overline{g(x)}\,\mathrm{d}x \qquad \text{for } f,\, g \in L^2([0,2\pi[).$$

The classical book of the theory of Fourier series was written by *Zygmund* [54]. We shall here follow this book in the introduction of the necessary definitions, needed in this chapter.

Based on (4.1) we define

Definition 4.1 *A real* trigonometric series *is a series of the form*

$$\frac{1}{2}\,a_0 + \sum_{n=1}^{+\infty}\left\{a_n\cos(nx) + b_n\sin(nx)\right\},$$

where x is a real variable, and where a_0, a_1, b_1, $\dots$, a_n, b_n, $\dots$, are real or complex constants.

We mention without proof that there exist convergent trigonometric series, which are not Fourier series for any $f \in L_1([0,2\pi[)$. On the other hand, every Fourier series must by definition be a trigonometric series.

We notice that pointwise convergence is not mentioned at all in Definition 4.1. In fact, pointwise convergence is far from a natural concept for Fourier series. Unfortunately, the elementary teaching of Fourier series has often laid misleadingly too much emphasis on the pointwise convergence. Although one often gets some pointwise results on Fourier series, one should rather think of their convergence as "convergence in energy" of "in L^2".

Definition 4.2 *A (real)* trigonometrical polynomial of order n *is a finite sum of the form*

$$T(x)=\frac{1}{2}\,a_0+\sum_{k=1}^{n}\left\{a_k\cos(kx)+b_k\sin(bx)\right\}.$$

If the coefficients of the highest order terms satisfy $(a_n,b_n)\neq(0,0)$*, then* $T(x)$ *is* strictly *of order* n*.*

In the classical theory of Fourier series we have two important families of trigonometric *kernels* namely the *Dirichlet kernels*, $D_n(t)$, and the *Fejer kernels*, $K_n(t)$. We shall in particular work with the Fejer kernel's, which however, are easily derived from the Dirichlet's kernels. Therefore, we start with the definition of the *Dirichlet's kernels.*

Definition 4.3 *The* Dirichlet's kernels*,* $D_n(t)$*, are given by the following trigonometric polynomials of strict order* $n\in\mathbb{N}_0$*,*

$$D_0(t):=\frac{1}{2}\qquad\textit{and}\qquad D_n(t):=\frac{1}{2}+\sum_{k=1}^{n}\cos(kt),\qquad n\in\mathbb{N}.$$

We derive another and more compact formula for the Dirichlet kernel $D_n(t)$. We use the following familiar antilogarithmic formula

$$2\sin u\cdot\cos v=\sin(u+v)-\sin(u-v),$$

and $\sin\frac{t}{2}\neq 0$ for $0<t<2\pi$. If $n\in\mathbb{N}$ and $t\in\,]0,2\pi[$, then

$$\begin{aligned}
D_n(t)&=\frac{1}{2}+\frac{1}{2\sin\frac{t}{2}}\sum_{k=1}^{n}2\sin\frac{t}{2}\,\cos kt\\
&=\frac{1}{2}+\frac{1}{2\sin\frac{t}{2}}\sum_{k=1}^{n}\left\{\sin\left(kt+\frac{t}{2}\right)-\sin\left((k-1)t+\frac{t}{2}\right)\right\}\\
&=\frac{1}{2}+\frac{1}{2\sin\frac{t}{2}}\left\{\sin\left(n+\frac{1}{2}\right)t-\sin\frac{t}{2}\right\}=\frac{\sin\left(n+\frac{1}{2}\right)t}{2\sin\frac{t}{2}},
\end{aligned}$$

and we have proved that for $t\in\,]0,2\pi[$, $p\in\mathbb{Z}$

$$D_n(t):=\frac{1}{2}+\sum_{k=1}^{n}\cos kt=\frac{\sin\left(n+\frac{1}{2}\right)t}{2\sin\frac{t}{2}},\qquad\text{for } n\in\mathbb{N}_0. \tag{4.3}$$

When $t = 2p\pi$, $p \in \mathbb{Z}$, we get

$$D_n(2p\pi) = n + \frac{1}{2},$$

which is also obtained by taking the limit in (4.3).

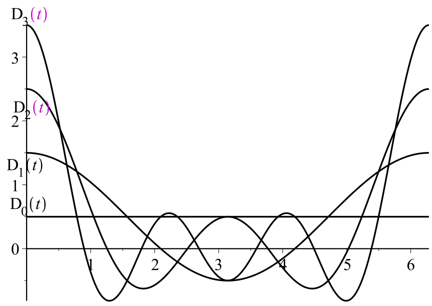

Figure 4.1: *The graphs of the Dirichlet kernels* $D_0(1)$, $D_1(t)$, $D_2(t)$, $D_3(t)$.

Let

$$\frac{1}{2}\,a_0+\sum_{n=1}^{+\infty}\{a_n\cos nx+b_n\sin nx\}$$

be a Fourier series of some $f \in L^1([0,2\pi[)$. We define the *n-th section* $S_n(x;f)$ of the Fourier series by

$$S_n(x;f) := \frac{1}{2}\,a_0+\sum_{k=1}^{n}\{a_k\cos kx+b_k\sin kx\}\,.$$

Then notice that the *"convolution"* of f and D_n over the interval $[0,2\pi[$ is

$$\begin{aligned}\frac{1}{\pi}\int_0^{2\pi} f(t)D_n(x-t)\,\mathrm{d}t &= \frac{1}{2\pi}\int_0^{2\pi} f(t)\cos(0(x-t))\,\mathrm{d}t+\sum_{k=1}^{n}\frac{1}{\pi}\int_0^{2\pi} f(t)\cos(kx-kt)\,\mathrm{d}t\\ &= \frac{1}{2}\,a_0+\sum_{k=1}^{n}\left\{\frac{1}{\pi}\int_0^{2\pi} f(t)\cos kt\,\mathrm{d}t\cdot\cos kx+\frac{1}{\pi}\int_0^{2\pi} f(t)\sin kt\,\mathrm{d}t\cdot\sin kx\right\}\\ &= \frac{1}{2}\,a_0+\sum_{k=1}^{n}\{a_k\cos kx+b_k\sin kx\}=S_n(x;f).\end{aligned}$$

Accordingly, the n-th section of the Fourier series can also be written

$$(4.4)\qquad S_n(x;f)=\frac{1}{\pi}\int_0^{2\pi} f(t)D_n(x-t)\,\mathrm{d}t.$$

If we in particular choose $f \equiv 1$, then $S_n(x;1)=1$ for all n and x, and it follows from (4.4) that

$$(4.5)\qquad \frac{1}{\pi}\int_0^{2\pi} D_n(x-t)\,\mathrm{d}t \equiv 1, \qquad \text{for all } n\in\mathbb{N}_0,$$

which of course also can be derived directly from Definition 4.3.

We mention without proof that if f is periodic of period 2π, and f is continuous and piecewise C^1 in $\mathbb{R}$, then we actually have the following *pointwise* result,

$$f(x)=\lim_{n\to+\infty} S_n(x;f),$$

We also mention that if instead the periodic function f is only piecewise continuous and piecewise C^1, then

$$\lim_{n\to+\infty} S_n(x;f)=\frac{f(x+)+f(x-)}{2}.$$

Therefore, for such functions f the sequence of sections $\{S_n(x;f)\}$ is *pointwise convergent* everywhere with *almost everywhere* $f(x)$ as its limit. This is the reason, why one usually adjusts a function $f(x)$ at every point of discontinuity by the value $\frac{1}{2}\{f(x+)+f(x-)\}$.

Definition 4.4 Fejer's kernel $K_n(t)$ *is the trigonometric polynomial strictly of order* $n \in \mathbb{N}_0$, *which is defined as the mean of the first* $n+1$ *Dirichlet kernels,*

$$K_n(t) := \frac{1}{n+1}\sum_{k=0}^{n} D_k(t) = \frac{1}{n+1}\sum_{k=0}^{n} \frac{\sin\left(k+\frac{1}{2}\right)t}{2\sin\frac{t}{2}}.$$

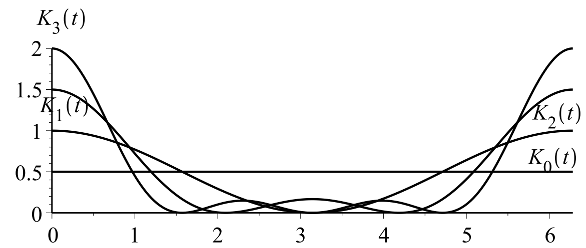

Figure 4.2: *The graphs of Fejer's kernels* $K_0(1)$, $K_1(t)$, $K_2(t)$, $K_3(t)$.

First notice that

$$\begin{aligned} K_n(t) &= \frac{1}{n+1}\sum_{k=0}^{n}\left\{\frac{1}{2}+\sum_{j=1}^{n}\cos jt\right\} = \frac{1}{2}+\frac{1}{n+1}\sum_{k=1}^{n}\sum_{j=1}^{k}\cos jt \\ &= \frac{1}{2}+\frac{1}{n+1}\sum_{j=1}^{n}\cos jt\left\{\sum_{k=j}^{n}1\right\} = \frac{1}{2}+\sum_{j=1}^{n}\left\{1-\frac{j}{n+1}\right\}\cos jt. \end{aligned}$$

Then apply the antilogarithmic trigonometric formula

$$2\sin u\sin v = \cos(v-u)-\cos(v+u)$$

for $t \in]0, 2\pi[$ to get

$$\begin{aligned}
K_n(t) &= \frac{1}{n+1}\sum_{k=0}^{n}\frac{\sin\left(k+\frac{1}{2}\right)t}{2\sin\frac{t}{2}} = \frac{1}{n+1}\sum_{k=0}^{n}\frac{2\sin\frac{t}{2}\sin\left(k+\frac{1}{2}\right)t}{2\sin\frac{t}{2}\cdot 2\sin\frac{t}{2}} \\
&= \frac{1}{n+1}\frac{1}{\left(2\sin\frac{t}{2}\right)^2}\sum_{k=0}^{n}\left\{\cos\left(\left(k+\frac{1}{2}\right)t-\frac{t}{2}\right)-\cos\left(\left(k+\frac{1}{2}\right)t+\frac{t}{2}\right)\right\} \\
&= \frac{1}{n+1}\frac{1}{\left(2\sin\frac{t}{2}\right)^2}\sum_{k=0}^{n}\{\cos kt-\cos(k+1)t\} \\
&= \frac{1}{n+1}\cdot\frac{1-\cos(n+1)t}{\left(2\sin\frac{t}{2}\right)^2} = \frac{2}{n+1}\left\{\frac{\sin\frac{n+1}{2}t}{2\sin\frac{t}{2}}\right\}^2.
\end{aligned}$$

If $t = 2p\pi$, $p \in \mathbb{Z}$, where the denominator above is 0, we get by e.g. taking the limit,

$$K_n(2p\pi) = \lim_{x\to 2p\pi}\frac{2}{n+1}\left\{\frac{\sin\frac{n+1}{2}t}{2\sin\frac{t}{2}}\right\}^2 = \frac{2}{n+1}\left\{\frac{n+1}{2}\right\}^2 = \frac{n+1}{2}.$$

We have proved with the same interpretations as in (4.3),

$$(4.6)\qquad K_n(t) := \frac{1}{n+1}\sum_{k=0}^{n}D_k(t) = \frac{1}{n+1}\frac{1-\cos(n+1)t}{\left(2\sin\frac{t}{2}\right)^2} = \frac{2}{n+1}\left\{\frac{\sin\frac{n+1}{2}t}{2\sin\frac{t}{2}}\right\}^2.$$

It follows from (4.6) that $K_n(t) \geq 0$ for every $t \in \mathbb{R}$. Furthermore, (4.5) implies that

$$(4.7)\qquad \frac{1}{\pi}\int_0^{2\pi}K_n(t)\,\mathrm{d}t = \frac{1}{n+1}\sum_{k=0}^{n}\frac{1}{\pi}\int_0^{2\pi}D_k(t)\,\mathrm{d}t = 1.$$

Finally, a trivial consideration of the graphs on Figure 4.3 shows that

$$(4.8)\qquad \frac{2}{\pi}t \leq \sin t \leq t, \qquad \text{for all } t \in \left[0, \frac{\pi}{2}\right].$$

We therefore get the following pointwise estimate for $t \in]0, \pi]$ and $n \in \mathbb{N}$,

$$(4.9)\qquad 0 \leq K_n(t) = \frac{2}{n+1}\left\{\frac{\sin\frac{n+1}{2}t}{2\sin\frac{t}{2}}\right\}^2 \leq \frac{2}{n+1}\left\{\frac{1}{2\cdot\frac{2}{\pi}\cdot\frac{t}{2}}\right\}^2 \leq \frac{2}{2\pi t^2},$$

so in the O-notation, $K_n(t) = O\left(\frac{1}{nt^2}\right)$. In particular, $K_n(t) \to 0$ for $n \to +\infty$ for every fixed $t \in]0, \pi]$.

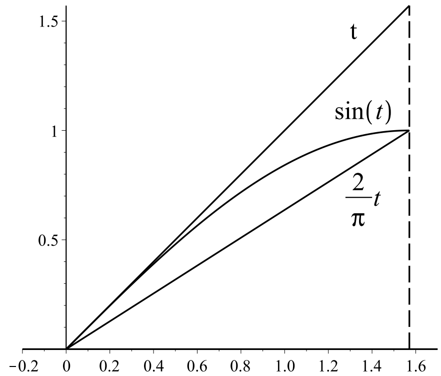

Figure 4.3: *The graphical estimate of* $\sin t$ *for* $t \in \left[0, \frac{\pi}{2}\right]$.

4.2 Gibbs's phenomenon

Gibbs's phenomenon describes an "overshoot" of the finite sections of a Fourier series in the neighbourhood of a point of discontinuity of some function. To see what is meant by this, let us compute the Fourier series of the discontinuous periodic function (period 2π), which is specified by its definition in the half open interval $[0, 2\pi[$ by

$$f(x) := \begin{cases} \pi - x & \text{for } 0 < x < 2\pi, \\[2mm] 0 & \text{for } x = 0, \end{cases}$$

continued by periodicity.

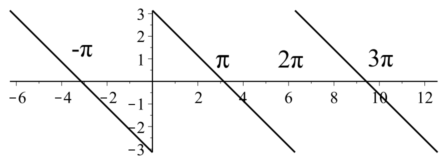

Figure 4.4: *The graph of the periodic function* $f(x) = \pi - x$ *for* $x \in\ ]0, 2\pi[$.

The function f is an odd function, thus $a_n = 0$ for all $n \in \mathbb{N}_0$. The coefficients b_n, $n \in \mathbb{N}$, are easily found,

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\,\mathrm{d}x = \frac{1}{\pi}\left[-(n-x)\cdot\frac{\cos nx}{n}\right]_0^{2\pi} - \frac{1}{\pi}\int_0^{2\pi}\frac{\cos nx}{n}\,\mathrm{d}x = \frac{2}{n},$$

so the Fourier series corresponding to f is given by

$$f \sim \sum_{n=1}^{+\infty} \frac{2\sin nx}{n}.$$

The function f is of class C^∞, except at the points of discontinuity, $2p\pi$, $p \in \mathbb{Z}$, so from a previous remark we conclude that the sum of the Fourier series is everywhere given by $\frac{1}{2}\{f(x+)+f(x-)\} = f(x)$, because $f(x)$ is already adjusted. Hence, we have pointwise

$$(4.10) \qquad f(x) = \sum_{n=1}^{+\infty} \frac{2\sin nx}{n} \qquad \text{for all } x \in \mathbb{R}.$$

Even though the Fourier series is pointwise convergent everywhere with the right sum function, we shall nevertheless take a closer look on how this convergence of (4.10) takes place. Let us consider the sequence of sections of the Fourier series,

$$S_n(x) := S_n(x; f) = 2\sum_{k=1}^{n} \frac{\sin kx}{k}, \qquad \text{for } n \in \mathbb{N}.$$

All these sections $S_n(x)$ are continuous, while the sum function $f(x)$ is discontinuous at all points $2p\pi$, $p \in \mathbb{Z}$. We therefore conclude that the sequence of sections, $\{S_n(x)\}_{n\in\mathbb{N}}$ does *not* converge *uniformly* towards f, because this would imply that f was also continuous.

We say that *Gibbs's phenomenon* describes how much $S_n(x)$ differs from $f(x)$ in the uniform norm, $\|\cdot\|_\infty$. We shall find the extrema of the difference function $S_n(x) - f(x)$ in the interval $]0, 2\pi[$. Using (4.3) in the latter two equalities below we find by termwise differentiation that

$$S_n'(x) - f'(x) = \frac{d}{dx}\left\{2\sum_{k=1}^{n} \frac{\sin kx}{k} - (\pi - x)\right\} = 2\sum_{k=1}^{n} \cos kx + 1 = 2D_n(x) = \frac{\sin\left(n + \frac{1}{2}\right)x}{\sin \frac{x}{2}}.$$

In the interval $]0, 2\pi[$ this expression is 0, when

$$x = \frac{2p\pi}{2n+1}, \qquad \text{for } p = 1, 2, \dots, 2n.$$

We choose the smallest one of these, $x = \frac{2\pi}{2n+1}$. Then we get by insertion

$$S_n\left(\frac{2\pi}{2n+1}\right) = 2\sum_{k=1}^{n} \frac{1}{k} \cdot \sin\left(k \cdot \frac{2\pi}{2n+1}\right) = 2\sum_{k=1}^{n} \frac{2\pi}{2n+1} \cdot \frac{\sin\left(k \cdot \frac{2\pi}{2n+1}\right)}{k \cdot \frac{2\pi}{2n+1}}.$$

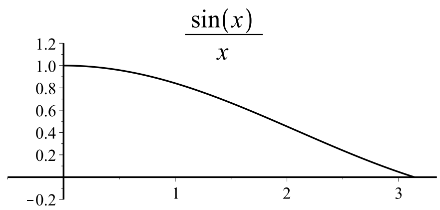

Figure 4.5: *The graph of* $\dfrac{\sin x}{x}$ *for* $x \in [0, \pi]$.

We interpret the latter sum as a *Riemann sum*, which approximates the integral of the continuous function

$$\varphi(x) := \begin{cases} \dfrac{\sin x}{x} & \text{for } x \neq 0, \\ 1 & \text{for } x = 0, \end{cases}$$

over the interval $[0, \pi]$, where each of the subintervals

$$\left[(k-1) \cdot \frac{2\pi}{2n+1}, k \cdot \frac{2\pi}{2n+1}\right], \qquad k = 1, 2, \dots, n,$$

has the length $\frac{2\pi}{2n+1}$, and where we choose the value $\varphi(\pi) = 0$ over the last subinterval $\left[\frac{2n\pi}{2n+1}, \pi\right]$. By taking the limit $n \to +\infty$,

$$\text{(4.11)} \qquad \lim_{n \to +\infty} S_n\left(\frac{2\pi}{2n+1}\right) = \lim_{n \to +\infty} \sum_{k=1}^{n} \frac{2\pi}{2n+1} \frac{\sin\left(k \cdot \dfrac{2\pi}{2n+1}\right)}{k \cdot \dfrac{2\pi}{2n+1}} = 2 \int_0^{\pi} \frac{\sin x}{x} \, \mathrm{d}x.$$

We define

$$c_n := \int_{n\pi}^{(n+1)\pi} \mathrm{d}x = (-1)^n |c_n| \qquad \text{for } n \in \mathbb{N}_0,$$

so (c_n) is alternating, and $0 < |c_{n+1}| < |c_n|$ for all $n \in \mathbb{N}_0$, and

$$|c_n| = \int_{n\pi}^{(n+1)\pi} \left|\frac{\sin x}{x}\right| \mathrm{d}x \leq \int_{n\pi}^{(n+1)\pi} \frac{1}{x} \, \mathrm{d}x = \ln\left(1 + \frac{1}{n}\right),$$

and $(|c_n|) \searrow 0$ decreasingly.

Summing up, the series $\sum_{n=0}^{+\infty} c_n = \sum_{n=0}^{+\infty} (-1)^k |c_n|$ is alternating, and $(|c_n|) \searrow 0$, hence the *improper Riemann integral* exists,

$$\int_0^{+\infty} \frac{\sin x}{x} \, \mathrm{d}x = \lim_{p \to +\infty} \int_0^{(p+1)\pi} \frac{\sin x}{x} \, \mathrm{d}x = \lim_{p \to +\infty} \sum_{n=0}^{p} c_n.$$

Remark 4.1 Notice that the function $\frac{\sin x}{x}$ is *not* Lebesgue integrable over $\mathbb{R}_+$. One easily proves (left to the reader) that

$$\int_0^{+\infty} \left|\frac{\sin x}{x}\right| \mathrm{d}x = +\infty.$$

Therefore, we here use the improper Riemann integral instead. $\Diamond$

Using that $|c_{2n-1}| - |c_{2n}| > 0$, we get

$$\int_0^{+\infty} \frac{\sin x}{x} \, \mathrm{d}x = c_0 - (|c_1| - |c_2|) - (|c_3| - |c_4|) - \cdots < c_0 = \int_0^{\pi} \frac{\sin x}{x} \, \mathrm{d}x,$$

so

$$\int_0^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x < c_0 = \int_0^{\pi} \frac{\sin x}{x}\,\mathrm{d}x.$$

It is well-known – or easily shown by e.g. residuum calculus from Complex Functions Theory – that the value of the improper Riemann integral is

$$\int_0^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}.$$

When we return to (4.11) we finally get the estimate

$$\lim_{n\to+\infty} S_n\left(\frac{2\pi}{2n+1}\right) = 2\int_0^{\pi} \frac{\sin x}{x}\,\mathrm{d}x > 2\int_0^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x = \pi.$$

Hence, there is an $N \in \mathbb{N}$, such that the *section* $S_n(x)$ for all $n \geq N$ is $> \pi$ at some points (an "overshoot"), while obviously $\sup f(x) = \pi$. This is, what we call *Gibbs's phenomenon.*

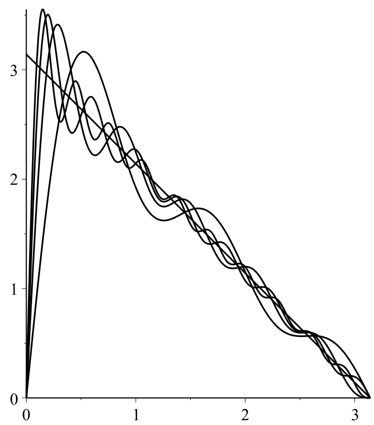

Figure 4.6: *Gibbs's phenomenon. The function* $\pi - x$ *and the sections* $S_5(x)$, $S_{10}(x)$, $S_{15}(x)$ *and* $S_{20}(x)$.

A numerical integration in e.g. MAPLE gives

$$2\int_0^{\pi} \frac{\sin x}{x}\,\mathrm{d}x \approx 3.7039, \qquad \text{and} \qquad \frac{2}{\pi}\int_0^{\pi} \frac{\sin x}{x}\,\mathrm{d}x \approx 1.1790,$$

so for large n the section $S_n(x)$ overshoots the function at some points with approximately $18/2 = 9$ %, because the 18 % describes the relative difference between the undershoot and the nearby overshoot.

Remark 4.2 According to Wikipedia *J. Willard Gibbs* [23], [24] only rediscovered this result in 1899, It was first found – and published in 1848 - by *Henry Wilbraham* [53]. Before it was understood in Mathematics it did some fuss in Physics. In fact, the phenomenon was also observed by experimental physicists, like e.g. *Albert A. Michelson* [40], who believed that it was due to some imperfections in his measuring apparatuses. Well, it was not! This is again a case where one should not have jumped to conclusions before the theory was fully understood. $\Diamond$

4.3 The sequence of functions $\left(\dfrac{\sin nx}{\pi x}\right)_{n\in\mathbb{N}}$ converges towards Dirac's delta "function" δ for $n \to +\infty$.

The unprecise headline is often met in the literature without any hint of what is meant by it. The type of convergence is not specified. In this section we make it precise by proving the following theorem.

Theorem 4.1 *We consider functions $f \in L^2(\mathbb{R})$, i.e.*

$$\|f\|_2^2 = \int_{-\infty}^{+\infty} |f(x)|^2\,dx < +\infty, \tag{4.12}$$

where we furthermore assume that f is also continuous *at 0. Then*

$$\lim_{n\to+\infty} \int_{-\infty}^{+\infty} \frac{\sin nx}{\pi x}\,f(x)\,dx = f(0) \quad [= \delta(f)]. \tag{4.13}$$

Formula (4.13) is one way to express the incorrect way of writing

$$\text{“}\int_{-\infty}^{+\infty} f(x)\,\delta(x)\,\mathrm{d}x = f(0)\text{”},$$

which is often seen in the literature. The problem with this notation is that Dirac's $\delta(x)$ is not a function, and we may easily get into trouble, if we are not very careful. Therefore, (4.13) is a better alternative, because it does not have the same flaws as the integral with "$\delta(x)$". There are other alternatives, based on e.g. sequences of probability functions, the mass of which in the limit are squeezed into $x = 0$.

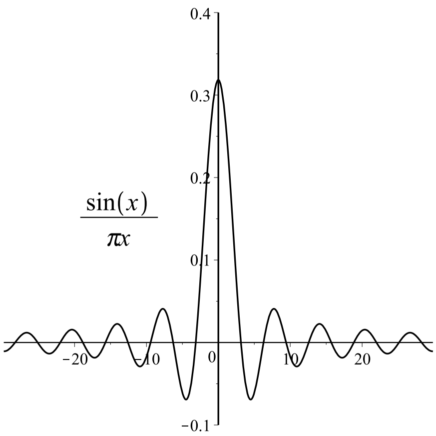

Figure 4.7: *The graph of the function $\varphi(x) = \dfrac{\sin x}{\pi x}$. (Not to scale.)*

Theorem 4.1 is not an easy one, and we shall in the proof take some advanced results on e.g. the Fourier transform for granted.

PROOF. The basic function is of course

$$\varphi(x)=\begin{cases}\dfrac{\sin x}{\pi x} & \text{for } x\in\mathbb{R}\setminus\{0\},\\[2ex] \dfrac{1}{\pi} & \text{for } x=0,\end{cases}\tag{4.14}$$

which is continuous and belongs to $L^p(\mathbb{R})$ for all $p>1$.

As already mentioned in Section 4.2 it does not belong to $L^1(\mathbb{R})$, because φ is not absolutely integrable over $\mathbb{R}$. We mention for completeness that

$$\begin{cases}\int_{-\infty}^{+\infty}|\varphi(x)|\,\mathrm{d}x=\frac{1}{\pi}\int_{-\infty}^{+\infty}\left|\frac{\sin x}{x}\right|\,\mathrm{d}x=+\infty & \text{for } p=1,\\[2ex] \int_{-\infty}^{+\infty}|\varphi(x)|^p\,\mathrm{d}x=\frac{1}{\pi}\int_{-\infty}^{+\infty}\left|\frac{\sin x}{x}\right|^p\,\mathrm{d}x<+\infty. & \end{cases}\tag{4.15}$$

In fact, if $p>1$, then

$$\begin{aligned}\int_{-\infty}^{+\infty}\left|\frac{\sin x}{x}\right|^p\mathrm{d}x &= \int_{-\pi}^{\pi}\left|\frac{\sin x}{x}\right|^p\mathrm{d}x+2\int_{\pi}^{+\infty}\left|\frac{\sin x}{x}\right|^p\mathrm{d}x\\ &\le \int_{-\pi}^{\pi}1\,\mathrm{d}x+2\int_{\pi}^{+\infty}\frac{1}{x^p}\,\mathrm{d}x=2\pi+\frac{1}{p-1}\left(\frac{1}{\pi}\right)^{p-1}<+\infty,\end{aligned}$$

and if $p=1$, then

$$\int_{-\infty}^{+\infty}\left|\frac{\sin x}{x}\right|\mathrm{d}x\ge\sum_{p=0}^{+\infty}\int_{\frac{\pi}{4}+p\pi}^{\frac{3\pi}{4}+p\pi}\frac{|\sin x|}{x}\,\mathrm{d}x\ge\sum_{p=0}^{+\infty}\frac{1}{\sqrt{2}}\cdot\frac{\pi}{2}\cdot\frac{1}{(p+1)\pi}=+\infty.$$

By using either residuum calculus or a partial integration followed by an application of the formula

$$\int_{-\infty}^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x = \pi,$$

we get for $p = 2$ that

$$\int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2}\,\mathrm{d}x = \pi. \tag{4.16}$$

We proved above that φ is not Lebesgue integrable. However, it was proved in Section 4.2 that the improper Riemann integral of φ exists, and it was noticed that using e.g. residuum calculus it could be proved that

$$\int_{-\infty}^{+\infty} \frac{\sin x}{\pi x}\,\mathrm{d}x = \lim_{n\to+\infty} \int_{-n}^{n} \frac{\sin x}{\pi x}\,\mathrm{d}x = 1.$$

We then apply the *Fourier transform* on $L^2(\mathbb{R})$. If $f \in L^2(\mathbb{R})$, then we define its *Fourier transform* $\mathfrak{F}f$ by

$$\mathfrak{F}f(t) := \int_{-\infty}^{+\infty} e^{-itx}\, f(x)\,\mathrm{d}x = \text{l.i.m.}_{n\to+\infty} \int_{-n}^{n} e^{-itx}\, f(x)\,\mathrm{d}x,$$

where l.i.m. is short for "limit in mean", i.e. $L^2(\mathbb{R})$ convergence. It can be proved that also $\mathfrak{F}f \in L^2(\mathbb{R})$, cf. e.g. *Riesz & Sz.-Nagy* [47], or any other standard theoretical books on the Fourier transform.

If we put

$$\langle f, g\rangle := \int_{-\infty}^{+\infty} f(x)\overline{g(x)}\,\mathrm{d}x \qquad \text{for } f,\, g \in L^2(\mathbb{R}),$$

then $\langle\cdot,\cdot\rangle$ is the usual *inner product* in $L^2(\mathbb{R})$. It can be proved that

$$\int_{-\infty}^{+\infty} f(x)\overline{g(x)}\,\mathrm{d}x = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathfrak{F}f(x)\overline{\mathfrak{F}g(x)}\,\mathrm{d}x, \qquad \text{for } f,\, g \in L^2(\mathbb{R}), \tag{4.17}$$

i.e.

$$\langle f, g\rangle = \frac{1}{2\pi}\langle \mathfrak{F}f, \mathfrak{F}g\rangle, \qquad \text{for } f,\, g \in L^2(\mathbb{R}).$$

Choose in particular

$$\psi_n(x) := \chi_{[-n,n]}(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}),$$

i.e. $\psi_n(x) = 1$ for $x \in [-n, n]$, and $\psi_n(x) = 0$ for $|x| > n$. Then

$$\mathfrak{F}\psi_n(t) = \int_{-n}^{n} e^{-itx}\,\mathrm{d}x = -\frac{1}{it}\left\{e^{-itn} - e^{itn}\right\} = 2\,\frac{\sin nt}{t} = 2\pi\varphi_n(t),$$

where we have put

$$\varphi_n(x) := \frac{\sin nx}{\pi x}.$$

Then by *Fourier's inversion formula* in $L^2(\mathbb{R})$,

$$\mathfrak{F}\varphi_n(t) = \psi_n(t) = \chi_{[-n,n]}(t) = \begin{cases} 1 & \text{for } t \in [-n,n] \\ 0 & \text{for } |x| > n. \end{cases}$$

Finally, if $f \in L^2(\mathbb{R})$ is continuous at $x = 0$, then it follows from the above that

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{\sin nx}{\pi x} f(x)\,\mathrm{d}x &= \int_{-\infty}^{+\infty} f(x)\overline{\varphi_n(x)}\,\mathrm{d}x = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathfrak{F}f(t)\overline{\mathfrak{F}\varphi_n(t)}\,\mathrm{d}t \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathfrak{F}f(t)\chi_{[-n,n]}(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_{-n}^{n} \mathfrak{F}f(t)\,\mathrm{d}t \\ &\to \text{l.i.m.}_{n\to+\infty}\frac{1}{2\pi}\int_{-n}^{n} \mathfrak{F}f(t)\,e^{i\cdot 0\cdot t}\,\mathrm{d}t = f(0) \qquad \text{for } n \to +\infty, \end{aligned}$$

and (4.13) – accordingly also the theorem – is proved. □

4.4 A continuous periodic function, where the corresponding Fourier series is divergent at some points

The first example of such a function was given by *du Bois-Reymond* in 1876. The example below is due to Fejér from 1911.

We proved in Section 4.2 that the function

$$g(x) = \begin{cases} \pi - x, & \text{for } x \in]0, 2\pi[, \\ \\ 0, & \text{for } x = 0, \end{cases}$$

continued periodically had the Fourier series

$$g(x) = \sum_{n=1}^{+\infty} \frac{2\sin nx}{n}, \qquad \text{for } x \in \mathbb{R},$$

where the Fourier series is pointwise convergent everywhere with the sum function $g(x)$.

Let

$$S_n(x; g) := \sum_{k=1}^{n} \frac{2\sin kx}{k}, \qquad \text{for } x \in \mathbb{R} \text{ and } n \in \mathbb{N},$$

denote its n-th section. We shall first prove that we can find a constant $C > 0$, such that

$$(4.18) \qquad |S_n(x; g)| \leq C \qquad \text{for all } x \in \mathbb{R} \text{ and all } n \in \mathbb{N}.$$

This was almost proved in Section 4.2. We shall here give a direct proof. Let K_n denote *Fejér's kernel*, introduced in Definition 4.4. We define

$$(4.19) \qquad s_n(x) := \frac{1}{\pi} \int_0^{2\pi} g(t) K_n(x - t)\, \mathrm{d}t.$$

Then we get by a small calculation

$$\begin{aligned}(4.20) \quad s_n(x) &= \frac{1}{n+1} \sum_{k=0}^{n} S_k(x) = \frac{1}{n+1} \sum_{k=0}^{n} \left\{ \sum_{j=1}^{k} \frac{2\sin jx}{j} \right\} = \frac{1}{n+1} \sum_{j=1}^{n} \frac{2\sin jx}{j} \left\{ \sum_{k=j}^{n} 1 \right\} \\ &= \sum_{j=1}^{n} \frac{n+1-j}{n+1} \cdot \frac{2\sin jx}{j} = \sum_{j=1}^{n} \left\{ 1 - \frac{j}{n+1} \right\} \cdot \frac{2}{j} \sin jx.\end{aligned}$$

Using (4.20, (4.19), (4.6) and (4.7) we get for every $x \in \mathbb{R}$ and every $n \in \mathbb{N}$ the following estimates,

$$\begin{aligned}|S_n(x; g)| &\leq |S_n(x; g) - s_n(x)| + |s_n(x)| \leq \left| \sum_{k=1}^{n} \frac{2}{n+1} \sin kx \right| + \frac{1}{\pi} \int_0^{2\pi} |f(t)| K_n(x - t)\, \mathrm{d}t \\ &\leq \frac{2}{n+1} \cdot n + \pi \cdot \frac{1}{pi} \int_0^{2\pi} K_n(x - t)\, \mathrm{d}t < 2 + \pi,\end{aligned}$$

and we have proved (4.18(with the constant $C = 2 + \pi$.

If p, $n \in \mathbb{N}$ and $x \in \mathbb{R}$, we define

$$
\begin{aligned}
(4.21)\quad f_{p,n}(x) &:= \sin px \cdot S_n(x;g) = \sum_{k=1}^{n} \frac{2\sin px \cdot \sin kx}{k} \\
&= \sum_{k=1}^{n} \frac{\cos(p-k)x - \cos(p+k)x}{k} = \sum_{k=1}^{n} \frac{\cos(p-k)x}{k} - \sum_{k=1}^{n} \frac{\cos(p+k)x}{k}.
\end{aligned}
$$

The trigonometric polynomial $f_{p,n}(x)$ is continuous, and it follows from (4.18) that we have the estimate

$$
(4.22)\qquad |f_{p,n}(x)| = |\sin px| \cdot |S_n(x;g)| < 2 + \pi \qquad \text{for } p,\, n \in \mathbb{N} \text{ and } x \in \mathbb{R}.
$$

On the other hand, the first sum in the latter expression of (4.21) has the value $\sum_{k=1}^{n} 1/k$ for $x = 0$. This sum tends towards $+\infty$ for $n \to +\infty$. We shall exploit this property in Fejér's construction.

Choose a sequence (n_k) of numbers $n_k \in \mathbb{N}$, such that

$$
(4.23)\qquad \sum_{j=1}^{n_k} \frac{1}{j} > 2^{k-1} \qquad \text{for all } k \in \mathbb{N}.
$$

Once the sequence (n_k) has been chosen according to (4.23) we choose another sequence (p_k) of numbers $p_k \in \mathbb{N}$, such that

$$
(4.24)\qquad n_1 < p_1, \text{ and } p_{k-1} + n_{k-1} < p_k - n_k \qquad \text{for all } k \in \mathbb{N} \setminus \{1\}.
$$

The requirement on the p_k can also be written $p_n > p_{k-1} + n_{k-1} + n_k$, and then it is obvious that such sequences (n_k) and (p_k) exist.

Define

$$
(4.25)\qquad f(x) := \sum_{k=1}^{+\infty} \frac{1}{2^{k-1}}\, f_{p_k,n_k}(x).
$$

According to (4.22),

$$
|f(x)| \le \sum_{k=1}^{+\infty} \frac{1}{2^{k-1}}\,(2+\pi) = 2(2+\pi),
$$

and it follows that the series in (4.25) is uniformly convergent. Every trigonometric polynomial $f_{p,n}$ is continuous, so by the uniform convergence f is also continuous. Every $f_{p,n}$ is periodic of period 2π, hence f is also periodic of period 2π.

We insert the expressions of $f_{p,n}$ into (4.25) to get

$$
\begin{aligned}
f(x) \quad = \quad & \left\{\frac{\cos(p_1-n_1)\,x}{n_1}+\cdots+\frac{\cos(p_1-1)\,x}{1}-\frac{\cos(p_1+1)\,x}{1}-\cdots-\frac{\cos(p_1+n_1)\,x}{n_1}\right\} \\
& +\left\{\frac{\cos(p_2-n_2)\,x}{2n_2}+\cdots+\frac{\cos(p_2-1)\,x}{2}-\frac{\cos(p_2+1)\,x}{2}-\cdots-\frac{\cos(p_2+n_2)\,x}{2n_2}\right\} \\
& +\cdots \\
& +\left\{\frac{\cos(p_k-n_k)\,x}{2^{k-1}n_k}+\cdots+\frac{\cos(p_k-1)\,x}{2^{k-1}}-\frac{\cos(p_k+1)\,x}{2^{k-1}}-\cdots-\frac{\cos(p_k+n_k)\,x}{2^{k-1}n_k}\right\} \\
& +\cdots,
\end{aligned}
\tag{4.26}
$$

where by (4.25)

$$0<p_1-n_1<p_1+n_1<p_2-n_2<p_2+n_2<\cdots<p_{k-1}+n_{k-1}<p_k-n_k<\cdots.$$

This means that if a term of the type $a_m \cos mx$ occurs in (4.26), then it only occurs once, so the terms do not interfere.

We shall prove that the Fourier series of f exists and that it is obtained by removing the parentheses in (4.26) followed by an addition of the missing terms as zero terms.

First notice that the series

$$\sum_{k=1}^{+\infty}\frac{1}{2^{k-1}}\,f_{p_k,n_k}(x)\cos nx \qquad \text{and} \qquad \sum_{k=1}^{+\infty}\frac{1}{2^{k-1}}\,f_{p_k,n_k}(x)\sin nx$$

are both uniformly convergent, so the integral and the sum can be interchanged. This gives

$$a_n=\frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,\mathrm{d}x=\sum_{k=1}^{+\infty}\frac{1}{2^{k-1}}\cdot\frac{1}{\pi}\int_0^{2\pi} f_{p_k,n_k}(x)\cos nx\,\mathrm{d}x,$$

and

$$b_n=\frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\,\mathrm{d}x=\sum_{k=1}^{+\infty}\frac{1}{2^{k-1}}\cdot\frac{1}{\pi}\int_0^{2\pi} f_{p_k,n_k}(x)\sin nx\,\mathrm{d}x.$$

If some $a_n \neq 0$, then there exists a $k \in \mathbb{N}$, such that $p_k-n_k \leq n \leq p_k+n_k$ and $n \neq p_k$, in which case a_n is equal to the corresponding element of (4.26).

Summing up, we have constructed a continuous and periodic function f with a Fourier series, which is obtained by removing the parentheses of (4.26). We shall then only show that this Fourier series is divergent at some points. We shall prove this for $x=0$.

We shall show that the sequence of sections $\{S_n(0)\}$ of the Fourier series is divergent at $x=0$. If $n=2n_1+2n_2+\cdots+2n_k$, then (4.26) implies that $s_n(0)=0$. If instead $n=2n_1+2n_2+\cdots+2n_{k-1}+n_k$, then it follows from (4.23) that

$$s_n(0)=0+\cdots+0+\frac{1}{2^{k-1}n_k}+\frac{1}{2^{k-1}(n_k-1)}+\cdots+\frac{1}{2^{k-1}\cdot 1}=\frac{1}{2^{k-1}}\sum_{j=1}^{n_k}\frac{1}{j}>1.$$

Hence, the sequence $\{s_n(0)\}$ is divergent, i.e. the Fourier series is divergent at $x = 0$, and the claim is proved.

Remark 4.3 It was a long time an open question, if it was possible to construct a *continuous* periodic function f, where the corresponding Fourier series is divergent everywhere. This conjecture was partly supported by the example above, and partly by *Kolmogorov's* examples, cf. Section 4.5 and Section 4.6. Kolmogorov's examples do not solve the problem, because the functions are not continuous. However, they are absolutely integrable, $f \in L^1([0, 2\pi[)$, and the condition $\int_0^{2\pi} |f(t)|\,\mathrm{d}t < +\infty$ guarantees that all Fourier coefficients are well-defined.

The problem above was solved by *Carleson* [11] in 1966, and his result was extended in 1967 by *Hunt* [26]. Summing up, they proved the following theorem

Carleson-Hunt's theorem. *For every $p > 1$ and every function $f \in L^p([0, 2\pi[)$ there exists a Lebesgue nullset $A \subset [0, 2\pi[$, such that*

$$\lim_{n \to +\infty} S_n(x; f) = f(x) \qquad \text{for all } x \in [0, 2\pi[\setminus A.$$

In other words, this theorem assures for $p > 1$ that every $f \in L^p([0, 2\pi[)$ is pointwise convergent almost everywhere. We notice that the function f given by (4.25) is continuous and periodic, so its restriction to the interval $[0, 2\pi[$ belongs to everyone of the spaces $L^p([0, 2\pi[)$, $p \geq 1$. Then it follows from *Carleson-Hunt's theorem* that its Fourier series is pointwise convergent almost everywhere. It was shear luck that the choice $x = 0$ actually belongs to the exceptional nullset. $\diamondsuit$

We notice that Kolmogorov's two functions defined in Section 4.5 and Section 4.6 necessarily must belong to the set

$$L^1([0,2\pi[) \setminus \bigcup_{p>1} L^p([0,2\pi[).$$

The interested reader may find a proof of *Carleson-Hunt's theorem* in *Jørsboe & Mejlbro* [28]. This is not an easy proof either (121 pages), and even today I feel that the result in some way is not "natural". In all useful applications in the physical or technical sciences the Fourier series are convergent in $L^2([0,2\pi[)$, i.e. in *limit in mean, l.i.m.* or in *energy.* It is in such a context a very strange request that it should also be pointwise convergent, and even more strange that this is actually true almost everywhere in most cases of practical interest.

It should be added that there exist continuous and periodic function, where *Carleson-Hunt's theorem* applies, and where the Fourier series diverges at all points from a *non-countable set*, which then necessarily is a nullset. And even more general, to every nullset $N \subset [0,2\pi[$ we can find a continuous function, such that its Fourier series is divergent at all points $x \in N$.

4.5 A Fourier series which is divergent almost everywhere

We shall in this section construct a function $f \in L^1([0,2\pi[)$, such that its Fourier series

$$f \sim \frac{1}{2}\,a_0 + \sum_{n=1}^{+\infty} \{a_n \cos nx + b_n \sin nx\}$$

is (pointwise) divergent almost everywhere. We first notice that it follows from the assumption that $f \in L^1([0,2\pi[)$ and an application of *Lebesgue's majorizing theorem* that all the coefficients a_n and b_n do exist. In fact,

$$\left|\frac{1}{\pi}\int_0^{2\pi} f(t)\left\{\begin{array}{c}\sin nt\\ \cos nt\end{array}\right\}\mathrm{d}t\right| \le \frac{1}{\pi}\|f\|_1 < +\infty.$$

Kolmogorov [31] first constructed a function of this type in 1923. Three years later, in 1926, *Kolmogorov* [32] gave the far more difficult construction of a function $f \in L^1([0,2pi[)$ with a Fourier series, which is (pointwise) *divergent* everywhere. When we compare with *Carleson-Hunt's theorem* from Section 4.4, it follows that none of *Kolmogorov's* functions can belong to any $L^p([0,2\pi[)$ for $p > 1$. This observation almost solves the classical problem concerned with pointwise convergence of Fourier series.

We shall in the present construction of such a weird function more or less follow *Zygmund* [54]. As usual, we put all the real difficulties into one preliminary lemma.

Lemma 4.1 *There exists a sequence of nonnegative trigonometric polynomials f_n, all of which have the constant term $\frac{1}{2}\,a_0 = \frac{1}{2}$, satisfying the following;*
To every $m \in \mathbb{N}$ there exist a constant $A_n > 0$ and a measurable subset $E_n \subset [0, 2\pi[$, such that

1) $A_n \to +\infty$ for $n \to +\infty$

2) $|E_n| \to 2\pi$ for $n \to +\infty$, where $|\cdot|$ denotes the Lebesgue measure of measurable sets.

3) For every fixed $n \in \mathbb{N}$ and every $x \in E_n$ there exists a constant $m = m_x \in \mathbb{N}$, such that

$$\left|\frac{1}{2} + \sum_{k=1}^{m} \{a_k(f_n)\cos kx + b_n(f_n)\sin kx\}\right| > A_n,$$

where $a_n(f_n)$ and $b_k(f_n)$ denote the Fourier coefficients of f_n.

PROOF. For any given $n \in \mathbb{N}$ we put

$$c_j = \frac{4\pi j}{2n+1}, \qquad j = 0, 1, \dots, n,$$

and choose $m_1, \dots, m_n \in \mathbb{N}$, such that

$$\text{(4.27)} \qquad \begin{cases} m_1 \geq n^4 \text{ and } m_{j+1} > 2m_j, \\[2mm] 2m_j + 1 = (2n+1)p_j \text{ for some } p_j \in \mathbb{N}. \end{cases}$$

We define the trigonometric polynomial f_n as a mean of some *Fejér kernels*, cf. Definition 4.4,

$$\text{(4.28)} \qquad f_n(x) := \frac{1}{n}\left\{K_{m_1}(x - c_1) + K_{m_2}(x - c_2) + \dots + K_{m_n}(x - c_n)\right\}.$$

It follows from the general property $K_n(t) \geq 0$ that also $f_n(x) \geqq 0$. Each of the functions $K_{m_j}(x - c_j)$ contains the constant term $\frac{1}{2}$. Consequently, $f_n(x)$ has also the constant term $\frac{1}{2}$.

Then we prove (1)–(3) of Lemma 4.1. We start by calculating the m_j-th section $S_{m_j}(x; f_n)$ of the Fourier series of the function f_n given by (4.28). The first j terms of (4.28) are the same. However, when $i \in \{j+1, \dots, n\}$ we must delete all terms of order $> m_j$ in $K_{m_i}(x - c_i)$. Therefore, by the definitions of K_m and D_k from Section 4.1 and a small count,

$$S_{m_j}(x; f_n) = \frac{1}{n}\sum_{i=1}^{j} K_{m_i}(x - c_i) + \frac{1}{n}\sum_{i=j+1}^{n}\left\{\frac{1}{2} + \sum_{p=1}^{m_j} \frac{m_i - p + 1}{m_i + 1}\cos p\,(x - c_i)\right\}.$$

It follows from

$$K_{m_j}(t) = \frac{1}{m_j + 1}\sum_{k=0}^{m_j} D_k(t) = \frac{1}{2} + \sum_{p=1}^{m_j} \frac{m_i - p + 1}{m_i + 1}\cos pt,$$

and

$$m_i - p + 1 = (m_i - m_j) + (m_j - p + 1),$$

that

$$S_{m_j}(x;f_n) = \frac{1}{n}\sum_{i=1}^{j} K_{m_i}(x-c_i) + \frac{1}{n}\sum_{i=j+1}^{m_j} \frac{m_j-p+1}{m_i+1}\cos p\,(x-c_i) + \sum_{p=1}^{m_j}\frac{m_i-m_j}{m_i+1}\cos p\,(x-c_i)$$

$$= \frac{1}{n}\sum_{i=1}^{j} K_{m_i}(x-c_i) + \frac{1}{n}\sum_{i=j+1}^{n}\left\{\frac{m_j+1}{m_i+1}\left(\frac{1}{2}+\sum_{p=1}^{m_j}\frac{m_j-p+1}{m_j+1}\cos p\,(x-c_i)\right)\right.$$

$$\left. + \frac{m_i-m_j}{m_i+1}\left(\frac{1}{2}+\sum_{p=1}^{m_j}\cos p\,(x-c_i)\right)\right\}$$

$$= \frac{1}{n}\sum_{i=1}^{j} K_{m_i}(x-c_i) + \frac{1}{n}\sum_{i=j+1}^{n}\frac{m_j+1}{m_i+1}K_{m_j}(x-c_i) + \frac{1}{n}\sum_{i=j+1}^{n}\frac{m_i-m_j}{m_i+1}D_{m_j}(x-c_i).$$

We define the intervals A_i and A_i' by

$$A_i = \left]c_{i-1},c_i\right[\text{ and } A_i' = \left]c_i - \frac{1}{n^2}.c_i+\frac{1}{n^2}\right[, \qquad \text{for } i = 1,2,\dots,n.$$

We noticed in (4.9) that $K_m(t) = O\left(\dfrac{1}{mt^2}\right)$, i.e. the expression $mt^2K_m(t)$ is bounded in t. Since $m_i \geq n^4$m it follows that $K_{m_i}(x-c_i)$ is uniformly bounded outside A_i'. Hence, outside the union $\bigcup_{i=1}^n A_i'$ the contribution from the first two sums in the latter expression of $S_{m_j}(x;f_n)$ is numerically smaller than some absolute constant, which we call A, and we get the estimate,

$$S_{m_j}(x;f_n) \geq \frac{1}{n}\left|\sum_{i=j+1}^{n}\frac{m_i-m_j}{m_i+1}D_{m_j}(x-c_i)\right| - A, \qquad \text{for } x \notin A_i' \text{ and } j = 1,2,\dots,n. \tag{4.29}$$

It follows from

$$(2m_j+1)\cdot\frac{1}{2}\cdot c_i = \frac{2m_j+1}{2n+1}\cdot\frac{1}{2}\cdot i\cdot\pi \equiv 0 \pmod{2\pi}$$

that

$$D_{m_j}(x-c_i) = \frac{\sin\left(m_j+\dfrac{1}{2}\right)(x-c_i)}{2\sin\dfrac{1}{2}(x-c_i)} = \frac{\sin\left(m_j+\dfrac{1}{2}\right)x}{2\sin\dfrac{1}{2}(x-c_i)}.$$

Let $[x]$ denote the integer part of $x\in\mathbb{R}$, and let $x \in A_j$, where $1 \leq j \leq [n-\sqrt{n}]$. It follows from $m_{i+1} \geq 2m_i+1$ that the coefficients of D_{m_j} in (4.29) are all at least $\frac{1}{2}$. Furthermore, the denominators, $2\sin\frac{1}{2}(x-c_i)$, have all constant sign for $i > j$. Hence, for n sufficiently large,

$$\frac{1}{n}\left|\sum_{i=j+1}^{n}\frac{m_i-m_j}{m_i+1}D_{m_j}(x-c_i)\right| \geq \frac{1}{2n}\left|\sin\left(m_j+\frac{1}{2}\right)x\right|\sum_{i=j+1}^{n}\frac{1}{c_i-c_{j-1}}$$

$$= \frac{1}{2n}\left|\sin\left(m_j+\frac{1}{2}\right)x\right|\cdot\frac{2n+1}{4\pi}\sum_{k=2}^{n-j+1}\frac{1}{k} \geq \frac{1}{4\pi}\left|\sin\left(m_j+\frac{1}{2}\right)x\right|\cdot\{-1+\ln(n-j)\}$$

$$\geq \frac{1}{4\pi}\left|\sin\left(m_j+\frac{1}{2}\right)x\right|\cdot\left\{-1+\frac{1}{2}\ln n\right\} \geq \frac{1}{9\pi}\left|\sin\left(m_j+\frac{1}{2}\right)x\right|\cdot\ln n.$$

If therefore

$$x \in \left\{ x \in [0, 2\pi[\setminus \bigcup_{i=1}^{n} A_i' \;\middle|\; \left| \sin \left(m_j + \frac{1}{2} \right) x \right| \geq \frac{9\pi}{\sqrt{\ln n}} \right\},$$

then

$$\left| S_{m_j}(x; f_n) \right| \geq \sqrt{\ln n} - a := A_n.$$

The measure of the set

$$\left\{ x \in [0, 2\pi[\;\middle|\; \left| \sin \left(m_j + \frac{1}{2} \right) x \right| < \frac{9\pi}{\sqrt{\ln n}} \right\}$$

is of the size $O\left(\dfrac{1}{\sqrt{\ln n}} \right)$. If we therefore define the set

$$E_n := \left\{ x \in \left[0, c_{[n-\sqrt{n}]} \right[\setminus \bigcup_{i=1}^{n} A_i' \;\middle|\; \left| \sin \left(m_j + \frac{1}{2} \right) x \right| \geq \frac{9\pi}{\sqrt{\ln n}} \right\},$$

then

$$2\pi - |E_n| = O\left(\frac{1}{\sqrt{\ln n}} \right) + O\left(\frac{n}{n^2} \right) + O\left(\frac{\sqrt{n}}{n} \right) = o(1),$$

which means that $2\pi - |E_n| \to 0$ for $n \to +\infty$.

Finally, for every $x \in E_n$ there is a $j = j(x) \in \mathbb{N}$, such that

$$\left| S_{m_j}(x; f_n) \right| \geq \sqrt{\ln n} - A = A_n,$$

and the lemma is proved. □

Theorem 4.2 Kolmogorov [31], 1923. *There exists a function $f \in L^1([0,2\pi[)$, such that its Fourier series*

$$f \sim \frac{1}{2}\,a_0 + \sum_{n=1}^{+\infty} \{a_n \cos nx + b_n \sin nx\}$$

is divergent almost everywhere.

Proof. Let (f_n) be a sequence of nonnegative trigonometric polynomials, all of constant term $\frac{1}{2}$ and satisfying Lemma 4.1. Assume that the constant $m = m_x$ of Lemma 4.1 (3) satisfies the estimates $1 \le m \le \nu_n$, where ν_n is the order of f_n. The sequence (A_n) of Lemma 4.1 (1) tends towards $+\infty$. Thus there exists a sequence (n_k), where $n_k \in \mathbb{N}$, such that

$$\sum_{k=1}^{+\infty} \frac{1}{\sqrt{A_{n_k}}} < +\infty.$$

The adjusted trigonometric polynomial $f_{n_k}(x) - \dfrac{1}{2}$ has the constant term 0. Therefore, we can choose another sequence (q_k) of numbers $q_k \in \mathbb{N}$, which increases so fast that the order of the trigonometric polynomial $f_{n_k}(q_k x) - \frac{1}{2}$ is smaller than the smallest occurring order of the terms in $f_{n_{k+1}}(q_{k+1}x) - \frac{1}{2}$. We may e.g. choose (q_k), such that $q_k \nu_{n_k} < q_{k+1}$. In particular, any term of the type $a \cdot \cos nx + b \cdot \sin nx$, where $(a,b) \neq (0,0)$, is at most a member of one of the $f_{n_k}(q_k x) - \frac{1}{2}$. We say that the terms of the sequence $\left(f_{n_k}(q_k x) - \frac{1}{2}\right)$ do not mutually overlap.

Define

$$f(x) := \sum_{k=1}^{+\infty} \frac{1}{\sqrt{A_{n_k}}} \left\{ f_{n_k}(q_k x) - \frac{1}{2} \right\}.$$

Every f_n is nonnegative and has constant term $\frac{1}{2}$. Since $\int_0^{2\pi} f_n(x)\,\mathrm{d}x = \pi$ for every n, we get the estimate

$$\sum_{k=1}^{+\infty} \frac{1}{\sqrt{A_{n_k}}} \int_0^{2\pi} \left| f_{n_k}(q_k x) - \frac{1}{2} \right| \mathrm{d}x \le \sum_{k=1}^{+\infty} \frac{1}{\sqrt{A_{n_k}}} \int_0^{2\pi} \left\{ f_{n_k}(q_k x) + \frac{1}{2} \right\} \mathrm{d}x = 2\pi \sum_{k=1}^{+\infty} \frac{1}{\sqrt{A_{n_k}}} < +\infty.$$

We therefore conclude that $f \in L^1([0,2\pi[)$.

The Fourier series of f is formally obtained by adding all Fourier series of each term. We replace each term

$$\varphi_k(x) := \frac{1}{\sqrt{A_{n_k}}} \left\{ f_{n_k}(q_k x) - \frac{1}{2} \right\}$$

by its full expression as a finite sum of trigonometric functions.

If we define

$$\mathcal{E}_k := \{ x \in [0,2\pi[\mid q_k x \pmod{2\pi} \ \in E_{n_k} \},$$

then obviously, $|\mathcal{E}_k| = |E_{n_k}|$.

It follows from Lemma 4.1 (3) that whenever $x \in \mathcal{E}_k$, then some partial sum of $\varphi_k(x)$, i.e. some successive terms in a block of the Fourier series of f, in absolute value exceeds

$$\frac{A_{n_k} - \frac{1}{2}}{\sqrt{A_{n_k}}}, \qquad \text{which tends towards } +\infty \text{ for } n \to +\infty.$$

We therefore conclude that the Fourier series is divergent at every point x, which lies in infinitely many of the $\mathcal{E}_k$. Since $|\mathcal{E}_k| \to 2\pi$ for $k \to +\infty$, it is not hard to prove that the set Ω of all points, which at most lie in finitely many of the $\mathcal{E}_k$ is a nullset. In fact, let $\varepsilon > 0$ be any positive constant, and define

$$\Omega_n := \left(\bigcup_{k=1}^{n} \mathcal{E}_k \right) \setminus \left(\bigcup_{k=n+1}^{+\infty} \mathcal{E}_k \right).$$

Then clearly, $\Omega = \bigcup_{n=1}^{+\infty}$. It follows from the definition of Ω_n that $\Omega_n \cap \mathcal{E}_k = \emptyset$ for all $k \geq n+1$. Then $|\mathcal{E}_k| \to 2\pi$ for $k \to +\infty$ implies that we can choose $k \geq n+1$, such that

$$\mathcal{E}_k > 2\pi - 2^{-n}\varepsilon, \qquad \text{i.e.} \qquad |\Omega_n| < \varepsilon \cdot 2^{-n}.$$

We finally get by a very crude estimate that

$$|\Omega| \leq \sum_{n=1}^{+\infty} |\Omega_n| < \varepsilon \sum_{n=1}^{+\infty} 2^{-n} = \varepsilon.$$

This holds for every $\varepsilon > 0$, which is only possible, if Ω is a nullset. We therefore conclude that the Fourier series of $f \in L^1([0, 2\pi[)$ is divergent almost everywhere. □

4.6 A Fourier series which diverges everywhere

We shall in this section use a modified version of a proof given by *Zygmund* [54] of *Kolmogorov's* result, [32], from 1926.

Define for fixed $n \in \mathbb{N}$,

$$x_j := \frac{2\pi j}{2n+1}, \qquad \text{for } j = 0, 1, \ldots, 2n.$$

Notice in particular that $x_{2j} = c_j$, where the c_j were defined in Section 4.5.

Recall the definition of *Fejér's kernel* in Definition 4.4. Then put

$$\text{(4.30)} \qquad \varphi_n(x) := K_{2n((2n+1))}(x) = \begin{cases} n + \dfrac{1}{2}, & \text{for } x = x_j, \\[2ex] \dfrac{2}{2n+1} \left\{ \dfrac{\sin\left(\frac{1}{2}(2n+1)^2 x\right)}{2\sin\left(\frac{1}{2}(2n+1)x\right)} \right\}^2, & \text{for } x \neq x_j, \end{cases}$$

where $x \in [0, 2\pi[$ and $j = 0, 1, \ldots, 2n$. It follows from the continuity of φ_n that we can find a $\delta \in \left]0, \frac{\pi}{2n+1}\right[$, such that $\varphi_n(x) \geq n$ for all $x \in [x_j - \delta, x_j + \delta]$. Choose $\delta > 0$ so small that also

$$\text{(4.31)} \qquad D_{m_0}(x) = D_{2n(2n+1)}(x) \geq 0, \qquad \text{for all } x \in [-\delta, \delta],$$

where $D_p(x)$ is *Dirichlet's kernel*, cf. Definition 4.3, and where $m_0 := 2n(2n+1)$ is the order of $\varphi_n(x)$ of its trigonometric polynomials. In this way we fix such a $\delta = \delta_n > 0$ for every $n \in \mathbb{N}$.

Define the intervals

$$I_j' = [x_j - \delta, x_j + \delta] \qquad \text{and} \qquad I_j = [x_j + \delta, x_{j+1} - \delta]\,.$$

Then

$$\varphi_n(x) \geq n \qquad \text{for } x \in I_j', \text{ where } j = 0, 1, \dots, 2n.$$

As usual we put all the difficult parts of the proof of the result into some preliminary lemmata.

Lemma 4.2 *For given $n \in \mathbb{N}$ let $\delta > 0$ and I_j be defined as above, and write for short $m_0 = 2n(2n+1)$. There exists a finite sequence of strictly increasing natural numbers,*

$$m_j \in \mathbb{N} \qquad \textit{and} \qquad m_0 < m_1 < m_2 < \cdots < m_n < \cdots\,,$$

such that for every $x \in I_{2j} \cup I_{2kj+1}$, $j = 0, 1, \dots, n-1$, there exists an $h = h_{x,j} \in \mathbb{N}$, satisfying

1) $2k + 1 = \varrho(2n+1)$ *for some* $\varrho \in \mathbb{N}$.

2) $m_j \leq k < \frac{1}{2}\, m_{j+1}$,

3) $\sin\left(k + \frac{1}{2}\right) x < -\frac{1}{2}$.

Proof. By induction. Assume that we already have fixed $m_0, m_1, \dots, m_j$ with the properties described in the lemma. We shall first prove that there exists an $m_j' \in \mathbb{N}$, $m_j' > m_j$, such that (1) and (3) hold for some $k = k_x \in \mathbb{N}$ between m_j and m_j'. Then later on we shall adjust this choice, such that (2) also holds.

Fix some $j \in \{0, 1, \dots, n-2\}$. Condition (1) tells us to check the set of numbers

$$2k + 1 = \varrho(2n+1),$$

where ϱ is an odd integer.

Using that

$$x \in I_{2j} \cup I_{2j+1} = [x_{2j} + \delta, x_{2j+1} - \delta] \cup [x_{2j+1} + \delta, x_{2j+2} - \delta]\,,$$

we see that

$$0 < \delta \leq x_{2j+2} - x \leq x_{2j+2} - x_{2j} = \frac{4\pi}{2n+1}.$$

Therefore, there exists a $\theta \in]0,1[$, such that

$$x_{2j+2} - x = \frac{4\pi\theta}{2n+1}.$$

Notice that

$$\left(k + \frac{1}{2}\right) x_{2j+2} = \frac{\varrho(2n+1)}{2} \cdot \frac{2\pi(2j+2)}{2n+1} = 2\pi\varrho(j+1),$$

which is a multiple of 2π, so we get

$$\text{(4.32)} \qquad -\sin\left(k+\frac{1}{2}\right)x = \sin\left(k+\frac{1}{2}\right)(x_{2j+2}-x) = \sin\left(\frac{2k+1}{2}\cdot\frac{4\pi\theta}{2n+1}\right) = \sin 2\pi\varrho\theta.$$

We see that $x \in I_{2j} \cup I_{2j+1}$, if and only if

$$x_{2j+2} - x = \frac{4\pi\theta}{2n+1} \in \left[\delta, \frac{2\pi}{2n+1} - \delta\right] \cup \left[\frac{2\pi}{2n+1} + \delta, \frac{4\pi}{2n+1} - \delta\right],$$

i.e. if and only if

$$\begin{aligned} \theta \in S \quad &:= \left[\frac{(2n+1)\delta}{4\pi}, \frac{1}{2} - \frac{(2n+1)\delta}{4\pi}\right] \cup \left[\frac{1}{2} + \frac{(2n+1)\delta}{4\pi}, 1 - \frac{(2n+1)\delta}{4\pi}\right] \\ &= \left[\eta, \frac{1}{2} - \eta\right] \cup \left[\frac{1}{2} + \eta, 1 - \eta\right], \qquad \text{where } \eta := \frac{(2n+1)\delta}{4\pi}. \end{aligned}$$

Let ϱ_0 be any odd integer. We shall prove that to every $\theta \in S$ there exists another odd integer $\varrho = \varrho(\theta) \geq \varrho_0$, such that

$$\text{(4.33)} \qquad \sin 2\pi\varrho\theta > \frac{1}{2}.$$

If $\theta \in S$, and $(2k+1)\theta \pmod 1$ lies in the interval $J = \left]\frac{1}{12}, \frac{5}{12}\right[$ for some $k \in \mathbb{N}$, then (4.33) holds for $\varrho = 2k+1$. Unfortunately, ϱ_0 can be very large, so we are only guaranteed the existence of such a $\varrho \geq \varrho_0$, if furthermore the sequence $\{(2k+1)\theta \pmod 1\}_{k \in \mathbb{N}_0}$ has infinite many elements in the interval J for fixed θ, so we must investigate the various possibilities of θ.

If θ is irrational, then $\{(2k+1)\theta \pmod 1\}_{k\in\mathbb{N}_0}$ obviously has infinitely many elements in J.

Then assume that $\theta = \dfrac{p}{q} \in S \cap \mathbb{Q}$, where p and q do not have common factors $\neq 1$. Then we have two possibilities:

(i) q is an odd integer (ii) q is an even integer.

(i) If q is an odd integer, then the q numbers

$$\varrho_0 p,\ (\varrho_0+2)\,p,\ \dots,\ (\varrho_0+2q-2)\,p \qquad (\mathrm{mod}\ q)$$

are mutually different, because $\dfrac{p}{q} \in \mathbb{Q}$ cannot be further reduced. Then divide by q, and we get, modulo 1, all numbers in the sequence

$$0,\ \frac{1}{q},\ \dots,\ \frac{q-1}{q}, \tag{4.34}$$

possibly in a different order. If $q > 3$, then at least one of the numbers of (4.34) lies in the interior of J, because $|J| = \dfrac{1}{3}$. If $q = 3$, then clearly $\dfrac{1}{3} \in J$.

Hence it follows that the sequence $\left\{(2k+1)\dfrac{p}{q}(\mathrm{mod}\ 1)\right\}_{k\in\mathbb{N}_0}$ has infinitely many elements lying in J, when the denominator q is an odd number.

(ii) If q is even, then p must be an odd integer. The $\dfrac{q}{2}$ numbers

$$\varrho_0 p,\ (\varrho_0+2)\,p,\ (\varrho_0+4)\,p,\ \dots,\ (\varrho_0+q-2)\,p,$$

are all odd and mutually different modulo q. When we divide by q, we obtain, modulo 1, all numbers

$$\frac{1}{q},\ \frac{3}{q},\ \dots,\ \frac{q-1}{q}.$$

If $\dfrac{2}{q} < \dfrac{1}{3}$, i.e. if $q > 6$, then at least one of these numbers lies in the interval J. If either $q = 4$ or $q = 6$, then clearly $\dfrac{1}{4}$ and $\dfrac{1}{6}$ both lie in $J = \left]\frac{1}{12}, \frac{5}{12}\right[$. Finally, $\dfrac{1}{2} \notin S$, so $q = 2$ is not relevant.

Summing up, we conclude that if $\theta \in S$, then infinitely many elements of the sequence

$$\{(2k+1)\theta\,(\mathrm{mod}\ 1)\}_{k\in\mathbb{N}_0}$$

lie in J.

Let $\theta \in S$, and let $\varrho = \varrho(\theta) \in \mathbb{N}$ be the smallest odd number $\geq \varrho_0$, such that $\varrho \cdot \theta$ (mod 1) belongs to J. Then (4.33) holds, so

$$\sin\left(2\pi\varrho\left(\theta_0\right)\theta_0\right) > \frac{1}{2}, \qquad \text{for } \theta_0 \in S.$$

Because the function $\theta \to \sin\left(2\pi\varrho\left(\theta_0\right)\cdot\theta\right)$ is continuous, there exists an $\varepsilon = \varepsilon\left(\theta_0\right) > 0$, such that

$$\sin\left(2\pi\varrho\left(\theta_0\right)\cdot\theta\right) > \frac{1}{2}, \qquad \text{for } \theta\ \left]\theta_0-\varepsilon, \theta_0+\varepsilon\right[.$$

Hence, the system of intervals $\{]\theta-\varepsilon(\theta),\theta+\varepsilon(\theta)[\mid \theta\in S\}$ is an open covering of the compact (i.e. closed and bounded) set S. According to *Heine-Borel's covering theorem* it contains a *finite* covering, i.e. there exist finitely many $\theta_1,\dots,\theta_s\in S$, such that

$$S\subseteq\bigcup_{j=1}^{s}\,]\theta_j-\varepsilon(\theta_j),\theta_j+\varepsilon(\theta_j)[\,.$$

It follows from the above that to every interval $]\theta_j-\varepsilon(\theta_j),\theta_j+\varepsilon(\theta_j)[$ there exists a corresponding odd number $\varrho(\theta_j)$, satisfying the above. Then put

$$\varrho_S:=\max\{\varrho(\theta_j)\mid j=1,\dots,s\}\,,$$

and it follows that for every $\theta\in S$ there exists a $\varrho(\theta)\le\varrho_S$, such that $\varrho(\theta)$ has the right properties.

Choose $\varrho_0=m_j$. To every $x\in I_{2j}\cup I_{2j+1}$ there is a constant $k=k_{x,j}\in\mathbb{N}$, lying between m_j and $m_j':=\varrho_S$, such that (1) and (3) both hold. If we finally choose $m_{j+1}>2m_j=2\varrho_S$, then (2) is also fulfilled, and the lemma is proved. $\square$

Let $m_0=2n(2n+1)$, cf. (4.31), and let $m_0<m_1<\cdots<m_n$ be determined according to Lemma 4.2. Then we define in the same way as in Section 4.5,

$$f_n(x):=\sum_{i=1}^{n}K_{m_i}(x-x_{2i})\,, \tag{4.35}$$

where the $K_m(t)$ are *Fejér's kernels*. One should, however, notice that the sequence (m_i) is not chosen in the same way as in Section 4.5. It is obvious that every f_n in nonnegative with $\frac{1}{2}$ as constant term.

Assume that $m_j\le k<m_{j+1}$. Then (cf. Section 4.5),

$$S_k(x;f_n)=\frac{1}{n}\sum_{i=1}^{k}jK_{m_i}(x-x_{2i})+\frac{1}{n}\sum_{i=j+1}^{n}\left\{\frac{1}{2}+\sum_{p=1}^{k}\frac{m_i-p+1}{m_i+1}\cos p\,(x-x_{2i})\right\}.$$

Then we get, using that $m_i-p+1=(m_i-k)+(k-p+1)$, and that $K_m\ge 0$,

$$S_k(x;f_n)\ge\frac{1}{n}\sum_{i=j+1}^{n}\frac{m_i-k}{m_i+1}D_k(x-x_{2i})\,,\qquad\text{for } m_j\le k<m_{j+1}. \tag{4.36}$$

If also $2k+1$ is a multiple of $2n+1$, then

$$S_k(x;f_n)\ge-\frac{1}{n}\left\{\sin\left(k+\frac{1}{2}\right)x\right\}\cdot\sum_{i=j+1}^{n}\frac{m_i-k}{m_i+1}\,\frac{1}{2\sin\left(\frac{1}{2}(x_{2i}-x)\right)}. \tag{4.37}$$

If we choose k as in Lemma 4.2, then it follows from (4.35) that for some constant $C>0$,

$$S_k(x;f_n)>\frac{1}{4n}\sum_{i=j+1}^{n}\frac{1}{x_{2i}-x}>\frac{2n+1}{16\pi n}\sum_{i=j+1}^{n}\frac{1}{i-j}>2C\cdot\ln(n-j).$$

Therefore, if $j\le n-\sqrt{n}$, then

$$S_k(x;f_n)>C\cdot\ln n.$$

Summing up, we have almost proved

Lemma 4.3 *Let φ_n be given by (4.30), and f_n by (4.35). Put $F_n := f_n + \varphi_n$.*
To every $j \in \mathbb{N}$, where $0 \le j \le n - \sqrt{n}$, and every $x \in I_{2j} \cup I_{2j+1}$, there exists $k \in \mathbb{N}$, such that $\frac{1}{2} m_{j+1} > k \ge m_j > m_0$, and such that

$$S_k(x; F_n) \ge C \cdot \ln n. \tag{4.38}$$

PROOF. We know that $\varphi_n(x) \ge n$, so when k is chosen as above, then

$$S_k(x; F_n) = S_k(x; f_n) + S_k(x; \varphi_n) + \varphi_n \ge C \cdot \ln n. \qquad \square$$

Lemma 4.4 *There exists an $N \in \mathbb{N}$, such that if $n \ge N$ and $m_0 = 2n(2n+1)$, then*

$$S_{m_0}(x; F_n) > \frac{1}{2} n \qquad \textit{for all } x \in \bigcup_{j=0}^{2n} I_j'. \tag{4.39}$$

PROOF. It follows from the definition (4.30) of φ_n that it is a trigonometric polynomial of degree $2n(2n+1) = m_0$. Hence,

$$S_{m_0}(x; F_n) = S_{m_0}(x; f_n) + S_{m_0}(x; \varphi_n) = S_{m_0}(x; f_n) + \varphi_n.$$

We defined previously δ and I_j', such that

$$\varphi_n(x) \ge n \qquad \text{for every } x \in \bigcup_{j=0}^{2n} I_j'. \tag{4.40}$$

Therefore, it suffices to prove the existence of a constant C, such that

$$S_{m_0}(x; f_n) > -C \cdot \ln n, \qquad \text{for every } x \in \bigcup_{j=0}^{2n} I_j' \text{ and } n > 1. \tag{4.41}$$

In fact, it follows from (4.40) and (4.41) that for $N \in \mathbb{N}$ sufficiently large,

$$S_{m_0}(x; F_n) \ge n - C \cdot \ln n > \frac{1}{2} n, \qquad \text{for all } n \ge N,$$

and (4.39) follows.

Finally we turn to the proof of (4.41). Let p be fixed, and choose any $x \in I_p'$. We put $k = m_0$ and $j = 0$ in (4.36) to get

$$S_{m_0}(x; f_n) \ge \frac{1}{n} \sum_{i=1}^{n} \frac{m_i - m_0}{m_i + 1} D_{m_0}(x - x_{2i}).$$

When $x \in I_p'$, then the distance between x and x_{2i} is at least $\pi \cdot \dfrac{|p - 2i|}{2n+1}$. Therefore, if p is an odd number, then

$$S_{m_0}(x; f_n) \ge -\frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} \cdot \frac{\pi}{|x - x_{2i}|} \ge -\frac{2n+1}{2n} \sum_{i=1}^{n} \frac{1}{|p - 2i|} \ge -C \cdot \ln n.$$

If instead p is even, then $x - x_p \in I_0'$. By (4.31) we have $D_{m_0}(x - x_p) \geq 0$, so we conclude that

$$S_{m_0}(x; f_n) \geq -\frac{2n+1}{2n} \sum_{\substack{i=1 \\ 2i \neq p}}^{n} \frac{1}{|p - 2i|} \geq -C \cdot \ln n,$$

and (4.41) is proved. As argued above, this finishes the proof. □

Lemma 4.5 *There exists a sequence of nonnegative trigonometric polynomials* (F_n) *of order* (ν_n), *where* $\nu_1 < \nu_2 < \cdots$, *and all of constant term* $\frac{1}{2}$, *such that the following conditions are fulfilled:*
To every $n \in \mathbb{N}$ *there exist a constant* $A_n \in \mathbb{R}$ *and a set* $E_n \subseteq [0, 2\pi[$, *and another constant* $\lambda_n \in \mathbb{N}$, *such that*

1) $A_n \to +\infty$ *for* $n \to +\infty$,

2) $E_1 \subset E_2 \subset \cdots$ *and* $\bigcup_{n=1}^{+\infty} E_n = [0, 2\pi[$,

3) $\lambda_n \to +\infty$ *for* $n \to +\infty$,

4) to every $x \in E_n$ *there exists a* $k = k(x) \in \mathbb{N}$, *such that* $\lambda_n \le k \le \nu_n$ *and*

$$S_k(x; F_n) > A_n.$$

PROOF. Choose $F_n := f_n + \varphi_n$, where f_n is given by (4.35) and φ_n by (4.30). Since every f_n and every φ_n have the constant term $\frac{1}{2}$, every F_n has the constant term 1. Furthermore, the orders of the F_n as trigonometric polynomials are strictly increasing.

We define

$$E_n := \left[0, \frac{4\pi(n - \sqrt{n})}{2n+1}\right[.$$

Then it follows from (4.38) of Lemma 4.3, and (4.39) of Lemma 4.4 that to every $n \ge N$ and every $x \in E_{2n}$ there exists a $k > 2n(2n+1)$, such that for some constant $C > 0$,

$$S_k(x; F_n) \ge C \cdot \ln n.$$

The remaining conditions of the lemma follow by putting $A_n = C \cdot \ln n$ for $n \ge N$ (and an appropriate A_n for $n < N$) and $\lambda_n = 2n(2n+1)$. $\square$

Theorem 4.3 *There exists a function* $f \in L^1([0, 2\pi[)$, *such that its Fourier series is divergent everywhere.*

PROOF. We choose the sequence of functions (F_n) as in Lemma 4.5. Then choose n_1, such that $A_{n_1} > 0$. Determine by induction a sequence (n_i) of natural numbers $n_i \in \mathbb{N}$, such that

$$(a) \quad \lambda_{n_i} > \nu_{n_{i-1}}, \qquad (b) \quad A_{n_i} > 4A_{n_{i-1}}, \qquad (c) \quad \sqrt{A_{n_i}} > \nu_{n_{i-1}},$$

where the constants λ_n, ν_n and A_n are given by Lemma 4.5.

If we put

$$f(x) := \sum_{k=1}^{+\infty} \frac{1}{\sqrt{A_{n_k}}} F_{n_k}(x), \tag{4.42}$$

then $f \in L^1([0, 2\pi[)$. In fact, $F_n \ge 0$ has the constant term 1, so

$$\int_0^{2\pi} |F_n(x)| \, \mathrm{d}x = \int_0^{2\pi} F_n(x) \, \mathrm{d}x = 2\pi \qquad \text{for all } n \in \mathbb{N}.$$

It follows from (b) that

$$A_{n_k} > 4\,A_{n_{k-1}} > \cdots > 4^{k-1}\,A_{n_1},$$

hence,

$$\|f\|_1 = \sum_{k=1}^{+\infty} \frac{1}{\sqrt{A_{n_k}}} \int_0^{2\pi} F_{n_k}(x)\,\mathrm{d}x < \frac{2\pi}{\sqrt{A_{n_1}}} \sum_{k=1}^{+\infty} \frac{1}{2^{k-1}} = \frac{4\pi}{\sqrt{A_{n_1}}} < +\infty,$$

and the claim is proved.

If $x \in E_{n_i}$, we split the expression of $f(x)$ in the following way,

$$F(x) = u+v+w = \sum_{k=1}^{i-1} \frac{1}{\sqrt{A_{n_k}}}\,F_{n_k}(x) + \frac{1}{\sqrt{A_{n_i}}}\,F_{n_i}(x) + \sum_{k=i+1}^{+\infty} \frac{1}{\sqrt{A_{n_k}}}\,F_{n_k}(x),$$

where u is the sum of all terms preceding

$$v := \frac{1}{\sqrt{A_{n_i}}}\,F_{n_i}(x),$$

and w is the sum of all terms following after v.

Due to the linearity, the section $S_p(x;f)$ is split in the same way,

$$S_p(x;f) = S_p(x,u) + S_p(x;v) + S_p(x,w). \tag{4.43}$$

When we apply Lemma 4.5 (4), we find $k = k(x,i)$, such that $\lambda_{n_i} \le k \le \nu_{n_i}$, and such that

$$S_k(x;v) = \frac{1}{\sqrt{A_{n_i}}}\,S_k(x;F_{n_i}) > \sqrt{A_{n_i}}. \tag{4.44}$$

Furthermore, it follows from (a) that

$$S_k(x;u) = u(x) \ge 0. \tag{4.45}$$

We notice that if $g \in L^1([0,2\pi[)$, then clearly every Fourier coefficient a_n and b_n corresponding to g in absolute value is at most equal to $\dfrac{1}{\pi}\,\|g\|_1$. Since the k-th section of the Fourier series at most contains $2k+1$ terms, we get the simple estimate

$$|S_k(x;g)| = \left|\frac{1}{2}\,a_0 + \sum_{j=1}^{k}\{a_j \cos jx + b_j \sin jx\}\right| \le \frac{2k+1}{\pi}\,\|g\|_1.$$

When this estimate is applied on $g = w \in L^1([0,2\pi[)$, it follows from (b) and (c) that

$$\begin{aligned}
|S_k(x;w)| &\le \frac{2k+1}{\pi} \int_0^{2\pi} \sum_{j=i+1}^{+\infty} \frac{1}{\sqrt{A_{n_j}}}\,F_{n_j}(x)\,\mathrm{d}x = \frac{2k+1}{\pi}\cdot 2\pi \sum_{j=i+1}^{+\infty} \frac{1}{\sqrt{A_{n_j}}} \\
&\le \frac{2(2k+1)}{\sqrt{A_{n_{i+1}}}} \sum_{j=0}^{+\infty} \frac{1}{2^j} < \frac{12k}{\sqrt{A_{n_{i+1}}}} \le \frac{12k}{\sqrt{A_{n_{i+1}}}} < 12.
\end{aligned}$$

When we use this estimate combined with (4.33), (4.34) and (4.35), we get

$$S_k(x; f) \geq 0 + \sqrt{A_{n_i}} - 12 = \sqrt{A_{n_i}} - 12. \tag{4.46}$$

Finally, we notice that for every $x \in [0, 2\pi[$ there exists an $i_0 \in \mathbb{N}$, such that $x \in E_{n_i}$ for all $i \geq i_0$. Then it follows from (4.36) and Lemma 4.5 (1) that the Fourier series of f is divergent everywhere in $[0, 2\pi[$, and the theorem is proved. $\square$

It is obvious that it is not possible to specify all coefficients a_n, b_n explicitly for such divergent Fourier series for a function from $L^1([0, 2\pi[)$.

4.7 Some "wild" Hilbert spaces

In most introductory descriptions of *Functional Analysis* one only considers *separable Hilbert spaces*, i.e. Hilbert spaces which contains a countable dense subset. The reason is that a Hilbert space has a countable orthonormal basis, if and only if it is separable. The prototype of a separable Hilbert space is $L^2([0, 2\pi[)$, or equivalently – and more convenient for our purposes in this section – the space $L^2([-\pi, \pi[)$ where the interval $[-\pi, \pi[$ has the center $x = 0$. We shall start with

4.7.1 Short review of the standard Hilbert space $L^2([-\pi, \pi[)$

We shall for convenience assume that we are dealing with complex separable Hilbert spaces over the field of complex numbers $\mathbb{C}$. It can be proved that there is essentially only one such separable Hilbert space, because they are all mutually isomorphic. We can therefore restrict ourselves to the standard example. Notice for the geometrical mind, that the Hilbert space in some sense is the generalisation of the *Euclidean spaces* E_n of dimension $n \in \mathbb{N}$ to infinite dimension. We may therefore expect some definitions which have a ring of Geometry, like e.g. inner product and orthogonality, and these concepts are indeed essential in the theory of Hilbert spaces.

We shall in this section consider the standard example of a separable Hilbert space over $\mathbb{C}$, namely $L^2([-\pi, \pi[)$. Recall that an element $f \in L^2([-\pi, \pi[)$ is a (measurable) function on $[-\pi, \pi[)$, for which the *norm* $\|f\|_2$ of f is finite,

$$\|f\|_2 := \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2 \, \mathrm{d}t}.$$

Remark 4.4 In practice, *all* functions in daily life are measurable, so we shall not worry here about his extra assumption. We shall in Chapter 6 return to the non-measurable sets and show that they have some really strange properties, which are impossible to realize in the ordinary world. They only exist in the theoretical minds of the mathematicians, including the author's. $\Diamond$

We define the *inner product* $\langle \cdot, \cdot \rangle$ in $L^2([-\pi, \pi[)$ by

$$\langle f, g \rangle := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \overline{g(t)} \, \mathrm{d}t \qquad \text{for } f,\, g \in L^2([-\pi, \pi[),$$

where we have chosen the notation from *Quantum Mechanics*, where Hilbert spaces also enter the physical models. The similar object in *Euclidean space* is the dot product, $\mathbf{u} \cdot \mathbf{v}$ for $\mathbf{u}$, $\mathbf{v} \in E_n$. It should be no surprise for the reader that the properties are the same. It also follows from *Cauchy-Schwarz's inequality*, cf. Section 2.1, that $\langle f, g\rangle$ is well-defined for $f,\, g \in L^2([-\pi,\pi[)$, and that

$$|\langle f,g\rangle| \leq \|f\|_2\|g\|_2, \qquad \text{and} \qquad \langle f,f\rangle = \|f\|_2^2.$$

Furthermore, it is easy to prove that the double infinite sequence of functions

$$\text{(4.47)} \qquad \{e_n \mid n \in \mathbb{Z}\}\,, \qquad \text{where } e_n(t) := e^{int} \text{ for } t \in \mathbb{R},$$

with respect to this inner product is an *orthonormal* sequence in $L^2([-\pi,\pi[)$. This means that

$$\langle e_m, e_n\rangle = \delta_{m,n} = \begin{cases} 1 & \text{for } m = n, \\ 0 & \text{for } m \neq n, \end{cases} \qquad m,\, n \in \mathbb{Z},$$

where $\delta_{m,n}$ is called the *Kronecker symbol*.

That (4.47) is an *orthonormal basis* means that every function $f \in L^2([-\pi,\pi[)$ has a (unique) representation

$$\text{(4.48)} \qquad f \sim \sum_{n=-\infty}^{+\infty} c_n\, e^{int}$$

as a complex Fourier series, where the coefficients c_n are determined by

$$c_n := \langle f, e_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\, e^{-int}\,\mathrm{d}t, \qquad \text{for } n \in \mathbb{Z}.$$

All this is well-known from the *Theory of Fourier series*, and this is actually the reason why we focus so much on the Fourier series. They provide us, when they are equipped with the right generalisation of the inner product, with a natural generalisation of the Euclidean spaces to a space of infinite dimensions, preserving as many as possible of the nice properties of the Euclidean spaces.

The sign $\sim$ in (4.48) indicates that the two expressions are not necessarily equal in a *pointwise* sense. They are, however, equal in L^2-sense. Formula (4.48) corresponds to the usual expansion of a vector in an Euclidean space with respect to an orthonormal basis. In the present situation the functions $\{e_n \mid n \in \mathbb{Z}\}$ play the same role as the basis vectors in the Euclidean space. We notice that *Pythagoras's theorem* is here generalised to *Parseval's equation*

$$\|f\|_2^2 = \sum_{n=-\infty}^{+\infty} |c_n|^2 = \|(e_n)\|_2^2,$$

hence by the definitions,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2 \,\mathrm{d}t = \sum_{n=-\infty}^{+\infty} \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\, e^{-int}\,\mathrm{d}t\right|^2.$$

Recall that according to *Carleson-Hunt's theorem*, cf. Section 4.4 or *Jørsboe & Mejlbro* [28], we have for the Hilbert space $L^2([-\pi,\pi[)$ pointwise equality almost everywhere in $[-\pi,\pi[$ in (4.8). This is very convenient from a physical and engineering point of view. Mathematicians may still consider *Carleson-Hunt's theorem* as a very strange phenomenon.

4.7.2 Kadec's $\frac{1}{4}$-theorem

It was emphasized above that the theory of Fourier series is very important in the sciences of Mathematics, Physics and Engineering. However, from the point of view of practical applications it is often neglected that signals only seldom can be expanded into oscillations of the form of linear combinations of e^{int}, where $n \in \mathbb{Z}$, or similarly. One simple example is the *vibrating string* with inhomogeneous density. Such a situation occurs e.g. when one tries to tune a worn-out string of a guitar. Fortunately it is still possible to deal with such a situation. Here we again have a strange result, namely

Kadec's $\frac{1}{4}$-theorem. *Let $(\lambda_n)_{n\in\mathbb{Z}}$ be a sequence of complex numbers, for which there exists a constant $c \in \left]0, \frac{1}{4}\right[$, such that for all $n \in \mathbb{Z}$,*

$$|\lambda_n - n| \le c.$$

For every $f \in L^2([-\pi,\pi[)$ there exists an uniquely determined sequence $(c_n)_{n\in\mathbb{Z}}$ of complex numbers, such that

$$f \sim \sum_{n=-\infty}^{+\infty} c_n\, e^{i\lambda t}, \qquad \textit{i.e. such that} \qquad \lim_{n\to+\infty}\left\| f - \sum_{k=-n}^{n} c_n\, e^{i\lambda_k t}\right\|_2 = 0.$$

It is possible to prove that the constant $\frac{1}{4}$ is the best possible. The proof of the theorem is easy, if we furthermore assume that $c < \dfrac{\ln 2}{\pi}$, cf. e.g. *Riesz & Sz.-Nagy* [47], and I have even once used it in one of my exams. However, if $\dfrac{\ln 2}{\pi} \le c < \dfrac{1}{4}$, then the proof becomes very hard.

A mathematical model of the vibrating string with inhomogeneous density describe the impurities of the materials by replacing numbers from $\mathbb{Z}$ by perturbed numbers $\{\lambda_n \mid n \in \mathbb{Z}\}$, which may even be a set of complex numbers. It can be proved that in the case of the vibrating string then all λ_n are real.

4.7.3 The Besicovitch space

Kadec's $\frac{1}{4}$-theorem makes it possible to use other oscillations than the usual ones. It also inspires one to allow even orthogonal systems of higher cardinality. We shall here sketch, how we introduce the *Besicovich space* $\mathbf{B}^2$, which is a *non-separable Hilbert space*. More precisely, the non-countable set of functions

$$\{e_\lambda \mid \lambda \in \mathbb{R}\}, \qquad \text{where } e_\lambda : \mathbb{R} \to \mathbb{C},\ e_\lambda(t) := e^{i\lambda t} \text{ for } \lambda \in \mathbb{R},$$

becomes a non-countable *orthonormal* set. In particular, the *Besicovitch space* contains all elements of the form

$$f \sim \sum_{n=-\infty}^{+\infty} c_n\, e^{i\lambda_n t}, \qquad \text{where } \sum_{n=-\infty}^{+\infty} |c_n|^2 < +\infty,$$

and where the real sequence (λ_n) is not necessarily the familiar sequence $(n)_{n\in\mathbb{Z}}$.

Let $\mathfrak{V} = \operatorname{span}\left\{e^{i\lambda t} \mid \lambda \in \mathbb{R}\right\}$. This means that every element of $\mathfrak{V}$ is a *finite* linear combination of functions of the type $e^{i\lambda t}$, $\lambda \in \mathbb{R}$. It is obvious that $\mathfrak{V}$ with the usual operations of calculation becomes a complex vector space.

We define a bilinear function $\langle \cdot, \cdot \rangle : \mathfrak{V} \times \mathfrak{V} \to \mathbb{C}$ by

$$(4.49) \qquad \langle f, g \rangle := \limsup_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} f(t)\overline{g(t)}\, \mathrm{d}t \qquad \text{for } f,\ g \in \mathfrak{V}.$$

When we compare this bilinear function with the *inner product* of $L^2([-\pi, \pi[)$, we notice that we in both cases integrate over a symmetric interval and then divide by the length of the interval of integration. In the present case we furthermore let the length of the interval tend towards $+\infty$. Some very simple calculations, which are left to the reader, show that $\mathfrak{V}$ indeed becomes a vector space over $\mathbb{C}$ with an *inner product* given by (4.49). We shall take all this for granted in the following. By the usual process of completion we get $\mathbf{B}^2 := \overline{\mathfrak{V}}$.

We shall finally show some simple properties of the *Besicovitch space* $\mathbf{B}^2$.

For fixed $\lambda \in \mathbb{R}$ we get

$$\langle e_\lambda, e_\lambda \rangle = \limsup_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} e^{i\lambda t} \cdot e^{-i\lambda t}\, \mathrm{d}t = \lim_{T\to+\infty} \frac{1}{2T} \cdot 2T = 1,$$

so $\|e_\lambda\|_2 = 1$.

Then let λ, $\mu \in \mathbb{R}$, where $\lambda \neq \mu$. In this casse we get

$$\begin{aligned} \langle e_\lambda, e_\mu \rangle &= \limsup_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} e^{i\lambda t}\, e^{-i\mu t}\, \mathrm{d}t = \lim_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} e^{i(\lambda-\mu)t}\, \mathrm{d}t \\ &= \lim_{T\to+\infty} \frac{1}{2T} \frac{1}{i(\lambda-\mu)} \left\{ e^{i(\lambda-\mu)T} - e^{-i(\lambda-\mu)T} \right\} = \lim_{T\to+\infty} \frac{1}{T} \cdot \frac{\sin(\lambda-\mu)T}{\lambda-\mu} = 0. \end{aligned}$$

It follows that $\{e_\lambda \mid \lambda \in \mathbb{R}\}$ is a non-countable orthonormal set of functions from $\mathbf{B}^2$ with respect to this inner product. Since we have found a non-countable set of basis vectors, the space $\mathbf{B}^2$ cannot be a separable Hilbert space.

It is possible by using *Parseval's equation* to prove that to every $f \in \mathbf{B}^2$ there exists a countable sequence $(\lambda_n)_{n\in\mathbb{N}}$ of *real* numbers, and another countable sequence $(c_n)_{n\in\mathbb{N}}$ of *complex* numbers, such that

$$f \sim \sum_{n=1}^{+\infty} c_n\, e^{i\lambda_n t} \qquad \text{and} \qquad \sum_{n=1}^{+\infty} |c_n|^2 < +\infty, \tag{4.50}$$

and we shall never get into the situation, where we have to sum over a noncountable set of λ-s.

Notice that we have defined a totally different inner product in $\mathbf{B}^2$ than in $L^2([-\pi,\pi[)$. Therefore, this result does not contradict *Kadec's* $\frac{1}{4}$*-theorem*.

Historical remarks

Obviously, if $(\lambda_n)_{n\in\mathbb{Z}} = (n)_{n\in\mathbb{Z}}$ in (4.50), then we get the usual Fourier series.

The structure of these usual Fourier series inspired *Harald Bohr* [8], [9], [10], in 1923 to generalise the Fourier series by allowing more general exponents of the exponential function $e^{i\lambda t}$. In total *Bohr* wrote 54 papers on this subject, where we here only refer to the first three of them.

One must here realise that in 1923 Lebesgue's theory of integration from 1907 had not yet been applied to form the Hilbert spaces, so these were not known to Bohr. His idea was instead to consider the closed hull of all *finite* sums of the form

$$\sum_{j=1}^{n} a_j\, e^{i\lambda_j t}, \qquad \lambda_1, \dots, \lambda_n \in \mathbb{R}, \tag{4.51}$$

with respect to the well-known uniform norm

$$\|f\|_\infty := \sup_{t\in\mathbb{R}} |f(t)|.$$

Well, (4.51) above is not Bohr's original definition, but it gives a shortcut to explain his idea without being too technical. The advantage of this choice is of course that everyone of Bohr's almost periodic functions are bounded and continuous, because they can be written as a limit of a uniformly convergent sequence of continuous functions of the form (4.51).

It is noteworthy to see that *Bohr* actually uses Hilbert space methods *before* the Hilbert space theory was known. In fact, *Bohr* first introduces the *mean value operator*,

$$\mathfrak{M}\{f(t)\} := \limsup_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} f(t)\,\mathrm{d}t,$$

and considers in particular,

$$\mathfrak{M}\left\{|f(t)|^2\right\} = \limsup_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} |f(t)|^2\,\mathrm{d}t. \tag{4.52}$$

Bohr even proves the *Cauchy-Schwarz inequality*,

$$|\mathfrak{M}\{f(t)g(t)\}| \le \sqrt{\mathfrak{M}\left\{|f(t)|^2\right\}}\sqrt{\mathfrak{M}\left\{|g(t)|^2\right\}},$$

for the mean value operator given above.

Furthermore, he proves *Parseval's equation*, which is a very deep result. More precisely, if $f(t)$ is almost periodic in the sense of Bohr, then we have the expansion

$$f(t) \sim \sum_{\lambda\in\mathbb{R}} a_\lambda\, e^{i\lambda t},$$

and

$$\mathfrak{M}\left\{|f(t)|^2\right\} = \sum_{\lambda} |a_\lambda|^2 < +\infty.$$

The inequality to the right is only possible, if all $a_\lambda = 0$, except for at most countably many λ.

With all the clues above it is a pity, though one cannot blame Bohr, that he misses to see that if f and g are almost periodic functions in his sense, then

$$\langle f, g\rangle := \mathfrak{M}\{f(t)\overline{g(t)}\}$$

is actually (almost) an inner product and the extension to the full Hilbert space is today an easy exercise. This was left to *Besicovitch*, who introduced the general Banach spaces $\mathbf{B}^p$, $p \geq 1$. In these spaces we define a seminorm $\|\cdot\|_p$ by

$$\|f\|_p^p := \limsup_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} |f(t)|^p \,\mathrm{d}t, \qquad f \in \mathbf{B}^p,$$

whenever $\|f\|_p < +\infty$. This is, however, not a norm. Let e.g. $f \neq 0$ be continuous and 0 outside a bounded set. Then clearly $\|f\|_p = 0$, so we have to restrict ourselves to the quotient space with respect to the nullspace $\{f \mid \|f\|_p = 0\}$. Two functions f and g then belong to the same equivalence class, if $\|f-g\|_p = 0$, in which case they are considered as equal. The set of all equivalence classes with respect to this nullspace becomes our Banach space $\mathbf{B}^p$. It can be proved that when $p = 2$, then we precisely get $\mathbf{B}^2$ with the norm given by (4.52), i.e.

$$\|f\|_2^2 := \mathfrak{M}\left\{|f(t)|^2\right\} = \limsup_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} |f(t)|^2 \,\mathrm{d}t, \qquad f \in \mathbf{B}^2.$$

In the present context only the non-separable Hilbert space $\mathbf{B}^2$ is of interest. The main result here is that we have an uncountable orthonormal set of basis vectors, but due to *Parseval's equation* the expansion

$$f(t) \sim \sum_{\lambda\in\mathbb{R}} a_\lambda\, e^{i\lambda t} = \sum_{n=1}^{+\infty} a_{\lambda_n}, \qquad f \in \mathbf{B}^2,$$

is *always* countable, because the sum is finite, $\sum_\lambda |a_\lambda|^2 < +\infty$.

5 Continuous curves

5.1 Introduction

We shall in this chapter go through two classical examples of *space filling curves*. We consider in Section 5.2 the famous *Peano's curve*,, and in Section 5.3 we modify Peano's curve to get *Osgood's curve.*. When we finally in Section 5.4 combine Peano's curve with *Cantor's function*, introduced in Section 3.3, we get a curve passing though every point of the unit square, i.e. its graph has positive area 1, and yet this curve is "almost everywhere" (in the parameter domain) "almost nowhere" (in the range, i.e. it lies almost nowhere in the unit square). In particular, it follows that nullsets are not necessarily mapped into nullsets by a continuous map.

Obviously, we must reconsider what is meant by a continuous curve. From time to time one comes across in the literature the following heuristic definition of a continuous curve,

Pseudo definition. *"The graph of a continuous curve is an infinitely thin line, which can be drawn without lifting the pencil from the paper."*

This "definition" is of course very convenient in most practical cases, because it contains the idea of a curve. However, it implicit assumes that the curve is smooth in some sense and not just continuous. E.g. piecewise differentiable. Once *Weierstraß* in 1861 constructed a continuous function, which was not differentiable anywhere, cf. Section 1.6, his result implicitly also implied that the pseudo definition above was far too imprecise, though the mathematical society still did not react. The next step forward in the direction of a reconsideration was given by *Peano* [43] in 1890, when he constructed a continuous curve in the plane, which went through every point of the unit square, in which case the pseudo definition above clearly does not apply.

We shall in Section 5.2 go through the classical construction of Peano without using the alternative construction of David Hilbert.

Peano's curve contained a lot of double points. Then in 1903 *Osgood* [42] modified Peano's construction and constructed a curve without double points, such that the area of its graph is $\geq 1-\varepsilon$, where $\varepsilon \in]0,1[$ is any given constant. The editor of *Trans. Amer. Math. Soc.* must in 1903 have thought that this example was very extraordinary, because Osgood's contribution was supplied with attached figures in colour to describe his procedure. This must have been very expensive in those days. We shall here not go that far.

Osgood's curve is also classical stuff, so we describe the construction in Section 5.3. This is a curve without double points in the unit square, and yet its graph has an area $\geq 1-\varepsilon$, where $\varepsilon \in]0,1[$ can be chosen as small as we want.

Instead of the pseudo definition above we shall in the following use the more precise

Definition 5.1 *Let $I \subset \mathbb{R}$ be an interval, and let $\varphi : I \to \mathbb{R}^2$ be a continuous map. We then call φ a* continuous curve *in the plane.*

We shall in the beginning also assume that I is bounded and closed, i.e. compact, because then it follows from one of the main theorems of continuous functions that then the range $\varphi(I)$ is compact, i.e. bounded and closed.

We mention that if I is not bounded, or not closed, then $\varphi(I)$ may be unbounded.

Then recall another main theorem from the elementary mathematical analysis.

Theorem 5.1 *Assume that* $\mathbf{f}_n : I \to \mathbb{R}^2$ *is an uniform Cauchy sequence of continuous functions. Then the limit function* $\mathbf{f} : I \to \mathbb{R}^2$ *of* $(\mathbf{f}_n)$ *exists, and it is continuous.*

5.2 Peano's curve

The idea of the construction of Peano's curve is to construct an uniform Cauchy sequence of continuous functions

$$\mathbf{f}_n : [0,1] \to [0,1] \times [0,1] = [0,1]^2, \qquad \text{for } n \in \mathbb{N}_0,$$

such that

a) The sequence of functions $(\mathbf{f}_n)$ converges uniformly towards a continuous function $\mathbf{f} : [0,1] \to [0,1]^2$.

b) For every point $(x,y) \in [0,1]^2$ there exists a parameter $t \in I$, such that $\mathbf{f}(t) = (x,y)$.

The limit function $\mathbf{f} : [0,1] \to [0,1]^2$, which is called *Peano's curve*, is continuous and its graph is passing through every point of the unit square, so this will solve our problem.

We define $(\mathbf{f}_n)$ inductively. First put

$$\mathbf{f}_0(t) := (t,t), \qquad \text{for } t \in [0,1].$$

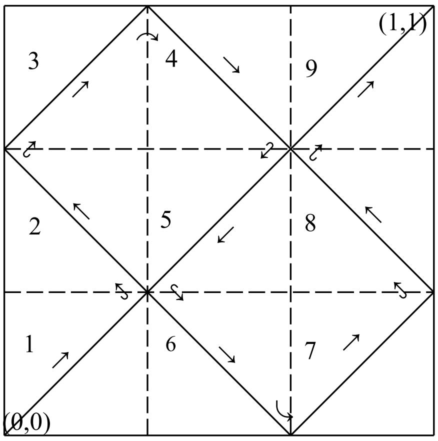

Figure 5.1: *The structure of the graph of the function* $\mathbf{f}_1 : [0,1] \to [0,1]^2$. *The ordering of the subsquares is indicated by the numbers* $1, 2, \dots, 9$.

We then define $\mathbf{f}_1 : [0,1] \to [0,1]^2$, based on Figure 5.1.

For every subinterval $\left[\dfrac{k-1}{9}, \dfrac{k}{9}\right]$, $k \in \{1, \dots, 9\}$, of the parameter interval I choose subsquare number k, defined by Figure 5.2.

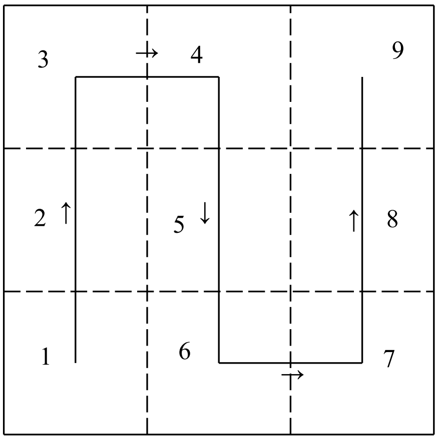

Figure 5.2: *The order of the numbering of the subsquares*

Choose the diagonal of this subsquare and let it be traversed in the same direction as given by the order of the subsquares of Figure 5.2. If $\mathbf{f}_1$ maps the subinterval $\left[\frac{k-1}{9}, \frac{k}{9}\right]$ onto this oriented diagonal of constant speed, then we obtain a continuous curve $\mathbf{f}_1 : [0,1] \to [0,1]^2$ as given on Figure 5.1

Then notice that every subsquare is a micro cosmos, which is similar to $\mathbf{f}_0 : [01] \to [0,1]^2$. In both cases we have a square with a directed diagonal traversed at a constant speed. If we for $\mathbf{f}_1$ consider the subinterval $\left[\frac{k-1}{9}, \frac{k}{9}\right]$ with its corresponding subsquare Q_k of edge length $\frac{1}{3}$, then divide Q_k into 9 smaller subsquares of edge length $\frac{1}{9} = \frac{1}{3^2}$, and order them according to Figure 5.2 and the direction of the diagonal, specified by $\mathbf{f}_1$. Similarly, divide the subinterval $\left[\frac{k-1}{9}, \frac{k}{9}\right]$ into 9 smaller intervals, and let the m-th of these be the parameter interval of the directed diagonal of the m-th subsquare of Q_k. This is just a repetition of the previous construction at a smaller scale, so when we glue all pieces together, we get a continuous map $\mathbf{f}_2 : [0,1] \to [0,1]^2$.

For later use we notice that $\mathbf{f}_2$ maps the interval $\left[\frac{k-1}{9}, \frac{k}{9}\right]$ connected with the previous map $\mathbf{f}_1$ into precisely the same subsquare Q_k. So when we iterate this construction to get $\mathbf{f}_3$, $\mathbf{f}_4$, ..., all these will always map $\left[\frac{k-1}{9}, \frac{k}{9}\right]$ into the same subsquare Q_k as $\mathbf{f}_2$. In particular, for every $n \in \mathbb{N}$ and every $m \in \{1, \dots, 3^n\}$,

$$\mathbf{f}_n\left(\frac{m}{3^n}\right) = \mathbf{f}_{n+1}\left(\frac{m}{3^n}\right) = \cdots ,$$

so the points $\mathbf{f}_n\left(\frac{m}{3^n}\right) \in [0,1]^2$ are not changed by the following maps.

Remark 5.1 Notice that the construction above, where one repeats the same geometrical pattern, is very similar to the methods much later used in the construction of fractals. $\Diamond$

We prove that the sequence $(\mathbf{f}_n)$ is uniformly convergent, i.e. for every given $\varepsilon > 0$ there exists a constant $N \in \mathbb{N}$, such that

$$\|\mathbf{f}_m - \mathbf{f}_n\|_\infty < \varepsilon, \qquad \text{for all } m,\, n \in \mathbb{N},$$

where the uniform norm $\|\cdot\|_\infty$ is defined by

$$\|\mathbf{f}\|_\infty := \sup_{t\in[0,1]} |\mathbf{f}(t)| = \sup_{t\in[0,t]} \sqrt{\{f_1(t)\}^2 + \{f_2(t)\}^2},$$

and where $|\mathbf{f}(t)|$ for every fixed $t \in [0,1]$ denotes the length of the vector $\mathbf{f}(t) = (f_1(t), f_2(t)) \in [0,1]^2$.

Choose $N \in \mathbb{N}$, such that $3^{-N}\sqrt{2} < \varepsilon$. Then we can for every $t \in [0,1]$ find $k \in \left\{1, \dots, 3^N\right\}$, such that

$$t \in \left[\frac{k-1}{3^N}, \frac{k}{3^N}\right].$$

We have two possible choices, when $t = k \cdot 3^{-N}$, $k \neq 3^N$. In such a case we always choose the interval

$$\left[\frac{k-1}{3^N}, \frac{k}{3^N}\right], \qquad k \in \left\{1, \dots, 3^N - 1\right\}.$$

If $m, n \geq N$, then it follows from the analysis above that the images

$$\mathbf{f}_m\left(\left[\frac{k-1}{3^N}, \frac{k}{3^N}\right]\right) \qquad \text{and} \qquad \mathbf{f}_m\left(\left[\frac{k-1}{3^N}, \frac{k}{3^N}\right]\right), \qquad \text{for } m,\, n \geq N,$$

must be subsets of the same subsquare of number k from level N. The edge length of this subsquare is 3^{-N}, hence

$$|\mathbf{f}_m(t) - \mathbf{f}_n(t)| := \sqrt{\{f_{m,1}(t) - f_{n,1}(t)\}^2 + \{f_{m,2}(t) - f_{n,2}(t)\}^2} \leq 3^{-N}\sqrt{2} < \varepsilon.$$

Since this estimate holds for every $t \in [0,1]$, it follows for $m, n \geq N$,

$$\|\mathbf{f}_m - \mathbf{f}_n\|_\infty := \sup_{t\in[0,1]} |\mathbf{f}_m(t) - \mathbf{f}_n(t)| < \varepsilon,$$

and we have proved that $(\mathbf{f}_n)$ is an uniform Cauchy sequence. We infer from Theorem 5.1 that it has a continuous limit function, i.e. a curve, $\mathbf{f} : [0,1] \to [0,1]^2$.

The interval $[0,1]$ is compact, i.e. closed and bounded, and the limit function $\mathbf{f}$ is continuous, so by one of the main theorems of continuous functions the range $\mathbf{f}([0,1]) \subseteq [0,1]^2$ is also closed and bounded. We shall only prove that $\mathbf{f}([0,1]) = [0,1]^2$. It suffices to prove that to every point $(x,y) \in [0,1]^2$ and to every $\varepsilon > 0$ there exists a point $(x_0, y_0) \in \mathbf{f}([0,1])$, such that

$$|(x,y) - (x_0 - y_0)| = \sqrt{(x-x_0)^2 + (y-y_0)^2} < \varepsilon. \tag{5.1}$$

In fact, if (5.1) holds, then choose the sequence (ε), $\varepsilon_n := \dfrac{1}{n}$, to obtain by (5.1) a sequence $((x_n, y_n))_{n\in\mathbb{N}}$ from the closed set $\mathbf{f}([0,1])$ converging towards the point (x,y) under consideration. The point (x,y) must lie in the closure of $\mathbf{f}([0,1])$, which is $\mathbf{f}([0,1])$ itself.

We turn to the proof of (5.1). Let $(x,y) \in [0,1]^2$ and $\varepsilon > 0$ be given. We choose as above $N \in \mathbb{N}$, such that $3^{-N}\sqrt{2} < \varepsilon$. Then choose the subsquare Q at level N, for which $(x,y) \in I$. The image of the function $\mathbf{f}_N$ includes the diagonal of I, and $k \cdot 3^{-N}$ is mapped into the same point by all functions $\mathbf{f}_m$, $m \geq N$, accordingly also by the limit function $\mathbf{f}$.

Each corner $(x_0, y_0) \in \mathbf{f}([0,1])$ of the subsquare Q has a distance $< \varepsilon$ from any other point in Q, so we conclude that

$$|(x,y) - (x_0, y_0)| < \varepsilon,$$

and (5.1) is proved. □

Summing up, we have proved

Theorem 5.2 *There exists a continuous curve, which runs through every point of the unit square.*

It is not hard to construct a generalised Peano's curve which runs through all points of the plane. We only have to allow the parameter interval to be $\mathbb{R}$. That I previously was chosen bounded was only due to some technicalities in the basic construction.

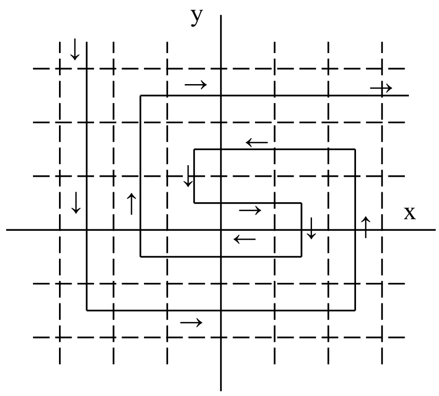

Figure 5.3: *The ordering of all squares in the plane. Peano's curve runs through a diagonal, so one must check that the diagonals match each others at their endpoints. This is possible by the ordering in the double spiral on the figure.*

The basic Peano curve goes from one diagonal point to the opposite diagonal point. We shall just be convinced that it is possible to glue together infinitely many Peano curves to get an infinite continuous Peano curve, which runs through every point of the plane $\mathbb{R}^2$. The simplest possibility is shown on Figure 5.3. The double spiral indicates the ordering of the subsquares, such that every endpoint of one chosen diagonal of one subsquare is identical with the initial point of the chosen diagonal in the next subsquare. This construction is obviously possible.

Another extension is to a space filling curve $\mathbf{f} : [0,1] \to [0,1]^3$ in the 3-dimensional space $\mathbb{R}^3$. The basic construction is easily illustrated by *Rubik's cube*. We consider the unit cube and start by moving along the diagonal from the point $(0,0,0)$ to the point $(1,1,1)$. Then apply the 2-dimensional Peano construction, described above with the only modification that we instead always use the 3-dimensional diagonal. So if a diagonal of a subcube starts from a point at the lower face of a layer of Rubik's cube, then it will always end at the upper face of this layer, and *vice versa*. Since we move through an odd number of subcubes, namely 9, and start at a point on the lower face, we shall end at the upper face of the lowest layer of Rubik's cube. This is, however, the lower face of the middle layer, so the process can be repeated, until we end at the upper face of the middle layer, which also is the lower face of the upper layer, and the procedure is then obvious, because we just repeat this construction of al subcubes of the next level of division, etc.. The proof itself is just a repetition of the previously given concerning the plane Peano curve with obvious modifications.

It is left as an exercise to the reader to find an ordering of all unit cubes in $\mathbb{R}^3$, such that we by this construction obtain a continuous curve, which runs through all points of $\mathbb{R}^3$.

The extension to a similar construction to general $\mathbb{R}^N$ is almost obvious. It is also left to the reader.

An immediate consequence of Theorem 5.2 is that it is possible *continuously* to map a *line segment* (of area zero) onto a set of *positive area.* One might here object that the parameter interval $I = [0,1]$ has the 1-dimensional measure 1 in the space $\mathbb{R}$, and that we are dealing with two different spaces. Unfortunately, it is not so simple that images of continuous maps and of positive area stem from sets of positive area in the parameter interval. A counterexample is given in Section 5.4, in which we construct a curve, where a nullset of the parameter interval is mapped continuously onto a set of area 1. So measure is *not* preserved by continuous functions. We must add some extra conditions to obtain such a desirable result.

As a preparation of Section 5.3 we repeat the fact that already the function $\mathbf{f}_1$ has double points, thus also following functions $\mathbf{f}_n$, as well as the limit function $\mathbf{f}$. In particular, $\mathbf{f}$ is not injective. It is possible to prove that if the range $\mathbf{f}([0,1])$ of a continuous curve contains interior points, then $\mathbf{f}$ cannot be injective. Notice that this is by no way contradicting the construction in Section 5.3. Although *Osgood's curve* has a positive area, it cannot have inner points, because it is injective, so its interior is the empty set.

5.3 Osgood's curve

In this section we show that using a simple modification of the construction of Peano's curve we are able to define an injective continuous curve of positive area. More precisely, to every $\varepsilon > 0$ there exists a continuous curve $\mathbf{f} : [0,1] \to [0,1]^2$, such that

a) The map $\mathbf{f} : [0,1] \to [0,1]^2$ is injective, i.e. if $\mathbf{f}(s) = \mathbf{f}(t)$ for $s, t \in [0,1]$, then $s = t$.

b) $\text{area}(\mathbf{f}([0,1]) \geq 1 - \varepsilon$.

We shall here more or less follow *Osgood* [42], but since much of the construction is the same as for the *Peano curve* in Section 5.2, we may at some points be brief.

Let $\varepsilon \in]0,1[$ be given, and put $\varepsilon_n = 2^{-n}\varepsilon$. Then

$$\sum_{n=1}^{+\infty} \varepsilon_n = \sum_{n=1}^{+\infty} \frac{1}{2^n}\,\varepsilon = \varepsilon.$$

In the following iterative process we prescribe the constant $\varepsilon_n > 0$ to the n-th level.

Divide $[0,1]$ into 17 subintervals of equal length,

$$\left[\frac{k-1}{17}, \frac{k}{17}\right], \qquad \text{for } k = 1, \dots, 17.$$

The unit square is at first divided into 9 subsquares of equal size like in Peano's construction, but since the first level has been given the constant $\varepsilon_1 = \varepsilon/2$, we furthermore cut out a "tartan" of $[0,1]^2$ along the division lines of a total area of $\varepsilon_1 = \varepsilon/2$, such that we get 9 smaller squares, each of which of area $\frac{1}{9}\left(1 - \frac{\varepsilon}{2}\right)$, and such that these remaining closed subsquares are mutually disjoint. We order the subsquares according to Peano's pattern, through the nine odd numbers, 1, 3, 5, ..., 15, 17.

We map the interval $\left[\frac{k-1}{17}, \frac{k}{17}\right]$ for k odd into the diagonal of subsquare number k as shown on Figure 5.4, like in the construction of Peano's curve. If k is even, this part of the curve is the connection line segment between two squares like on Figure 5.4.

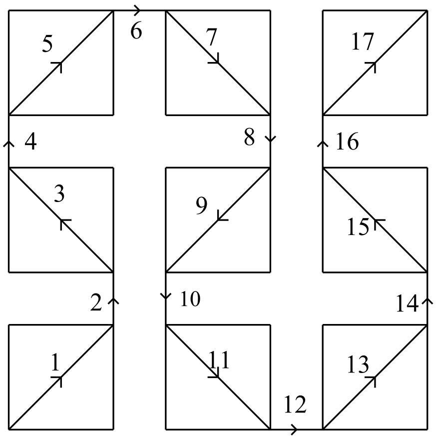

Figure 5.4: *The ordering of the subsquares for* $\mathbf{f}_1 : [0,1] \to [0,1]^2$ *in the construction of Osgood's curve.*

By this process we obtain that all double points in Peano's construction are removed, so $\mathbf{f}_1$ does not have any double point. On the other hand, we have in the following lost an area $\le \varepsilon_1$.

In the construction of $\mathbf{f}_2 : [0,1] \to [0,1]^2$ we remove a "tartan" from each of the nine subsquares of Figure 5.4, of a total area ε_2, so we remove an area of the size $\dfrac{\varepsilon}{9 \cdot 4} = \dfrac{\varepsilon}{36}$ from each subsquare.

We repeat the process already given for $\mathbf{f}_1$ on each of the subsquares following the same pattern as for Peano's curve, with the following modification. If for some $t \in [0,1]$, the point $\mathbf{f}_1(t)$ lies on a line segment between two of the subsquares, then

$$\mathbf{f}_1(t) = \mathbf{f}_2(t) = \mathbf{f}_3(t) = \cdots = \mathbf{f}(t) \qquad \text{in the limit.}$$

In general, if $\mathbf{f}_N(t)$ lies on such a connecting line segment, then

$$\mathbf{f}_N(t) = \mathbf{f}_{N+1}(t) = \mathbf{f}_{N+2}(t) = \cdots = \mathbf{f}(t) \qquad \text{in the limit.}$$

Proceed in this way.

In the n-th step of this iteration we remove a "tartan" from each of the 3^n subsquares of a total area of ε_n, and then construct the continuous curve $\mathbf{f}_n : [0,1] \to [0,1]^2$, following the iterative description above.

We prove almost in the same way as in Section 5.2 that $(\mathbf{f}_n)$ is uniformly convergent. The only modification is that on the connecting line segments this uniform convergence is trivial, because the sequence $(\mathbf{f}_n(t))$ for fixed t is constant eventually. We therefore conclude that the limit curve $\mathbf{f} : [0,1] \to [0,1]^2$ exists and is a continuous function.

The range of the limit curve $\mathbf{f}([0,1])$ consists of all points of the square $[0,1]^2$, which do not belong to any of the removed "tartans". Hence,

$$\operatorname{area}(\mathbf{f}([0,1])) = 1 - \sum_{n=1}^{+\infty} \varepsilon_n = 1 - \varepsilon > 0.$$

Finally, we shall prove that the map $\mathbf{f}$ is injective. If $t_1 < t_2$, then there exists an $n \in \mathbb{N}$, such that

$$t_2 - t_1 > \frac{1}{17^n}. \tag{5.2}$$

If $\mathbf{f}(t_1)$ and $\mathbf{f}(t_2)$ belong to the same line segment between two squares, then they must be different, $\mathbf{f}(t_1) \neq \mathbf{f}(t_2)$, because the line segment does not contain double points.

If $\mathbf{f}(t_1)$ and $\mathbf{f}(t_2)$ belong to different line segments, or one point a line segment and the other one some square from level n, or two different subsquares from level n, then obviously $\mathbf{f}(t_1) \neq \mathbf{f}(t_2)$.

The remaining possibility is that $\mathbf{f}(t_1)$ and $\mathbf{f}(t_2)$ belong to the same k-th square from level n. This is, however, not possible, because this would require that $t_1,\ t_2 \in \left[\frac{k-1}{17^n}, \frac{k}{17^n}\right]$, contradicting the assumption (5.2).

We conclude that $\mathbf{f}$ is injective. Summing up we have proved

Theorem 5.3 *To every given $\varepsilon \in]0,1[$ there exists a continuous and injective curve $\mathbf{f} : [0,1] \to [0,1]^2$, such that*

$$\operatorname{area}(\mathbf{f}([0,1])) \geq 1 - \varepsilon.$$

The curve constructed above is called *Osgood's curve*. It can be used to illustrate the difficulties of the following classical theorem. We shall first introduce some terminology. Every injective curve, i.e. without double points, is called a *simple curve*. A *Jordan curve* is a simple curve, except for its coinciding endpoints, i.e. a Jordan curve is a closed curve. Every Jordan curve in the plane divides $\mathbb{R}^2$ into three mutually disjoint subsets, namely the Jordan curve itself, the open interior of the curve, and the open exterior of the curve. The classical result is the following theorem,

Jordan's curve theorem. *The complementary set of a Jordan curve in the plane has a bounded and an unbounded connected component.*

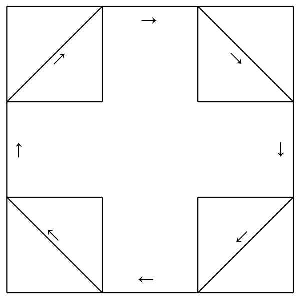

Figure 5.5: *Sketch of a nontrivial Jordan curve of positive area. The four copies of Osgood's curve are indicated by squares with a diagonal to show the direction of the curve.*

The complications of Jordan's curve theorem may be illustrated in the following way. If we combine four copies of Osgood's curve as indicated on Figure 5.5 and with the connecting line segments, we obtain a Jordan curve of positive area contained in the unit square $[0,1]^2$ and of area $1-\varepsilon$, where $\varepsilon \in]0,1[$ can be any constant. One can easily imagine that the proof of Jordan's curve theorem must be difficult, because how do we decide whether a point in $[0,1]^2$ not lying on the Jordan curve itself is an interior or an exterior point? Notice that the Jordan curve is closed, so the complementary set is open. In particular, the bounded component, i.e. the interior of the curve, is open, hence even curve connected. This means that every two points inside the Jordan curve can be connected by a continuous curve, which does not intersect the Jordan curve. When one considers the construction of Osgood's curve it does not look like an easy job to connect two such points in the interior.

As already mentioned above, another strange observation is that one can define a Jordan curve in $[0,1]^2$, such that its area is $\geq 1-\varepsilon$ for any given $\varepsilon \in]0,1[$. By a remark in Section 5.2 no *injective* curve can contain interior points, so the interior of the Jordan curve itself is empty, and it must be the boundary of the much smaller set in area of all points lying inside the curve.

5.4 A space filling curve, which is almost everywhere almost nowhere, and vice versa

The headline of this section is more precisely interpreted in the following way

Theorem 5.4 *There exist a continuous curve $\varphi : [0,1] \to [0,1]^2$ with range $\varphi([0,1]) = [0,1]^2$, and nullsets $N_1 \subset [0,1]$ and $N_2 \subset [0,1]^2$, such that*

$$\varphi\left([0,1] \setminus N_1\right) = N_2 \qquad \textit{and} \qquad \varphi\left(N_1\right) \supseteq [0,1]^2 \setminus N_2.$$

Furthermore, φ is differentiable almost everywhere with the derivative $\mathbf{0}$.

PROOF. Let $\Phi : [0,1] \to [0,1]^2$ be *Peano's curve* as described in Section 5.2, and let f denote the *Cantor function* as constructed in Section 3.3. We define a continuous curve $\varphi : [0,1] \to [0,1]^2$ by the composition

$$\varphi(x) := \Phi(f(x)) \qquad \text{for } x \in [0,1].$$

The function $f : [0,1] \to [0,1]$ is surjective, so it follows from Section 5.1 that

$$\varphi([0,1]) = \Phi([0,1]) = [0,1]^2,$$

proving that the range of φ is the unit square.

Let C denote the *Cantor set* of Section 3.2. If $x \in [0,1] \setminus C$, then we can find an open interval I, such that

$$x \in I \subset [0,1] \setminus C.$$

The *Cantor function* is constant on I, hence $\Phi(y)$ is constant for $y \in I$. In particular, φ is differentiable almost everywhere with the derivative

$$\varphi'(x) = \mathbf{0}, \qquad \text{for } x \in [0,1] \setminus C.$$

It follows from the construction of the *Cantor function* that the range of $[0,1] \setminus C$ is countable, hence a nullset, so the curve is "almost nowhere" in the unit square, almost everywhere in the parameter interval. It therefore follows that

$$\varphi(C) \supseteq \varphi([0,1]) \setminus \varphi([0,1] \setminus C).$$

This means that the *Cantor set* C, which is a nullset in the parameter interval, is mapped onto a set $\subseteq [0,1]^2$ of area 1, and the theorem is proved. $\square$

The traditional error of jumping to conclusions is of course, "when $\varphi' = \mathbf{0}$ almost everywhere, then the range is a nullset". The example above shows that this claim is far from true.

6 Nonmeasurable sets

6.1 Introduction

The nonmeasurable sets are very strange objects, which will be demonstrated in this chapter. We start in Section 6.2, where we show the standard example of a non-measurable set. We are in this construction forced to apply the dubious *axiom of choice*, which implies that it cannot be constructed in a physical way. For the same reason it is not possible to illustrate any nonmeasurable set on a figure. In other words, nonmeasurable sets in practice only exist in the minds of mathematicians. Whenever we consider a theory reflecting something from the real world, then the functions are *always* measurable. We mention here *Hausdorff's paradox*, which will be given in Section 6.4. It implies that *if* one *explicitly* could construct (i.e. physically) nonmeasurable sets, then one would be able cut the unit ball into 10 non-measurable pieces, where the four of them could be joined together to form another full unit ball, just as the remaining six can be joined together to make up a full unit ball, so we would have a means to double everything without making holes, i.e. without leaving out any point. Practice from Physics shows that this is not possible in the real world.

6.2 A nonmeasurable set

We shall in this section "construct" a subset of $\mathbb{R}$, which is not Lebesgue measurable. The doubtful of this "construction" is that it applies the plausible *axiom of choice*. However, even if the *axiom of choice* is intuitively "obvious", its consequences are *not!* We shall later give some other weird examples relying only on the *axiom of choice*, such like dividing the solid unit ball into jigsaw pieces, which are then put together in a different way to form *two* solid unit balls without missing any point!

For the time being we shall assume the *axiom of choice*, cf. Section 6.3.

Let $\sim$ denote the *relation* in $\mathbb{R}$, which is defined by

$$x \sim y \qquad \text{if and only if} \qquad x - y \in \mathbb{Q}.$$

Then $\sim$ is *reflexive*, i.e. $x \sim x$ (in fact, $x - x = 0 \in \mathbb{Q}$) and *symmetric* (in fact, $x \sim y$ implies that $y - x \in \mathbb{Q}$, hence $y \sim x$), and *transitive*. In fact, if $x \sim y$ and $y \sim z$, then $x \sim z$, because $x - y \in \mathbb{Q}$ and $y - z \in \mathbb{Q}$ implies that $x - z = (x - y) + (y - z) \in \mathbb{Q}$.

A reflexive, symmetric and transitive relation $\sim$ on $\mathbb{R}$ is an *equivalence relation* on $\mathbb{R}$, so $\mathbb{R}$ can be written as a *disjoint* union of equivalence classes. A point $x \in \mathbb{R}$ belongs to the equivalence class

$$X := \{y \in \mathbb{R} \mid x \sim y\} = \{y \in \mathbb{R} \mid y = x + q \text{ for some } q \in \mathbb{Q}\} = \{x\} + \mathbb{Q}.$$

Every $x \in \mathbb{R}$ belongs to some class X, and since two classes, X and Y, with respect to $\sim$ are either disjoint or identical, we can write

$$\mathbb{R} = \bigcup_{i \in I} X_i,$$

where I is some index set, where each $X_i \neq \emptyset$, and where the X_i are mutually disjoint.

The set $\mathbb{Q}$ is dense in $\mathbb{R}$, and since $X_i = \{x_i\} + \mathbb{Q}$ for some $x_i \in \mathbb{R}$, every X_i is also dense in $\mathbb{R}$. We therefore conclude that

(6.1) To every $x \in I$ there exists an $x_i \in X_i \cap [0, 1]$.

At this step we apply the *axiom of choice*. Choose *one* representative x_i for each class X_i, such that (6.1) is fulfilled, and define the set

$$A := \{x_i \mid i \in I\} \subset [0,1]. \tag{6.2}$$

We mention here that the Lebesgue measure is translation invariant. This means that if the measure $|A|$ of a set A exists, and if we define

$$\{c\} + A := \{c + x \mid x \in A\},$$

then $\{c\} + A$ is also Lebesgue measurable, and $|\{c\} + A| = |A|$. Furthermore, if $B = \bigcup_n A_n$, where the A_n are mutually disjoint and Lebesgue measurable, then B is also Lebesgue measurable, and its measure is given by

$$|B| = \sum_n |A_n|\,,$$

where we allow the value $+\infty$. It is possible to *prove* these intuitive properties of the Lebesgue measure.

We shall prove that the set A given by (6.2) is not Lebesgue measurable. We have the following three possibilities, of which precisely one must hold,

a) A is Lebesgue measurable, and $|A| = 0$.

b) A is Lebesgue measurable, and $|A| > 0$.

a) A is not Lebesgue measurable.

We shall in the following show that neither **a)** nor **b)** is possible, inferring that we are in case c).

a) Assume that $|A| = 0$, i.e. A is a Lebesgue nullset. Then we can write $\mathbb{R}$ as a disjoint union in the following way,

$$\mathbb{R} = \bigcup_{i \in I} X_i = \bigcup_{i \in I} \{x_i + q \mid q \in \mathbb{Q}\} = \bigcup_{q \in \mathbb{Q}} \bigcup_{i \in I} \{x_i + q\} = \bigcup_{q \in \mathbb{Q}} (A + \{q\}).$$

The Lebesgue measure is translation invariant, thus $|A + \{q\}| = |A| = 0$, and we have written $\mathbb{R}$ as a countable union of nullsets. This implies the contradiction that $\mathbb{R}$ itself is a nullset. Hence, case a) is not possible.

b) Then assume that A is Lebesgue measurable and $|A| > 0$. Consider the set $B \subseteq [0,2]$ defined by

$$B := \bigcup_{q \in \mathbb{Q} \cap [0,1]} (A + \{q\}) = \bigcup_{n=1}^{+\infty} (A + \{q_n\}) = \bigcup_{n=1}^{+\infty} A_n,$$

where we have written $\mathbb{Q} \cap [0,1] = \{q_n \mid n \in \mathbb{N}\}$ as a sequence, and put $A_n := A + \{q_n\}$. It follows from the definition of A that if $m \neq n$, then $A_m \cap A_n = \emptyset$. Finally, since $B \subseteq [0,2]$ and the A_n are mutually disjoint, we get the following contradiction

$$2 \geq |B| = \sum_{n=1}^{+\infty} |A_n| = |A| \sum_{n=1}^{+\infty} 1 = +\infty.$$

We therefore conclude that A is not Lebesgue measurable, so we have proved

Theorem 6.1 *The set A given by (6.2) is not Lebesgue measurable.*

6.3 The axiom of choice

We have already pointed out that the questionable of the construction is the application of the *axiom of choice*, when we from (6.1) *choose* an "explicit" $x_i \in X_i$ from the equivalence class X_i. This assumption is extremely plausible, but nevertheless it can neither be proved nor disproved within the present system of axioms. We must either add this property or its negation as an extra axiom. We formulate

The axiom of choice. *Let $M \neq \emptyset$ be any nonempty set, and let $\mathfrak{D}(M)$ denote the system of all subsets of M. There exists a function*

$$u : \mathfrak{D}(M) \setminus \{\emptyset\} \to M,$$

such that

$$u(A) \in A \qquad \textit{for every } A \in \mathfrak{D}(M) \setminus \{\emptyset\}.$$

In other words, there is a function u, which explicitly selects an element from every subset $\neq \emptyset$. This is of course easy to construct, if M has a finite number n of elements, in which case $\mathfrak{D}(M) \setminus \{\emptyset\}$ consists of $2^n - 1$ nonempty subsets. We just line them all up and choose an element in each of them, defining in this way the function u.

If M, however, has infinitely many elements, the method above does no longer apply. It can be proved that $\mathfrak{D}(\mathbb{N})$ has as many elements as the set $\mathbb{R}$ of all real numbers, so it is no longer possible to line them all up in a table. It is even worse in the present case, where we constructed the nonmeasurable set A given by (6.2). We chose $M = \mathbb{R}$ and then tacitly applied the axiom of choice on the large set $\mathfrak{D}(\mathbb{R}) \setminus \{\emptyset\}$. The set $\mathfrak{D}(\mathbb{R})$ contains more elements then there are in $\mathbb{R}$ itself.

It can be proved that if the present system of mathematical axioms is consistent, it remains consistent whether we add the axiom of choice or its negation. Axioms are characterized by this property.

To illustrate what is meant by this we consider another famous axiom, the *parallel axiom* in the Euclidean geometry.

The parallel axiom. *Two parallel straight lines are either coinciding, or they are never intersecting each other.*

Its negation is

The negated parallel axiom. *There exist two different straight lines, which intersect each other,*

The latter axiom is the basic axiom of the non-Euclidean geometry which was developped in the 19th century.

Let us return to the *axiom of choice* in order to illustrate some of the problems involved with it. We have already seen that it implies the *existence* of really nasty sets. We emphasize the word *"existence"*, because there is no *constructive* way (e.g. an iterative method) to create such a nonmeasurable set. They exist in the mind of some mathematicians, but never in the real world.

The *axiom of choice* has far more strange consequences. The most famous is

The well-ordering theorem. *Every nonempty set M can be well-ordered.*

This means that there exists an ordering relation $\prec$ on M, such that every element $a \in M$ which is not the last element, if such an element exists, has a uniquely determined successor $b \in M$, i.e. there is no element $c \in M$, such that $a \prec c \prec b$, i.e. different from a and b and lying between them.

Clearly, $(\mathbb{R}, <)$ with the usual ordering relation $<$ is not well-ordered, and it is not possible to get any idea of how to construct a well-ordering $(\mathbb{R}, \prec)$ of $\mathbb{R}$.

In connection with the *axiom of choice* and the *well-ordering theorem* one usually also mentions *Zorn's lemma*, which mostly is used in mathematical proofs. We shall not formulate *Zorn's lemma*, only mention that the three mentioned statements are mathematically equivalent. Choose any of them as an axiom, and the other two follows as theorems.

Outside the circles of mathematicians it is most convenient just to accept the *axiom of choice*, even if it has some very strange consequences.

As a curiosity we mention that Alan Turing, who knew all the problems with the axiom of choice, in his thesis, which has been reprinted in [4], tried to develop a theory for at least some of the nonmeasurable sets. His ideas were to the author's knowledge never followed up by others, though one never can tell.

6.4 The Hausdorff paradox

In this section we shall assume the *axiom of choice* and show that it implies that the unit sphere in $\mathbb{R}^3$ – denoted S in the following – can be divided into ten disjoint subsets, where four of these by only translations and rotations in space can be put together, such that they form a unit sphere without missing any point, while the remaining six subsets also can be put together, using only translations and rotations in another jigsaw puzzle to form another unit sphere, again without missing any point. If the pieces in this jigsaw puzzle were all measurable sets, this would of course be impossible, because the measures then would add up in the expected way, and when we consider the areas, we would get $4\pi = 4\pi + 4\pi$, or put in another way, $1 = 1 + 1 = 2$, which clearly is nonsense. Here we claim that it is possible to double the area using only isometries by allowing non-measurable sets in the division of S. This is what is called *Hausdorff's paradox.*

The paradox is easily derived from a result, first proved by *Hausdorff* [25] in 1914. We shall here more or less follow the proof given by *Sierpinski* [50]. This proof can also be found in *Sierpinski, Klein, Runge & Dickson* [51].

It is remarkable that the proof only requires a little knowledge of elementary Group Theory, some knowledge of elementary Geometry, and Calculus at high school level. It is, however, very long, and the reader must be painstakingly accurate in the arguments. The *axiom of choice*, cf. Section 6.3, is used *only once* in the proof in an apparently very innocent way, and we nevertheless obtain this paradox.

We shall first introduce some terminology.

Definition 6.1 *A map $T : \mathbb{R}^3 \to \mathbb{R}^3$ is called an* isometry*, if it preserves Euclidean distances, i.e.*

$$\|Tx - Ty\|_2 = \|x - y\|_2, \qquad \text{for all } x,\, y \in \mathbb{R}^3.$$

It should be well-known that an isometry T in $\mathbb{R}^3$ is either a *translation*, or a *rotation*, or a *reflection*, or a composition of the above mentioned simple isometries. We shall not need the reflections in the following, and the translations are only used at the very end of the construction to put the jigsaw pieces together, so the proof will mostly be applications of rotations and their compositions.

Definition 6.2 *Two subsets A and B of $\mathbb{R}^3$ are said to be* congruent*, or* isometric*, if there exists an isometry $T : \mathbb{R}^3 \to \mathbb{R}^3$, such that $B = T(A)$. When A and B are isometric, we write $A \cong B$.*

Obviously the two definitions above apply in any Euclidean space. In this section we shall only use them in $\mathbb{R}^3$.

Definition 6.3 *Two subsets A and B of $\mathbb{R}^3$ are called n-*equivalent*,, and we write $A \underset{n}{=} B$, if there exist $2n$ subsets $A_1, \dots, A_n$ and $B_1, \dots, B_n$, such that*

1) $A = \bigcup_{j=1}^n A_j$ and $B = \bigcup_{j=1}^n B_j$.

2) Each of the two systems $\{A_j \mid j = 1, \dots, n\}$ and $\{B_j \mid j = 1, \dots, n\}$ consists of mutually disjoint sets.

3) $A_j \cong B_j$ for all $j = 1, \dots, n$.

In other words, $A \underset{n}{=} B$, if the two sets each can be divided into n mutually disjoint sets, such that each subset A_j of A is congruent with the subset B_j of B of the same index.

Given two subsets A and B. If $A \underset{n}{=} B$ for some $n \in \mathbb{N}$, we also say that A and B are equivalent by finite decomposition, and we may write $A \overset{f}{=} B$.

Furthermore, we can also derive the following theorem.

Theorem 6.2 *If $A \underset{n}{=} B$ and $B \underset{m}{=} C$, them $A \underset{mn}{=} C$.*

Remark. It is possible to prove that this result cannot be improved. $\Diamond$

Proof of Theorem 6.2. It follows from the assumption $A \underset{n}{=} B$ that there exist decompositions, such that

1) $A = A_1 \cup \cdots \cup A_n$ and $B = B_1 \cup \cdots \cup B_n$.

2) $A_k \cap A_\ell = \emptyset$ and $B_k \cap B_\ell = \emptyset$ for $1 \le k < \ell \le n$,

3) $A_k \cong B_k$ for $k = 1, 2, \dots, n$

Similarly, it follows from $B \underset{m}{=} C$ that there exist decompositions

$$B = B_1' \cup B_2' \cup \cdots \cup B_m' \text{ and } C = C_1 \cup C_2 \cup \cdots \cup C_m, \tag{6.3}$$

such that

$$B_k' \cap B_\ell' = \emptyset \qquad \text{and} \qquad C_k \cap C_\ell = \emptyset \qquad \text{for } 1 \le k < \ell \le m, \tag{6.4}$$

and

$$B_\ell' \cong C_\ell \qquad \text{for } \ell = 1, 2, \dots, m. \tag{6.5}$$

It is natural to define

$$B_{k,\ell} := B_k \cap B_\ell' \qquad \text{for } 1 \le k \le n \text{ and } 1 \le \ell \le m. \tag{6.6}$$

Then by 1) and (6.3) and (6.6),

$$B_k = B_{k,1} \cup B_{k,2} \cup \cdots \cup B_{k,m} \qquad \text{for } 1 \le k \le n, \tag{6.7}$$

and

$$B_k' = B_{1,\ell} \cup B_{2,\ell} \cup \cdots \cup B_{n,\ell} \qquad \text{for } 1 \le \ell \le m. \tag{6.8}$$

Let k, $k_1 \in \{1, \dots, n\}$ and ℓ, $\ell_1 \in \{1, \dots, m\}$. Then it follows from (6.6) and 2) and (6.4) that

$$B_{k,\ell} \cap B_{k_1,\ell_1} = \emptyset, \qquad \text{if either } k \ne k_1 \text{ or } \ell \ne \ell_1. \tag{6.9}$$

Using 3) and (6.7) and 2) and (6.9) we obtain the decomposition

$$A_k = A_{k,1} \cup A_{k,2} \cup \cdots \cup A_{k,m} \qquad \text{for } k = 1, \dots, n. \tag{6.10}$$

It therefore follows for $1 \le k$, $k_1 \le n$ and $1 \le \ell$, $\ell_1 \le m$ that

$$(6.11) \qquad A_{k,\ell} \cap A_{k_1,\ell_1} = \emptyset, \qquad \text{if either } k \neq k_1 \text{ or } \ell \neq \ell_1,$$

and that

$$(6.12) \qquad A_{k,\ell} \cong B_{k\ell}, \qquad \text{when } 1 \le k \le n \text{ and } 1 \le \ell \le m.$$

Analogously, it follows from (6.5) and (6.8) and (6.4) and (6.9) that there exist decompositions

$$(6.13) \qquad C_\ell = C_{1,\ell} \cup C_{2,\ell} \cup \cdots \cup C_{n,\ell} \qquad \text{for } \ell = 1, \dots, m,$$

such that

$$(6.14) \qquad C_{k,\ell} \cap C_{k_1,\ell_1} = \emptyset, \qquad \text{if either } k \neq k_1 \text{ or } \ell \neq \ell_1,$$

and

$$(6.15) \qquad B_{k,\ell} \cong C_{k,\ell} \qquad \text{for } 1 \le k \le n \text{ and } 1 \le \ell \le m.$$

Then by 1), (6.3) and (6.13),

$$(6.16) \qquad A = \bigcup_{k=1}^{n} \bigcup_{\ell=1}^{m} A_{k,\ell} \qquad \text{and} \qquad C = \bigcup_{k=1}^{n} \bigcup_{\ell=1}^{m} C_{k,\ell}.$$

We infer from (6.12) and (6.15) that

$$A_{k,\ell} \cong C_{k,\ell} \qquad \text{for } k = 1, \dots, n \text{ and } \ell = 1, \dots, n,$$

and it finally follows from (6.16 and (6.11) and (6.14) that $A \underset{mn}{=} = C$. □

Corollary 6.1 *The relation $\overset{f}{=}$ is transitive, i.e. if $A \overset{f}{=} B$ and $B \overset{f}{=} C$, then also $A \overset{f}{=} C$.*

We finally formulate

Hausdorff's paradox. *Let S denote the unit sphere of $\mathbb{R}^3$. There exist four disjoint sets A, B, C and D, where D is a countable set, such that*

1) $S = A \cup B \cup C \cup D$,

2) $A \cong B \cong C \cong B \cup C$.

We shall later use *Hausdorff's paradox* to prove that the unit sphere S can be divided into ten subsets, such that

$$S \cong S_1 \cong S_2 \qquad \text{and} \qquad S \underset{10}{=} S_1 \cup S_2,$$

where S_1 and S_2 are two mutually disjoint unit spheres.

Proof. Let Φ be a straight line in the XZ-plane through $\mathbf{0}$, forming the angle $\dfrac{\nu}{2} = 1$ radian with the Z-axis, i.e. $\nu = 2$, cf. Figure 6.1.

Let ψ denote the rotation in the XYZ-space with the Z-axis as its rotational axis, and with the rotational angle $\dfrac{2\pi}{3}$. Similarly, let φ denote the rotation in the XYZ-space with rotational axis Φ and rotational angle π. Finally, I denotes the identical map. Then

$$(6.17) \qquad \varphi^2 = \psi^3 = I.$$

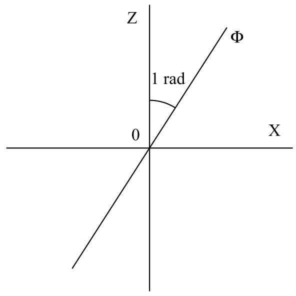

Figure 6.1: *The two rotational axes Z and* Φ *in the proof of Hausdorff's paradox. The angle between them is 1 radian.*

Let $\mathfrak{G}$ be the set of all transformations of $\mathbb{R}^3$, which are obtained by applying φ and ψ a finite number of times. it follows easily from (6.17) that $\mathfrak{G}$ is a group with the usual composition of maps as its composition. It follows also from (6.17) that every element of $\mathfrak{G}$ is generated by the four *fundamental factors*

$$I, \quad \varphi, \quad \psi \quad \text{and} \quad \psi^2,$$

in such a way that φ is never followed by φ (because then $\varphi \circ \varphi = \varphi^2 = I$), and ψ and ψ^2 are never followed by ψ or ψ^2, because the composition then can be reduced,

$$\psi \circ \psi = \psi^2, \qquad \psi \circ \psi^2 = \psi^2 \circ \psi = I, \qquad \psi^2 \circ \psi^2 = \psi.$$

Notice in particular that if $p \in \Phi \setminus \{0\}$ is a point $\neq 0$ on the straight line Φ, then $\varphi(\psi(p)) \neq \psi(\varphi(p))$, so $\varphi \circ \psi \neq \psi \circ \varphi$, and $\mathfrak{G}$ is not an Abelian group.

We divide $\mathfrak{G}$ into classes. We put $\mathfrak{G}_0 := \{I\}$, and let $\mathfrak{G}_n$ denote the class of all transformations from $\mathfrak{G}$, which can be written as precisely n fundamental factors $\neq I$. In particular,

$$\mathfrak{G}_1 = \left\{\varphi, \psi, \psi^2\right\}.$$

The identification of $\mathfrak{G}_n$ goes by induction. Assume that $\mathfrak{G}_n$ is known. Then we find the elements of $\mathfrak{G}_{n+1}$ in the following way. Every element from $\mathfrak{G}_{\mathfrak{n}}$, which starts with a φ, is multiplied by either ψ or ψ^2, from the left. Every element from $\mathfrak{G}_n$, which starts with either ψ or ψ^2, is multiplied by φ from the left. In this way we get

$$\mathfrak{G}_2 = \left\{\varphi\psi, \varphi\psi^2, \psi\varphi, \psi^2\varphi\right\}$$

and

$$\mathfrak{G}_3 = \left\{\varphi\psi\varphi, \varphi\psi^2\varphi, \psi\varphi\psi, \psi\varphi\psi^2, \psi^2\varphi\psi, \psi^2\varphi\psi^2\right\},$$

etc..

The next step is a fairly long proof of the claim that all elements of

$$\mathfrak{G} = \bigcup_{n=0}^{+\infty} \mathfrak{G}_n$$

are mutually disjoint, so $\{\mathfrak{G}_n \mid n \in \mathbb{N}_0\}$ is a division of $\mathfrak{G}$ into mutually disjoint classes.

With the exception of I and $\gamma = \varphi$, every other element of $\mathfrak{G}$ can be written in one of the following four possible ways,

$$(a) \qquad \alpha = \psi^{m_1}\varphi\psi^{m_2}\varphi\cdots\psi^{m_n}\varphi,$$

$$(b) \qquad \beta = \varphi\psi^{m_1}\varphi\psi^{m_2}\varphi\cdots\psi^{m_n},$$

$$(c) \qquad \gamma = \varphi\psi^{m_1}\varphi\psi^{m_2}\varphi\cdots\psi^{m_n}\varphi,$$

$$(d) \qquad \delta = \psi^{m_1}\varphi\psi^{m_2}\varphi\cdots\psi^{m_n},$$

where $n \in \mathbb{N}$ and $m_1, \dots, m_n \in \{1, 2\}$.

Let $\mathbf{N} \in \mathbb{R}^3$ be any point in space of the coordinates (x, y, z). Then $\psi(\mathbf{N}) = (x', y', z')$. Since ψ is a rotation of the angle $\dfrac{2\pi}{3}$ with respect to the Z-axis, we describe ψ in coordinates by

$$(\psi) \qquad \begin{cases} x' = -\dfrac{1}{2}\,x - \dfrac{\sqrt{3}}{2}\,y, \\[2ex] y' = \dfrac{\sqrt{3}}{2}\,x + \dfrac{1}{2}\,y, \\[2ex] z' = z. \end{cases}$$

Considering $\psi^2 = \psi^{-1}$, which is a rotation on $-\dfrac{2\pi}{3}$ with respect to the Z-axis, we get similarly, if we let $\psi^2(\mathbf{N})$ have the coordinates (x', y', z'),

$$(\psi^{-1}) \qquad \begin{cases} x' = -\dfrac{1}{2}\,x + \dfrac{\sqrt{3}}{2}\,y, \\[2ex] y' = -\dfrac{\sqrt{3}}{2}\,x + \dfrac{1}{2}\,y, \\[2ex] z' = z. \end{cases}$$

Summing up, we see that

$$(\psi^{\pm}) \qquad \begin{cases} x' = -\dfrac{1}{2}\,x \mp \dfrac{\sqrt{3}}{2}\,y, \\[2ex] y' = \pm\dfrac{\sqrt{3}}{2}\,x + \dfrac{1}{2}\,y, \\[2ex] z' = z. \end{cases}$$

We chose in the beginning $\dfrac{\nu}{2} = 1$ radian, so $\nu = 2$. If we put in coordinates $\mathbf{N} = (x, y, z)$ and $\varphi(\mathbf{N}) = (x', y', z')$, then we get for the rotation φ with respect to the axis Φ,

$$(\varphi) \qquad \begin{cases} x' = \{\cos 2\} \cdot x + \{\sin 2\} \cdot z, \\[2ex] y' = -y, \\[2ex] z' = \{\sin 2\} \cdot x + \{\cos 2\} \cdot z. \end{cases}$$

When we compose these maps, it follows from the above that $\psi^{\pm 1}\varphi$ is described by

$$(\psi^{\pm 1}\varphi) \qquad \begin{cases} x' = \dfrac{1}{2}\,\cos 2 \cdot x \pm \dfrac{\sqrt{3}}{2}\,y - \dfrac{1}{2}\,\sin 2 \cdot z, \\[2ex] y' = \mp\dfrac{\sqrt{3}}{2}\,\cos 2 \cdot x - \dfrac{1}{2}\,y \pm \dfrac{\sqrt{3}}{2}\,\sin 2 \cdot z, \\[2ex] z' = \sin 2 \cdot x + \cos 2 \cdot z. \end{cases}$$

Assume that $\alpha \in \mathfrak{G}$ is given in the form (a), i.e.

$$\alpha = (\psi^{m_1}\varphi)\,(\psi^{m_2}\varphi)\cdots(\psi^{m_n}\varphi)\,.$$

Choose $\mathbf{N} = (0, 0, 1)$ and write $\alpha(\mathbf{N}) = (\xi, \eta, \zeta)$. We shall prove that

$$(\alpha) \qquad \begin{cases} \xi = \sin 2 \cdot \left(a \cdot \cos^{n-1} 2 + \cdots\right), \\[2ex] \eta = \sin 2 \cdot \left(b \cdot \cos^{n-1} 2 + \cdots\right), \\[2ex] \zeta = c \cdot \cos^n 2 + \cdots\,, \end{cases}$$

where $a \cdot \cos^{n-1} 2 + \cdots$, $b \cdot \cos^{n-1} 2 + \cdots$, and $c \cdot \cos^{n} 2 + \cdots$, are polynomials in $\cos 2$ of degrees $n-1$, $n-1$ and n, resp., and where the coefficients are polynomials in $\lambda := -\dfrac{1}{2}$ and $\mu := \dfrac{\sqrt{3}}{2}$ of integer coefficients.

As mentioned above, $\psi^2 = \psi^{-1}$, so it follows for $n = 1$ that $\alpha = \psi^{\pm}\varphi$. We conclude that $\psi^{\pm 1}\varphi$ maps $\mathbf{N} = (0, 0, 1)$ into

$$\begin{cases} \xi = -\dfrac{1}{2} \sin 2 = \lambda \sin 2, \\ \eta = \pm\dfrac{\sqrt{3}}{2} \sin 2 = \pm\mu \sin 2, \\ \zeta = \cos 2, \end{cases}$$

which are of the right structure.

Then assume that (α) holds for some $n \in \mathbb{N}$, and let $\alpha' = \psi^{\pm 1}\varphi\alpha$. When we compose $\left(\psi^{\pm 1}\varphi\right)$ and (α), we obtain the following formulæ for $\alpha'(\mathbf{N}) = (x', y', z')$, using the shorthand $\lambda = -\dfrac{1}{2}$ and $\mu = \dfrac{\sqrt{3}}{2}$ in $\left(\psi^{\pm 1}\varphi\right)$,

$$\begin{aligned}
x' &= -\sin 2 \cdot \left\{a \cdot \cos^{n-1} 2 + \cdots\right\} \lambda \cos 2 \pm \sin 2 \cdot \left\{b \cdot \cos^{n-1} 2 + \cdots\right\} + \{c \cdot \cos^n 2 + \cdots\} \lambda \sin 2 \\
&= \sin 2 \cdot \{(c-a)\lambda \cos^n + \cdots\},
\end{aligned}$$

$$\begin{aligned}
y' &= \mp \sin 2 \cdot \left\{a \cdot \cos^{n-1} 2 + \cdots\right\} \mu \cos 2 - \sin 2 \left\{b \cos^{n-1} 2 + \cdots\right\} \pm \{c \cdot \cos^n 2 + \cdots\} \mu \sin 2 \\
&= \sin 2 \{\pm (c-a)\mu \cos^n 2 + \cdots\},
\end{aligned}$$

$$\begin{aligned}
z' &= \sin^2 2 \left\{a \cdot \cos^{n-1} 2 + \cdots\right\} + \{c \cdot \cos^n 2 + \cdots\} \cos 2 \\
&= (c-a) \cos^{n+1} 2 + \cdots,
\end{aligned}$$

hence

$$(\alpha') \qquad \begin{cases} x' = \sin 2 \cdot \{a' \cos^n 2 + \cdots\}, \\[2mm] y' = \sin 2 \cdot \{b' \cos^n 2 + \cdots\}, \\[2mm] z' = c' \cos^{n+1} 2 + \cdots, \end{cases}$$

where

$$(6.18) \qquad a' = (c-a)\lambda, \qquad b' = \pm(c-a)\mu, \qquad c' = c - a.$$

The claim then follows by induction for every α of the structure given by (a).

We notice that it follows from the shorthand $\lambda = -\dfrac{1}{2}$ and (6.18) that

$$c' - a' = (c-a)(1-\lambda) = \frac{3}{2}(c-a),$$

hence by recursion at level n,

$$c - a = \left(\frac{3}{2}\right)^n.$$

Clearly, $\cos 2$ is a transcendental number, i.e. it is not a root in any polynomial of algebraic coefficients. Thus it follows from the above that the point $\mathbf{N} = (0,0,1)$ is never mapped into itself by any transformation of the form (a), which means that no transformation of type (a) is the identity.

The next step of the proof is to show that the identity cannot be of any of the types (b), (c), (d) either. We here give an indirect proof.

Assume that some $\beta \in \mathfrak{G}$ of type (b) is the identity. This means that

$$\beta = \varphi\psi^{m_1}\varphi\psi^{m_2} \cdots \varphi\psi^{m_n} = I.$$

It follows from $\varphi\beta\varphi = \varphi I \varphi = \varphi^2 = I$ that

$$I = \varphi\beta\varphi = \psi^{m_1}\varphi\psi^{m_2}\varphi \cdots \varphi\psi^{m_n}\varphi,$$

so $I = \varphi\beta\varphi$ is of type (a), which we previously proved was not possible. Hence, no transformation of type (b) is equal to the identity I.

Then assume that some $\delta_n \in \mathfrak{G}$ of type (d) is the identity, i.e.

$$\delta_n = \psi^{m_1}\varphi\psi^{m_2}\varphi\cdots\varphi\psi^{m_n} = I.$$

Then

$$\psi^{3-m_1}\delta_n\psi^{m_1} = \psi^{3-m_1}I\psi^{m_1} = \psi^3 = I,$$

so

$$\sigma := \psi^{3-m_1}\delta_n\psi^{m_1} = \psi^3\varphi\psi^{m_2}\varphi\cdots\varphi\psi^{m_n}\psi^{M-1} = \varphi\psi^{m_2}\cdots\varphi\psi^{m_n+m_1} = I.$$

It follows from m_1, $m_n \in \{1,2\}$, that $m_1 + m_n \in \{2,3,4\}$. If $m_1 + m_n \in \{2,4\}$, then $\sigma = I$ is of type (b), which is not possible. If instead $m_1 + m_n = 3$, then

$$\sigma = \varphi\psi^{m_2}\varphi\cdots\psi^{m_{n-1}}\varphi,$$

and

$$\varphi\sigma\varphi = \psi^{m_2}\varphi\cdots\varphi\psi^{m_n-1} = \delta_{n-2} = I,$$

so $\varphi\sigma\varphi = \delta_{n-2} = I$ is again of type (d), only with fewer factors. Hence, by recursion,

$$\text{either } \delta_1 = \psi^{m_1} = I \qquad \text{or} \qquad \delta_2 = \psi^{m_1}\varphi\psi^{m_2} = I,$$

depending on whether n is an odd or an even number. When $m_1 \in \{1,2\}$, then obviously $\psi^{m_1} \neq I$, so n cannot be an odd number.

We turn to the other possibility,

$$\delta_2 = \psi^{m_1}\varphi\psi^{m_2} = I.$$

In this case

$$\varphi = \psi^{-m_1-m_2} \in \left\{I, \psi, \psi^2\right\},$$

which is not true. Hence, the assumption that I is of type (d) is wrong, and all transformations of type (d) are different from the identity I.

Finally, we assume that some transformation γ of type (c) is the identity, i.e.

$$\gamma = \varphi\psi^{m_1}\varphi\psi^{m_2}\cdots\varphi\psi^{m_n}\varphi = I.$$

Then

$$I = \varphi\gamma\varphi = \psi^{m_1}\varphi\psi^{m_2}\cdots\varphi\psi^{m_n},$$

and we have written I as a transformation of type (d), which is not possible according to the above.

Summing up, we have proved that I cannot be written in any of the forms (a), (b), (c), (d). This implies that they are all mutually different. In fact, assume that $\sigma = \tau$. Then $\sigma\tau^{-1} = I$. If $\tau = I$, then also $\sigma = I$, and neither σ nor τ can be written in one of the forms (a)–(d).

Then assume that $\tau \neq I$, thus also $\sigma = \tau \neq I$. We have

$$\begin{cases} \sigma = \varphi^k\psi^{m_1}\varphi\psi^{m_2}\cdots\varphi\psi^{m_n}\varphi^\ell, \\ \tau = \varphi^r\psi^{q_1}\varphi\psi^{q_2}\cdots\varphi\psi^{q_p}\varphi^s, \end{cases} \tag{6.19}$$

where $m_1, \dots m_n$, $q_1, \dots, q_p \in \{1,2\}$, and where k, ℓ, r, $s \in \{0,1\}$. It follows from $\sigma\tau^{-1} = I$ that

$$\varphi^k\psi^{m_1}\varphi\psi^{m_2}\cdots\varphi\psi^{m_n}\varphi^\ell\varphi^{-s}\psi^{-q_p}\varphi\psi^{-q_{p-1}}\cdots\varphi\psi^{-q_1}\varphi^{-r} = I.$$

The left hand side is equal to i, so it cannot be of any of the types (a)–(d). The only difference from (a)–(d) is the inner composition $\varphi^\ell\varphi^{-s}$, so we must have $\varphi^\ell\varphi^{-s} = I$, from which $\ell = s$. Then by reduction, $\psi^{m_n}I\psi^{-q_p} = \psi^{m_n-q_p} = I$, etc., and we conclude by recursion that

$$\ell = s, \quad p = n, \quad m_n = q_n,\, m_{n-1} = q_{n-1}, \dots, m_1 = q_1, \quad k = r.$$

It follows from (6.19) that σ and τ have the same expansions as products of fundamental factors.

Summing up, we have proved that every element of $\mathfrak{G} \setminus \{I\}$ has an unique expansion as a product of fundamental factors. This proves that all elements of $\mathfrak{G} = \bigcup_{n=0}^{+\infty} \mathfrak{G}_n$ are mutually different.

We show in the next step that $\mathfrak{G}$ can be split into three disjoint subsets

$$\mathfrak{G} = \mathfrak{G}^{(1)} \cup \mathfrak{G}^{(2)} \cup \mathfrak{G}^{(3)},$$

such that

$$\varphi\mathfrak{G}^{(1)} = \mathfrak{G}^{(2)} \cup \mathfrak{G}^{(3)}, \qquad \psi\mathfrak{G}^{(1)} = \mathfrak{G}^{(2)}, \qquad \psi^2\mathfrak{G}^{(1)} = \mathfrak{G}^{(3)}, \tag{6.20}$$

where we write

$$\varphi\mathfrak{G}^{(1)} := \left\{\varphi\sigma \mid \sigma \in \mathfrak{G}^{(1)}\right\},$$

and analogously in the other cases.

The proof of (6.20) is not difficult, but it is also fairly long. The idea is to split each $\mathfrak{G}_n$ into three disjoint subsets $\mathfrak{G}_n^{(1)}$, $\mathfrak{G}_n^{(2)}$, $\mathfrak{G}_n^{(3)}$, which is done inductively in the following way.

First we put

$$\mathfrak{G}_0^{(1)} := \{I\}, \qquad \mathfrak{G}_0^{(2)} := \emptyset, \qquad \mathfrak{G}_0^{(3)} := \emptyset.$$

Then assume that we already have defined $\mathfrak{G}_n^{(1)}$, $\mathfrak{G}_n^{(2)}$ and $\mathfrak{G}_n^{(3)}$ for some $n \in \mathbb{N}_0$. It follows from the above that to every $\sigma \in \mathfrak{G}_{n+1}$ there is an unique $\varrho \in \mathfrak{G}_n$, such that one of the following three possibilities occurs,

1) $\sigma = \varphi\varrho$, where the first factor of ϱ is either ψ or ψ^2,

2) $\sigma = \psi\varrho$, where the first factor of ϱ is φ,

3) $\sigma = \psi^2\varrho$, where the first factor of ϱ is φ.

Then define $\mathfrak{G}_{n+1}^{(1)}$, $\mathfrak{G}_{n+1}^{(2)}$, $\mathfrak{G}_{n+1}^{(3)}$ in the following way,

1a) If $\sigma = \varphi\varrho$ and $\varrho \in \mathfrak{G}_n^{(1)}$, then put $\sigma \in \mathfrak{G}_{n+1}^{(2)}$,

1b) If $\sigma = \varphi\varrho$ and $\varrho \in \mathfrak{G}_n^{(2)} \cup \mathfrak{G}_n^{(3)}$, then we put $\sigma \in \mathfrak{G}_{n+1}^{(1)}$,

2a) If $\sigma = \psi\varrho$ and $\varrho \in \mathfrak{G}_n^{(1)}$, then we put $\sigma \in \mathfrak{G}_{n+1}^{(2)}$,

2b) If $\sigma = \psi\varrho$ and $\varrho \in \mathfrak{G}_n^{(2)}$, then we put $\sigma \in \mathfrak{G}_{n+1}^{(3)}$,

2c) If $\sigma = \psi\varrho$ and $\varrho \in \mathfrak{G}_n^{(3)}$, then we put $\sigma \in \mathfrak{G}_{n+1}^{(1)}$,

3a) If $\sigma = \psi^2\varrho$ and $\varrho \in \mathfrak{G}_n^{(1)}$, then we put $\sigma \in \mathfrak{G}_{n+1}^{(3)}$,

3b) If $\sigma = \psi^2\varrho$ and $\varrho \in \mathfrak{G}_n^{(2)}$, then we put $\sigma \in \mathfrak{G}_{n+1}^{(1)}$,

3c) If $\sigma = \psi^2\varrho$ and $\varrho \in \mathfrak{G}_n^{(3)}$, then we put $\sigma \in \mathfrak{G}_{n+1}^{(2)}$.

When we prove (6.20) we shall repeatedly refer back to these eight possibilities. At this point we only notice that because $\mathfrak{G}_n^{(1)}$, $\mathfrak{G}_n^{(2)}$, $\mathfrak{G}_n^{(3)}$ are mutually disjoint, $\mathfrak{G}_{n+1}^{(1)}$. $\mathfrak{G}_{n+1}^{(2)}$, $\mathfrak{G}_{n+1}^{(3)}$ are also mutually disjoint, so they are all mutually disjoint.

We put

$$\mathfrak{G}^{(i)} := \bigcup_{n=0}^{+\infty} \mathfrak{G}_n^{(i)} \qquad \text{for } i = 1, 2, 3.$$

This gives us a decomposition of $\mathfrak{G} = \mathfrak{G}^{(1)} \cup \mathfrak{G}^{(2)} \cup \mathfrak{G}^{(3)}$ inbto three disjoint subsets.

We shall prove that this decomposition of $\mathfrak{G}$ satisfies (6.20).

Let $\varrho \in \mathfrak{G}^{(1)}$. Then $\varrho \in \mathfrak{G}_n^{(1)}$ for some $n \in \mathbb{N}_0$. If $n = 0$, then $\varrho = I$ is the only possibility, and we get

$$(6.21) \qquad \varphi\varrho = \varphi \in \mathfrak{G}_1^{(2)}, \qquad \psi\varrho = \varphi \in \mathfrak{G}_1^{(2)} \quad \text{and} \quad \psi^2\varrho = \psi^2 \in \mathfrak{G}_1^{(3)},$$

where we have applied 1a), 2a) and 3a). The result is satisfying (6.20).

It follows from $\mathfrak{G}_1 = \left\{\varphi, \psi, \psi^2\right\}$ and (6.21) that $\mathfrak{G}_1^{(1)} = \emptyset$. Therefore, if $n \neq 0$ and $\varrho \in \mathfrak{G}_n^{(1)}$ we must have $n \geq 2$. Then there exists a $\varrho_1 \in \mathfrak{G}_{n-1}$, such that one of th following three possibilities is fulfilled,

I.1) $\varrho = \varphi\varrho_1$ and the first factor of ϱ_1 is either ψ or ψ^2,

I.2) $\varrho = \psi\varrho_1$, and the first factor of ϱ_1 is φ,

I.3) $\varrho = \psi^2\varrho_1$, and the first factor of ϱ_1 is φ.

The Roman numeral I refers to $\varrho \in \mathfrak{G}^{(1)}$. Later we shall let II denote that $\varrho \in \mathfrak{G}^{(2)}$, and III that $\varrho \in \mathfrak{G}^{(3)}$.

Let us consider each of the three possibilities successively.

I.1) We assume that $\varrho = \varphi\varrho_1 \in \mathfrak{G}_n^{(1)}$. It follows from 1b) that $\varrho_1 \in \mathfrak{G}_{n-1}^{(2)} \cup \mathfrak{G}_{n-1}^{(3)}$, hence

$$\varphi\varrho = \varphi^2\varrho_1 = \varrho_1 \in \mathfrak{G}_{n-1}^{(2)} \cup \mathfrak{G}_{n-1}^{(3)}.$$

Analogously, using 2a) and 3a) instead,

$$\pi\varrho \in \mathfrak{G}_{n+1}^{(2)} \qquad \text{and} \qquad \psi^2\varrho \in \mathfrak{G}_{n+1}^{(3)}.$$

I.2) Assume that $\varrho = \psi\varrho_1 \in \mathfrak{G}_n^{(1)}$. It follows from 1a) that $\varphi\varrho \in \mathfrak{G}_{n+1}^{(2)}$, and from 2c) that $\varrho_1 \in \mathfrak{G}_{n-1}^{(3)}$, hence by 3c),

$$\psi\varrho = \pi^2\varrho_1 \in \mathfrak{G}_n^{(2)}.$$

Finally,

$$\psi^2\varrho = \psi^3\varrho_1 = \varrho_1 \in \mathfrak{G}_{n-1}^{(3)}.$$

I.3) Assume that $\varrho = \psi^2\varrho_1 \in \mathfrak{G}_n^{(1)}$. It follows from 3b) that $\varrho_1 \in \mathfrak{G}_{n-2}^{(2)}$, hence by 1a),

$$\varphi\varrho \in \mathfrak{G}_{n+1}^{(2)},$$

and

$$\psi\varrho = \psi^3\varrho_1 = \varrho_1 \in \mathfrak{G}_{n-1}^{(2)},$$

by 3b), and finally by 2b),

$$\psi^2\varrho = \psi^4\varrho_1 = \psi\varrho_1 \in \mathfrak{G}_n^{(3)}.$$

Summing up, we get for $\varrho \in \mathfrak{G}^{(1)}$ that

$$\varphi\varrho \in \mathfrak{G}^{(2)} \cup \mathfrak{G}^{(3)}, \qquad \psi\varrho \in \mathfrak{G}^{(2)} \quad \text{and} \quad \psi^2\varrho \in \mathfrak{G}^{(3)},$$

which implies that

$$(6.22) \qquad \varphi\mathfrak{G}^{(1)} \subseteq \mathfrak{G}^{(2)} \cup \mathfrak{G}^{(3)}, \qquad \psi\mathfrak{G}^{(1)} \subseteq \mathfrak{G}^{(2)}, \qquad \psi^2\mathfrak{G}^{(1)} \subseteq \mathfrak{G}^{(3)},$$

and we have proved one half of (6.20).

In the proof of the other half of (6.20) we treat $\mathfrak{G}^{(2)}$ and $\mathfrak{G}^{(3)}$ similarly.

If $\varrho \in \mathfrak{G}^{(2)}$, then $\varrho \in \mathfrak{G}_n^{(2)}$ for some $n \in \mathbb{N}$, and there exists a $\varrho_1 \in \mathfrak{G}_{n-1}$, such that one of the following three possibilities is fulfilled,

II.1) $\varrho = \varphi\varrho_1$, and the first factor of ϱ_1 is either ψ or ψ^2,

II.2) $\varrho = \psi\varrho_1$, and the first factor of ϱ_1 is φ,

II.3) $\varrho = \psi^2\varrho_1$, and the first factor of ϱ_1 is φ.

We go successively through the three possibilities.

II.1) If $\varrho = \varphi\varrho_1 \in \mathfrak{G}_n^{(2)}$, then it follows from 1a) that $\varrho_1 \in \mathfrak{G}_{n-1}^{(1)}$, hence

$$\varphi\varrho = \varrho_1 \in \mathfrak{G}_{n-1}^{(1)}, \qquad \text{by 1a)},$$

$$\psi\varrho \in \mathfrak{G}_{n+1}^{(3)}, \qquad \text{by 2b)},$$

$$\psi^2\varrho \in \mathfrak{G}_{n+1}^{(1)}, \qquad \text{by 3b)}.$$

II.2) If $\varrho = \psi\varrho_1 \in \mathfrak{G}_n^{(2)}$, then $\varrho_1 \in \mathfrak{G}_{n-1}^{(1)}$ by 2a), and

$$\varphi\varrho \in \mathfrak{G}_{n+1}^{(1)}, \qquad \text{by 1b)},$$

$$\psi\varrho = \psi^2\varrho_1 \in \mathfrak{G}_n^{(3)}, \qquad \text{by 3a)},$$

$$\psi^2\varrho = \varrho_1 \in \mathfrak{G}_{n-1}^{(1)}, \qquad \text{by 2a)}.$$

II.3) If $\varrho = \psi^2\varrho_1 \in \mathfrak{G}_n^{(2)}$, then $\varrho_1 \in \mathfrak{G}_{n-1}^{(3)}$ by 3c), and

$$\varphi\varrho \in \mathfrak{G}_{n+1}^{(1)}, \qquad \text{by 1b)},$$

$$\psi\varrho = \varrho_1 \in \mathfrak{G}_{n-1}^{(3)}, \qquad \text{by 3b)}$$

$$\psi^2\varrho = \psi\varrho_1 \in \mathfrak{G}_n^{(1)}, \qquad \text{by 2c)}.$$

Summing up, it follows in case II that

$$\text{(6.23)} \qquad \varphi\mathfrak{G}^{(2)} \subseteq \mathfrak{G}^{(1)}, \qquad \psi\mathfrak{G}^{(2)} \subseteq \mathfrak{G}^{(3)}, \qquad \psi^2\mathfrak{G}^{(2)} \subseteq \mathfrak{G}^{(1)}.$$

In particular,

$$\text{(6.24)} \qquad \mathfrak{G}^{(2)} = \psi^3\mathfrak{G}^{(2)} \subseteq \psi\mathfrak{G}^1 \qquad \text{and} \qquad \mathfrak{G}^{(2)} \subseteq \varphi\mathfrak{G}^{(1)}.$$

Finally, if $\varrho \in \mathfrak{G}^{(3)}$, then $\varrho \in \mathfrak{G}_n^{(3)}$ for some $n \in \mathbb{N}$. There exists a $\varrho_1 \in \mathfrak{G}_{n-1}$, such that one of the following three possibilities occurs,

III.1) $\varrho = \varphi\varrho_1$, where the first factor of ϱ_1 is either ψ or ψ^2,

III.2) $\varrho = \psi\varrho_1$, where the first factor of ϱ_1 is φ,

III.3) $\varrho = \psi^2\varrho_1$, where the first factor of ϱ_1 is φ.

III.1) If $\varrho = \varphi\varrho_1$, then it follows from 1b) and 1a) that $\varrho \in \mathfrak{G}_n^{(1)} \cup \mathfrak{G}_n^{(3)}$, contradicting that $\varrho \in \mathfrak{G}^{(3)}$. Hence, this case never occurs.

III.2) If $\varrho = \psi\varrho_1 \in \mathfrak{G}_n^{(3)}$, then it follows from 2b) that $\varrho_1 \in \mathfrak{G}_{n-1}^{(2)}$, so

$$\varphi\varrho \in \mathfrak{G}_{n+1}^{(1)}, \quad \text{by 1b)}, \qquad \text{and} \qquad \psi\varrho = \psi^2\varrho_1 \in \mathfrak{G}_n^{(1)} \quad \text{by 3b)}.$$

III.3) If $\varrho = \psi^2 \varrho_1 \in \mathfrak{G}_n^{(3)}$, then it follows from 3a) that $\varrho_1 \in \mathfrak{G}_{n-1}^{(1)}$, so

$$\varphi\varrho \in \mathfrak{G}_{n+1}^{(1)}, \quad \text{by 1b),} \qquad \text{and} \qquad \psi\varrho = \psi^3\varrho_1 = \varrho_1 \in \mathfrak{G}_{n-1}^{(1)} \quad \text{by 3a).}$$

Summing up, we get in case III that

$$(6.25) \qquad \varphi\mathfrak{G}^{(3)} \subseteq \mathfrak{G}^{(1)} \qquad \text{and} \qquad \psi\mathfrak{G}^{(3)} \subseteq \mathfrak{G}^{(1)},$$

hence

$$(6.26) \qquad \mathfrak{G}^{(3)} = \psi^3\mathfrak{G}^{(3)} \subseteq \psi^2\mathfrak{G}^{(1)} \qquad \text{and} \qquad \mathfrak{G}^{(3)} = \varphi^2\mathfrak{G}^{(3)} \subseteq \varphi\mathfrak{G}^{(1)}.$$

When we combine (6.22), (6.24) and (6.26), we get

$$\varphi\mathfrak{G}^{(1)} \subseteq \mathfrak{G}^{(2)} \cup \mathfrak{G}^{(3)} \subset \varphi\mathfrak{G}^{(1)} \cup \varphi\mathfrak{G}^{(1)} = \varphi\mathfrak{G}^{(1)},$$

and

$$\psi\mathfrak{G}^{(1)} \subseteq \mathfrak{G}^{(2)} \subseteq \psi\mathfrak{G}^{(1)}, \qquad \psi^2\mathfrak{G}^{(1)} \subseteq \mathfrak{G}^{(3)} \subseteq \psi^2\mathfrak{G}^{(1)},$$

so we have equality everywhere, and we have proved (6.20), i.e.

$$\varphi\mathfrak{G}^{(1)} = \mathfrak{G}^{(2)} \cup \mathfrak{G}^{(3)}, \qquad \psi\mathfrak{G}^{(1)} = \mathfrak{G}^{(2)}, \qquad \psi^2\mathfrak{G}^{(1)} = \mathfrak{G}^{(3)}.$$

Recall the following theorem from elementary Geometry.

Theorem 6.3 *If a body is transformed isometrically in $\mathbb{R}^3$, such that one of its points is a fixpoint, then this transformation is equivalent to a rotation with respect to an axis going through this fixpoint.*

Theorem 6.3 implies that to every transformation from $\mathfrak{G} \setminus \{I\}$ we can find a rotational axis and a rotation with respect to this axis. The rotational axis cuts the given sphere in two points, which are called the *poles* corresponding to this transformation. The set of all transformations from $\mathfrak{G}$ is countable, so the set D of all poles is also countable.

Consider a rotation, the angle of which is not a multiple of 2π. Then the corresponding two poles are the only fixpoints of this rotation. Hence, no point from $S \setminus D$ is a fixpoint for any transformation from $\mathfrak{G} \setminus \{I\}$.

We shall in the following prove that

$$(6.27) \qquad \text{if } \sigma,\, \sigma' \in \mathfrak{G} \text{ and } \sigma \neq \sigma', \text{ then } \sigma(p) \neq \sigma'(p) \text{ for all } p \in S \setminus DE.$$

In fact, assume that $\sigma(p) = \sigma'(p)$. Then $\sigma'\sigma^{-1}(p) = p$. Since $p \neq D$ is not a pole, we conclude that $\sigma'\sigma^{-1} = I$, from which we get that $\sigma' = \sigma$, and (6.27) follows by contraposition.

Let $p \in S \setminus D$ be a point on the unit sphere, which is not a pole. We define

$$\mathfrak{G}(p) := \{\sigma(p) \in S \mid \sigma \in \mathfrak{G}\}.$$

It follows from the above that all points of $\mathfrak{G}(p)$ are mutually different.

Assume that p, $p' \in S \setminus D$ and $p \neq p'$. We shall prove that

$$\text{(6.28)} \qquad \text{either } \mathfrak{G}(p) = \mathfrak{G}(p') \qquad \text{or} \qquad \mathfrak{G}(p) \cap \mathfrak{G}(p') = \emptyset.$$

In the proof we assume that $\mathfrak{G}(p) \cap \mathfrak{G}(p') \neq \emptyset$. This means that we can find two elements σ_0, $\sigma_0' \in \mathfrak{G}$, such that $\sigma_0(p) = \sigma_0'(p')$. If $\sigma \in \mathfrak{G}$, then

$$\sigma(p) = \sigma\sigma_0^{-1}\sigma_0(p) = \sigma\sigma_0^{-1}\sigma_0'(p').$$

If we put $\sigma' := \sigma\sigma_0^{-1}\sigma_0'$, then $\sigma' \in \mathfrak{G}$, and $\sigma(p) = \sigma'(p')$. This shows that $\sigma(p) \in \mathfrak{G}(p')$ for all $\sigma \in \mathfrak{G}$, and we have proved that $\mathfrak{G}(p) \subseteq \mathfrak{G}(p')$.

Due to the symmetry we also get $\mathfrak{G}(p') \subseteq \mathfrak{G}(p)$, so $\mathfrak{G}(p) = \mathfrak{G}(p')$, and (6.28) is proved.

We have come to the crucial point of the proof. We introduce an equivalence relation $\sim$ on the set $S \setminus D$ by

$$p \sim p' \quad \text{for } p,\, p' \in S \setminus D, \qquad \text{if and only if} \qquad \mathfrak{G}(p) = \mathfrak{G}(p').$$

Then the set $S \setminus D$ is split up into (countably many) disjoint equivalence classes.

It is at this step that *we apply the axiom of choice*, cf. Section 6.3. We choose one element from each equivalence class and let M denote the set of all these choices. Notice that this is the only place, where we apply the axiom of choice. Notice also that

$$S \setminus D = \bigcup_{\sigma \in \mathfrak{G}} \sigma(M).$$

It is easy to see that $\sigma(M) \cap \sigma'(M) = \emptyset$, if $\sigma \neq \sigma'$. In fact, assume that $\sigma(M) \cap \sigma'(M) \neq \emptyset$. Then there are elements p, $p' \in M$, such that $\mathfrak{G}(p) = \mathfrak{G}(p')$, and it follows from the definition above of the set M that $p = p'$, so $\sigma(p) = \sigma'(p)$. Since $p \in S \setminus D$ is not a pole, (6.27) implies that $\sigma = \sigma'$.

Finally, we define

$$(6.29) \qquad A := \bigcup_{\sigma \in \mathfrak{G}^{(1)}} \sigma(M), \qquad B := \bigcup_{\sigma \in \mathfrak{G}^{(2)}} \sigma(M), \qquad C := \bigcup_{\sigma \in \mathfrak{G}^{(3)}} \sigma(M).$$

We proved previously that $\mathfrak{G} = \mathfrak{G}^{(1)} \cup \mathfrak{G}^{(2)} \cup \mathfrak{G}^{(3)}$ is a decomposition of $\mathfrak{G}$ into three disjoint classes, hence

$$(6.30) \qquad S \setminus D = A \cup B \cup C$$

is a decomposition of $S \setminus D$ into three disjoint sets.

It follows from (6.29) and (6.30) that

$$(6.31) \qquad \varphi(A) = B \cup C, \qquad \psi(A) = B, \qquad \psi^2(A) = C,$$

hence by Definition 6.2 that

$$(6.32) \qquad A \cong B \cong C \cong B \cup C.$$

Finally, $\varphi(D) = \psi(D) = D$, because D is the set of all poles, and *Hausdorff's paradox* is proved. □

Notice that we only used rotations in the proof of *Hausdorff's paradox*.

Let $S = A \cup B \cup C \cup D$ be the decomposition as defined above in the proof of Hausdorff's paradox. Then it follows from (6.32) that we can find decompositions

$$A = A_1 \cup A_2, \qquad A_1 \cap A_2 = \emptyset, \qquad A_1 \cong A_2 \cong A,$$

$$B = B_1 \cup B_2, \qquad B_1 \cap B_2 = \emptyset, \qquad B_1 \cong B_2 \cong B,$$

$$C = C_1 \cup C_2, \qquad C_1 \cap C_2 = \emptyset, \qquad C_1 \cong C_2 \cong C.$$

From (6.31) we get $\varphi(A) = B \cup C$, so it follows from $\varphi^2 = I$ that

$$A = \varphi(B) \cup \varphi(C),$$

and we can define $A_1 := \varphi(B)$ and $A_2 = \varphi(C)$, and similarly in the other cases. We therefore have the following disjoint decomposition

$$S = A_1 \cup A_2 \cup B_1 \cup B_2 \cup C_1 \cup C_2 \cup D.$$

Define

$$S_1 := A_1 \cup B_1 \cup C_1 \cup D_1 \qquad \text{and} \qquad S_2 := A_2 \cup B_2 \cup C_2.$$

Then

$$S_1 \underset{4}{=} S \qquad \text{and} \qquad S_2 \underset{3}{=} S \setminus D.$$

In fact, $A_1 \cong A$ and $B_1 \cong B$ and $C_1 \cong C$ and $D = D$, so we get $S_1 \underset{4}{=} S$ by Definition 6.3. Similarly we prove that $S_2 \underset{3}{=} S \setminus D$.

Summing up, we have proved that S can be divided into seven disjoint sets, which by isometric transformations (translations and rotations) can be put together to form the entire unit sphere S_1 and at the same time also another unit sphere with the exception of a nullset. We shall now improve this construction by using the following

Theorem 6.4 *If D is a finite or countable subset of the unit spere S, then $S \setminus D \underset{2}{=} S$.*

Proof. Let the unit sphere S be given by the equation $x^2 + y^2 + z^2 = 1$. The set D is assumed to be at most countable, hence there exist straight lines through the centre of the sphere, which do not contain points from D. We may without loss of generality assume that no point from D lies on the Z-axis.

For every $p \in D$ we let $\alpha(p) \in [0, 2\pi[$ denote the angle between the projection of the vector defined by p onto the XY-plane and the positive X-axis. Since no point from D lies on the Z-axis, the angle $\alpha(p)$ is uniquely defined for every $p \in D$.

Let β be a given angle, and denote by $D(\beta)$ the set, which is obtained, when D is rotated the angle β with respect to the Z-axis. We claim that we can choose the angle β, such that the sets

$$(6.33) \qquad D(n\beta) \qquad \text{for } n \in \mathbb{N}_0,$$

are mutually disjoint.

Let k, $\ell \in \mathbb{N}_0$, where $k < \ell$, and assume that there exists a $q \in S$, such that $q \in D(k\beta) \cap D(\ell\beta)$. Then there exist points p_1, $p_2 \in D$, such that

$$\alpha(q) = \alpha(p_1) + k\beta - 2k_1\pi, \qquad \text{and} \qquad \alpha(q) = \alpha(p_2) + \ell\beta - 2\ell_1\pi,$$

where k_1, $\ell_1 \in \mathbb{Z}$. It follows from $k < \ell$ that

$$(6.34) \qquad \beta = \frac{\alpha(p_1) - \alpha(p_1) + 2(\ell_1 - k_1)\pi}{\ell - k}.$$

The set

$$(6.35) \qquad \left\{ \frac{\alpha(p_1) - \alpha(p_2) + 2s\pi}{n} \;\middle|\; p_1, p_2 \in D, s \in \mathbb{Z}, n \in \mathbb{N} \right\}$$

is countable. Hence, there exists a $\beta \in]0, 2\pi[$, which does not belong to the set (6.35), i.e. it is not of the form given by (6.34). For any such β the system (6.33) consists of sets, which are mutually disjoint.

We define

$$\Omega := \bigcup_{n=0}^{+\infty} D(n\beta) \qquad \text{and} \qquad R := S \setminus \Omega.$$

Then clearly $\Omega(\beta) = \Omega \setminus D$, so $\Omega \setminus D \cong \Omega$.

Finally, it follows from

$$S = \Omega \cup R, \qquad \Omega \cap R = \emptyset, \qquad S \setminus D = (\Omega \setminus D) \cup R, \qquad (\Omega \setminus D) \cap R = \emptyset,$$

that $S \setminus D \underset{2}{=} S$. $\square$

Recall that we just before Theorem 6.4 decomposed the unit sphere S into seven disjoint subsets

$$S = A_1 \cup A_2 \cup B_1 \cup B_2 \cup C_1 \cup C_2 \cup D,$$

where D is a nullset, and where

$$S_4 \underset{4}{=} A_1 \cup B_1 \cup C_1 \cup D \qquad \text{and} \qquad S_2 \setminus D \underset{3}{=} A_2 \cup B_2 \cup C_2.$$

It follows from Theorem 6.4 that $S_2 \setminus \underset{2}{=} S$, hence, we can – using Theorem 6.2 – write as a disjoint union,

$$A_2 \cup B_2 \cup C_2 = A_{21} \cup A_{22} \cup B_{21} \cup B_{22} \cup C_{21} \cup C_{22} \underset{6}{=} S_2.$$

Therefore, we have the disjoint union

$$S = A_1 \cup A_{21} \cup A_{22} \cup B_{21} \cup B_{22} \cup C_{21} \cup C_{22} \cup D,$$

where

$$S_1 = A_1 \cup B_1 \cup C_1 \cup D \underset{4}{=} S,$$

and

$$S_2 = A_{21} \cup A_{22} \cup B_{21} \cup B_{22} \cup C_{21} \cup C_{22} \underset{6}{=} S.$$

We see that the unit sphere S can be decomposed into precisely ten disjoint sets, where four of these can be put together by isometric transformations to form an entire unit sphere, while the remaining six subsets by other isometric transformations can be joined to form another entire unit sphere.

The area has clearly been doubled by this process, where we only use isometries, so we conclude that some of the ten subsets must be nonmeasurable.

It is left to the reader to prove that when $n \geq 2$ we can decompose S into $10 + 36(n-2)$ disjoint subsets, which afterwards by isometries can be put together to form n copies of the unit sphere.

6.5 Banach-Tarski's paradox

We shall apply *Hausdorff's paradox*, cf. Section 6.4, to derive

Banach-Tarski's paradox. *If E_1 and E_2 are two bounded sets in $\mathbb{R}^3$, both with interior points, i.e. $E_1^\circ \neq \emptyset$ and $E_2^\circ \neq \emptyset$, then $E_1 \overset{f}{=} E_2$.*

In other words, if we assume the *axiom of choice*, then e.g. a lump of gold can theoretically be split into finitely many disjoint pieces, which then can be put together in a different way to form another lump of gold of the double size. Such a procedure would of course be devastating for the world's economy, and the worth of gold would fall to nothing, so luckily no one has yet *constructed* a non-measurable set.

The theorem was first proved by *Banach & Tarski* [5] in 1924. The proof given here is simpler. We shall more or less follow *Sierpinski* [50], and we shall use the notation already given in Section 6.4.

We shall split the proof into a series of simpler results. The first one is the following theorem.

Theorem 6.5 *Given three sets, A, B and E in $\mathbb{R}^3$, where $A \supseteq E \supset B$ and $A \underset{n}{=} B$, then $A \underset{n+1}{=} E$.*

Proof. By assumption, $A \underset{n}{=} B$, so there exists $2n$ sets, $A_1, \dots, A_n$ and $B_1, \dots, B_n$, such that

1) $A = A_1 \cup \dots \cup A_n$ and $B = B_1 \cup \cdots \cup B_n$.

2) $A_k \cap A_\ell = \emptyset$ and $B_k \cap B_\ell = \emptyset$ for $1 \le k < \ell \le n$.

3) $A_k \cong B_k$ for all $k = 1, 2 \dots, n$.

Using 3) we get that for every $k = 1, 2, \dots, n$ there exists an isometry $\varphi_k : A_k \to B_k$, i.e.

(6.36) $$B_k = \varphi(A_k), \qquad \text{for } k = 1, 2, \dots, n.$$

Define

(6.37) $$\varphi(x) := \varphi_k(x), \qquad \text{if } x \in A_k \text{ and } k = 1, 2, \dots, n.$$

Then $\varphi : A \to B$ is bijective, and φ *realises* the equivalence $A \overset{=}{n} B$.

Using that $A \supseteq E \supseteq B$ we get $\varphi(a) = B \subseteq A$, and in particular $\varphi^k(A) \subseteq A$ for all $k \in \mathbb{N}$, where we define $\varphi^k := \varphi \circ \cdots \circ \varphi$ (with k "factors"). Then also

$$\varphi^k(A \setminus E) \subseteq E \qquad \text{for all } k \in \mathbb{N}.$$

If we define

(6.38) $$A' := \bigcup_{k=0}^{+\infty} \varphi^k(A \setminus E),$$

then $A' \setminus A$, so we define

(6.39) $$A := A \setminus A',$$

and

(6.40) $$E' := \varphi(A').$$

It follows from $A' \subseteq A$ and (6.39) that

(6.41) $$A = A' \cup A'' \qquad \text{and} \qquad A' \cap A'' = \emptyset.$$

Then, using (6.40) and $A' \subseteq A$ and $\varphi(a) = B \subseteq E$,

(6.42) $$E' = \varphi(A') \subseteq \varphi(A) = B \subseteq E.$$

It follows from (6.38) that

$$\varphi(A') = \bigcup_{k=1}^{+\infty} \varphi^k(A \setminus E),$$

so we get from (6.38), (6.40) and (6.42), that

(6.43) $$A' = (A \setminus E) \cup \varphi(A') = (A \setminus E) \cup E' = A \setminus (E \setminus E').$$

Then it follows from $E \subseteq A$ and (6.39) that

$$E \setminus E' = A \setminus A' = A'',$$

hence

(6.44) $$E = E' \cup A'' \qquad \text{and} \qquad E' \cap A'' = \emptyset.$$

We conclude from (6.41) that $A = (A' \cap A) \cup A''$, so it follows from 1) and 2) that we have a decomposition of A into $n+1$ sets given by

$$A = (A' \cap A_1) \cup (A' \cap A_2) \cup \cdots \cup (A' \cap A_n) \cup A''. \tag{6.45}$$

The function φ is injective on A, hence by (6.41), (6.40) and (6.36),

$$\varphi(A' \cap A_k) = \varphi(A') \cap \varphi(A_k) = E' \cap B_k \qquad \text{for } k = 1, \dots, n. \tag{6.46}$$

We infer, cf. (6.37), that $\varphi(A' \cap A_k) = \varphi_k(A' \cap A_k)$, and since $\varphi_k : A_k \to B_k$ is an isometry, we also have $\varphi(A' \cap A_k) = A' \cap A_k$, and it follows from (6.46) that

$$A' \cap A_k \cong E' \cap B_k \qquad \text{for all } k = 1,\, 2,\, \dots,\, n. \tag{6.47}$$

The sets $\{E' \cap B_k \mid k = 1, \dots, n\}$ are disjoint, cf. 2).

Then (6.42) and $E \subseteq B$ imply that $E' \subseteq B$, so $E' = E' \cap B$, and if follows from 1) and (6.44) that

$$E \setminus (E' \cap \{B_1 \cup B_2 \cup \cdots \cup B_n\}) = E \setminus (E' \cap B) = E \setminus E' = A'',$$

or, put in another way,

$$E = (E' \cap B_1) \cup (E' \cap B_2) \cup \cdots \cup (E' \cap B_n) \cup A''. \tag{6.48}$$

The subsets of this decomposition are mutually disjoint, cf. 2) and (6.44).

Finally, collecting (6.45), (6.47) and (6.48) it follows that $A \underset{n+1}{=} E$, and the theorem is proved. □

An immediate consequence of Theorem 6.5 is

Corollary 6.2 *If $A \supseteq E \supseteq B$ and $A \overset{f}{=} B$, then $A \overset{f}{=} E$.*

To end the proof of *Banach-Tarski's paradox* we finally prove

Lemma 6.1 *Assume that a bounded set E in $\mathbb{R}^3$ contains a solid ball $B\,(\subseteq E)$. Then $E \overset{f}{=} B$.*

Proof. We apply *Hausdorff's paradox* to get

$$S \overset{f}{=} S_1 \overset{f}{=} S_2 \overset{f}{=} S_1 \cup S_2,$$

where S and S_1 and S_2 are disjoint spheres, all of radius 1.

Let αS, $\alpha \in]0,1]$ denote the sphere of radius α and same centre as S, and similarly for S_1 and S_2. Then

$$\alpha S \overset{f}{=} \alpha S_1 \overset{f}{=} \alpha S_2 \overset{f}{=} \alpha S_1 \cup \alpha S_2,$$

where we can use the scaled decompositions, defined already for $\alpha = 1$. If B, B_1 and B_2 are the solid balls inside S, S_1 and S_2, we get

$$B \overset{f}{=} B_1 \overset{f}{=} B_2 \overset{f}{=} B_1 \cup B_2.$$

Strictly speaking, we do not get the centres in the first place in this construction, but they can be added by e.g. a variant of Theorem 5.4.

Let $r > 0$ be the radius of B. The set E is bounded, so $E \subseteq \bigcup_{j=1}^{s} Q_j$, where the Q_j are e.g. disjoint cubes of diagonal $2r$. Let $B_1, 6s, B_s$ be disjoint solid balls, each of which $\cong B$. Then by the result above,

$$B \overset{f}{=} B_1 \cup B_2 \cup \cdots \cup B_s. \tag{6.49}$$

Using that $E \subseteq Q_1 \cup \cdots Q_s$ it follows that E can be divided into s disjoint sets $E_1, \dots, E_s$, where $E_i \subseteq Q_i$ for $i = 1, \dots, s$.

The cube Q_i, $i \in \{1, 2, \dots, s\}$, of diagonal $2r$ is congruent with a part of the ball B_i (of radius r), hence the set $E_i \subseteq Q_i$ is also congruent with a subset of B_i. This means that E is $\overset{f}{=}$ a subset of $B_1 \cup \cdots \cup B_s$, Then it follows from (6.49) and Corollary 6.1 that the set E is $\overset{f}{=}$ a subset H of B. If finally follows from $H \subseteq B \subseteq E$ and $E \overset{f}{=}$ and Corollary 6.2 that $E \overset{f}{=} B$, and the lemma is proved. □

Proof of Banach-Taerski's paradox. Let B_1 and B_2 be two solid balls, contained in E_1 and E_2, resp.. By choosing the same radius in B_1 and B_2 we may assume that $B_1 \cong B_2$. Then it follows from Lemma 6.1 that $E_1 \overset{f}{=} B_1$ and $E_2 \overset{f}{=} B_2$, hence by Corollary 6.1, $E_1 \overset{f}{=} E_2$. □

7 Miscellaneous

7.1 Introduction

In this chapter we collect some examples, which do not belong to the classification of any of the other chapters. It is on purpose that we have put it before Chapter 8, which must be considered to be the most difficult chapter in this book. Here we first consider the method of induction in Section 7.2. This section should not be taken too seriously, because in practice most applications of the method of induction are straight forward. Then in the following Section 7.3 we play a little with series. The theory becomes more serious in Section 7.4, in which we organise the set of irrational numbers $\mathbb{R} \setminus \mathbb{Q}$ as a complete set without changing the topology inherited from $\mathbb{R}$. Finally, we show in Section 3.8 an alternative way of solving the *heat equation*, where the method applied here is not commonly known in the literature.

7.2 The method of induction

We shall in this section first give an anecdote, in which a mathematician, a physicist and an engineer are making fools of themselves by applying the method of proof by induction in a wrong way. Then follows an example, in which the standard proof of induction is successfully demonstrated, and then finally we demonstrate that there may even in the standard method of proof by induction be some unexpected pitfalls.

7.2.1 An anecdote

There are lots of anecdotes describing the difference in thinking of mathematicians, physicists and engineers. They are all more or less very silly, and yet they often hit the head of the nail, concerning the characteristics of these three scientific tribes, where the author himself feels that he is a part of all three of them.

A mathematician, a physicist and an engineer are sitting together discussing how to "prove" that all odd numbers, except for 1, are primes.

First the mathematician says: "This is easy, because 3 is a prime, 5 is a prime and 7 is a prime, and then the claim follows by induction!"

The physicist is not convinced. "No, no!", he objects. "You must reason in the following way: 3 is a prime, 5 is a prime, 7 is a prime, 9 is an experimental error, 11 is a prime, 13 is a prime, and the claim follows!"

"Well, well, well!", says the engineer. "It is much easier to prove, because 3 is a prime, 5 is a prime, 7 is a prime, 9 is a prime, 11 is a prime, etc.."

The usual reaction is that mathematicians cannot see that this is a joke. Physicists are in general really amused, while engineers are embarrassed.

All three proofs are wrong, including the mathematician's application of the proof by induction. We shall in the next section in a simple example give a standard method of the proof by inductions, and finally in Section 7.2.3 follow this standard procedure in a case, where it cannot be applied, due to a hidden pitfall.

7.2.2 The standard method of proof by induction

We shall only shortly recall this method, because is should be well-known to the reader. Therefore, we only demonstrate it in the proof of the formula

$$\text{(7.1)} \qquad \sum_{j=1}^{n} j^2 = \frac{1}{6}\, n(n+1)(2n+1).$$

Let us assume that we do not know the formula (7.1), so we first have to derive a reasonable formula. The argument may go as follows. Summation is analogous to integration, and as $\int x^2\,\mathrm{d}x = \frac{1}{3}\,x^3$, we may expect that

$$\sum_{j=1}^{n} j^2 = \frac{1}{3}\,n^3 + a\cdot n^2 + b\cdot n + c,$$

where a, b and c are unknown coefficients. To find these we set up the equations

$$n=1, \qquad \sum_{j=1}^{1} j^2 = 1 = \frac{1}{3} + a + b + c,$$

$$n=2, \qquad \sum_{j=1}^{2} j^2 = 5 = \frac{8}{3} + 4a + 2b + c,$$

$$n=3, \qquad \sum_{j=1}^{3} j^2 = 14 = 9 + 9a + 3b + c,$$

which are easily solved by Cramèr's formulæ,

$$a = \frac{1}{2}, \qquad b = \frac{1}{6} \quad \text{and} \quad c = 0,$$

so the only possibility of such a formula is

$$\sum_{j=1}^{2} \quad \text{is equal to} \quad \frac{1}{3}\,n^3 + \frac{1}{2}\,n^2 + \frac{1}{6}\,n = \frac{1}{6}\,n(n+1)(2n+1),$$

which is (7.1). Since this expression is derived from the cases $n = 1, 2, 3$, we know that (7.1) is correct for $n = 1, 2, 3$.

Assume (7.1) for some fixed value $n \in \mathbb{N}$, i.e.

$$\sum_{j=1}^{n} j^2 = \frac{1}{6}\,n(n+1)(2n+1).$$

We shall prove that it is also correct, when n is replaced bu $n+1$, i.e.

$$\sum_{j=1}^{n+1} j^2 = \frac{1}{6}\,(n+1)(n+2)(2n+3)?$$

We get for the left hand side,

$$
\begin{aligned}
\sum_{j=1}^{n+1} j^2 &= \sum_{j=1}^{n} j^2 + (n+1)^2 = \frac{1}{6}\, n(n+1)(2n+1) + (n+1)^2 \\
&= \frac{1}{6}\,(n+1)\{n(2n+1) + 6(n+1)\} = \frac{1}{6}\,(n+1)\left\{2n^2 + 7n + 6\right\} \\
&= \frac{1}{6}\,(n+1)(n+2)(2n+3) = \frac{1}{6}\,(n+1)(\{n+1\}+1)(2\{n+1\}+1),
\end{aligned}
$$

which is (7.1) with n replaced by $n+1$. If therefore (7.1) holds for some n, it does for the first numbers, $n = 1, 2, 3$, then it also holds for the successor $n+1$, and then for $n+2$, etc., so it holds by induction for all $n \in \mathbb{N}$.

The above is a standard application of the principle of induction. We assume that some formula holds for some number n and then prove that it also holds for the following number $n+1$. Finally, we check that it also holds in the beginning, i.e. for $n = 1$ (sufficient here), and for $n = 2$ (to be absolutely sure).

7.2.3 Wrong application of the principle of induction

We shall in this section by induction "prove" that given any bundle $\mathfrak{L}_n$ of n mutually nonparallel straight lines in the plane, they will all pass through the same point in the plane. This claim is obviously wrong.

The "proof" goes as follows. The claim is trivial for $n = 1$, because then we only have one line. For $n = 2$ the bundle $\mathfrak{L}_2$ contains two straight lines which are not parallel. They must therefore intersect each other at a point, and the claim is also true for $n = 2$.

Then assume that given any bundle $\mathfrak{L}_n$, consisting of n mutually nonparallel straight lines from $\mathfrak{L}_n$ these will all intersect each other at a point $\mathbf{x}$, say. This is at least true for $n = 1$ and $n = 2$.

Consider any such bundle $\mathfrak{L}_{n+1}$ of $n+1$ mutually nonparallel straight lines. If we remove one of these lines, we reduce the situation to a bundle $\mathfrak{L}_n$ consisting of n lines. By assumption, they intersect each other at a point $\mathbf{x}$. Then we remove another line to get another bundle $\mathfrak{L}'_n$, consisting of n lines. By assumption they must also intersect at a point $\mathbf{y}$. Then the bundle $\mathfrak{L}_{n-1}$ in which we have removed both lines above, must intersect each other at both $\mathbf{x}$ and $\mathbf{y}$. Since the lines of $\mathfrak{L}_{n-1}$ are mutually nonparallel, this is only possible if $\mathbf{x} = \mathbf{y}$, and the claim "follows" by induction.

This is an example of what happens, if one argues too formally and jumps to conclusions. The proof follows the usual scheme of a proof by induction. Assuming that some claim is right for some n, we prove that it also holds for the successor $n + 1$. Furthermore, we checked that the claim was true for both $n = 1$ and $n = 2$, so it should follow by this bootstrap method. However, it does *not* hold for $n = 3$.

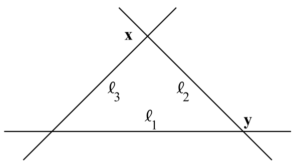

Figure 7.1: *A bundle $\mathfrak{L}_3$ consisting of three mutually nonparallel straight lines in the plane and not passing through a common point.*

If we in the bundle $\mathfrak{L}_3$ of Figure 7.1 remove ℓ_1, the intersection point is $\mathbf{x}$. If we instead remove ℓ_2, then the intersection point is $\mathbf{y}$. There is only one line in $\mathfrak{L}_3 \setminus \{\ell_1, \ell_2\} = \{\ell_3\}$, and it surely passes through both $\mathbf{x}$ and $\mathbf{y} \neq \mathbf{x}$. We see that the induction step from $n = 2$ to $n = 2 + 1 = 3$ does not apply.

7.3 Two decreasing sequences (a_n) and (b_n) of positive numbers, which satisfy $\sum_{n=1}^{+\infty} a_n = \sum_{n=1}^{+\infty} b_n = +\infty$ and $\sum_{n=1}^{+\infty} \min\{a_n, b_n\} < +\infty$.

This problem occurs from time to time in *American Mathematical Monthly*, so it should just be mentioned here.

We put $n_1 = 1$ and then define inductively

$$n_{p+1} := 1 + \sum_{k=1}^{p} n_k^2 = n_p + n_p^2 = n_p\left(n_p + 1\right),$$

identities which can be derived in the following way,

$$n_p = 1 + \sum_{k=1}^{p-1} n_k^2 = 1 + \sum_{k=1}^{p} n_k^2 - n_p^2 = n_{p+1} - n_p^2,$$

and the formulæ follow by some rearrangements.

We define the sequences (a_n) and (b_n) in the following way,

$$\begin{cases} a_n := \dfrac{1}{n_p^2} \text{ and } b_n := \dfrac{1}{n^2} & \text{for } n_p \le n < n_{p+1} \text{ and } p \text{ odd} \\[2ex] a_n := \dfrac{1}{n^2} \text{ and } b_n := \dfrac{1}{n_p^2} & \text{for } n_p \le n < n_{p+1} \text{ and } p \text{ even.} \end{cases}$$

Then (a_n) and (b_n) are non-increasing sequences, and

$$c_n := \min\{a_n, b_n\} = \frac{1}{n^2} \qquad \text{for all } n \in \mathbb{N},$$

hence, the series

$$\sum_{n=1}^{+\infty} c_n = \sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

is convergent, while

$$\sum_{n=1}^{n_{p+1}} a_n = \sum_{k=1}^{p} \sum_{n=n_k}^{n_k+n_k^2} a_n \ge \frac{p}{2} \quad \text{and} \quad \sum_{n=1}^{n_{p+1}} b_n = \sum_{k=1}^{p} \sum_{n=n_k}^{n_k+n_k^2} b_n \ge \frac{p}{2},$$

for all $p \in \mathbb{N}$, so

$$\sum_{n=1}^{+\infty} a_n = \lim_{p\to+\infty} \sum_{n=1}^{n_{p+1}} a_n \ge \lim_{p\to+\infty} \frac{p}{2} = +\infty \quad \text{and} \quad \sum_{n=1}^{+\infty} b_n = \lim_{p\to+\infty} \sum_{n=1}^{n_{p+1}} b_n \ge \lim_{p\to+\infty} \frac{p}{2} = +\infty,$$

and both series are divergent.

A small modification, which is left to the reader, shows that the sequences (a_n) and (b_n) can be chosen strictly decreasing.

Conversely, by using the same technique as above we may choose any decreasing sequence (c_n), for which $\sum_n c_n$ is convergent, and then construct decreasing sequences (a_n) and (b_n), such that

$$c_n = \min\{a_n, b_n\}, \quad \sum_{n=1}^{+\infty} c_n < +\infty, \quad \text{and} \quad \sum_{n=1}^{+\infty} a_n = \sum_{n=1}^{+\infty} b_n = +\infty.$$

7.4 The irrational numbers organised as a complete metric space without changing the subspace topology inherited from $\mathbb{R}$

This pathological example is mostly of interest for mathematicians. Let us first recall some topological notions.

Let $M \neq \emptyset$ be a nonempty set. A *metric* on M is a map $d : M \times M \to \mathbb{R}$, which fulfils the following four conditions:

m1) The *distance* between two points x, $y \in M$ is always nonnegative, i.e.

$$d(x,y) \geq 0 \qquad \text{for all } x,\ y \in M.$$

m2) The *distance* between two points x, $y \in M$ is 0, if and only if the points are identical, $x = y$, i.e.

$$d(x,y) = 0 \quad \text{if and only if} \quad x = y.$$

m3) The distance from x to y is the same as the distance from y to x, i.e.

$$d(x,y) = d(y,x) \qquad \text{for all } x,\ y \in M.$$

m4) *The triangle inequality.* The distance between two points x and y is smaller than or equal to the distance from x to z plus the distance from z to y, i.e.

$$d(x,y) \leq d(x,z) + d(z,y),$$

i.e. a detour is always longer (or at least equal to) than the straight way.

When d is a metric on M, we call (M,d) a *metric space.*
We define the *open ball* of centre $x \in M$ and radius $r > 0$ in the metric space (M,d) as the set

$$B_d(x,r) := \{y \in M \mid d(x,y) < r\}.$$

Similarly,

$$B_d[x,r] := \{y \in M \mid d(x,y) \leq r\}$$

denotes the *closed ball* of centre $x \in M$ and radius $r > 0$ in the metric d.

Assume that d_1 and d_2 are two metrics on M. We say that the two metric spaces (M,d_1) and (M,d_2) – or just the metrics d_1 and d_2 – are *equivalent*, if the following two conditions are fulfilled.

e1) For every $x \in M$ and every $r > 0$ there exists an $R > 0$, such that

$$B_{d_1}(x,R) \subseteq B_{d_2}(x,r).$$

e2) For every $x \in M$ and every $r > 0$ there exists an $R > 0$, such that

$$B_{d_2}(x,R) \subseteq B_{d_1}(x,r).$$

In other words, every ball in the d_2-metric contains a ball in the d_1-metric, and *vice versa*.

A sequence (x_n) of points from the metric space (M, d) is called *convergent* in the metric d, if there exists a point $x \in M$, such that for any given $\varepsilon > 0$ there is a constant $N \in \mathbb{N}$, such that

$$d\left(x_n, x\right) < \varepsilon \qquad \text{for all } n \geq N.$$

In other words, the elements of (x_n) eventually belong to the open ball $B_d(x, \varepsilon)$

When (x_n) is convergent in the metric d, then the point x is uniquely determined. In fact, if the elements of (x_n) eventually belong to both $B_d(x, \varepsilon)$ and $B_d(y, \varepsilon)$, and $x \neq y$, then we obtain a contradiction by choosing $\varepsilon < d(x, y)/2$, because then $B_d(x, \varepsilon) \cap B_d(y, \varepsilon) = \emptyset$. Hence, x is uniquely determined, when (x_n) is convergent in (M, d), and we call x the *limit* of (x_n), and we write

$$\lim_{n \to +\infty} x_n = x.$$

A sequence (x_n) from (M, d) is called a *Cauchy sequence* (with respect to the metric d), if for every $\varepsilon > 0$ there is an $N \in \mathbb{N}$, such that

$$d(x_m, x_n) < \varepsilon, \qquad \text{whenever } m,\, n \geq N.$$

It is easy to prove that every convergent sequence is also a Cauchy sequence. In fact, choose $N \in \mathbb{N}$, such that

$$d(x_m, x) < \frac{\varepsilon}{2} \quad \text{and} \quad d(x_n, x) < \frac{\varepsilon}{2} \qquad \text{for all } m,\, n \geq N.$$

Then it follows from the triangle inequality for $m,\, n \geq N = N(\varepsilon)$,

$$d(x_m, x_n) \leq d(x_m, x) + d(x, x_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

On the other hand, a Cauchy sequence is not always convergent. An example is given in $(\mathbb{Q}, |\cdot|)$ by

$$x_1 = 3; \quad x_2 = 3.1; \quad x_3 = 3.14; \quad x_4 = 3.141; \quad x_5 = 3.1415; \quad x_6 = 3.14159; \quad \dots,$$

where we approximate π by the sequence of its decimal fractions. Since the limit $\pi \notin \mathbb{Q}$, this sequence is not convergent in $(\mathbb{Q}, |\cdot|)$, though it is in the larger space $(\mathbb{R}, |\cdot|)$, so it is clearly a Cauchy sequence.

If a metric space (M, d) has the property that every Cauchy sequence is also convergent with a limit in M, then we call (M, d) a *complete metric space*.

The prototype of a complete space is the space of all real numbers, $(\mathbb{R}; |\cdot|)$ in the usual metric, $d(x, y) := |x - y|$. In general we define the *usual metric* on a subset $A \subseteq \mathbb{R}$ by

$$d(x, y) := |x - y| \qquad \text{for } x,\, y \in A, \qquad A \subseteq \mathbb{R},$$

and we denote the metric space by $(A, |\cdot|)$.

As mentioned above, $(\mathbb{R}, |\cdot|)$ is complete, but neither $(\mathbb{Q}, |\cdot|)$ nor $(\mathbb{R} \setminus \mathbb{Q}, |\cdot|)$ are complete. In the latter case, just consider the sequence $(2^{-n}\pi)$ from the set of irrational numbers $\mathbb{R} \setminus \mathbb{Q}$. It is obviously a Cauchy sequence, which in the larger space $(\mathbb{R}, |\cdot|)$ of formally the same metric converges towards $0 \notin \mathbb{R} \setminus \mathbb{Q}$, and we know already that the limit is unique, when it exists.

After all these preparations the following theorem should indeed be surprising.

Theorem 7.1 *There exists a metric d on the irrational numbers $\mathbb{R} \setminus \mathbb{Q}$, such that*

1) $(\mathbb{R} \setminus \mathbb{Q}, d)$ and $(\mathbb{R} \setminus \mathbb{Q}, |\cdot|)$ are equivalent metric spaces.

2) $(\mathbb{R} \setminus \mathbb{Q}, d)$ is a complete metric space, while $(\mathbb{R} \setminus \mathbb{Q}, |\cdot|)$ is not.

PROOF. The set of all rational numbers $\mathbb{Q}$ is countable. Hence, its elements can be arranged as a sequence

$$\mathbb{Q} = \{r_n \mid n \in \mathbb{N}\}.$$

Let $x,\, y \in \mathbb{R} \setminus \mathbb{Q}$ be two irrational numbers. Then define

$$d(x, y) := |x - y| + \sum_{n=1}^{+\infty} \frac{1}{2^n} \min\left\{1, \left| \max_{m \le n} \frac{1}{|x - r_m|} - \max_{m \le n} \frac{1}{|y - r_m|} \right| \right\}. \tag{7.2}$$

Then $d : \mathbb{R} \setminus \mathbb{Q} \times \mathbb{R} \setminus \mathbb{Q} \to \mathbb{R}$ is well-defined.

We shall first prove that $(\mathbb{R} \setminus \mathbb{Q}, d)$ is a metric space, i.e. we shall prove that d given by (7.2) is a metric. The first three conditions, m1), m2) and m3), are trivial, so we shall only prove the triangle inequality m4).

We first notice that if $a,\, b,\, c > 0$, then

$$\min\{1, |a - b|\} \le \min\{1, |a - c| + |c - b|\} \le \min\{1, |a - c|\} + \min\{1, |c - b|\}.$$

Then applying this inequality with

$$a := \max_{m \le n} \frac{1}{|x - r_m|}, \qquad b := \max_{m \le n} \frac{1}{|y - r_m|}, \qquad c := \max_{m \le n} \frac{1}{|z - r_m|},$$

we get

$$\begin{aligned} d(x,y) &= |x-y| + \sum_{n=1}^{+\infty} \frac{1}{2^n} \min\left\{1, \left|\max_{m\le n} \frac{1}{|x-r_m|} - \max_{m\le n} \frac{1}{|y-r_m|}\right|\right\} \\ &\le |x-z| + \sum_{n=1}^{+\infty} \frac{1}{2^n} \min\left\{1, \left|\max_{m\le n} \frac{1}{|x-r_m|} - \max_{m\le n} \frac{1}{|z-r_m|}\right|\right\} \\ &\quad +|z-y| + \sum_{n=1}^{+\infty} \frac{1}{2^n} \min\left\{1, \left|\max_{m\le n} \frac{1}{|z-r_m|} - \max_{m\le n} \frac{1}{|y-r_m|}\right|\right\} \\ &= d(x,z)+d(z,y), \end{aligned}$$

and we have proved m4), so d is a metric, and $(\mathbb{R}\setminus\mathbb{Q}, d)$ is a metric space.

Then we prove that the two metrics are equivalent.

It is trivial that $|x-y| \le d(x,y)$. If therefore $d(x,y) < r$, then also $|x-y| < r$, and e1) follows, i.e. for all $x \in \mathbb{R}\setminus\mathbb{Q}$ and all $r > 0$ we can choose $R = r$, such that

$$B_d(x,r) \subseteq B_{|\cdot|}(x,r).$$

Then we shall prove e2). Let $x \in \mathbb{R}\setminus\mathbb{Q}$ and $r > 0$ be given. Choose $N \in \mathbb{N}$, such that $2^{-N} < \frac{r}{2}$. Then we get for all $y \in \mathbb{R}\setminus\mathbb{Q}$ that

$$\sum_{k=N+1}^{+\infty} \frac{1}{2^k} \min\left\{1, \left|\max_{m\le k} \frac{1}{|x-r_m|} - \max_{m\le k} \frac{1}{|y-x_m|}\right|\right\} \le \sum_{k=N+1}^{+\infty} \frac{1}{2^k} = \frac{1}{2^N} < \frac{r}{2}.$$

Hence,

$$\begin{aligned} d(x,y) &= |x-y| + \sum_{k=1}^{+\infty} \frac{1}{2^k} \min\left\{1, \left|\max_{m\le k} \frac{1}{|x-r_m|} - \max_{m\le k} \frac{1}{|y-r_m|}\right|\right\} \\ &\le |x-y| + \sum_{k=1}^{N} \frac{1}{2^k} \min\left\{1, \left|\max_{m\le k} \frac{1}{|x-r_m|} - \max_{m\le k} \frac{1}{|y-r_m|}\right|\right\} + \frac{r}{2}, \end{aligned} \tag{7.3}$$

Then define

$$c := \min_{n\le N} |x-r_n| > 0, \tag{7.4}$$

where x and N are the same as above. It follows that if $y \in \mathbb{R}\setminus\mathbb{Q}$ and $|x-y| \le \frac{c}{2}$, then

$$\min_{n\le N} |y-r_n| \ge \frac{c}{2}.$$

Given any $y \in \mathbb{R}\setminus\mathbb{Q}$ for which $|x-y| \le \frac{c}{2}$, and any $n \le N$. There exist constants m_1, $m_2 \in \{1,2,\dots,n\}$, such that for all $m \in \{1,2,\dots,n\}$,

$$|x-r_{m_1}| \le |x-r_m| \qquad \text{and} \qquad |y-r_{m_1}| \le |y-r_m|.$$

Then we get the following estimate, whenever $|x-y| \le \dfrac{c}{2}$,

$$\left|\max_{m \le n} \frac{1}{|x-r_m|} - \max_{m \le n} \frac{1}{|y-r_m|}\right| = \left|\frac{1}{|x-r_{m_1}|} - \frac{1}{|y-r_{m_2}|}\right|$$

$$= \frac{||y-r_{m_2}| - |x-r_{m_1}||}{|x-r_{m_1}| - |y-r_{m_2}|} \le \frac{2}{c^2}\,|x-y|,$$

where we have used that

$$|x-r_{m_1}| - |y-r_{m_2}| \le |x-r_{m_2}| - |y-r_{m_2}| \le |x-y|,$$

and

$$|y-r_{m_2}| - |x-r_{m_1}| \le |y-r_{m_1}| - |x-r_{m_1}| \le |x-y|.$$

Assume that $|x-y| \le \min\left\{\dfrac{c}{2}, \dfrac{c^2}{2}\right\}$. Then $\dfrac{2}{c^2}\,|x-y| \le 1$, and then by insertion into (7.3),

$$\begin{aligned} d(x,y) &\le |x-y| + \frac{r}{2} + \sum_{k=1}^{N} \frac{1}{2^k} \min\left\{1, \left|\max_{m \le k} \frac{1}{|x-r_m|} - \max_{m \le k} \frac{1}{|y-r_m|}\right|\right\} \\ &\le |x-y| + \frac{r}{2} + \frac{2}{c^2}\,|x-y| \sum_{k=1}^{N} \frac{1}{2^k} \le \frac{r}{2} + \left(1 + \frac{2}{c^2}\right)|x-y|. \end{aligned}$$

We finally choose the radius

$$R := \min\left\{\frac{c}{2}, \frac{c^2}{2}, \frac{c^2 r}{2\,(2+c^2)}\right\},$$

where $c > 0$ is given by (7.4). If $|x-y|$, then $d(x,y) < r$, and we have proved for all $x \in \mathbb{R} \setminus \mathbb{Q}$ and all $r > 0$ that we can find $R > 0$, such that

$$B_{|\cdot|}(x,R) \subseteq B_d(x,r), \tag{7.5}$$

and we have proved e2).

Summing up, we have proved that the two metric spaces $(\mathbb{R} \setminus \mathbb{Q}, |\cdot|)$ and $(\mathbb{R} \setminus \mathbb{Q}, d)$ are equivalent.

We shall finally prove that $(\mathbb{R} \setminus \mathbb{Q}, d)$, contrarily to $(\mathbb{R} \setminus \mathbb{Q}, |\cdot|)$, is complete.

Let (x_n) be a Cauchy sequence from $(\mathbb{R} \setminus \mathbb{Q}, d)$. It follows from the trivial estimate

$$|x_m - x_n| \le d(x_m, x_n),$$

that (x_n) is also a Cauchy sequence in $(\mathbb{R} \setminus \mathbb{Q}, |\cdot|)$ in the usual metric, hence also in the larger space $(\mathbb{R}, |\cdot|)$. Since $(\mathbb{R}, |\cdot|)$ is complete, (x_n) has a limit $x \in \mathbb{R}$ in this usual metric,

$$|x - x_n| \to 0 \qquad \text{for } n \to +\infty.$$

There are two possibilities,

either i) $x \in \mathbb{R} \setminus \mathbb{Q}$, or ii) $x \in \mathbb{Q}$.

Assume that i) $x \in \mathbb{R} \setminus \mathbb{Q}$. We shall prove that $d(x, x_n) \to 0$ for $n \to +\infty$, i.e. for every given $\varepsilon > 0$ there is an $N = N(\varepsilon) \in \mathbb{N}$, such that

$$(7.6) \qquad x_n \in B_d(x, \varepsilon) \qquad \text{for all } n \geq N.$$

Using (7.5) we find an $R = R(x, \varepsilon) > 0$ corresponding to $x \in \mathbb{R} \setminus \mathbb{Q}$ and $r = \varepsilon > 0$, such that

$$B_{|\cdot|}(x, R) \subseteq B_d(x, \varepsilon).$$

Since $x_n \to x$ for $n \to +\infty$ in $(\mathbb{R} \setminus \mathbb{Q}, |\cdot|)$, there exists an $N \in \mathbb{N}$, such that

$$x_n \in B_{|\cdot|}(x, R) \qquad \text{for all } n \geq N,$$

which implies (7.6)

Hence, if (x_n) is a Cauchy sequence in $(\mathbb{R} \setminus \mathbb{Q}, d)$, and $x_n \to x \in \mathbb{R} \setminus \mathbb{Q}$ in $(\mathbb{R}, |\cdot|)$, then also $x_n \to x$ in the metric space $(\mathbb{R} \setminus \mathbb{Q}, d)$.

Then assume that ii) $x_n \to x \in \mathbb{Q}$ for $n \to 0 + \infty$ in the usual metric in $\mathbb{R}$, i.e. in $(\mathbb{R}, |\cdot|)$. We shall prove that (x_n) is *not* a Cauchy sequence in the metric space $(\mathbb{R} \setminus \mathbb{Q}, d)$. In other words, we shall prove that there exists an $\varepsilon > 0$, such that for every $N \in \mathbb{N}$ we can find m, $n \geqq N$, such that

$$(7.7) \qquad d(x_m, x_n) \geq \varepsilon,$$

It follows from the assumption $x \in \mathbb{Q}$ that the limit $x = r_{n_0}$ for some $n_0 \in \mathbb{N}$. We choose $\varepsilon := 2^{-n_0}$ and consider any $N \in \mathbb{N}$ in (7.7). If we also choose $m = N$, then we see that

$$c := \max_{k \leq n_0} \frac{1}{|x_m - r_k|} > 0.$$

From $|x_n - r_{n_0}| = |x_n - x| \to 0$ for $n \to +\infty$ follows that there exists an $n > m$, such that

$$|x_n - r_{n_0}| \leq \frac{1}{c+1}.$$

If we choose any such $n > m$, we get the estimate

$$d(x_m, x_n) \geq \frac{1}{2^{n_0}} \min\left\{1, \left|\max_{k \leq n_0} \frac{1}{|x_m - r_k|} - \max_{k \leq n_0} \frac{1}{|x_n - r_k|}\right|\right\} = \frac{1}{2^{n_0}} = \varepsilon,$$

and we have proved (7.7).

Summing up, we have proved that $(\mathbb{R} \setminus \mathbb{Q}, d)$ is a complete metric space. □

7.5 The diffusion equation

In 1995 I was asked to give a practical solution of the *diffusion equation*, also called the *heat equation* and *Fick's second law*, depending on in which connection the equation is met. It was a big surprise to me that although a solution formula has existed for a long time, it was very far from being of any practical use. In fact, consider the special initial/boundary problem

$$(7.8) \qquad \begin{cases} \dfrac{\partial C}{\partial t} = \dfrac{\partial^2 C}{\partial x^2}, & \text{for } x,\, t > 0, \\[2mm] C(x, 0) = 0, & \text{for } x > 0, \\[2mm] C(0, t) = \varphi(t), & \text{for } t > 0, \end{cases}$$

where $C(x, t)$ e.g. denotes the concentration of chloride ions in an half-infinite slab of concrete with chloride ions penetrating into the concrete from the left at $x = 0$, and where the boundary condition is a given function $\varphi(t)$.

If we furthermore assume that $\varphi \in C^1(\mathbb{R})$, then the unique bounded solution has been known for over a century,

$$(7.9) \qquad C(x, t) = \int_0^t \varphi'(t - \tau)\, \mathrm{erfc}\left(\frac{x}{\sqrt{4\tau}}\right) \mathrm{d}\tau = \int_0^t \varphi'(\tau)\, \mathrm{erfc}\left(\frac{x}{\sqrt{4(t-\tau)}}\right) \mathrm{d}\tau,$$

where $\mathrm{erfc}(u)$ denotes the complementary error function given by

$$\mathrm{erfc}(u) := \frac{2}{\sqrt{\pi}} \int_u^{+\infty} \exp\left(-s^2\right) \mathrm{d}s.$$

It is an easy exercise (left to the reader) to check that (7.9) indeed is a solution of (7.8).

All this was quite satisfactory, because formula (7.9) is exact. There was, however, only one drawback. It was not in general possible to calculate the convolution integrals of (7.9). It was only possible to calculate the solution, when $\varphi(t) = 1$ for $t > 0$ and $\varphi(t) = 0$ for $t < 0$, and in this case we must first apply a partial integration, because $\varphi \notin C^1(\mathbb{R})$. It is not continuous at $t = 0$. The result is the well-known

$$C(x, t) = \mathrm{erfc}\left(\frac{x}{\sqrt{4t}}\right) \qquad \text{for } t > 0, \text{ where } \varphi(t) \equiv 1.$$

The structure of (7.9) is as mentioned above a convolution integral, but even to date one is forced to use very powerful computers to solve (7.9), and programs like MATHEMATICA or MAPLE are not sufficient.

The original problem had to be solved, so I developed a different approach in [37], later repeated in [38]. Frankly speaking, I more or less stumbled over this alternative method, which I believed had been known for a long time. Asking recently a professor in linear partial differential operators I learned that it is apparently not known. In other words, one may still find alternative and simpler proofs than the usual ones known for e.g. more than a century.

In the following I shall give a proof diverging from the one given in [37], this time trying to argue why this method is quite natural.

Let us consider the heat equation alone without initial conditions, i.e.

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}. \tag{7.10}$$

Already *Boltzmann* was very well aware of that $x/\sqrt{4t}$ played a special role in the problem of diffusion, so this expression was therefore later called the *Boltzmann variable*. What I do not understand is why no one earlier has got the idea to change variables to

$$z = \frac{x}{\sqrt{4t}} \qquad \text{and} \quad \tau = t. \tag{7.11}$$

The Boltzmann variable invites to such a procedure, and the inverse is simple

$$x = 2z\sqrt{\tau} \qquad \text{and} \qquad t = \tau.$$

Then by the chain rule,

$$\frac{\partial}{\partial t} = \frac{\partial z}{\partial t}\,\frac{\partial}{\partial z} + \frac{\partial \tau}{\partial t}\,\frac{\partial}{\partial \tau} = -\frac{x}{2t\sqrt{4t}}\,\frac{\partial}{\partial z} + \frac{\partial \tau}{\partial t}\,\frac{\partial}{\partial \tau} = -\frac{z}{2\tau}\,\frac{\partial}{\partial z} + \frac{\partial}{\partial \tau},$$

and

$$\frac{\partial}{\partial x} = \frac{\partial z}{\partial x}\,\frac{\partial}{\partial z} + \frac{\partial \tau}{\partial x}\,\frac{\partial}{\partial \tau} = \frac{1}{\sqrt{4t}}\,\frac{\partial}{\partial z} + 0\cdot\frac{\partial}{\partial \tau} = \frac{1}{2\sqrt{\tau}}\,\frac{\partial}{\partial z},$$

from which

$$\frac{\partial^2}{\partial x^2} = \frac{1}{4\tau}\,\frac{\partial^2}{\partial z^2}.$$

By this change of variables equation (7.10) is transferred into

$$-\frac{z}{2\tau}\,\frac{\partial}{\partial z} + \frac{\partial}{\partial \tau} = \frac{1}{4\tau}\,\frac{\partial^2}{\partial z^2},$$

hence, by a rearrangement,

$$\frac{\partial u}{\partial \tau} = \frac{z}{2\tau}\,\frac{\partial u}{\partial z} + \frac{1}{4\tau}\,\frac{\partial^2 u}{\partial z^2}. \tag{7.12}$$

If we require that the variables are separated in some solution u, then it follows from (7.12) that the only possible structure of u is

$$u(\tau, z) := \tau^p\,\Psi_p(z), \qquad p \in \mathbb{R}, \tag{7.13}$$

where p is some constant, and $\Psi_p(z)$ is a corresponding unknown function. When (7.13) is inserted into (7.12), we get

$$p\,\tau^{p-1}\,\Psi_p(z) = \frac{z}{2}\cdot\tau^{p-1}\,\Psi_p'(z) + \frac{1}{4}\,\tau^{p-1}\,\Psi_p''(z),$$

hence, by cancelling the common factor τ^{p-1}, followed by a rearrangement,

$$\Psi_p''(a) + 2z\,\Psi_p'(z) - 4p\,\Psi_p(z) = 0. \tag{7.14}$$

This is an ordinary linear differential equation of second order and with polynomial coefficients, so it can be solved by insertion of a power series and then identifying the coefficients. The relevant solutions are (after some tedious computations)

$$\Psi_p(z) = \sum_{n=0}^{+\infty}\frac{1}{(2n)!}\,p^{(n)}(2z)^{2n} - \frac{\Gamma(p+1)}{\Gamma(p+\frac{1}{2})}\sum_{n=0}^{+\infty}\frac{1}{(2n+1)!}\left(p-\frac{1}{2}\right)^{(n)}(2z)^{2n+1}, \tag{7.15}$$

where we have used the notation of descending factors for $q \in \mathbb{R}$,

$$q^{(0)} := 1, \quad \text{and} \quad q^{(n)} := q(q-1)\cdots(q-n+1) \qquad (n \text{ factors}), \quad n \ge 1.$$

We get in particular for $p = 0$, $\frac{1}{2}$ and 1, that

$$\Psi_0(z) = \operatorname{erfc}(z),$$

$$\Psi_{0.5}(z) = \exp\left(-z^2\right) - \sqrt{\pi}\,z\operatorname{erfc}(z),$$

$$\Psi_1(z) = \left(1+2z^2\right)\operatorname{erfc}(z) - \frac{2}{\sqrt{\pi}}\,z\,\exp\left(-z^2\right),$$

supplied with

$$\Psi_{-0.5}(z) = \exp\left(-z^2\right),$$

because the functions Ψ_p are defined for $p > -1$, though in practice only relevant for $p \ge 0$.

It can alternatively be shown that if

$${}_1F_1[a;b;x] := 1 + \frac{a}{b}\,x + \frac{a(a+1)}{b(b+1)}\,\frac{x^2}{2!} + \frac{a(a+1)(a+2)}{b(b+1)(b+3)}\,\frac{x^3}{3!} + \cdots$$

denotes the *hypergeometric function* (*Kummer's function*), then we have in general for $p > -1$,

$$\Psi_p(z) = {}_1F_1\left[-p;\frac{1}{2};-z^2\right] - 2z\,\frac{\Gamma(p+1)}{\Gamma(p+\frac{1}{2})}\,{}_1F_1\left[-p+\frac{1}{2};\frac{3}{2};-z^2\right],$$

a formula, which has the advantage that Kummer's function is programmed in e.g. MAPLE, so in practice these formulæ are easily to apply.

Using the functions (7.15) and the special solutions (7.13) it is easy to build an approximate solution of the problem (7.8), when $\varphi(t)$ is continuous in some bounded half-open interval $]0,T]$, and where $\varphi(t) \to a_0$ for $t \to 0+$. In fact, if we put $\tilde{\varphi}(0) = a_0$, and $\tilde{\varphi}(t) = \varphi(t)$ for $t \in]0,T]$, then $\tilde{\varphi}(t)$ is continuous in a compact interval $[0,T]$. Then it follows from *Weierstraß's approximation theorem*, that for every $\varepsilon > 0$ there exists a polynomial $P_\varepsilon(t)$, such that

$$|\tilde{\varphi}(t) - P_\varepsilon(t)| < \varepsilon \qquad \text{for all } t \in [0,T].$$

If

$$P_\varepsilon(t) = \sum_{j=0}^{n} a_j\, t^j, \tag{7.16}$$

it can be proved that if $C(x,t)$ is the unique solution of (7.8), and

$$C_\varepsilon(x,t) := \sum_{j=0}^{n} a_j\, t^t\, \Psi_j\left(\frac{x}{\sqrt{4t}}\right), \tag{7.17}$$

then

$$|C(x,t) - C_\varepsilon(x,t)| \le \varepsilon \qquad \text{for all } x > 0 \text{ and all } t \in [0,T],$$

so $C_\varepsilon(x,t)$ is an uniform approximation of the solution $C(x,t)$ of error $< \varepsilon$ for all $x > 0$ and $t \in [0,T]$.

Since the family of functions $\{t^p \mid p \ge 0\}$ contains the smaller family of monomials $\{t^p \mid p \in \mathbb{N}_0\}$, we may also approximate $\tilde{\varphi}(t)$ with sums of the type

$$\tilde{P}_\varepsilon(t) := \sum_{j=0}^{m} a_j\, t^{p_j},$$

whenever this is more convenient.

We shall not here go into the proofs of these approximations. We only notice that they rely on the fact that the operation going from (7.16) to (7.17) is a contraction, so the error is largest at the boundary $x = 0$, where we already know that it is $< \varepsilon$.

Although this result is nice, it is for the time being not the most useful one. The reason is that at present the problem is usually a different one, namely given a set of data, how can we approximate this set of data by linear combinations of solutions of the type (7.17), where we tacitly have included the diffusion coefficient in the problem (7.8)? The first attempt in that direction was made by *Collepardi* [12], [13] and [14], from 1970 and the following years. He suggested to approximate a set of measurements by the complementary error function, i.e. by

$$c_0 \operatorname{erfc}\left(a \frac{x}{\sqrt{4t}}\right) = c_0 \Psi_0\left(a \frac{x}{\sqrt{4t}}\right), \tag{7.18}$$

with the two unknown constants a and c_0.

By using the statistical method of least squares to fix the two unknown constants a and c_0, he obtained a lot of excellent results, which were not known before, so this was really a big step forward. In the 1980s, however, one discovered some examples where the Collepardi method did not apply very well. This was why there was a need of a better model.

In line with the above, which was not known before 1970, the next approach would be to model with

$$c_0 \Psi\left(a \frac{x}{\sqrt{4t}}\right) + c_1 t \Psi_1\left(a \frac{x}{\sqrt{4t}}\right) \tag{7.19}$$

with three unknowns a, c_0 and c_2 and then use the method of the least square approximation. Neither of the two approximations are fully correct. In (7.18) we tacitly approximate the boundary value $\tilde{\varphi}(t)$ with a constant c_0, and in the second case (7.19) we approximate the boundary value $\tilde{\varphi}(t)$ with the affine map $c_0 + c_1 t$. Since the latter one includes the former one by choosing $c_1 = 0$, we must necessarily get a better approximation by using (7.19) instead of (7.18).

Notice that if the same set of measurements is modelled by both (7.18) and (7.19), the two constants a and c_1 are not the same in the two cases, when $c_1 \neq 0$.

One should also notice that the approximation (7.18) is always *decreasing* in x, when t is fixed, while (7.19) in some cases for fixed t is first *increasing* in x to some maximum, before it decreases. This is apparently against common sense, although it was predicted in 1995. Recently, in 2015, this phenomenon was actually observed in some data measured at the Swedish research station at Träslövsläge on the Swedish west coast. Here slaps of concrete of different composition had been exposed to salty water for 20 years. In the measurements they discovered that in some cases the concentration of chloride ions reached a maximum at some distance from the surface. This can never be modelled by the complementary error function solution alone. But it can by the more general model of (7.19).

It is obvious that we can extend this method by including more terms of the special structure in the approximation. Each time we add a new unknown constant, which should be found by the method of least squares. There is no need here to go further into this theory.

8 Measures in $\mathbb{R}^n$ and in infinite dimensional Hilbert space

8.1 Introduction

We shall in Section 8.2 consider the geometry and measure of balls in $\mathbb{R}^n$ for large n. This theory is the stepping stone for Section 8.3, in which we consider a Gaussian measure on an infinitely dimensional separable Hilbert space, for which no differentiation theorem exists with respect to the systems of small balls of contracting radii (and fixed centre). We shall use this result in Section 8.4 to derive an unexpected positivity principle in Hilbert space.

8.2 Surface measure and volume measure in $\mathbb{R}^n$

Most people have a vague sense of that the Euclidean geometry of $\mathbb{R}^n$ resembles the ordinary geometry of the usual space $\mathbb{R}^3$. However, it should not be a surprise that we also here may find some unexpected results. We shall here only deal with two closely connected results, none of which can be called obvious.

We shall in the following also vary the dimension $n \in \mathbb{N}$, so let us first fix the notation,

A point in $\mathbb{R}^n$ is denoted by $\mathbf{x} = (x_1, \dots, x_n)$, or similarly, and we define the *Euclidean norm* of $\mathbf{x}$ by

$$\|\mathbf{x}\| := \sqrt{x_1^2 + \cdots + x_n^2}.$$

Notice that we shall use the same notation of the norm, no matter the dimension $n \in \mathbb{N}$ of the space.

The n-dimensional *volume* of a body $L \subset \mathbb{R}^n$ is denoted by $|L|_n$. If $\mathfrak{F}$ is some $(n-1)$-dimensional "surface" in $\mathbb{R}^n$, then we denote the *surface area* of $\mathfrak{F}$ by $\sigma_{n-1}(\mathfrak{F})$.

We shall prove the following theorem.

Theorem 8.1 *Let $0 \le a < b < 1$ be given constants, and denote by*

$$\omega_n := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1,\, a \le x_1 \le b\}, \quad \textit{and} \quad \Omega_n := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1,\, b \le x_1 \le 1\}$$

two slices of the surfaces of the unit sphere in $\mathbb{R}^n$.
To every constant $c > 0$ we can find $n_0 = n_0(a, b, c) \in \mathbb{N}$, such that for all dimensions $n \ge n_0$,

$$\sigma_{n-1}(\omega_n) \ge c \cdot \sigma_{n-1}(\Omega_n).$$

Remark. The result must have been known for a long time, although I have not been able to find any reference. I guess that it is at latest from the 1920s, though something in the spirit of the example indicates that it is much older. The following proof is my own reconstruction, and it is probably more or less different from the original proof. $\Diamond$

We consider Figure 8.1 as a model of the general case. We see that Theorem 8.1 claims that the $(n-1)$-dimensional surface area of the small slab ω_n can be much larger than the $(n-1)$-dimensional surface area of Ω_n, which is lying further away from the centre of the ball, provided that the dimension n is chosen sufficiently large, Notice that we may choose the length of the projection $b - a$ of the slab onto the X_1-axis much smaller than the length of the projection $1 - b$ of the remainder part of the sphere bounded away from the centre.

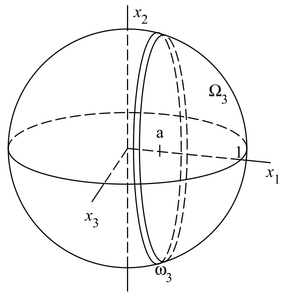

Figure 8.1: *Visualisation of Theorem 8.1 in the case of $n = 3$, i.e. in $\mathbb{R}^3$. For higher dimensions, $n > 3$, it is difficult to imagine, what is going on.*

We shall start with some simple definitions and results.

Definition 8.1 *A continuous positive real function on an interval I in $\mathbb{R}_+$, i.e. $f : I \to \mathbb{R}$, is called* logarithmic concave, *if $\ln f$ is concave on I. This means that for every a, $b \in I$ and every $0 < \lambda < 1$,*

$$\lambda \cdot \ln f(a) + (1 - \lambda) \ln f(b) \leq \ln f(\lambda a + (1 - \lambda)b).$$

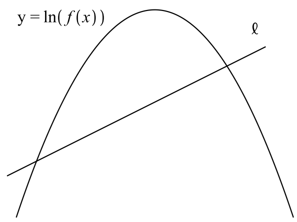

Figure 8.2: *Graph of $\ln f(x)$ for a logarithmic concave function $f : I \to \mathbb{R}_+$.*

The geometric interpretation of this definition is the following. For any points a, $b \in I$, where $a < b$, the straight line ℓ through the points $(a, \ln f(a))$ and $(b, \ln f(b))$ lies below the curve $y = \ln f(x)$ for $x \in]a, b[$, and above the curve for $x \in I \setminus [a, b]$. We shall in the following use this geometrical observation.

If $f \in C^2$, then f is logarithmic concave, whenever

$$\frac{d^2}{dx^2} \ln f(x) \leq 0, \qquad \text{for all } x \in I.$$

Lemma 8.1 *Assume that $f : [a, b] \to [0, +\infty[$ is continuous and strictly increasing, and that f is logarithmic concave in $]a, b$. Then*

$$\int_x^b f(t)\, dt \geq \left\{ \frac{f(b)}{f(x)} - 1 \right\} \int_a^x f(t)\, dt, \qquad \textit{for all } x \in]a, b[.$$

PROOF. Let $t \in]a, b[$. Then we can write

$$t = \frac{b-t}{b-x} \cdot x + \frac{t-x}{b-x} \cdot b,$$

which is easily checked. By assumption, $\ln f$ is concave, so outside the interval $[x, b]$, i.e. in $]a, x[$,

$$\ln f(t) \le \frac{b-t}{b-x} \ln f(x) + \frac{t-x}{b-x} \ln f(b) = \ln f(b) + \frac{b-t}{b-x} \ln\left(\frac{f(x)}{f(b)}\right), \qquad \text{for } t \in]a, x[.$$

If instead $t \in]x, b[$, then the straight line through $(x, \ln f(x))$ and $(b, \ln f(b))$ lies below the graph of $y = \ln f(t)$, so

$$\ln f(t) \ge \frac{b-t}{b-x} \ln f(x) + \frac{t-x}{b-x} \ln f(b) = \ln f(b) + \frac{b-t}{b-x} \ln\left(\frac{f(x)}{f(b)}\right), \qquad \text{for } t \in]x, b[.$$

If we introduce the shorthand

$$k := \frac{1}{b-x} \qquad \text{and} \qquad c := \frac{f(b)}{f(x)} > 1,$$

then these estimates can be written

$$f(t) \le f(b) \cdot e^{-kb} \cdot c^{kt}, \qquad \text{for } t \in]a, x[,$$

$$f(t) \ge f(b) \cdot e^{-kb} \cdot c^{kt}, \qquad \text{for } t \in]x, b[.$$

It follows from the first estimate that if we use the definition of k in the latter equality, then

$$\begin{aligned} c\int_a^x f(t)\,\mathrm{d}t &\le f(b)\cdot c^{-kb}\cdot c\int_a^x c^{kt}\,\mathrm{d}t = \frac{f(b)}{k\cdot\ln c}\cdot c\cdot c^{-kb}\left(c^{kx}-c^{ka}\right) \\ &\le \frac{f(b)}{k\cdot\ln c}\cdot c\cdot c^{-k(b-x)} = \frac{f(b)}{k\cdot\ln c}. \end{aligned}$$

Similarly we get from the second estimate,

$$\int_x^b f(t)\,\mathrm{d}t \ge \frac{f(b)}{k\cdot\ln c}\cdot c^{-kb}\left(c^{kb}-c^{kx}\right) = \frac{f(b)}{k\cdot\ln c}\left\{1-\frac{1}{c}\right\}.$$

When we combine these estimates, we finally get by insertion of the definition of $c = \dfrac{f(b)}{f(x)}$,

$$\int_x^b f(t)\,\mathrm{d}t \ge \left(1-\frac{1}{c}\right)c\int_a^x f(t)\,\mathrm{d}t = \left\{\frac{f(b)}{f(x)}-1\right\}\int_a^x f(t)\,\mathrm{d}t,$$

and the lemma is proved. □

Finally, we turn to

Proof of Theorem 8.1. The trick is to apply spherical coordinates in $\mathbb{R}^n$, i.e.

$$\begin{cases} x_1 &= r\cdot\cos\varphi_1, \\ x_2 &= r\cdot\sin\varphi_1\cdot\cos\varphi_2, \\ x_3 &= r\cdot\sin\varphi_1\cdot\sin\varphi_2\cdot\cos\varphi_3, \\ \vdots & \qquad\qquad \vdots \\ x_{n-1} &= r\cdot\sin\varphi_1\cdot\sin\varphi_2\cdot\sin\varphi_3\cdots\sin\varphi_{n-2}\cdot\cos\varphi_{n-1}, \\ x_n &= r\cdot\sin\varphi_1\cdot\sin\varphi_2\cdot\sin\varphi_3\cdots\sin\varphi_{n-2}\cdot\sin\varphi_{n-1}. \end{cases}$$

It is possible to prove, e.g. by induction (left to the reader), that the Jacobian J_n corresponding to this change of coordinates is given by

$$J_n = r^{n-1}\cdot\sin^{n-2}\varphi_1\cdot\sin^{n-3}\varphi_3\cdots\sin^2\varphi_{n-3}\cdot\sin\varphi_{n-2}.$$

Notice in particular that the variable φ_{n-1} does not occur in J_n. It is proved in the elementary calculus that $J_2 = r$ (polar coordinates in the plane) and $J_3 = r^2\sin\varphi_1$ (spherical coordinates in the 3-dimensional space), which are in agreement with the general formula above.

We have expressed in the general spherical coordinates in $\mathbb{R}^n$,

$$\hat{\Omega}_n = \left\{\mathbf{x}\in\mathbb{R}^n \mid r=1 \text{ and } 0\le\varphi_1\le\theta\right\},$$

and

$$\hat{\omega}_n = \left\{\mathbf{x}\in\mathbb{R}^n \mid r=1 \text{ and } \theta\le\varphi_1\le\theta+\alpha\right\},$$

where

$$\theta = \arccos b \in \left]0,\frac{\pi}{2}\right[\quad \text{and} \quad \theta+\alpha = \arccos\alpha \in \left]0,\frac{\pi}{2}\right[.$$

Then

$$\text{(8.1)} \qquad \frac{\sigma_{n-1}(\omega_n)}{\sigma_{n-1}(\Omega_n)} = \frac{|\hat{\omega}_n|_n}{\left|\hat{\Omega}_n\right|_n} = \frac{\int_\theta^{\theta+\alpha} \sin^{n-2}\varphi_1 \, \mathrm{d}\varphi_1}{\int_0^\theta \sin^{n-2}\varphi_1 \, \mathrm{d}\varphi_1},$$

because when we compute the volumes $|\hat{\omega}_n|$ and $\left|\hat{\Omega}_n\right|$ the variables can be separated and then some of the factors cancel each other.

When $n > 2$, the function $f(\varphi) := \sin^{n-2}\varphi$ is logarithmic concave in the interval $\left]0, \frac{\pi}{2}\right[$, because

$$\frac{d^2}{d\varphi^2} \ln f(\varphi) = -\frac{n-2}{\sin^2\varphi} < 0.$$

Clearly, $f(\varphi)$ is strictly increassing in $\left[0, \frac{\pi}{2}\right]$, so it follows from Lemma 8.1 that

$$\text{(8.2)} \qquad \int_0^{\theta+\alpha} \sin^{n-2}\varphi \, \mathrm{d}\varphi \geq \left\{ \left(\frac{\sin(\theta+\alpha)}{\sin\theta} \right)^{n-2} - 1 \right\} \int_0^\theta \sin^{n-2}\varphi \, \mathrm{d}\varphi.$$

Finally, the trivial estimate $\sin(\theta+\alpha) > \sin\theta$ implies that we can find a constant n_0, such that

$$\left(\frac{\sin(\theta+\alpha)}{\sin\theta} \right)^{n-2} - 1 \geq c \qquad \text{for all } n \geq n_0,$$

and the theorem follows from (8.1) and (8.2). □

A simple consequence of Theorem 8.1 is the following

Corollary 8.1 *Given constants* $0 \leq a < b < 1$, *define the bodies*

$$\tilde{\omega}_n := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq 1, \, a \leq x_1 \leq b\} \quad \text{and} \quad \tilde{\Omega}_n := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq 1, \, b \leq x_1 \leq 1\}.$$

Given any constant $c > 0$. *There exists* $n_0 = n_0(a, b, c) \in \mathbb{N}$, *such that for all* $n \geq n_0$,

$$|\tilde{\omega}_n|_n \geq c \cdot \left|\tilde{\Omega}_n\right|_n.$$

Intuitively one would guess that the measure of $\tilde{\Omega}_n$ would be the larger, when $b - a \ll 1 - b$. Nevertheless, the geometric position of the thin slab $\tilde{\omega}_n$ closer to the centre implies that the measure of $\tilde{\omega}_n$ for fixed a and b can be bigger than $c \cdot \left|\tilde{\Omega}_n\right|_n$ for any constant $c > 0$, by choosing the dimensions n sufficiently large.

Proof of Corollary 8.1. When we consider a shell in $\mathbb{R}^n$, given by $\|\mathbf{x}\| = r$, where $r \in [b, 1]$, it follows that we can apply the same proof as in Theorem 8.1 with the modifications, that the limits are now given by

$$\theta_r = \arccos\frac{b}{r} \qquad \text{and} \qquad \theta_r + \alpha_r = \arccos\frac{a}{r}.$$

In particular,

$$\begin{aligned}\frac{\sin(\theta_r+\alpha_r)}{\sin\theta_r} &= \sqrt{\frac{1-\dfrac{a^2}{r^2}}{1-\dfrac{b^2}{r^2}}} = \sqrt{\frac{1-a^2}{1-b^2}+\frac{\left(b^2-a^2\right)\left(1-r^2\right)}{\left(r^2-b^2\right)\left(1-b^2\right)}} \\ &\geq \sqrt{\frac{1-a^2}{1-b^2}} = \frac{\sin(\theta_1+\alpha_1)}{\sin\theta_1} = \frac{\sin(\theta+\alpha)}{\sin\theta},\end{aligned}$$

where $\theta=\theta_1$ and $\alpha=\alpha_1$ were already given in the proof of Theorem 8.1. Let n_0 be the constant defined in Theorem 8.1. Then for all $n\geq n_0$,

$$\sigma_{n-1}(\omega_{r,n-1})\geq c\cdot\sigma_{n-1}(\Omega_{r,n-1}), \qquad \text{for } r\in[b,1].$$

Separating the variables, we obtain for $n\geq n_0$, that

$$\begin{aligned}\frac{\left|\tilde{\omega}_n\right|_n}{\left|\tilde{\Omega}_n\right|_n} &\geq \frac{\int_b^1\int_{\theta_r}^{\theta_r+\alpha_r}\sin^{n-2}\varphi_1\,\mathrm{d}\varphi_1\,\mathrm{d}r}{\int_b^1\int_0^{\theta_r}\sin^{n-2}\varphi_1\,\mathrm{d}\varphi_1\,\mathrm{d}r} \geq \frac{\int_b^1\left\{\left(\dfrac{\sin(\theta_r+\alpha_r)}{\sin\theta_r}\right)^{n-2}-1\right\}\int_0^{\theta_r}\sin^{n-2}\varphi_1\,\mathrm{d}\varphi_1\,\mathrm{d}r}{\int_b^1\int_0^{\theta_r}\sin^{n-2}\varphi_1\,\mathrm{d}\varphi_1\,\mathrm{d}r} \\ &\geq \min\left\{\left(\frac{\sin(\theta_r+\alpha_r)}{\sin\theta_r}\right)^{n-2}-1 \;\middle|\; r\in[b,1]\right\}\geq c,\end{aligned}$$

and the corollary is proved. □

8.3 Negative differentiation results

8.3.1 The main result and some notation

It was proved in Section 3.8 that it is possible to construct a metric space (M,d) with two different probability measures, which nevertheless agree on every ball in (M,d). However, Hilbert spaces are nicer, because it can be proved that a measure is uniquely determined by its values on a family of small balls, e.g. on all balls of radius ≤ 1. Unfortunately, this property is not sufficient for also having a *differentiation theorem*, or just a density theorem,, so we cannot in general conclude that if μ is a measure on a Hilbert space, then

$$\lim_{r\to 0+}\left\{\frac{1}{\mu B[x,r]}\int_{B[x,r]}f(y)\,\mathrm{d}\mu(y)\right\}=f(x) \qquad \text{almost everywhere.} \tag{8.3}$$

It is known from classical Measure Theory that (8.3) holds for all Euclidean spaces, i.e. for all *finite dimensional* Hilbert spaces. But in infinite dimensional Hilbert spaces this is not always true, which is the result of the following theorem.

Theorem 8.2 *There exist a Gaussian measure γ on a separable Hilbert space H and a nonnegative function $f\in L^1(\gamma)$, such that*

$$\liminf_{r\to 0+}\left\{\frac{1}{\gamma B[x,s]}\int_{B[x,s]}f(y)\,d\gamma(y) \;\middle|\; x\in H,\ 0<s\leq r\right\}=+\infty.$$

The idea is simple, but the technical proof itself is very complicated. We choose a sequence of Euclidean spaces of higher and higher dimensions and supply each of them with a (finite dimensional) *Gaussian measure*, where we "squeeze" the parameters in an appropriate way and then put all spaces together, such that the wanted result follows. We shall even construct the function $f \in L^1(\gamma)$, $f \geq 0$, such that we have the uniform estimate

$$\frac{1}{\gamma B} \int_B f \, \mathrm{d}\gamma \geq \frac{1}{\ln 2} \ln\left(\frac{1}{r^2}\right) - 2,$$

whenever the radius r of the ball B is smaller than 2.

The first negative results in this direction were obtained by [45]. We shall here follow a more streamlined version, also given by D. Preiss (private communication).

Let γ be a Gaussian measure on a separable Hilbert space H, and let $\sigma := \left(\sigma_j^2\right)_{j\in\mathbb{N}}$ be the sequence of eigenvalues of the covariance operator of the measure γ. The covariance operator is nuclear, so

$$\sum_{j=1}^{+\infty} \sigma_j^2 < +\infty,$$

and we may introduce the inner product

$$(x,y)_\sigma := \sum_{j=1}^{+\infty} \sigma_j^2\, x_j\, y_j$$

in another Hilbert space

$$\ell^2(\sigma) := \left\{ x \in \mathbb{R}^{\mathbb{N}} \;\middle|\; \sum_{j=1}^{+\infty} \sigma_j^2\, x_j^2 < +\infty \right\}.$$

Then notice that the isomorphism

$$I : \ell^2 \to \ell^2(\sigma), \qquad I\left(x_j\right) = \left(\sigma_j^{-1}\, x_j\right)$$

maps the measure γ onto the measure $\mu = I[\gamma]$, the latter being the countable product $\nu^{\mathbb{N}}$ of standard one-dimensional Gaussian measures ν.

We shall in the following choose the sequence $\left(\sigma_j^2\right)$ piecewisely constant. This is done in the following way. Let (n_j) be an increasing sequence of integers to be fixed later. We define

$$p_0 := 0, \qquad \text{and} \qquad p_i := \sum_{j=1}^{i} n_j \quad \text{for } i \in \mathbb{N},$$

$$\sigma_j^2 := \frac{1}{2^i\, n_i}, \qquad \text{whenever } p_{i-1} < j \le p_i, \quad i \in \mathbb{N}.$$

Then the space $\ell^2(\sigma)$ is given by

$$\ell^2(\sigma) = \left\{ x \in \prod_{j=1}^{+\infty} \mathbb{R}^{n_j} \;\middle|\; \sum_{j=1}^{+\infty} \frac{1}{2^j\, n_j}\, \|x\|_{n_j}^2 < +\infty \right\},$$

where the symbol $\|\cdot\|_{n_j}$ denotes the Euclidean norm in $\mathbb{R}^{n_j}$.

Strictly speaking, every $x \in \prod_{j=1}^{+\infty} \mathbb{R}^{n_j}$ is a sequence of elements $x_j \in \mathbb{R}^{n_j}$, so instead of $\|x\|_{n_j}$ we should write $\|x_j\|_{n_j}$, but this inconsistency will hardly confuse anynone.

The corresponding norm $\|\cdot\|_\sigma$ on $\ell^2(\sigma)$ is defined by

$$\|x\|_\sigma^2 := \sum_{j=1}^{+\infty} \frac{1}{2^j\, n_j}\, \|x\|_{n_j}^2.$$

We shall also define the following Hilbert spaces,

$$H_k := \prod_{j=1}^{k} \mathbb{R}^{n_j} \qquad \text{(actually finite dimensional, i.e. a Euclidean space)}$$

equipped with the norm $\|\cdot\|_{\sigma,k}$, given by

$$\|x\|_{\sigma,k}^2 := \sum_{j=1}^{k} \frac{1}{2^j\, n_j}\, \|x\|_{n_j}^2,$$

and the infinite separable Hilbert space

$$H^k := \left\{ x \in \prod_{j=k}^{+\infty} \mathbb{R}^{n_j} \;\middle|\; \sum_{j=k}^{+\infty} \frac{1}{2^j\, n_j}\, \|x\|_{n_j}^2 < +\infty \right\},$$

with the norm $|\cdot|_{\sigma,k}$ given by

$$|x|_{\sigma,k}^2 := \sum_{j=k}^{+\infty} \frac{1}{2^j\, n_j}\, \|x\|_{n_j}^2.$$

The standard *Gaussian measure* in $\mathbb{R}^{n_j}$ is denoted by ν_{n_j}. We also introduce the product measures

$$\mu := \prod_{j=1}^{+\infty} \nu_{n_j}, \qquad \mu_k := \prod_{j=1}^{k} \nu_{n_j} \quad \text{and} \quad \mu^k := \prod_{j=k}^{+\infty} \nu_{n_j},$$

as the corresponding measuress on H, H_k and H^k, respectively. We notice that

$$H = H_k \times H^{k-1}, \qquad \|x\|_\sigma^2 = \|x\|_{\sigma,k}^2 + |x|_{\sigma,k+1}^2, \quad \text{and} \quad \mu = \mu_k \times \mu^{k+1}.$$

Finally, we use the following notations for closed balls,

$$K_n[x,r] := \{y \in \mathbb{R}^n \mid \|y-x\|_n \le r\} \subset \mathbb{R}^n,$$

$$B_k[x,r] := \{y \in H_k \mid \|y-x\|_{\sigma,k} \le r\} \subset H_k,$$

$$B^k[x,r] := \left\{y \in H^k \mid |y-x|_{\sigma,k} \le r\right\} \subset H^k,$$

$$B[x,k] \;:= \{y \in H \mid \|y-x\| \le r\} \subset H.$$

The proof of Theorem 8.2 is split into three parts, each described in one of the following subsections. We first investigate Gaussian measures in (finite dimensional) Euclidean spaces of high dimension. In the following subsection we apply the obtained finite dimensional results to also get some technical results in the infinite dimensional case, including the specification of the sequence (n_j), which is the crucial step. Finally, we put all things together and construct a nonnegative function $f \in \ell^2(H)$, which satisfies the condition of Theorem 8.2.

8.3.2 Gaussian measures in Euclidean spaces of high dimension

We shall in this subsection derive some properties concerning the geometry of balls in Euclidean spaces of high dimension. The first results were already given in Section 8.2, where we focussed on the simple part of the proof, which can be read independently. We might of course just refer to Section 8.2, but in order not to interrupt the context we here repeat the construction in a slightly different way and with a slightly different notation.

We say that a function $f : [a,b] \to \mathbb{R}_+$ is *logarithmic concave*, if $\ln f : [a,b] \to \mathbb{R}$ is concave.

Lemma 8.2 *Given constants $a \in \mathbb{R} \cup \{-\infty\}$ and $b \in \mathbb{R} \cup \{+\infty\}$. Let $f : [a, b] \to \mathbb{R}_+$ be a logarithmic concave function, and let $x \in]a, b[$.*

1) If $f(x) < f(b)$, then

$$\int_x^b f(t)\, dt \geq \left\{ \frac{f(b)}{f(x)} - 1 \right\} \int_a^x f(t)\, dt.$$

2) If $f(x) < f(a)$, then

$$\int_a^x f(t)\, dt \geq \left\{ \frac{f(a)}{f(x)} - 1 \right\} \int_x^b f(t)\, dt.$$

PROOF. We shall only prove 1), since the proof of 2) is analogous. Let $t \in]a, b[$. Then we can write

$$t = \frac{b-t}{b-x}\, x + \frac{t-x}{b-x}\, b, \qquad \text{where} \qquad \frac{b-t}{b-x} + \frac{t-x}{b-x} = 1,$$

Since $\ln f$ is assumed to be concave, this implies after a small calculation that

$$\ln f(t) \leq \ln f(b) + \frac{b-t}{b-x} \ln\left(\frac{f(x)}{f(b)} \right) \qquad \text{for } t \in]a, x[,$$

and

$$\ln f(t) \geq \ln f(b) + \frac{b-t}{b-x} \ln\left(\frac{f(x)}{f(b)} \right) \qquad \text{for } t \in]x, b[.$$

We write for short

$$\lambda = \frac{1}{b-x} \qquad \text{and} \qquad c = \frac{f(b)}{f(x)} > 1,$$

where the latter inequality follows from the assumption $f(x) < f(b)$ in 1).

The estimates above are then rewritten as

$$f(t) \le f(b) \cdot c^{-\lambda b} \cdot c^{\lambda t} \qquad \text{for } t \in]a, x[,$$

and

$$f(t) \ge f(b) \cdot c^{-\lambda b} \cdot c^{\lambda t} \qquad \text{for } t \in]x, b[.$$

Then integrate the first inequality over the interval $]a, x[$ to get

$$\begin{aligned} c \int_a^x f(t)\,\mathrm{d}t \quad &\le \quad f(b) \cdot c \cdot e^{-\lambda b} \int_a^x c^{\lambda t}\,\mathrm{d}t = \frac{f(b)}{\lambda \ln c} \cdot c \cdot c^{-\lambda t}\left(c^{\lambda x} - c^{\lambda a}\right) \\ &\le \quad \frac{f(b)}{\lambda \ln c} \cdot c \cdot c^{-\lambda(b-x)} = \frac{f(b)}{\lambda \ln c}, \end{aligned}$$

where we have used the definition of λ, so

$$\int_a^x f(t)\,\mathrm{d}t \le \frac{f(b)}{\lambda \cdot c \ln c}.$$

Similarly, by integrating the second inequality above over the interval $]x, b[$, we get

$$\int_x^b f(t)\,\mathrm{d}t \ge \frac{f(b)}{\lambda \cdot \ln c} \cdot c^{-\lambda b}\left(c^{\lambda b} - c^{\lambda x}\right) = \frac{f(b)}{\lambda \ln c}\left\{1 - \frac{1}{c}\right\}.$$

When we combine these two estimates and use that $c = f(b)/f(x)$, we finally get

$$\int_x^b f(t)\,\mathrm{d}t \ge \left\{1 - \frac{1}{c}\right\} c \int_a^x f(t)\,\mathrm{d}t = \left\{\frac{f(b)}{f(x)-1}\right\} \int_a^x f(t)\,\mathrm{d}t,$$

and 1) is proved. □

We shall use Lemma 8.2 to prove the following result on surface areas in Euclidean spaces of high dimension. This result was also proved in Section 8.2 with a slightly different notation.

We shall here include Figure 8.1 from Section 8.2 (in three dimensions) as Figure 8.3, so the reader can get an idea of the content of Lemma 8.3 below.

We first need some new notation. Let $\mathcal{H}^n$ denote the n-dimensional *Hausdorff measure*, and let S^{n-1} denote the unit sphere in $\mathbb{R}^n$, i.e.

$$S^{n-1} := \{x \in \mathbb{R}^n \mid \|x\|_n = 1\}.$$

Let $\theta \in [0, \pi]$ be some constant. Then we define

$$S_\theta^{n-1} := \left\{x = (x_1, \dots, x_n) \in S^{n-1} \mid x_n \ge \cos\theta\right\},$$

i.e. S_θ^{n-1} is the subset of the unit sphere S^{n-1}, which lies "above" the hyperplane defined by $x_n = \cos\theta$.

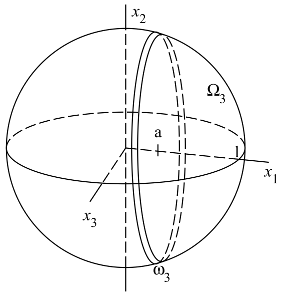

Figure 8.3: *Visualisation of Lemma 8.3 in the case of* $n = 3$*, i.e. in* $\mathbb{R}^3$*. For higher dimensions,* $n > 3$*, it is difficult to imagine, what is going on.*

Lemma 8.3 *Assume that* $0 \le \theta < \vartheta \le \frac{\pi}{2}$*. Then*

$$\mathcal{H}^{n-1}\left(S_\theta^{n-1} \setminus S_\vartheta^{n-1}\right) \ge \left\{\left(\frac{\sin\vartheta}{\sin\theta}\right)^{n-2} - 1\right\} \mathcal{H}^{n-1}\left(S_\theta^{n-1}\right) \qquad \textit{for } n > 2.$$

Clearly, Lemma 8.3 contains Theorem 8.1.

Notice that the relation between the $\mathcal{H}^{n-1}$ measures of the subsets of the unit sphere does not depend on the radius r of the sphere ($= 1$ for the unit sphere). Since we shall consider spheres of general radius $r > 0$, we here rewrite the statement of Lemma 8.3 in a more elaborate way, using the previous introduced notation,

$$\begin{aligned}&\mathcal{H}^{n-1}\left\{x \in \mathbb{R}^n \mid \|x\|_n \cos\vartheta \le x_n \le \|x\|_n \cos\theta\right\}\\ &\qquad \ge \left\{\left(\frac{\sin\vartheta}{\sin\theta}\right)^{n-2} - 1\right\} \mathcal{H}^{n-1}\left(\left\{x \in \mathbb{R}^n \mid \|x\|_n = v,\ \|x\|_n \cos\theta\right\}\right)\end{aligned} \tag{8.4}$$

Proof of Lemma 8.3. The slab $S_\vartheta^{n-1} \setminus S_\theta^{n-1}$ can be considered as a part of the *graph* of the function

$$f\left(x_1, \dots, x_{n-1}\right) := \left\{\sum_{i=1}^{n-1} x_i^n\right\}^{\frac{1}{2}}, \qquad \text{for } \left(x_1, \dots, x_n\right) \in S^{n-1},$$

(the projection of S^{n-1} onto the hyperplane $\mathbb{R}^{n-1}$). This implies by elementary calculus that the measure is given by

$$\mathcal{H}^{n-1}\left(S_\vartheta^{n-1} \setminus S_\theta^{n-1}\right) = \int_D \left\{1 + \sum_{i=1}^{n-1}\left(\frac{\partial f}{\partial x_i}\right)^2\right\}^{\frac{1}{2}} \mathrm{d}\mathcal{L}^{n-1},$$

where $\mathcal{L}^{n-1}$ denotes the $(n-1)$-dimensional Lebesgue measure, and where the domain D is given by

$$D = \left\{x \in \mathbb{R}^{n-1} \;\middle|\; \sin^2\theta \le \sum_{i=1}^{n-1} x_i^2 \le \sin^2\vartheta\right\}.$$

When we change to spherical coordinates in R^{n-1}, we get by using the substitution $\varrho := \sin t$,

$$\mathcal{H}^{n-1}\left(S_\vartheta^{n-1} \setminus S_\theta^{n-1}\right) = \mathcal{H}^{n-2}\left(S^{n-2}\right) \int_{\sin\theta}^{\sin\vartheta} \frac{\varrho^{n-2}}{\sqrt{1-\varrho^2}}\, \mathrm{d}\varrho = \mathcal{H}^{n-2}\left(S^{n-2}\right) \int_\theta^\vartheta \sin^{n-2} t\, \mathrm{d}t.$$

When $n > 2$, the function $\sin^{n-2} t$ is logarithmic concave on the interval $]0, \frac{\pi}{2}[$, so it follows from Lemma 8.2 (1) that

$$\begin{aligned}\mathcal{H}^{n-1}\left(S_\vartheta^{n-1} \setminus S_\theta^{n-1}\right) &\geq \left\{\left(\frac{\sin\vartheta}{\sin\theta}\right)^{n-2} - 1\right\} \mathcal{H}^{n-2}\left(S^{n-2}\right) \int_0^\theta \sin^{n-2} t\, \mathrm{d}t \\ &= \left\{\left(\frac{\sin\vartheta}{\sin\theta}\right)^{n-2} - 1\right\} \mathcal{H}^{n-1}\left(S_\theta^{n-1}\right),\end{aligned}$$

and the lemma is proved. □

Remark. Let $\varepsilon > 0$ be a small positive constant, and let

$$\left\{(x_1, \dots, x_n) \in S^{n-1} \mid |x_n| \leq \varepsilon\right\},$$

be an ε-slab containing the origo in $\mathbb{R}^n$. Then for large n most of the $\mathcal{H}^{n-1}$ measure of the unit sphere S^{n-1} will lie inside this slab. In fact, it follows from Lemma 8.3 that

$$\frac{\mathcal{H}^{n-1}\left(\left\{(x_1, \dots, x_n) \in S^{n-1} \mid |x_n| \leq \varepsilon\right\}\right)}{\mathcal{H}^{n-1}\left(S^{n-1}\right)} \to 1 \qquad \text{for } n \to +\infty. \qquad \Diamond$$

The results above were purely geometrical. We shall in the following extend these to some estimates, which also involve Gaussian measures ν_n on $\mathbb{R}^n$. First we recall that we have in spherical coordinates in $\mathbb{R}^n$

$$\text{(8.5)} \qquad \nu_n\left(K_n[0, R]\right) = c_n \int_0^R r^{n-1} \exp\left(-\frac{1}{2}\, r^2\right) \mathrm{d}r.$$

It is left as an exercise to the reader to prove that

$$c_n = \frac{\mathcal{H}^{n-1}\left(S^{n-1}\right)}{(2\pi)^{n/2}}.$$

We shall not use this exact value in the following.

We shall, however, use the simple fact that the function $f_n(r) := r^{n-1} \exp\left(-\frac{1}{2}\, r^2\right)$, $r > 0$, is logarithmic concave, and that $f_n(r)$ is increasing in $]0, \sqrt{n-1}[$, attaining it global maximum at $r = \sqrt{n-1}$, and decreasing in $]\sqrt{n-1}, +\infty[$.

In the next step we compare the Gaussian measures of two balls of same centre x, where $\|x\|_n \leq \sqrt{bn}$, and of different radii r and R. Cf. the sketch in Figure 8.4.

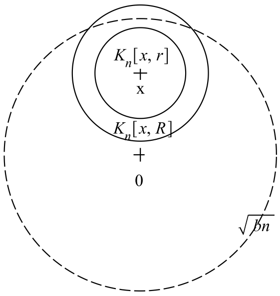

Figure 8.4: *Sketch of the geometry of Lemma 8.4.*

Lemma 8.4 *Let $b, c > 1$ and $\varepsilon \in]0,1[$ be given constants. There exists an $n_0 \in \mathbb{N}$, such that whenever $n \geq n_0$ and $x \in \mathbb{R}^n$ and $r,\ R > 0$ satisfy the conditions*

$$\|x\|_n^2 \leq bn, \quad \text{and} \quad r^2 \leq (1-\varepsilon)n + \|x\|_n^2, \quad \text{and} \quad R^2 \geq r^2 + \varepsilon n,$$

then

$$\nu_n\left(K_n[x,R]\right) \geq c \cdot \nu_n\left(K_n[x,r]\right).$$

The proof of this technical lemma is fairly long and complicated. Notice that we always can choose R as small as possible, i.e. in the proof we may always assume that $R^2 = r^2 + \varepsilon n$.

PROOF. For given b, $c > 1$ and $\varepsilon \in]0,1[$ we define $\varphi_0 \in \left]0, \frac{\pi}{2}\right[$ by the relation

$$\cos\varphi_0 := \frac{\varepsilon}{4(b+\sqrt{b(1-\varepsilon+b)})},$$

where this strange definition will first be justified later in the proof.

Then choose $n_0 \in \mathbb{N}$, such that at least

$$(8.6) \qquad \left(\frac{1}{\sin\varphi_0}\right)^{n-2} \geq 2c+1, \qquad \text{for all } n \geq n_0.$$

Since $(1-t)e^t < 1$ for all $t \in]0,1[$, we can furthermore choose n_0, such that also

$$(8.7) \qquad \sqrt{1-\frac{\varepsilon}{2}} \cdot \left\{\left(1-\frac{\varepsilon}{2}\right)e^{\varepsilon/2}\right\}^{-\frac{n}{2}} \geq 4c+1, \qquad \text{for all } n \geq n_0.$$

Notice that if R as already mentioned above is chosen as small as possible, then

$$R^2 = r^2 + \varepsilon n \leq n + \|x\|_n^2 \qquad (\leq (b+1)n).$$

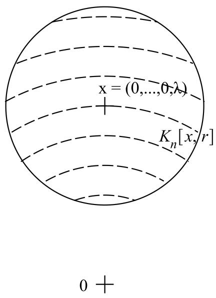

Figure 8.5: *Calculation of the Gaussian measure $\nu_n(K_n[x,r])$ by cutting $K_n[x,r]$ with spheres of centre 0 and radius v, where we have put $x = (0,\dots,0,\lambda)$, $0 \leq \lambda \leq \sqrt{bn}$.*

When we divide $K_n[x,r]$ into shells of the distance $v > 0$ from the origo we compute the Gaussian measure of $K_n[x,r]$ in the following way,

$$\nu_n(K_n[x,r]) = \frac{1}{(2\pi)^{n/2}} \int_0^{+\infty} \exp\left(-\frac{1}{2}v^2\right) \mathcal{H}^{n-1}\left(K_n[x,r], \cap \{y \in \mathbb{R}^n \mid \|y\|_n = v\}\right) \mathrm{d}v.$$

We shall more closely analyze the expression

$$\mathcal{H}^{n-1}\left(K_n[x,r], \cap \{y \in \mathbb{R}^n \mid \|y\|_n = v\}\right).$$

Since the Gaussian measure is invariant with respect to rotations, we may for simplicity assume that $x = (0,\dots,0,\lambda)$, where $0 \leq \lambda \leq \sqrt{bn}$, cf. also Figure 8.5. The set

$$(8.8) \qquad K_n[x,r] \cap \{y \in \mathbb{R}^n \mid \|y\|_n = v\}$$

is nonempty, if and only if

$$(\lambda - r)^+ \le v \le \lambda + r,$$

where for $c \in \mathbb{R}$ a real number, $c^+ := \max\{c, 0\}$, cf. Figure 8.5.

We introduce for brevity an auxiliary parameter $\tau \ge 0$ by

$$\tau^2 := \left(r^2 - \lambda^2 + \frac{\varepsilon n}{2}\right)^+ .$$

Then notice that by assumption,

$$r^2 \le (1-\varepsilon)n + \|x\|_n^2 = (1-\varepsilon)n + \lambda^2,$$

so

$$r^2 - \lambda^2 + \frac{\varepsilon n}{2} \le (1-\varepsilon)n + \frac{\varepsilon n}{n} = \left(1 - \frac{\varepsilon}{2}\right) n, \qquad \text{hence } \tau^2 \le \left(1 - \frac{\varepsilon}{2}\right) n. \tag{8.9}$$

We split the proof into two cases, a) and b).

a) We first assume that

$$\max\left\{\tau, (\lambda - r)^+\right\} \le v \le \lambda + r.$$

It follows from Figure 8.5 that $\lambda > 0$ (this was an assumption) and

$$\{y \in \mathbb{R}^n \mid \|x - y\|_n = r\} \cap \{y \in \mathbb{R}^n \mid \|y\|_n = v\} \ne \emptyset. \tag{8.10}$$

When we write $y = (y_1, y_2) \in \mathbb{R}^{n-1} \times \mathbb{R}$, we see that the set given in (8.10) is determined by the equations

$$\|y_1\|_{n-1}^2 + (y_2 - \lambda)^2 = r^2 \qquad \text{and} \qquad \|y_1\|_{n-1}^2 + y_2^2 = v^2.$$

When we eliminate $\|y_1\|_{n-1}^2$ from these two equations we get

$$-(y_2 - \lambda)^2 + r^2 = -y_2^2 + v^2, \qquad \text{thus} \qquad 2\lambda y_2 - \lambda^2 + r^2 = v^2,$$

from which

$$y_2 = \frac{v^2 - r^2 + \lambda^2}{2\lambda}.$$

We introduce the angle φ between the point $y \in \mathbb{R}^n$ and the positive part of the n-th coordinate axis, i.e.

$$\cos\varphi = \frac{v^2 - r^2 + \lambda^2}{2\lambda v}.$$

Then we get

$$K_n[x, r] \cap \{y \in \mathbb{R}^n \mid \|y\|_n = v\} = \{y \in \mathbb{R}^n \mid \|y\|_n = v,\ y_2 \ge \|y\|_n \cos\varphi\}. \tag{8.11}$$

We shall find the lower bound of $\cos\varphi$. We first see that by the assumption of the lemma,

$$v \le \lambda + r \le \sqrt{n}\left(\sqrt{b} + \sqrt{1 - \varepsilon + b}\right).$$

By the assumption a) above we have $v^2 \geq \tau^2$.

If $\tau = 0$, then it follows from the definition of τ that

$$r^2 + \frac{\varepsilon n}{2} \leq \lambda^2,$$

so

$$v^2 - r^2 + \lambda^2 \geq -r^2 + \lambda^2 \geq \frac{\varepsilon n}{2}.$$

If instead $\tau > 0$, then $v^2 \geq \tau^2 = r^2 - \lambda^2 + \dfrac{\varepsilon n}{2}$, hence again

$$v^2 - r^2 + \lambda^2 \geq \frac{\varepsilon n}{2}.$$

Therefore,

$$\begin{aligned}\cos\varphi &= \frac{v^2 - r^2 + \lambda^2}{2\lambda v} \geq \frac{\varepsilon n}{2} \cdot \frac{1}{\sqrt{b}\,\sqrt{n}} \cdot \frac{1}{2\sqrt{n}\left(\sqrt{b} + \sqrt{1-\varepsilon+b}\right)} \\ &= \frac{\varepsilon}{4\left(b + \sqrt{b(1-\varepsilon+b)}\right)} = \cos\varphi_0,\end{aligned}$$

where $\varphi_0 \in \left]0, \frac{\pi}{2}\right[$ was defined in the beginning of the proof.

It follows in particular that

$$0 \leq \varphi \leq \varphi_0 \leq \frac{\pi}{2}.$$

Still assuming that $\max\left\{\tau, (\lambda - r)^+\right\} \leq v \leq \lambda + r$ we then consider the set

$$K_n[x,R] \cap \left\{y \in \mathbb{R}^n \mid \|y\|_n = v,\, y_2 \geq 0\right\}. \tag{8.12}$$

Since $\varphi < \pi/2$, we get

$$K_n[x,r] \cap \left\{y \mid \|y\|_n = v\right\} \subset K_n[x,R] \cap \left\{y \mid \|y\|_n = v,\, y_2 \geq 0\right\}.$$

Similarly, as in (8.11), we can write

$$K_n[x,R] \cap \left\{y \in \mathbb{R}^n \mid \|y\|_n = v,\, y_2 \geq 0\right\} = \left\{y \in \mathbb{R}^n \mid \|y\|_n = v,\, y_2 \geq \|y\|_n \cos(\varphi + \eta)\right\}, \tag{8.13}$$

where we still have to determine the angle η. Clearly, the formula for $\cos(\varphi+\eta)$ must be analogous to the formula for $\cos\varphi$. There are, however, two modifications. 1) The radius r is replaced by the larger radius R, and 2) the restriction in (8.12) to the upper half-sphere implies that $\cos(\varphi + \eta)$ cannot be negative. Therefore

$$\cos(\varphi+\eta) = \left(\frac{v^2 - R^2 + \lambda^2}{2\lambda v}\right)^+ = \left(\frac{v^2 - r^2 + \lambda^2}{2\lambda v} - \frac{\varepsilon n}{2\lambda v}\right)^+ = \left(\cos\varphi - \frac{\varepsilon n}{2\lambda v}\right)^+. \tag{8.14}$$

If $\cos\varphi \leq \dfrac{\varepsilon n}{2\lambda v}$, then $\eta = \dfrac{\pi}{2} - \varphi$. Otherwise, $\cos\varphi > \dfrac{\varphi n}{2\lambda v}$, and we get the following estimate of η,

$$\begin{aligned}\cos(\varphi+\eta) &= \cos\varphi - \frac{\varepsilon n}{2} \leq \cos\varphi - \frac{\varepsilon n}{4\lambda(\lambda + r)} \\ &\leq \cos\varphi - \frac{\varepsilon}{4\left(b + \sqrt{b(1-\varepsilon+b)}\right)} = \cos\varphi - \cos\varphi_0.\end{aligned}$$

We shall find a lower bound for the function

$$\frac{\sin(\varphi+\eta)}{\sin\varphi}, \qquad \varphi \in \,]0,\varphi_0]\,.$$

If $\eta = \pi/2 - \varphi$, then

$$\frac{\sin(\varphi+\eta)}{\sin\varphi} = \frac{1}{\sin\varphi} \geq \frac{1}{\sin\varphi_0}.$$

Otherwise, we can write

$$\frac{\sin(\varphi+\eta)}{\sin\varphi} \geq \frac{1}{\sin\varphi}\sqrt{1-(\cos\varphi-\cos\varphi_0)^2} = \sqrt{1+\frac{2\cos\varphi}{\sin^2\varphi}\cdot\cos\varphi_0 - \frac{\cos^2\varphi_0}{\sin^2\varphi_0}}.$$

By elementary calculus, the latter expression is decreasing in $\varphi \in \,]0,\varphi_0]$, so its minimum is attained at $\varphi = \varphi_0$, hence,

$$\text{(8.15)} \qquad \frac{\sin(\varphi+\eta)}{\sin\varphi} \geq \frac{1}{\sin\varphi_0} > 1 \qquad \text{for } \varphi \in \,]0,\varphi_0]\,.$$

We shall apply Lemma 8.3 in the form (8.4). Using also (8.15) we get

$$\mathcal{H}^{n-1}\left(\{y \in \mathbb{R}^n \mid \|y\|_n = v,\, y_2 = \|y\|_n \cos(\varphi+\eta)\}\right)$$

$$\geq \left\{\left(\frac{1}{\sin\varphi_0}\right)^{n-2} - 1\right\}\mathcal{H}^{n-1}\left(K_n[x,R] \cap \{y \in \mathbb{R}^n \mid \|y\|_n = v,\, y_2 \geq 0\}\right),$$

whenever a) $\max\left\{\tau, (\lambda-r)^+\right\} \leq v \leq \lambda + r$, where $\tau = \left(r^2 - \lambda^2 + \frac{\varepsilon n}{2}\right)^+$.

Summing up, we get in this case, cf. also (8.6),

$$\begin{aligned}
&\nu_n\left(K_n[x,r]\setminus K_n[0,\tau]\right)\\
&\qquad = \frac{1}{(2\pi)^{n/2}}\int_\tau^{\tau+\lambda}\exp\left(-\frac{1}{2}v^2\right)\mathcal{H}\left(K_n[x,r]\cap\{y\mid \|y\|_n\}\right)\,\mathrm{d}v\\
&\qquad \le \frac{1}{2c}\cdot\frac{1}{(2\pi)^{n/2}}\int_\tau^{R+\lambda}\exp\left(-\frac{1}{2}v^2\right)\mathcal{H}^{n-1}\left(K_n[x,R]\cap\{y\mid \|y\|_n=v\}\right)\,\mathrm{d}v\\
&\qquad = \frac{1}{2c}\cdot\nu_n\left(K_n[x,R]\setminus K_n[0,\tau]\right),
\end{aligned}$$

i.e.

$$\nu_n\left(K_n[x,r]\setminus K_n[0,\tau]\right)\le\frac{1}{2c}\,\nu_n\left(K_n[x,R]\setminus K_n[0,\tau]\right). \tag{8.16}$$

b) We assume that $(\lambda-r)^+\le v\le\min\{\tau,\lambda+r\}$, and notice that $\max\{\tau,(\lambda-r)^+\}\le\min\{\tau,\lambda+r\}$, so joined together a) and b) cover all possibilities.

Clearly, $\tau>0$. If

$$v_0^2=\tau^2+\frac{\varepsilon n}{2}=R^2-\lambda^2,$$

then it follows by a geometrical consideration that

$$\mathcal{H}^{n-1}\left(K_n[x,R]\cap\{y\mid\|y\|_n=v\}\right)\ge\frac{1}{2}\,\mathcal{H}^{n-1}\left(\{y\mid\|y\|_n=v\}\right),\qquad\text{for all } v\in\,]0,v_0]\,. \tag{8.17}$$

In particular, this inequality is true for all v, for which

$$(\lambda-r)^+\le v\le\min\{\tau,\lambda+r\}.$$

Recalling that $f_n(r)=r^{n-1}\exp\left(-\frac{1}{2}r^2\right)$ we get from (8.9) and (8.7) that

$$\begin{aligned}
\frac{f_n\left(\sqrt{R^2-\lambda^2}\right)}{f_n(\tau)}&=\left(\frac{\tau^2+\frac{\varepsilon n}{2}}{\tau^2}\right)^{\frac{n-1}{2}}\exp\left(-\frac{\varepsilon n}{4}\right)=\left(1+\frac{\varepsilon n}{2\tau^2}\right)^{\frac{n-1}{2}}\exp\left(-\frac{\varepsilon n}{4}\right)\\
&\ge\left(1+\frac{\varepsilon}{2-\varepsilon}\right)^{\frac{n-1}{2}}\exp\left(-\frac{\varepsilon n}{4}\right)=\sqrt{1-\frac{\varepsilon}{2}}\cdot\left\{\left(1-\frac{\varepsilon}{2}\right)\exp\left(\frac{\varepsilon}{2}\right)\right\}^{-n/2}\ge 4c+1
\end{aligned}$$

for all $n \geq n_0$. It therefore follows from Lemma 8.2 (1) and (8.17) that

$$
\begin{aligned}
&\nu_n\left(K_n[x,r] \cap K_n[0,\tau]\right) \leq \nu_n\left(K_n[0,\tau]\right) = c_n \int_0^\tau v^{n-1} \exp\left(-\frac{1}{2}v^2\right) \mathrm{d}v \\
&\leq \left(\frac{f_n\left(\sqrt{R^2\lambda^2}\right)}{f_n(\tau)} - 1\right)^{-1} c_n \int_\tau^{\sqrt{R^2-\lambda^2}} v^{n-1} \exp\left(-\frac{1}{2}v^2\right) \mathrm{d}v \\
&= \frac{1}{4} \cdot \frac{1}{(2\pi)^{n/2}} \int_\tau^{\sqrt{R^2-\lambda^2}} \exp\left(-\frac{1}{2}v^2\right) \mathcal{H}^{n-1}\left(\{y \mid \|y\|_n = v\}\right) \mathrm{d}v \\
&\leq \frac{1}{2c} \cdot \frac{1}{(2\pi)^{n/2}} \int_{tau}^{R} \exp\left(-\frac{1}{2}v^2\right) \mathcal{H}^{n-1}\left(K_n[x,R] \cap \{y \mid \|y\|_n = v\}\right) \mathrm{d}v \\
&\leq \frac{1}{2c}\, \nu_n\left(K_n[x,R] \setminus K_n[0,\tau]\right),
\end{aligned}
$$

and we conclude summing up that

$$
\nu_n\left(K_n[x,r] \cap K_n[0,\tau]\right) \leq \frac{1}{2c}\, \nu_n\left(K_n[x,R] \setminus K_n[0,\tau]\right). \tag{8.18}
$$

Finally, using the two estimates (8.16) and (8.18), we get

$$
\begin{aligned}
\nu_n\left(K_n[x,r]\right) &= \nu_n\left(K_n[x,r] \setminus K_n[0,\tau]\right) + \nu_n\left(K_n[x,r] \cap K_n[0,\tau]\right) \\
&\leq \frac{1}{c}\, \nu_n\left(k_n[x,R] \setminus K_n[0,\tau]\right) \leq \frac{1}{c}\, \nu_n\left(K_n[x,R]\right),
\end{aligned}
$$

and the lemma is proved. □

The last lemma of this section shows that most of the ν_n of $\mathbb{R}^n$ actually lies in an "ε-shell" of centre x, provided that n is sufficiently large, and $\|x\|_n$ is not too big. More precisely,

Lemma 8.5 *Given constants ε, $\in]0,1[$ and $b > 1$. There exists an $n_0 \in \mathbb{N}$, such that whenever $n \geq n_0$, and $x \in \mathbb{R}^n$ is bounded by $\|x\|_n^2 \leq bn$, then we have the estimate*

$$
\begin{aligned}
&\nu_n\left(\left\{y \in \mathbb{R}^n \mid (1-\varepsilon)n + \|x\|_n^2 \leq \|y-x\|_n^2 \leq (1+\varepsilon)n + \|x\|_n^2\right\}\right) \\
&= \nu_n\left(K_n\left[x, \sqrt{(1+\varepsilon)n + \|x\|_n^2}\right] \setminus K_n\left[x, \sqrt{(1-\varepsilon)n + \|x\|_n^2}\right]\right) \geq 1-\delta.
\end{aligned}
$$

Proof. Choose $\delta_0 \in]0,1[$, such that

$$
1-\delta = (1-\delta_0)^3, \qquad \text{i.e.} \qquad \delta_0 = 1 - \sqrt[3]{1-\delta}. \tag{8.19}
$$

Choosing

$$
R^2 = (1+\varepsilon)n + \|x\|_n^2 \qquad \text{and} \qquad r^2 = (1-\varepsilon)n + \|x\|_n^2,
$$

we see that the assumptions of Lemma 8.4 are fulfilled, so there exists an $n_1 \in \mathbb{N}$, such that for all $n \geq n_1$,

$$
\nu_n\left(K_n\left[x, \sqrt{(1+\varepsilon)n + \|x\|_n^2}\right]\right) \geq \frac{1}{\delta_0}\, \nu_n\left(K_n\left[x, \sqrt{(1-\varepsilon)n + \|x\|_n^2}\right]\right).
$$

By a rearrangement,

$$\begin{aligned}
&\nu_n\left(K_n\left[x,\sqrt{(1+\varepsilon)n+\|x\|_n^2}\right]\setminus K_n\left(x,\sqrt{(1-\varepsilon)n+\|x\|_n^2}\right)\right) \\
&\qquad \geq (1-\delta_0)\,\nu_n\left(K_n\left[x,\sqrt{(1+\varepsilon)n+\|x\|_n^2}\right]\right).
\end{aligned} \tag{8.20}$$

Due to the rotational invariance we may assume that $x=(0,\dots,0,\lambda$, where $0\leq\lambda\leq\sqrt{bn}$. We then define a set A by

$$A:=\left\{y\in\mathbb{R}^n\;\left|\;\|y\|_n^2\leq\left(1+\frac{\varepsilon}{2}\right)n \text{ and } |y_n|\leq\sqrt{\left(1+\frac{\varepsilon}{2}\right)n}\cdot\cos\varphi\right.\right\},$$

where y_n denotes the n-th coordinate of $y=(y_1,\dots,y_n)$, and where

$$\cos\varphi:=\frac{\sqrt{\lambda^2+\dfrac{\varepsilon n}{2}}-\lambda}{\sqrt{\left(1+\dfrac{\varepsilon n}{2}\right)n}}.$$

If $y\in A$, then

$$\begin{aligned}
\|y-x\|_n^2 &\leq (y_n-\lambda)^2+\left(1+\frac{\varepsilon}{2}\right)n\leq(\lambda+|y_n|)^2+\left(1+\frac{\varepsilon}{2}\right)n \\
&\leq \lambda^2+\frac{\varepsilon n}{2}+\left(1+\frac{\varepsilon}{2}\right)n=(1+\varepsilon)n+\lambda^2,
\end{aligned}$$

which implies that

$$A\subset K_n\left[x,\sqrt{(1+\varepsilon)n+\|x\|_n^2}\right].$$

We shall estimate $\nu_n(A)$. We conclude from Lemma 8.3 that there exists an $n_2 \in \mathbb{N}$, such that whenever $n \geq n_2$, then

$$\mathcal{H}^{n-1}\left(S_{\pi/2}^{n-1} \setminus S_{\varphi}^{n-1}\right) \geq \frac{1}{\delta_0}\,\mathcal{H}^{n-1}\left(S_{\varphi}^{n-1}\right),$$

hence

$$\mathcal{H}^{n-1}\left(S_{\pi/2}^{n-1} \setminus S_{\varphi}^{n-1}\right) \geq (1-\delta_0)\,\mathcal{H}^{n-1}\left(S_{\pi/2}^{n-1}\right).$$

If $0 \leq v \leq \sqrt{\left(1+\frac{\varepsilon}{2}\right)n}$, then

$$\begin{aligned}
&\mathcal{H}^{n-1}\left(A \cap \{y \mid \|y\|_n = v\}\right) \geq 2\mathcal{H}^{n-1}\left(\{y \mid \|y\|_n = v,\, 0 \leq y_n \leq \|y\|_n \cos\varphi\}\right)\\
&\quad = 2v^{n-1}\mathcal{H}^{n-1}\left(S_{\pi/2}^{n-1} \setminus S_{\varphi}^{n-1}\right) \geq 2v^{n-1}(1-\delta_0)\,\mathcal{H}^{n-1}\left(S_{\pi/2}^{n-1}\right)\\
&\quad = (1-\delta_0)\,\mathcal{H}^{n-1}\left(\{y \mid \|y\|_n = v\}\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\nu_n\left(K_n\left[x, \sqrt{(1+\varepsilon)n + \|x\|_n^2}\right]\right) \geq \nu_n(A)\\
(8.21)\qquad &\quad = c_n \int_0^{\sqrt{(1+\varepsilon/2)n}} \exp\left(-\frac{v^2}{2}\right)\mathcal{H}^{n-1}\left(A \cap \{y \mid \|y\|_n = v\}\right) \mathrm{d}v\\
&\quad \geq (1-\delta_0)\,c_n \int_0^{\sqrt{(1+\frac{\varepsilon}{2})n}} \exp\left(-\frac{v^2}{2}\right)\mathcal{H}^{n-1}\left(\{y \mid \|y\|_n = v\}\right) \mathrm{d}v\\
&\qquad\qquad - (1-\delta_0)\,\nu_n\left(K_n\left[0, \sqrt{(1+\varepsilon/2)\,n}\right]\right).
\end{aligned}$$

The function $f_n(r) = r^{n-1}\exp\left(-\frac{1}{2}\,r^2\right)$ is logarithmic concave, so by using Lemma 8.2, (2),

$$\begin{aligned}
&\nu_n\left(K_n\left[0, \sqrt{\left(1+\frac{\varepsilon}{2}\right)n}\right]\right) \geq c_n \int_{\sqrt{n}}^{\sqrt{(1+\varepsilon/2)n}} f_n(r)\,\mathrm{d}r\\
&\quad \geq c_n\left(\frac{f_n\left(\sqrt{n}\right)}{f_n\left(\sqrt{(1+\varepsilon/2)n}\right)} - 1\right)\int_{\sqrt{(1+\varepsilon/2)n}}^{+\infty} f_n(r)\,\mathrm{d}r\\
&\quad = \left(\frac{n}{(1+\varepsilon/2)n}\right)^{\frac{n-1}{2}} \exp\left(\frac{\varepsilon n}{4}\right)\nu_n\left(\mathbb{R}^n \setminus K_n\left[0, \sqrt{(1+\varepsilon)n}\right]\right)\\
&\quad \geq \left(\frac{e^{\varepsilon/2}}{1+\varepsilon/2}\right)^{n/2}\left(1 - \nu_n\left(K_n\left[0, \sqrt{\left(1+\frac{\varepsilon}{2}\right)n}\right]\right)\right).
\end{aligned}$$

Since $e^t/(1+t) > 1$ for all $t \in]0,1[$, there exists an $n_3 \in \mathbb{N}$, such that for all $n > n_3$,

$$\nu_n\left(K_n\left[0, \sqrt{(1+\varepsilon/2)n}\right]\right) \geq \frac{1-\delta_0}{\delta_0}\left(1 - \nu_n\left(K_n\left[\sqrt{(1+\varepsilon/2)n}\right]\right)\right),$$

so by a rearrangement,

$$(8.22)\qquad \nu_n\left(K_n\left[0,\sqrt{(1+\varepsilon/2)n}\right]\right)\geq 1-\delta_0.$$

Combining (8.20), (8.21), (8.22) and (8.19) we get the desired estimate for $n \geq \max\{n_1,n_2,n_3\}$, and the lemma is proved. □

Corollary 8.2 *Given constants $\delta>0$ and $c>1$ there exists an $n_0\in\mathbb{N}$, such that for all $n\geq n_0$,*

$$\nu_n\left(K_n\left[0,c\sqrt{n}\right]\right)\geq 1-\delta.$$

Proof. Choose $\varepsilon\in]0,1[$, such that $c\geq\sqrt{1+\varepsilon}$. Then by Lemma 8.5 there exists an n_0, such that for all $n\geq n_0$,

$$\nu_n\left(K_n\left[0,c\sqrt{n}\right]\right)\geq\nu_n\left(K_n\left[0,\sqrt{(1+\varepsilon)n}\right]\setminus K_n\left[0,\sqrt{(1-\varepsilon)n}\right]\right)\geq 1-\delta,$$

and the corollary is proved. □

8.3.3 The construction of the Gaussian measure on the Hilbert space

So far we have proved that for given constants ε, $\delta\in]0,1[$ and b, $c>0$ there exists an $n_0\in\mathbb{N}$, such that for all $n\geq n_0$ and every $x\in\mathbb{R}^n$, where $\|x\|_n^2\leqq bn$,

$$\nu_n\left(\left\{y\in\mathbb{R}^n \mid (1-\varepsilon)n+\|x\|_n^2\leq\|y-x\|_n^2\leq(1+\varepsilon)n+\|x\|_n^2\right\}\right)\geq 1-\delta.$$

Furthermore, if $r^2\leq(1-\varepsilon)n+\|x\|_n^2$ and $R^2\geq r^2+\varepsilon n$, then

$$\nu_n\left(K_n[x,R]\right)\geq c\,\nu_n\left(K_n[x,r]\right),$$

where everything takes place in the Euclidean space $\mathbb{R}^n$.

The idea is to concatenate infinitely many spaces of this type of increasing dimension n. We shall let the sequence (n) govern the corresponding constants b_n, c_n and δ_n.

For technical reasons we shall also need a sequence (A_j) of compact subsets of the Euclidean space H_j $\left(=\prod_{k=1}^{j}\mathbb{R}^{n_j}\right)$ introduced in Section 8.3.1, such that

$$\mu_j\left(H_j\setminus A_j\right)<2^{-j},\qquad\qquad \text{where } \mu_j:=\prod_{k=1}^{j}\nu_{n_k},$$

and a sequence (d_j) of positive real numbers, such that

$$\frac{1}{d_j}\,r^{p_j}\leq\mu_j\left(B_j[x,r]\right)\leq d_j\,r^{n_j}.$$

Assume for the time being that the sequences (n_j) and (d_j) are known for at least $j=1,\dots,j_0$. then we define

$$(8.23)\qquad \left\{\begin{array}{l} p_j:=\sum_{i=1}^{j}n_i,\qquad b_j:=6\cdot 2^j,\qquad c_j:=d_{j-1}^2\left(\dfrac{2^j}{\varepsilon}\right)^{p_{j-1}} \\[4mm] \delta_j:=\dfrac{1}{2}\cdot\dfrac{\left(2^{-j}\varepsilon\right)^{\frac{1}{2}\,p_{j-1}}}{d_{j-1}^2+\left(2^{-j}\varepsilon\right)^{\frac{1}{2}\,p_{j-1}}}\qquad \text{for } j\geq 1. \end{array}\right.$$

It follows immediately that $b_j = 6 \cdot 2^j > 1$ is defined for all $j \in \mathbb{N}_0$. Furthermore, $c_j > 1$ and $\delta_j \in \left]0, \frac{1}{2}\right[$ only require knowledge of p_{j-1}. So, it suffices in the proof of the lemma below "only" to construct (n_j), (A_j) and (d_j) by induction. The other constants follow from the definitions in (8.23).

Lemma 8.6 *Given $\varepsilon \in]0,1[$. There exist a sequence (n_j) of integers, a sequence (A_j) of compact subsets of H_j, and a nondecreasing sequence (d_j), $d_j > 1$, defining the sequences (p_j), (b_j), (c_j) and (δ_j) of (8.23), such that the following assertions hold.*

1) If $x \in \mathbb{R}^{n_j}$ satisfies $\|x\|_{n_j}^2 \le b_j n_j$, tehn

$$\nu_{n_j}\left(\left\{y \in \mathbb{R}^{n_j} \mid (1-\varepsilon)n_j + \|x\|_{n_j}^2 \le \|y-x\|_{n_j}^2 \le (1+\varepsilon)n_j + \|x\|_{n_j}^2\right\}\right) \ge 1 - \delta_j.$$

2) If $\|x\|_{n_j}^2 \le b_j n_j$, and $r^2 \le (1-\varepsilon)n_j + \|x\|_{n_j}^2$, and $R^2 \ge r^2 + \varepsilon n_j$, then

$$\nu_{n_j}\left(K_{n_j}[x,R]\right) \ge c_j\, \nu_{n_j}\left(K_{n_j}[x,r]\right).$$

3) $\mu_j\left(H_j \setminus A_j\right) < 2^{-j}$.

4) If $B_j[x,1] \cap A_j \ne \emptyset$ and $r \in]0,1]$, then

$$\frac{1}{d_j}\, r^{p_j} \le \mu_j\left(B_j[x,r]\right) \le d_j\, r^{p_j}.$$

PROOF. As mentioned previously it suffices only to construct (n_j), (A_j) and (d_j) by induction. Choose $\varepsilon \in]0,1[$, and put formally $d_0 = 1$ and $p_0 = 0$. Then $c_1 = 1$ and $\delta_1 = \frac{1}{4}$.

Apply Lemma 8.4 and Lemma 8.5 to find $n_1 \ge 3$, such that if r, $R > 0$ and $x \in \mathbb{R}^{n_1}$ satisfy

$$\|x\|_{n_1}^2 \le b_1 n_1, \qquad r^2 \le (1-\varepsilon)n_1 + \|x\|_{n_1}^2, \qquad R^2 \ge r^2 + \varepsilon n_1,$$

then

$$\nu_{n_1}\left(\left\{y \in \mathbb{R}^{n_1} \mid (1-\varepsilon)n_1 + \|x\|_{n_1}^2 \le \|y-x\|_{n_1}^2 \le (1+\varepsilon)n_1 + \|x\|_{n_1}^2\right\}\right) \ge 1 - \delta_1,$$

(if x is bounded in some sense, then some relatively thin shell of centre x contains most of the ν_{n_1} mass of $\mathbb{R}^{n_1}$), and

$$\nu_{n_1}\left(K_{n_1}[x,R]\right) \ge c_1 \cdot \nu_{n_1}\left(K_{n_1}[x,r]\right).$$

This shows that (1) and (2) hold for $j = 1$.

The space H_1 is the Euclidean space $\mathbb{R}^{n_1}$, provided with the norm defined by

$$\|x\|_{\sigma,1} := \frac{1}{\sqrt{2n_1}}\, \|x\|_n.$$

Choose any compact set $A_1 \subset H_1$, such that

$$\mu_1\left(H_1 \setminus A_1\right) < 2^{-1}.$$

The set $\{x \in H_1 \mid B_1[x,1] \cap A_1 \neq \emptyset\}$ is bounded, so there exists a constant $d_1 > 1$, such that

$$\frac{1}{d_1}\, r^{p_1} \leq \mu_1\left(B_1[x,r]\right) \leq d_1\, r^{p_1}$$

for every $r \in [0,1]$ and every $x \in H_1$, for which $B_1[x,1] \cap A_1 \neq \emptyset$.

This proves the first step of the induction.

In general, we assume that n_j, A_j and d_j, $1 \leq j \leq k-1$, have been determined, such that (1)–(4) hold. We shall prove that we can find n_k, A_k and d_k, such that (1)–(4) also hold at step number k.

It was noticed above that we can right away find c_k and δ_k, because they only rely on p_{k-1}, already found at level $k-1$. Applying Lemma 8.4 and Lemma 8.5, we find as previously $n_k > n_{k-1}$, such that if r, $R > 0$ and $x \in \mathbb{R}^{n_k}$ satisfy the conditions

$$\|x\|_{n_k}^2 \leq b_k\, n_k, \qquad r^2 \leq (1-\varepsilon)n_k + \|x\|_{n_k}^2, \qquad R^2 \geq r^2 + \varepsilon n_k,$$

where $b_k := 6 \cdot 2^k$, then

$$\nu_{n_k}\left(\left\{y \in \mathbb{R}^{n_k} \mid (1-\varepsilon)n_k + \|x\|_{n_k}^2 \leq \|y-x\|_{n_k}^2 \leq (1+\varepsilon)n_k + \|x\|_{n_k}^2\right\}\right) \geq 1 - \delta_k,$$

and

$$\nu_{n_k}\left(K_{n_k}[x,R]\right) \geq c_k \cdot \left(K_{n_k}[x,r]\right),$$

so (1) and (2) also hold at level k.

Recall that the Hilbert space (actually an Euclidean space) H_k is inductively given by

$$H_k = H_{k-1} \times \mathbb{R}^{n_k} \simeq \mathbb{R}^{p_k},$$

provided with the norm $\| \cdot \|_{\sigma,k}$, defined by

$$\|x\|_{\sigma,k}^2 ;= \|x_1\|_{\sigma,k-1}^2 + \frac{1}{2^k \, n_k} \, \|x_2\|_{n_k}^2 , \qquad \text{where } x = (x_1, x_2) \in H_{k-1} \times \mathbb{R}^{n_k}.$$

We attach inductively the measure $\mu_k := \mu_{k-1} \times \nu_{n_k}$ to $H_k = H_{k-1} \times \mathbb{R}^{n_k}$. Then there exists a compact set $A_k \subset H_k$, such that

$$\mu_k \left(H_k \setminus A_k \right) \leq 2^{-k}.$$

In particular, A_k is bounded, so there exists a constant $d_k > d_{k-1}$, such that

$$\frac{1}{d_k} \, r^{p_k} \leq \mu_k \left(B_k[x, r] \right) \leq d_k \, r^{p_k} , \qquad \text{for } r \in]0, 1],$$

whenever $x \in H_k$ and $B_k[x, 1] \cap A_k \neq \emptyset$. We have proved that (3) and (4) also hold at level k, and the lemma then follows by induction. $\square$

Let the constants be as in Lemma 8.6. Notice in particular that

$$c_k := d_{j-1}^2 \left(\frac{2^j}{\varepsilon} \right)_j^p .$$

We shall prove the following lemma concerning H_k.

Lemma 8.7 *Given $\varepsilon \in]0, 1[$. Choose any $x \in H_{k-1} \times \mathbb{R}^{n_k} = H_k$, and any $r \in]0, 1]$, for which*

$$\|x\|_{n_k}^2 \leq b_k n_k \qquad \text{and} \qquad r^2 \leq 2^{-k} (1 - \varepsilon) + \frac{1}{2^k n_k} \, \|x\|_{n_k}^2 ,$$

and such that $B_{k-1}[x, 1] \cap A_{k-1} \neq \emptyset$. Then

$$\mu_k \left(B_k \left[x, \sqrt{r^2 + 2^{-k} \cdot 2\varepsilon} \right] \right) \geq \sqrt{c_k} \cdot \nu_{n_k} \left(K_{n_k} \left[x, \sqrt{2^k n_k r^2} \right] \right) .$$

In particular,

$$\mu_k \left(B_k \left[x, \sqrt{r^2 + 2^{-(k-1)} \varepsilon} \right] \right) \geq \sqrt{c_k} \cdot \mu_k \left(B_k[x, r] \right) . \tag{8.24}$$

Proof. It is not hard to check that

$$B_{k-1} \left[x, \sqrt{2^{-k} \varepsilon} \right] \times K_{n_k} \left[x, \sqrt{2^k n_k r^2 + \varepsilon n_k} \right] \subset B_k \left[x, r^2 + 2^{-(k-1) \varepsilon} \right] .$$

In fact, if $(y_1, y_2) \in B_{k-1} \left[x, \sqrt{2^{-k} \varepsilon} \right] \times K_{n_k} \left[x, \sqrt{2^k n_k r^2 + \varepsilon n_k} \right]$, then for $x = (x_1, x_2)$ and $y = (y_1, y_2) \in H_{k-1} \times H^{n_k}$,

$$\|y_1 - x_1|_{\sigma,k-1}^2 = \sum_{j=1}^{k-1} \frac{1}{2^j n_j} \, \|y_1 - x_1\|^2 \leq 2^{-k} \varepsilon,$$

and

$$\|y_2 - x_2\|_{n_k}^2 \le 2^k n_k r^2 + \varepsilon n_k,$$

so

$$\|y-x\|_{\sigma,k}^2 \le 2^{-k}\varepsilon + \frac{1}{2^k n_k}\left(2^k n_k r^2 + \varepsilon n_k\right) = r^2 + 2^{-k}\varepsilon + 2^{-k}\varepsilon = r^2 + 2^{-(k-1)}\varepsilon,$$

proving that $y \in B_k\left[x, \sqrt{r^2 + 2^{-(k-1)}\varepsilon}\right]$. From this inclusion we get

$$\mu_k\left(B_k\left[x,\sqrt{r^2+2^{-(k-1)}\varepsilon}\right]\right) \ge \mu_{k-1}\left(B_{k-1}\left[x,\sqrt{2^{-k}\varepsilon}\right]\right)\cdot \nu_{n_k}\left(K_{n_k}\left[x,\sqrt{2^k n_k r^2 + \varepsilon n_k}\right]\right).$$

It follows by choosing $R^2 = 2^k n_k r^2 + \varepsilon n_k$ and $r_1^2 = 2^k n_k r^2$ that

$$r_1^2 \le (1-\varepsilon) n_k + \|x\|_{n_k}^2 \qquad \text{and} \qquad R^2 = r_1^2 + \varepsilon n_k.$$

We therefore conclude from Lemma 8.6 (2) and (4) and the definition of c_k that

$$\begin{aligned}
\mu_k\left(B_k\left[x,\sqrt{r^2+2^{-(k-1)}\varepsilon}\right]\right) &\ge \frac{1}{d_{k-1}}\left(2^{-k}\varepsilon\right)^{p_k/2}\, c_k\cdot \nu_{n_k}\left(K_{n_k}\left[x,\sqrt{2^k n_k r^2}\right]\right)\\
&= \frac{1}{d_{k-1}}\left(\frac{\varepsilon}{2^k}\right)^{p_k/2} d_{k-1}^2\left(\frac{2^k}{\varepsilon}\right)^{p_k-1}\nu_{n_k}\left(K\left[x,\sqrt{2^k n_k r^2}\right]\right)\\
&= \sqrt{c_k}\cdot \nu_{n_k}\left(K_{n_k}\left[x,\sqrt{2^k n_k r^2}\right]\right) \ge \sqrt{c_k}\cdot \mu_k\left(B_k[x,r]\right),
\end{aligned}$$

and the lemma is proved. □

Let $x \in H$ and $k \in \mathbb{N}$, and define for fixed $\varepsilon \in]0,1[$,

$$\begin{aligned}
H^k(x) &:= \bigcap_{j=k}^{+\infty}\left\{y \in H^k \mid (1-\varepsilon)n_j + \|x\|_{n_j}^2 \le \|y-x\|_{n_j}^2 \le (1+\varepsilon)n_j + \|x\|_{n_j}^2\right\}\\
&= \prod_{j=k}^{+\infty}\left(K_{n_j}\left[x,\sqrt{(1+\varepsilon)n_j + \|x\|_{n_j}^2}\right] \setminus K_{n_j}\left(x,\sqrt{(1-\varepsilon)n_j + \|x\|_{n_j}^2}\right)\right).
\end{aligned}$$

We also define

$$M^k(x) := H_{k-1} \times H^k(x) \subset H.$$

If $x \in H$ satsifies $\|x\|_{n_j}^2 \le b_j n_j$ for all $j \ge k$, then it follows from Lemma 8.6 (1) that

$$\mu\left(M^k(x)\right) = \mu^k\left(H^k(x)\right) \ge \prod_{j=k}^{+\infty}(1-\delta_j) \ge 1 - \sum_{j=k}^{+\infty}\delta_k = 1 - \Delta_k,$$

where we have put

$$\Delta_k := \sum_{j=k}^{+\infty}\delta_j.$$

Notice that

$$\begin{aligned}\frac{\delta_{j+1}}{\delta_j} &= \frac{\left(2^{-j-1}\varepsilon\right)^{p_j/2}}{d_{j+1}^2+(2^{-j-1}\varepsilon)^{p_j/2}}\cdot\frac{d_j^2+\left(2^{-j}\varepsilon\right)^{p_{j-1}/2}}{(2^{-j}\varepsilon)^{p_{j-1}/2}}\\ &= \frac{d_j^2+\left(2^{-j}\varepsilon\right)^{p_{j-1}/2}}{d_{j+1}^2+(2^{-j-1}\varepsilon)^{p_j/2}}\cdot\frac{\left(2^{-j-1}\varepsilon\right)^{p_{j-1}/2}\cdot\left(2^{-j-1}\right)^{n_j/2}}{(2^{-j}\varepsilon)^{p_{j-1}/2}}\le\frac{1}{2},\end{aligned}$$

so $2\delta_{j+1}\le\delta_j$, hence also $\delta_j\le 2\delta_j-2\delta_{j+1}$, and we get by summing a telescoping series that

$$0<\Delta_k=\sum_{j=k}^{+\infty}\delta_k\le\sum_{j=k}^{+\infty}\{2\delta_j-2\delta_{j+1}\}=2\delta_k,$$

because $\delta_k\searrow 0$. We therefore conclude that

$$\frac{\Delta_k}{1-\Delta_k}\le\frac{2\delta_k}{1-2\delta_k}=\frac{1}{d_{k-1}^2}\cdot\left(2^{-k}\varepsilon\right)^{\frac{1}{2}p_{k-1}}. \tag{8.25}$$

We shall use this estimate in the proof of the following lemma.

Lemma 8.8 *Given $\varepsilon\in]0,1[$ and $x\in H$, and assume that*

$$B_k[x,1]\cap A_k\neq\emptyset\qquad\textit{and}\qquad 2^{-(k-1)}(1+2\varepsilon)+|x|_{\sigma,k}^2\le r^2\le 1.$$

Then

$$\mu(B[x,r])\le 2\mu\left([x,r]\cap M^k(x)\right).$$

PROOF. Assume that $y\in B_{k-1}\left[x,\sqrt{2^{-k}\varepsilon}\right]\times H^k(x)$, i.e. $\|x-y\|_{\sigma,k+1}^2\le 2^{-k}\varepsilon$ and

$$\|x-y\|_{n_j}^2\le(1-\varepsilon)n_j+\|x\|_{n_j}^2\qquad\text{for } j\ge k.$$

When we sum all these inequalities for all $j\ge k$, we get

$$|x-y|_{\sigma,k}^2\le 2^{-(k-1)}(1+\varepsilon)+|x|_{\sigma,k}^2,$$

so

$$\|x-y\|_\sigma^2=\|x-y\|_{\sigma,k-1}^2+|x-y|_{\sigma,k}^2\le 2^{-k}\varepsilon+2^{-(k-1)}(1+\varepsilon)+|x|_{\sigma,r}^2\le r^2,$$

and it follows that

$$B_{k-1}\left[x,\sqrt{2^{-k}\varepsilon}\right]\times H^k(x)\subset B[x,r]\cap M^k(x).$$

Then an application of Lemma 8.6 (4) gives

$$\begin{aligned}\mu\left(B[x,r]\cap M^k(x)\right) &\ge \mu_{k-1}\left(B_{k-1}\left[x,\sqrt{2^{-k}\varepsilon}\right]\right)\cdot\mu^k\left(H^k(x)\right)\\ &\ge \frac{1}{d_{k-1}}\left(2^{-k}\varepsilon\right)^{\frac{1}{2}p_{p-1}}\cdot\mu\left(M^k(x)\right).\end{aligned} \tag{8.26}$$

We shall also derive an esimate of $\mu\left(B[x,r]\setminus M^k(x)\right)$. We get

$$\begin{aligned} B[x,r]\setminus M^k(x) &\subset \left(B_{k-1}[x,r]\times H^k\right)\setminus M^k(x) \\ &= \left(B_{k-1}[x,r]\times H^k(x)\right)\setminus\left(H_{k-1}\times H^k(x)\right) = B_{k-1}[x,r]\times\left(H^k\setminus H^k(x)\right), \end{aligned}$$

from which

$$\mu\left(B[x,r]\setminus M^k(x)\right)\le\mu_{k-1}\left(B_{k-1}[x,r]\right)\cdot\mu^k\left(H^k\setminus H^k(x)\right).$$

By assumption, $B_{k-1}[x,1]\cap A_{k-1}\neq\emptyset$, so it follows from (8.25) that

$$\begin{aligned} \mu\left(B[x,r]\setminus M^k(x)\right) &\le d_{k-1}r^{p_{k-1}}\mu^k\left(H^k\setminus H^k(x)\right) \\ &\le d_{k-1}\mu^k\left(H^k\setminus H^k(x)\right) = d_{k-1}\cdot\frac{\mu^k\left(H^k\setminus H^k(x)\right)}{\mu^k\left(H^k(x)\right)}\cdot\mu^k(x) \\ &\le d_{k-1}\cdot\frac{\Delta_k}{1-\Delta_k}\cdot\mu\left(H^k(x)\right)\le\frac{1}{d_{k-1}}\left(2^{-k}\varepsilon\right)^{\frac{1}{2}p_{k-1}}\cdot\mu\left(M^k(x)\right), \end{aligned} \tag{8.27}$$

Finally, combining (8.26) and (8.27),

$$\mu(B[x,r])=\mu\left(B[x,r]\cap M^k(x)\right)+\mu\left(B[x,r]\setminus M^k(x)\right)\le 2\mu\left(B[x,r]\cap M^k(x)\right),$$

and the lemma is proved. □

In the remaining two lemmas we finally obtain estimates of $\mu(B[x,r])$ for different values of r.

Lemma 8.9 *Given $\varepsilon \in]0,1[$. Choose $x \in H$, such that $B_{k-1}[x,1] \cap A_{k-1} \neq \emptyset$, at let $r \in]0,1[$ for $k \geq 2$ fulfil the inequalities*

$$2^{-(k-1)}(1+2\varepsilon) + |x|^2_{\sigma,k} \leq r^2 \leq 2^{-(k+1)} + 2^{-(k-1)}(1+2\varepsilon) + |x|^2_{\sigma,k}.$$

Then

$$\mu(B[x,r]) \leq 2\mu\left(M^k(x)\right) \cdot \mu_{k-1}\left(B_{k-1}\left[x, \sqrt{2^{-(k+1)}(1+12\varepsilon)}\right]\right).$$

Proof. An application of Lemma 8.8 gives

$$\mu(B[x,r]) \leq 2\mu\left(B[x,r] \cap M^k(x)\right). \tag{8.28}$$

If $y \in B[x,r] \cap M^k(x)$, then

$$\begin{aligned}\|y-x\|^2_{\sigma,k-1} &= \|y-x\|^2 - |y-x|^2_{\sigma,k} \leq r^2 - \sum_{j=k}^{+\infty} \frac{1}{2^j n_j}\left((1-\varepsilon)n_k + \|x\|^2_{n_j}\right) \\ &\leq \left\{2^{-(k+1)} + 2^{-(k-1)}(1+2\varepsilon) + |x|^2_{\sigma,k}\right\} - \left\{2^{-(k-1)}(1-\varepsilon)n_j + \|x\|^2_{n_j}\right\} \\ &= 2^{-(k+1)} + 2^{-(k-1)} \cdot 3\varepsilon = 2^{-(k+1)}(1+12\varepsilon).\end{aligned}$$

We therefore conclude that

$$B[x,r] \cap M^k(x) \subset B_{k-1}\left[x, \sqrt{2^{-(k+1)}(1+12\varepsilon)}\right] \times H^k(x).$$

It follows from (8.28) that

$$\mu(B[x,r]) \leq 2\mu\left(M^k(x)\right) \cdot \mu_{k-1}\left(B_{k-1}\left[x, \sqrt{2^{-(k+1)}(1+12\varepsilon)}\right]\right),$$

proving the lemma. □

Lemma 8.10 *Given $\varepsilon \ ]0, \frac{1}{20}[$. Assume for $k \geq 2$ that a point $x \in H$ satisfies*

$$B_j[x,1] \cap A_j \neq \emptyset \quad \text{and} \quad \|x\|^2_{n_j} \leq b_j n_j \qquad \text{for } k-1 \leq j \leq k.$$

Then it follows for all $r \in]0,1]$, for which

$$2^{-(k+2)} + 2^{-k}(1+2\varepsilon) + |x|^2_{\sigma,k+1} \leq r^2 \leq 2^{-(k-1)}(1+2\varepsilon) + |x|^2_{\sigma,k},$$

that

$$\mu(B[x,r]) \leq 4\mu\left(M^{k+1}(x)\right) \cdot \mu_{k-1}\left(B_{k-1}\left[x, \sqrt{6 \cdot 2^{-k}\varepsilon}\right]\right).$$

PROOF. We shall apply Lemma 8.8 with k replaced by $k+1$, i.e.

$$\mu(B[x,r]) \le 2\mu\left(B[x,r] \cap M^{k+1}(x)\right) \le 2\int_{M^{k+1}(x)} \mu_k\left(B_k[x,r(y)]\right) \mathrm{d}\mu(y), \tag{8.29}$$

where $r^2(y) := r^2 - |x-y|^2_{\sigma,k+1}$.

Let $z \in B_k[x,r(y)]$, i.e.

$$\|z-x\|^2_{\sigma,k-1} + \frac{1}{2^k n_k}\|z-x\|^2_{n_k} \le r^2(y).$$

We shall consider two possibilities. Either

$$\frac{1}{2^k n_k}\|z-x\|^2_{n_k} < r^2(y) - 6\cdot 2^{-k}\varepsilon,$$

or

$$\frac{1}{2^k n_k}\|z-x\|^2_{n_k} \ge r^2(y) - 6\cdot 2^{-k}\varepsilon \qquad \text{and} \qquad \|z-x\|^2_{\sigma,k-1} \le 6\cdot 2^{-k}\varepsilon.$$

Notice that the choice of ε guarantees that the expression $r^2(y) - 6\cdot 2^{-k}\varepsilon > 0$ for every $y \in M^{k+1}(x)$.

When we split according to these two alternatives, we get

$$B_k[x,r(y)] \subset \left(H^{k-1} \times K_{n_k}\left[x, \sqrt{2^k n_k\left(r^2(y) - 6\cdot 2^{-k}\varepsilon\right)}\right]\right) \cup, \left(B_{k-1}\left[x, 6\cdot 2^{-k}\varepsilon\right] \times \mathbb{R}^{n_k}\right),$$

so

$$\mu_k\left(B_k[x,r(y)]\right) \le \nu_{n_k}\left(K_{n_k}\left[x, \sqrt{2^k n_k\left(r(y) - 6\cdot 2^{-k}\varepsilon\right)}\right]\right) + \mu_{k+1}\left(B_{k-1}\left[x, \sqrt{6\cdot 2^{-k}\varepsilon}\right]\right). \tag{8.30}$$

If $y \in M^{k+1}(x)$, then

$$\begin{aligned}
r^2(y) - 6\cdot 2^{-k}\varepsilon &= r^2 - |x-y|^2_{\sigma,k+1} \le r^2 \sum_{j=k+1}^{+\infty} \frac{1}{2^j n_j}\left\{(1-\varepsilon)n_j + \|x\|^2_{n_j}\right\} - 6\cdot 2^{-k}\varepsilon \\
&= r^2 - 2^{-k}(1-\varepsilon) - |x|^2_{\sigma,k+1} - 6\cdot 2^{-k}\varepsilon \\
&\le 2^{-(k-1)}(1+2\varepsilon) + |x|^2_{\sigma,k} - 2^{-k}(1-\varepsilon) - |x|^2_{\sigma,k+1} - 6\cdot 2^{-k}\varepsilon \\
&= 2^{-k}(1-\varepsilon) + \frac{1}{2^k n_k}\|x\|^2_{n_k}.
\end{aligned}$$

We have checked the assumptions of Lemma 8.7, so an application of this lemma gives

$$\nu_{n_k}\left(K_{n_k}\left[x, \sqrt{2^k n_k\left(r^2(y) - 6\cdot 2^{-k}\varepsilon\right)}\right]\right) \le \frac{1}{\sqrt{c_k}}\,\mu_k\left(B_k[x,r(y)]\right).$$

When we combine this estimate with (8.30), then we get

$$\mu_k\left(B_k[x,r(y)]\right) \le \frac{1}{\sqrt{c_k}}\,\mu_k\left(B_k[x,r(y)]\right) + \mu_{k-1}\left(B_{k-1}\left[x, 6\cdot 2^{-k}\varepsilon\right]\right),$$

hence, by a rearrangement,

$$\mu_k\left(B_k[x,r(y)]\right) \leq \frac{\sqrt{c_k}}{\sqrt{c_k}-1}\cdot\mu_{k-1}\left(B_{k-1}\left[x,6\cdot 2^{-k}\varepsilon\right]\right) \leq 2\mu_{k-1}\left(B_{k-1}\left[x,6\cdot 2^{-k}\varepsilon\right]\right),$$

where we have used that $c_k \geq 2$ for $k \geq 2$.

This estimate above holds for every $y \in M^{k+1}(x)$, so it follows from (8.29) that

$$\mu(B[x,r]) \leq 4\mu\left(M^{k+1}(x)\right)\cdot\mu_{k-1}\left(B_{k-1}\left[x,\sqrt{6\cdot 2^{-k}\varepsilon}\right]\right),$$

and the lemma is proved. □

Finally, we prove the following proposition, in which the first three statements only repeat some conclusions, which we have obtained previously.

Proposition 8.1 *Given* $\varepsilon \in \left]0,\frac{1}{36}\right[$ *and* $b_j := 6\cdot 2^j$ *for* $j \in \mathbb{N}$*. Then there exist*

1) an increasing sequence (n_j) *of integers,*

2) a sequence (A_j) *of compact subsets of* H_j*, and*

3) a nondecreasing sequence (d_j)*, where* $d_j > 1$ *for all* $j \in \mathbb{N}$*,*

such that

(i) $\nu_{n_j}\left(\mathbb{R}^{n_j}\setminus K_{n_j}\left[0,2\sqrt{n_j}\right]\right) \leq 2^{-j}$.

(ii) $\mu_j\left(H_j\setminus A_j\right) < 2^{-j}$.

(iii) *If* $B_j[x,1]\cap A_j \neq \emptyset$ *and* $r \in [0,1]$*, then*

$$\frac{1}{d_j\,r^{p_j}} \leq \mu_j\left(B_j[x,r]\right) \leq d_j\,r^{p_j}, \qquad \text{where } p_j := \sum_{i=1}^{j} n_i.$$

Define

$$c_j := d_{j-1}^2\left(\frac{2^j}{\varepsilon}\right)^{p_{j-1}} \qquad \text{and} \qquad d_j := \frac{1}{2}\cdot\frac{\left(2^{-j}\varepsilon\right)^{\frac{1}{2}p_{j-1}}}{d_{j-1}^2+\left(2^{-j}\varepsilon\right)^{\frac{1}{2}\,p_{j-1}}}.$$

Whenever $k \geq 3$*, there exists a positive number* $\varrho_k > 0$*, satisfying the following condition: If* $x \in H$ *and* $r \in]0,1]$ *satisfy*

$$B_j[x,1]\cap A_j \neq \emptyset, \qquad \|x\|_{n_j}^2 \leq b_j n_j \qquad \text{for } k-2 \leq j \leq k,$$

and

$$2^{-(k+2)}+2^{-k}(1+2\varepsilon)+|x|_{\sigma,k+1}^2 \leq r^2 \leq 2^{-(k+1)}+2^{-(k-1)}(1+2\varepsilon)+|x|_{\sigma,k}^2.$$

Then

(iv) $K_{n_k}\left[0,\varrho_k\right]\times H^{k+1}(x) \subset B^k[x,r]$.

(v) $\mu\left(\left\{y \in B[x,r] \mid \|y\|_{n_k} \le \varrho_k\right\}\right) \ge \dfrac{1}{4}\sqrt{c_{k-1}} \cdot \nu_{n_k}\left(K_{n_k}\left[0,\varrho_k\right]\right) \cdot \mu(B[x,r])$.

PROOF. Statement (i) follows immediately from Corollary 8.2. The next two statements,, (ii) and (iii), follows from Lemma 8.6 (3) and (4). Thus, we shall only prove (iv) and (v).

Choose $\varrho_k > 0$, such that for all $k \ge 3$,

$$\left(2\sqrt{b_k n_k} + \varrho_k\right)\varrho_k \le \varepsilon n_k \qquad \text{and} \qquad \varrho_k^2 \le 2^{-(k+2)} - 9 \cdot 2^{-k}\varepsilon.$$

Notice that since $0 < \varepsilon < \frac{1}{36}$, we have $2^{-(k+2)} - 9 \cdot 2^{-k}\varepsilon > 0$ for sure.

Assume that $\|x\|_{n_k}^2 \le b_k n_k$ and $\|y\|_{n_k} \le \varrho_k$. Then

$$\begin{aligned}\frac{1}{2^k n_k}\|y-x\|_{n_k}^2 &\le \frac{1}{2^k n_k}\left(\|x\|_{n_k} + \|y\|_{n_k}\right)^2 \\ &\le \frac{1}{2^k n_k}\|x\|_{n_k}^2 + \frac{1}{2^k n_k}\left(2\sqrt{b_k n_k} + \varrho_k\right)\varrho_k \le \frac{1}{2^k n_k}\|x\|_{n_k}^2 + 2^{-k}\varepsilon.\end{aligned}$$

If $y \in K_{n_k}\left[0,\varrho_k\right] \times H^{k+1}(x)$, then

$$\begin{aligned}|y-x|_{\sigma,k}^2 &= \frac{1}{2^k n_k}\|y-x\|_{n_k}^2 + \sum_{j=k+1}^{+\infty}\frac{1}{2^j n_j}\|y-x\|_{n_j}^2 \\ &\le \frac{1}{2^k n_k}\|x\|_{n_k}^2 + 2^{-k}\varepsilon + \sum_{j=k+1}^{+\infty}\frac{1}{2^j n_j}\left\{(1+\varepsilon)n_j + \|x\|_{n_j}^2\right\} \\ &= 2^{-k}(1+2\varepsilon) + |x|_{\sigma,k}^2 < r^2,\end{aligned} \tag{8.31}$$

so $y \in B^k[x,r]$, and (iv) holds.

The proof of (v) is split into two cases. If r fulfils the assumptions of Proposition 8.1, then either
(a) $2^{-(k-1)}(1+2\varepsilon)+|x|_{\sigma,k}^2 \le r^2 \le 2^{-(k+1)}+2^{-(k-1)}(1+2\varepsilon)+|x|_{\sigma,k-1}^2$,
or
(b) $2^{-(k+2)}+2^{-k}(1+2\varepsilon)+|x|_{\sigma,k}^2 \le r^2 \le 2^{-(k-1)}(1+2\varepsilon)+\|x\|_{\sigma,k}^2$.

Case (a). We conclude from (8.31) that if $y \in K_{n_k}[0,\varrho_k] \times H^{k+1}(x)$, then

$$r^2 - |y-x|_{\sigma,k}^2 \ge 2^{-k}(1+2\varepsilon) > 2^{-k},$$

hence

$$\{y \in B[x,r] \mid \|y\|_{n_k} \le \varrho_k\} \supset B_{k-1}\left[x,\sqrt{2^{-k}}\right] \times K_{n_k}[0,\varrho_k] \times H^{k+1}(x),$$

and consequently

$$\mu\left(\{y \in B[x,r] \mid \|y\|_{n_k} \le \varrho\}\right) \ge \mu_{k-1}\left(B_{k-1}\left[x,\sqrt{2^{-k}}\right]\right) \cdot \nu_{n_k}\left(K_{n_k}[0,\varrho_k]\right) \cdot \mu\left(M^{k+1}(x)\right). \tag{8.32}$$

Define $r_1^2 := (1+12\varepsilon)\cdot 2^{-(k+1)}$. Then

$$r_1^2 \le 2^{-(k-1)}(1-\varepsilon) \le 2^{-(k-1)}(1-\varepsilon) + \frac{1}{2^{k-1}n_{k-1}}\|x|_{n_k}^2,$$

where we again have used that $0 < \varepsilon < \frac{1}{36}$. We use the inequality (8.24) to obtain

$$\mu_{k-1}\left(B_{k-1}\left[x,\sqrt{r:12^2+2^{-(k-2)}\varepsilon}\right]\right) \ge \sqrt{c_{k-1}}\cdot \mu_{k-1}\left(B_{k-1}[x,r_1]\right). \tag{8.33}$$

Furthermore, $r_1^2 + 2^{-(k-2)}\varepsilon \le 2^{-k}$, so we get, using (8.32), (8.33) and Lemma 8.9,

$$\begin{aligned}
&\mu\left(\{y \in B[x,r] \mid \|y\|_{n_k} \le \varrho_k\}\right)\\
&\qquad \ge \mu_{k-1}\left(B_{k-1}\left[x,\sqrt{r_1^2+2^{-(k-1)}\varepsilon}\right]\right)\cdot \nu_{n_k}\left(K_{n_k}[0,\varrho_k]\right)\cdot \mu\left(M^{k+1}(x)\right)\\
&\qquad \ge \sqrt{c_{k-1}}\cdot \mu_{k-1}\left(B_{k-1}[x,r_1]\right)\cdot \nu_{n_k}\left(K_{n_k}[0,\varrho_k]\right)\cdot \mu\left(M^{k+1}(x)\right)\\
&\qquad \ge \frac{1}{2}\sqrt{c_{k-1}}\cdot \nu_{n_k}\left(K_{n_k}[0,\varrho_k]\right)\cdot \mu(B[x,r]),
\end{aligned}$$

and the lemma is proved in case (a).

Case (b). It follows from $0 < \varepsilon < \frac{1}{36}$ and (8.31) for $y \in K_{n_k}[0,\varrho_k] \times H^{k+1}(x)$ that

$$r^2 - |y-x|_{\sigma,k+1}^2 \ge r^2 - 2^{-k}(1+2\varepsilon) - |x|_{\sigma,k+1}^2 \ge 2^{-(k+2)} + 2^{-k}\varepsilon \ge 10\cdot 2^{-k}\varepsilon.$$

Recalling the definition of ϱ_k we get

$$\{y \in B[x,r] \mid \|y\|_{n_k} \le \varrho_k\} \subset B_{k-1}\left[x,\sqrt{10\cdot 1^{-k}\varepsilon}\right] \times K_{n_k}[0,\varrho_k] \times H^{k+1}(x),$$

so

$$\mu\left(\{y \in B[x,r] \mid \|y\|_{n_k} \le \varrho_k\}\right) \ge \mu_{k-1}\left(B_{k-1}\left[x,\sqrt{10\cdot 2^{-k}\varepsilon}\right]\right)\cdot \nu_{n_k}\left(K_{n_k}[0,\varrho_k]\right)\cdot \mu\left(M^{k+1}(x)\right). \tag{8.34}$$

When we replace k by $k-1$ in (8.24), we get

$$\mu_{k-1}\left(B_{k-1}\left(x,\sqrt{10\cdot 2^{-k}\varepsilon}\right]\right)\geq\sqrt{c_{k-1}}\cdot\mu_{k-1}\left(B_{k-1}\left[x,\sqrt{6\cdot 2^{-k}\varepsilon}\right]\right).$$

This inequality holds, whenever the assumptions of Lemma 8.7 are fulfilled. In the present case a routine check shows that this is true, so we get, using (8.34),

$$\mu\left(\{y\in B[x,r]\mid \|y\|_{n_k}\leq\varrho_k\}\right)\geq\sqrt{c_{k-1}}\cdot\mu_{k-1}\left(B_{k-1}\left[x,\sqrt{6\cdot 2^{-k}\varepsilon}\right]\right)\cdot\nu_k\left(K_{n_k}\left[0,\varrho_k\right]\right)\cdot\mu\left(M^{k+1}(x)\right).$$

Thus, in case (b) we can apply Lemma 8.10, and we finally obtain the estimate

$$\mu\left(\{y\in B[x,r]\mid \|y\|_{n_k}\leq\varrho_k\}\right)\geq\frac{1}{4}\sqrt{c_{k-1}}\cdot\nu_{n_k}\left(K_{n_k}\left[0,\varrho_k\right]\right)\cdot\mu(B([x,r])),$$

and the proposition is proved. □

8.3.4 Proof of Theorem 8.2.

Once we have proved Proposition 8.1 we are able to prove the main theorem, Theorem 8.2.

PROOF OF THEOREM 8.2. Let $\left(\ell^2(\sigma),\mu\right)$ be a representation of the Hilbert space H with the Gaussian measure γ, and choose any $\varepsilon\in\left]0,\frac{1}{36}\right[$. Using the notation from Proposition 8.1 we define a function $f\geq 0$ by

$$f(x):=\sum_{k=1}^{+\infty}\frac{1}{k^2}\,\frac{1}{\nu_{n_k}\left(K_{n_k}\left[0,\varrho_k\right]\right)}\cdot\chi_{K_{n_k}[0,\varrho_k]}(x)+\sum_{k=1}^{+\infty}k\left\{\chi_{H^k\setminus A_k}(x)+\chi_{\mathbb{R}^{n_k}\setminus K_{n_k}\left[0,2\sqrt{n_k}\right]}(x)\right\}.$$

It follows from (i) and (ii) of Proposition 8.1 that

$$\begin{aligned}\int_H f\,\mathrm{d}\mu &= \sum_{k=1}^{+\infty}\frac{1}{k^2}+\sum_{k=1}^{+\infty}k\cdot\mu_k\left(H_k\setminus A_k\right)+\sum_{k=1}^{+\infty}k\cdot\nu_{n_k}\left(\mathbb{R}^{n_k}\setminus K_{n_k}\left[0,2\sqrt{n_k}\right]\right)\\ &\leq \frac{\pi^2}{6}+\sum_{k=1}^{+\infty}k\cdot 2^{-k}+\sum_{k=1}^{+\infty}k\cdot 2^{-k}<+\infty,\end{aligned}$$

and it follows that $f\in L^1(\gamma)$.

Assume that $x\in\mathbb{R}^{n_k}$ satisfies the estimate $\|x\|_{n_k}^2>b_kn_k$, and that $y\in K_{n_k}\left[x,\sqrt{2^kn_k}\right]$. Then for every $k\geq 1$,

$$\|y\|_{n_k}\geq\|x\|_{n_k}-\sqrt{2^kn_k}>\left(\sqrt{6}-\frac{1}{\sqrt{8}}\right)\sqrt{2^kn_k}>2\sqrt{n_k}.$$

Thus, the estimate $\|x\|_{n_k}^2>b_kn_k$ implies that

$$K_{n_k}\left[x,\sqrt{2^kn_k}\right]\cap K_{n_k}\left[0,2\sqrt{n_k}\right]=\emptyset. \tag{8.35}$$

Assume that $x\in H$ and $r^2\in\left]0,2^{-(m+2)}\right]$, where $m\geq 3$. Then we have the following three possibilities.

(i) If $B_k[x,r] \cap A_k = \emptyset$ for some $k \geq m$, then obviously

$$f(y) \geq k \cdot \chi_{H_k \setminus A_k(y)} = k \geq m, \qquad \text{for all } y \in B[x,r],$$

and it follows that

$$\frac{1}{\mu(B[x,r])} \int_{B[x,r]} f(y)\,\mathrm{d}\mu(y) \geq m.$$

(ii) If $\|x\|_{n_k}^2 > b_k n_k$ for some $k \geq m$, we infer from (8.35) that

$$f(y) \geq k \cdot \chi_{\mathbb{R}^{n_k} \setminus K_{n_k}[0,2\sqrt{n_k}]}(y) = k \geq m, \qquad \text{for all } y \in B[x,r],$$

and it follows as above that

$$\frac{1}{\mu(B[x,r])} \int_{B[x,r]} f(y)\,\mathrm{d}\mu(y) \geq m.$$

(iii) Finally, if neither (i) nor (ii) hold, then

$$B_k[x,r] \cap A_k \neq \emptyset \qquad \text{and} \qquad \|x\|_{n_k}^2 \leq b_k n_k \quad \text{for all } k \geq m.$$

Let k be the smallest integer, for which

$$2^{-(k+2)} + 2^{-k}(1+2\varepsilon) + |x|_{\sigma,k+1}^2 \leq r^2 < 2^{-(k+1)} + 2^{-(k-1)}(1+2\varepsilon) + |x|_{\sigma,k}^2.$$

That such a k exists follows from $|x|_{\sigma,k} \to 0$ for $k \to +\infty$. Furthermore, since $r^2 < 2^{-(m+2)}$, we immediately find that this $k \geq m$.

We apply Proposition 8.1 (v) to get

$$\begin{aligned}
\frac{1}{\mu(B[x,r])} \int_{B[x,r]} f(y)\,\mathrm{d}\mu(y) &\geq \frac{1}{\mu(B[x,r])} \int_{B[x,r]} \frac{1}{k^2}\,\frac{1}{\nu_{n_k}}\left(K_{n_k}\left[0,\varrho_k\right]\right) \cdot \chi_{K_{n_k}[0,\varrho_k]}(y)\,\mathrm{d}\mu(y) \\
&= \frac{1}{\mu(B[x,r])} \cdot \frac{1}{k^2} \cdot \frac{1}{\nu_{n_k}\left(K_{n_k}\left[0,\varrho_k\right]\right)}\,\mu\left(\{y \in B[x,r] \mid \|y\|_{n_k} \leq \varrho_k\}\right) \\
&\geq \frac{1}{4k^2}\sqrt{c_{k-1}} = \frac{1}{4k^2}\,d_{k-2}\left(\frac{2^{k-1}}{\varepsilon}\right)^{\frac{1}{2}p_{k-1}} \geq \frac{1}{4k^2} \cdot \frac{2^{k-1}}{\varepsilon} \geq k \geq m,
\end{aligned}$$

where we have used that $k \geq 3$ and $\varepsilon \in \left]0, \frac{1}{36}\right[$.

Summing up, we have in all three cases proved that

$$\frac{1}{\mu(B[x,r])} \int_{B[x,r]} f(y)\,\mathrm{d}\mu(y) \geq m,$$

for all $x \in H$ and $r^2 \in \left]0, 2^{-(m+2)}\right]$. Hence, for every $x \in H$,

$$\liminf_{r \to 0+} \left\{ \frac{1}{\mu(B[x,s])} \int_{B[x,s]} f(y)\,\mathrm{d}\mu(y) \;\middle|\; x \in H,\, 0 < s \leq r \right\} = +\infty,$$

and the theorem is proved. □

Once Theorem 8.2 has been proved, it is not hard to obtain other strange examples. We here include a result which shows that the *density theorem* also fails, cf. Corollary 8.3 in the following, because its proof is so simple. Another and more difficult consequence is postponed to the next section.

Corollary 8.3 *Let H be a separable Hilbert space. There exist a finite measure τ on H and a Borel set $A \subset H$, such that $\tau(A) > 0$, and*

$$\limsup_{r \to 0+} \left\{ \frac{\tau(B[x,s] \cap A)}{\mu(B[x,s])} \;\middle|\; x \in H,\, 0 < s \leq r \right\} = 0.$$

Proof. Let γ be the Gaussian measure and f the function constructed in the proof of Theorem 8.2. We define another measure τ by

$$\frac{\mathrm{d}\tau}{\mathrm{d}\gamma} := 1 + f.$$

We may choose $c > 0$ and a corresponding set $A := \{x \in H \mid f(x) \leq c\}$, such that $\tau(A) > 0$. Then for every given $x \in H$,

$$\tau(B[x,r] \cap A) = \int_{B[x,r] \cap A} (1+f)\,\mathrm{d}\gamma \leq (1+c)\gamma(B[x,r]),$$

and

$$\tau(B[x,r]) = \int_{B[x,r]} (1+f)\,\mathrm{d}\gamma \geq \int_{B[x,r]} f\,\mathrm{d}\gamma \qquad (>0).$$

Dividing the latter estimate into the former one we get

$$\frac{\tau(B[x,r] \cap A)}{\tau(B[x,r])} \leq (1+c) \cdot \left\{\frac{1}{\gamma(B[x,r])} \int_{B[x,r]} f\,\mathrm{d}\gamma\right\}^{-1},$$

and the corollary follows immediately from Theorem 8.2. □

8.4 Positivity principles in Hilbert spaces

We shall in this section give some unexpected consequences of Theorem 8.2. The material is based on *Mejlbro, Preiss & Tišer* [39].

We first recall the definition of the positivity principle.

THE POSITIVITY PRINCIPLE. *Let $(X, \mathfrak{B})$ be a measurable space. We say that a subfamily (also called a* paving*) $\mathfrak{P} \subset \mathfrak{B}$ satisfies the* positivity principle, *if we can conclude for every signed measure μ on $(X, \mathfrak{B})$, for which $\mu(C) \geq 0$ for every $C \in \mathfrak{P}$, that also $\mu \geq 0$.*

It is classically known that the *Positivity Principle* is equivalent to the following computation formula,

$$\mu(A) = \inf\left\{\sum_i \mu(C_i) \;\middle|\; C_i \in \mathfrak{P} \text{ and } \mu\left(A \setminus \bigcup_i C_i\right) = 0\right\}. \tag{8.36}$$

This explains, why we are interested in the Positivity Principle.

We shall in most of this section consider separable Hilbert spaces, and only at the very end of this section also take a look at more general Banach spaces, Furthermore we shall be concerned with only two types of pavings, either a paving consisting of *"small balls"*, i.e. all balls in the Hilbert space of diameter ≤ 1, say, or the paving of all *"large balls"*,, i.e. all balls of diameter ≥ 1, say. The bound 1 can obviously be replaced by any other positive constant $c > 0$.

We proved in Theorem 8.2 that it was possible to find a Gaussian measure γ on H, and a nonnegative function $f \in L^2(\gamma)$, such that

$$\lim_{r \to 0+} \frac{1}{\gamma(B[x,r])} \int_{B[x,r]} f(y)\,\mathrm{d}\gamma(y) = +\infty,$$

uniformly for $x \in H$. Hence, there exists for every $s > 0$ a constant $\varepsilon(s) > 0$, such that for every ball B of diameter $< \varepsilon(s)$,

$$\frac{1}{\gamma(B)} \int_B f(y)\,\mathrm{d}\gamma(y) \geq s. \tag{8.37}$$

Theorem 8.3 *Given a constant $c > 0$, and let H be an infinite dimensional separable Hilbert space. Then we can find two probability measures γ and ν on H, such that*

$$\nu(B) \geq c \cdot \gamma(B)$$

for every "small ball" B in H, i.e. for every ball B in H of diameter ≤ 1, say.

PROOF. Let $\tilde{H}$ be the separable Hilbert space of Theorem 8.2, and let γ and $f \in L^1(\gamma)$, where $f \geq 0$, be chosen, such that (8.37) holds. Then γ is in particular a probability measure. In particular, we can find a positive constant $\alpha > 0$, such that

$$\text{(8.38)} \qquad \int_{\tilde{H}} \alpha\, f(y)\, \mathrm{d}\gamma(y) = 1.$$

Define ν as the measure, which har αf as its *Radon-Nikodym derivative* with respect to γ. This means that

$$\nu(A) = \int_A \alpha f(y)\, \mathrm{d}\gamma(y).$$

It follows from (8.38) that ν is a probability measure on $\tilde{H}$. Furthermore, (8.37) implies that

$$\frac{\nu(B)}{\gamma(B)} \geq c$$

for every ball of diameter $\leq \varepsilon! \left(\frac{c}{\alpha}\right)$. Hence,

$$\text{(8.39)} \qquad \nu(B) \geq c \cdot \gamma(B).$$

Finally, let H be the Hilbert space consisting of the same points as $\tilde{H}$, and with the rescaled inner product,

$$\langle x, y\rangle_H := \frac{4}{\left(\varepsilon\left(\frac{c}{\alpha}\right)\right)^2} \cdot \langle x, y\rangle_{\tilde{H}} \qquad \text{for } x,\, y \in H.$$

Then (8.39) holds for every ball in H of diameter ≤ 3, and Theorem 8.3 is proved. □

REMARK. It is well-known that all infinite dimensional and separable Hilbert spaces are equivalent. Therefore, Theorem 8.3 can be restated in the following way: Whenever H is an infinite dimensional, separable Hilbert space, there exist probability measures γ and ν on H, such that (8.39) holds. ◊

The classical result is

Theorem 8.4 *The positivity principle holds for the paving of "all small balls" in a separable Hilbert space H, if and only if H is finite dimensional.*

Before we prove Theorem 8.4 we give the following

REMARK. Theorem 8.4 may be reformulated in the following way. Assume that the signed measure μ on a separable Hilbert space is nonnegative on the paving of all "small balls". Then μ itself is nonnegative, $\mu \geq 0$, if and only if the Hilbert space H is finite dimensional, i.e. H is an Euclidean space. When this is the case, we can calculate the μ measure $\mu(A)$ of any measurable set A in H by formula (8.36). We notice, that this is not possible, if H is infinite dimensional, and μ is only known on all "small balls". $\Diamond$

PROOF OF THEOREM 8.4. Assume that H is an Euclidean space, i.e. of finite dimension. Then the statement follows from e.g. the *Besicovitch Covering Theorem.*

Concerning the "only if" part, we assume that the dimension is infinite. According to the remark following Theorem 8.3 there exist two probability measures ν and γ, such that $\nu(B) \geq 2\gamma(B)$, whenever B is a ball of diameter ≤ 1, say, i.e. when B is a "small ball". Define $\mu = \nu - 2\gamma$. Then μ is a measure, which is nonnegative on all "small balls", while clearly $\mu(H) = -1$, so μ is a truly signed measure. $\square$

A surprising result is the following.

Theorem 8.5 *The positivity principle holds for a paving of "large balls" in a separable Hilbert space H, if and only if H is infinite dimensional.*

Again we first give a remark, before the theorem is proved.

REMARK. We notice that one of the consequences of Theorem 8.5 is, that if H is infinite dimensional, and the signed measure μ is positive on all balls of diameter ≥ 1, then the measure $\mu(A)$ of a measurable set A is reconstructed by means of the formula (8.36), i.e.

$$\mu(A) = \inf \left\{ \sum_i \mu\left(B_i\right) \;\middle|\; B_i \text{ ball, diam } B_i \geq 1 \text{ and } \mu\left(A \setminus \bigcup_i B_i\right) = 0 \right\}. \qquad \Diamond$$

PROOF. "$\Rightarrow$". Let H be the Euclidean n-space. Let $\mathcal{L}_n$ denote the Lebesgue measure in H. Let B_1 be any fixed unit ball, and let B_ε denote the ball of same centre as B_1 and of radius $\varepsilon > 0$. For every $\varepsilon \in]0,1[$ we define

$$\mu := \mathcal{L}_{n|B_1} - 2\mathcal{L}_{n|B_\varepsilon}.$$

Then μ is a truly signed measure.

Choosing $\varepsilon > 0$ sufficiently small (depending on the dimension) we clearly obtain a truly signed measure, which is nonnegative on every "large ball" (of radius ≥ 1, say).

"$\Leftarrow$. When we prove the inverse implication, we shall check that every signed measure, which satisfies the assumption on an infinite dimensional Hilbert space, is in fact nonnegative.

It suffices to prove that every orthogonal projection Π_μ onto any finite dimensional subspace V is nonnegative. Fix any such finite dimensional subspace V.

Choose a sequence of n-dimensional subspaces $U_n \subset V^\perp$, such that

$$U_1 \subset U_2 \subset \cdots \subset U_n \subset \cdots,$$

and such that the union $\bigcup_{n=1}^{+\infty} U_n$ is dense in $V^\perp$. Define

$$V_n := U_n^\perp \cap V^\perp, \qquad \text{from which } \bigcap_{n=1}^{+\infty} = \{0\}.$$

Let $B[x,r]$ be any ball in H of centre x and radius $r \geq \frac{1}{2}$. Let $v := \mathrm{pr}_V\, x$ and $v_n := \mathrm{pr}_{V_n}\, x$ be the orthogonal projections of x onto V and V_n, respectively. For every $n \in \mathbb{N}$ define the function $g_n : H \to \mathbb{R}$ by

$$g_n(y) := \int_{U_n} \chi_{B[v+u+v_n,r]}(y)\, \mathrm{d}\mathcal{L}_n(u), \qquad y \in H.$$

Then by a geometrical consideration and a small calculation,

$$g_n(y) = \begin{cases} 0, & \text{if } \operatorname{dist}(y, x+U_n) \geq r, \\ \alpha(n)\left\{r^2 - \|v-w\|^2 - |v_n - w_n|^2\right\}^{n/2}, & \text{otherwise,} \end{cases}$$

where we have written $w := \mathrm{pr}_V\, y$ and $w_n := \mathrm{pr}_{V_n}, y$, and where $\alpha(n)$ denotes the volume of the unit ball in the n-dimensional Euclidean space.

Since $r \geq \frac{1}{2}$, every ball $B\left[v+u+v_n, r\right]$ is a "large ball". We assumed that μ is nonnegative on every "large ball". We therefore conclude that every function g_n has a nonnegative integral with respect to μ, i.e.

$$\int_H g_n(y)\, \mathrm{d}\mu(y) = \int_H \int_{U_n} \chi_{B[v+u+v_n,r]}(y)\, \mathrm{d}\mathcal{L}_n(u)\, \mathrm{d}\mu(y) = \int_{U_n} \mu\left(B\left[v+u+v_n, r\right]\right) \mathrm{d}\mathcal{L}_n(u) \geq 0.$$

Define

$$H_{n.r} := \{y \in H \mid \text{dist}\,(y, x + U_n) \le r\}\,.$$

Then it follows from the expression of g_n above that

$$\text{(8.40)} \qquad \int_{H_{n,r}} \left\{1 - \frac{\|v-w\|^2 + \|v_n - w_n\|^2}{r^2}\right\}^{n/2} \mathrm{d}\mu(y) \ge 0 \qquad \text{for all } n \in \mathbb{N}.$$

Fix a positive constant $\lambda > 0$, and put $r^2 := n/(2\lambda)$ for $n \ge \lambda$. We put $H_n := H_{n,r}$ for $r^2 = n(2\lambda)$. Then we rewrite (8.40) in the following way

$$\int_{H_n} \left\{1 - \lambda \cdot \frac{\|v-w\|^2 + \|v_n - w_n\|^2}{n/\lambda}\right\}^{n/2} \mathrm{d}\mu(y) \ge 0 \qquad \text{for } n \ge \lambda.$$

It follows from $\bigcap_n V_n = \{0\}$ that $\|v_n - w_n\| \to 0$ for $n \to +\infty$. By the same limit, $n \to +\infty$, we also get $H_n \uparrow H$, so we get by taking this limit that

$$0 \le \int_H \exp\left(-\lambda\|v-w\|^2\right) \mathrm{d}\mu(y) = \int_V \exp\left(-\lambda\|v-w\|^2\right) \mathrm{d}\Pi_\mu(w).$$

We have proved that the measure $\nu = \Pi$ on V satisfies

$$\text{(8.41)} \qquad \int_V \exp\left(-\lambda\|v-w\|^2\right) \mathrm{d}\nu(w) \ge 0 \qquad \text{for } \lambda > 0 \text{ and } v \in V.$$

Let γ_λ denote the Gaussian measure on V of density function

$$c_n \cdot \exp\left(-\lambda\|w\|^2\right), \qquad \text{where } c_n > 0 \text{ is the normalisation constant.}$$

Then the left hand side of (8.41 is the convolution of γ_λ and ν, so we infer that $\gamma_\lambda \star \nu \ge 0$. The sequence (γ_λ) of Gaussian measures converges weakly towards the Dirac measure δ_0 at the origin for $\lambda \to +\infty$. We therefore conclude by taking this limit that

$$0 \le \delta_0 \star \nu = \nu,$$

and we have proved that $\nu = \Pi$ is nonnegative. Since this is true for every finite dimensional subspace V, we conclude that μ must also be nonnegative, and the theorem is proved. □

We finally mention that the situation is totally different in general Banach spaces. We shall demonstrate this by the following

Theorem 8.6 *There exist a separable Banach space X and a truly signed measure μ on X, such that μ is nonnegative on* all *balls.*

REMARK. This theorem means that the Positivity Principle does not hold in general for the family of *all balls* in a Banach space. In particular, the μ measure of a measurable set A cannot always be reconstructed in a Banach spade by means of only the formula (8.36). ◊

PROOF. Let H be an infinite dimensional, separable Hilbert space, and ν and γ two probability measures as described in Theorem 8.3, i.e. such that for every ball B of diameter ≤ 1, say,

$$\nu(B) \ge 2\gamma(B).$$

We define the Banach space X as the ℓ^∞-sum of the real line $\mathbb{R}$ and the Hilbert space H, described above, i.e.

$$(8.42) \qquad \|(t,x)\|_X := \max\{|t|, \|x\|_H\}, \qquad \text{for } (t,x) \in \mathbb{R} \times H = X.$$

Furthermore, define on the Banach space X the measure

$$\mu := \delta_0 \otimes \left(\frac{1}{2}\nu - \gamma\right) + \left(\delta_{\frac{1}{2}} + \delta_{-\frac{1}{2}}\right) \otimes \gamma$$

First notice that $\mu(X) = \dfrac{3}{2}$ and $\mu(\{0\} \times H) = -\dfrac{1}{2}$, so the measure μ is truly a signed measure.

If B is a "small ball", i.e. $\operatorname{diam}(B) \leq 1$, or if B is a ball satisfying $B \cap (\{0\} \times H) = \emptyset$, then $\mu(B) \geq 0$, because all the remaining terms are nonnegative.

If instead B is a "large ball", $\operatorname{diam}(B) \geq 1$, for which $B \cap, (\{0\} \times H) \neq \emptyset$, then a small geometrical consideration shows that B must intersect either $\left\{-\dfrac{1}{2}\right\} \times H$, or $\left\{\dfrac{1}{2}\right\} \times H$. Without loss of generality we may assume that $B \cap \left(\left\{\dfrac{1}{2}\right\} \times H\right) \neq \emptyset$. Since X is the ℓ^∞-sum of $\mathbb{R}$ and H, cf. (8.42) above, it follows that

$$\{x \in H \mid (0,x) \in B\} = \left\{x \in H \;\middle|\; \left(\frac{1}{2}, x\right) \in B\right\}.$$

Finally, by using the definition of μ,

$$\begin{aligned}
\mu(B) &\geq \mu(B \cap (\{0\} \times H)) + \mu\left(B \cap \left(\left\{\frac{1}{2}\right\} \times H\right)\right) \\
&= \frac{1}{2}\nu(\{x \in H \mid (0,x) \in B\} - \gamma(\{x \in H \mid (0,x) \in B\} + \gamma\left(\left\{x \in H \;\middle|\; \left(\frac{1}{2}, x\right) \in B\right\}\right) \\
&= \frac{1}{2}\nu(\{x \in H \mid (0,x) \in B\}) \geq 0,
\end{aligned}$$

and the theorem is proved. □

Index

References

[1] **Abel, N.**, (1823) Solution de quelques problèmes a l'aide d'intégrales définies. *Oeuvres 1, Christiania (1881)*, 11–27.

[2] **Abel, N.**, (1823) Résolution d'un problème de mécanique, *Oeuvres 1, Christiania (1881)*, 97–101

[3] **Ampère, A. M.**, (1806) Recherches sur quelques points de la théorie des fonction dérivées .... *École Polytechnique, 6, fasc. 13.*

[4] **Appel, Andrew W.**, *Alan Turing's System of Logic. The Princeton Thesis, Princeton University Press, Princeton and Oxford, 2012, Second printing 2014.*

[5] **Banach, S. et Tarski, A.**, *(1924) Sur la décomposition des ensembles de points en parties respectivement congruentes*, *Fund. Math. 6*, 244–277

[6] **Bang, Th.**, (1946) Om quasi-analytiske Funktioner *(in Danish). Nyt Nordisk Forlag, Arnold Busck, Disputats (Dissertation)*

[7] **Besicovitch, A. S.** (1928) On Kakeya's problem and a similar one. *Math. Zeitschrift 27*, 312–320

[8] **Bohr, H.** Sur les fonctions presque périodiques. *C. R. Hebdom. des Séances de l'Académie des Sciences, Paris 177 (1923)*, 737–739

[9] **Bohr, H.** Sur l'approximation des fonctions presque périodiques par des sommes trigonométriques. *C. R. Hebdom. des Séances de l'Académie des Sciences, Paris 177 (1923)*, 1090–1092

[10] **Bohr, H.** Theorie der fast periodischen Funktionen. I. Eine Verallgemeinerung der Theorie der Fourierreihen. *Acta Math. 45 (1924)*, 29–127.

[11] **Carleson, L.**, (1966) Convergence and growth of partial sums of Fourier series. *Acta Math. 116*, 135–157

[12] **Collepardi, Marcialis & Turriziani**, (1970) The Kinetics of Chloride Ions Penetration in Concrete. (In Italian). *Il Cemento, Vol. 67, Italy.*

[13] **Collepardi, Marcialis & Turriziani**, (1972) Penetraion of Chloride Ions in Cement Pastes and in Concretes. *Journal American Ceramices Soc. Vol. 55, USA 1972.*

[14] **Collepardi**, (1995) Quick method to determine free and bound clorides in concrete. Chloride Penetration into Concrete. *Proceedings of the international RILEM Workshop, St-Rémy-les-Chevreuse, France 1995*

[15] **Cunningham, F.**, (1971) The Kakeya problem for simply connected and for star-shaped sets. *Amer. Math. Monthly 78*, 114–129

[16] **Cunningham, F.**, (1974) Three Kakeya problems, *Amer. Math. Monthly 81*, 582–592

[17] **Davies, R. O.**, (1953) Accessibility of plane sets and differentiation of functions of two variables *(Ph. D. Dissertation, Cambridge University)*

[18] **Davies, R. O.**, (1971) Measures not approximable or not specifiable by means of balls, *Mathematika 18*, 157–160

[19] **Davies, R. O.**, (1982) Some counter-examples in Measure Theory, *Proc. Conf. Topology and Measure III, Greifswald*, 49–55

[20] **du Bois-Reymond, P.** (1875) Versuch einer Classification der willkürlichen Funktionen reeller Argumente, *Journal für Math. 79*, 21–37

[21] **Falconer, R. J.** (1988) The geometry of fractal sets. *Cambridge University Press 85*

[22] **Gelbaum, B. R., and Olmsted, J. M. H.** (1964) Counterexamples in Analysis. *Holden-Day Inc., San Francisco, London, Amsterdam*

[23] **Gibbs, J. W.**, (1898) Fourier's Series, *Nature* **59**:1522, 200

[24] **Gibbs, J. W.**, (1898) Fourier's Series, *Nature* **59**:1539, 606

[25] **Hausdorff, F.**, (1914) Grundzüge der Mengenlehre, *Leipzig*, 469–472

[26] **Hunt, R. A.**, (1967) On the convergence of Fourier series. Orthogonal Expansions and Their Continuous Analogues. *Proc. Conf. Edwardsville, Ill.*, 235–255. *See also Southern Illinois Univ. Press, Carbondale, Ill. (1968)*

[27] **Hörmander, L.**, (1964) Linear partial differential operators. *Springer-Verlag, Berlin–Heidelberg–New York*

[28] **Jørsboe, O. G., and Mejlbro, L.**, (1982) The Carleson–Hunt theorem on Fourier series, *Lecture Notes in Math. 911, Springer–Verlag, Berlin–Heidelberg–New York*

[29] **Kahanpää, L., and Mejlbro, L.** (1984) Some new results on the Muckenhoupt conjecture concerning weighted norm inequalities connecting the Hilbert transform with the maximal function. *Proc. Second Finnish–Polish Summer School in complex analysis at Jyväskylä, Ber. Univ. Jyväskylä Math. Inst. 28*, 53–72

[30] **Kakeya, S.**, (1917) Some problems on maximum and minimum regarding ovals. *Tohoku Science Reports 6 (July 1917)*, 71–88

[31] **Kolmogorov, A. N.** (1923) Une série de Fourier–Lebesgue divergente presque partout. *Fund. Math. 4*, 324–328

[32] **Kolmogorov, A. N.** (1926) Une série de Fourier–Lebesque divergente partout. *Comptes Rendues 183*, 1327–1328

[33] **Lewy, H.** (1957) An example of a smooth linear partial differential equation without solution. *Ann. Math. (2) 66*, 155–158

[34] **Marcinkiewicz, J.** (1935) Sur les nombres dérivés, *Fund. Math. 24*, 305–308

[35] **Mejlbo, L. C.**, (1969) Noget om reelle funktioner og deres ejendommeligheder, *(in Danish) Matematisk Institut. Aarhus Universitet, Elementærafdeling Nr. 12*

[36] **Mejlbro, L.** (1976) Om ligelig kontinuitet i uendelig. *(in Danish) Nordisk Matematisk Tidskrift 24*, 71–74

[37] **Mejlbro, L.** (1996) The Complete Solution of Fick's Second Law of Diffusion with Time-dependent Diffusion Coefficient and Surface Concentration. In "Durability of Concrete in Saline Environment", CEMENTA Special Publication, Danderyd, Sweden 1996.

[38] **Mejlbro, L. & Poulsen, E.** (2006) Diffusion of Chloride in Concrete. Theory and Application. Taylor & Francis 2006.

[39] **Mejlbro, L. & Preiss, D. & Tišer, J.** (1992) Positivity principles in geometrical measure theory. *Measure Theory, Oberwolfach 1990, Proc. Conf. Oberwolfach, March 18-24, 1990.* Supplemento ai Rendiconti del Circolo Matrematico di Palermo, Serie II, numero 28, (1992), 163–167.

[40] **Michelson, A. A.; Stratton, S. W.** (1898) A new harmonic analyser, *Philosophical Magazine 5 (45), 85–91*

[41] **Nikodym, O.** (1927) Sur la mesure des ensembles plans dont tous les points sont rectilinéarement accessible. *Fund. Math. 10*, 116–168

[42] **Osgood, W. F.**, (1903) A Jordan curve of positive area. *Trans. Amer. Math. Soc. 4*, 104–112

[43] **Peano, G.**, (1890) Sur une courbe, qui remplit toute une aire plane. *Math. Ann. 36*, 157–160

[44] **Perron, O.**, (1928) Ueber einen Satz von Besicovitch. *Math. Zeitschrift 28*, 383–386

[45] **Preiss, D.**, (1982) Differentiation of measures in infinite dimensional spaces. *Proc. of the Conf. Top. and Measure III, Greifswald, Part 2*, 201–207

[46] **Rademacher, H.** (1962) A new construction of the Perron tree. *in Studies in Mathematical Analysis, Stanford*

[47] **Riesz, F., et Sz.–Nagy, B.**, (1965) Leçons d'analyse fonctionelle. *Paris Gauthier–Villars, Budapest, Akadémiai Kiado*

[48] **Schoenberg, I. J.**, (1962) On the Besicovitch–Perron solution of the Kakeya problem. *in Studies in Mathematical Analysis, Stanford*

[49] **Schwartz, L.** (1966) Théorie des distributions. *Hermann & Cie, Paris*

[50] **Sierpinski, W.**, (1954) On the congruence of sets and their equivalence by finite decomposition. *Lucknow University, Lucknow*

[51] **Sierpinski, W., Klein, F., Runge, C., Dickson, L. E.**, Congruence of sets andd other monographs. *Chelsea*

[52] **van der Waerden, B. L.**, (1930) Ein einfaches Beispiel einer nichtdifferenzierbaren stetigen Funktion. *Math. Zeitschrift 42*, 474–475

[53] **Wilbraham, H.**, (1848) On a certain periodic function, *The Cambridge and Dublin Math. Journal 3, 198–201*

[54] **Zygmund, A.**, (1959) Trigonometric series, Vol. 1. *Cambridge University Press*