Constructive analysis with witnesses

Helmut Schwichtenberg

Mathematisches Institut der LMU, Wintersemester 2019/20

Contents

Preface

The present text grew out of an attempt to give an exposition of basic parts of classical analysis from a constructive point of view, as pioneered by Brouwer and developed by Bishop (1967); Bishop and Bridges (1985). A special emphasis is on computational aspects of the constructive proofs, so in a sense it is an attempt to unify theoretical and numerical analysis. In the subjects covered and even in the exposition I have closely followed Otto Forster's well-known textbook (2004).

Part of the material in these notes was the subject of seminars at the Mathematics department of LMU, the last one "Constructive analysis" in Wintersemester 2018/19. I would like to thank the participating students for their useful contributions.

München, 29. Oktober 2019
Helmut Schwichtenberg

Introduction

We are interested in *exact real numbers*, as opposed to floating point numbers. The final goal is to develop the basics of real analysis in such a way that from a proof of an existence formula one can extract a program. For instance, from a proof of the intermediate value theorem we want to extract a program that, given an arbitrary error bound $\frac{1}{2^p}$, computes a rational x where the given function is zero up to the error bound.

Why should we be interested in logic in a study of constructive analysis? There are at least two reasons.

(1) Obviously we need to be aware of the difference of the classical and the constructive existential quantifier, and try to prove the stronger statements involving the latter whenever possible. Then one is forced to give "constructive" proofs, whose algorithmic content can be "seen" and then used as a basis to formulate a program for computing the solution. This was the point of view in Bishop's classic textbook Bishop (1967) (and its successor Bishop and Bridges (1985)), and more explicitely carried through in Andersson's Master's thesis Andersson (2001) (based on Palmgren's Palmgren (1996)), with Mathematica as the target programming language.

(2) However, one can go one step further and automatize the step from the (formalized) constructive proof to the corresponding program. This can be done by means of the so-called realizability interpretation, whose existence was clear from the beginnings of constructive logic. The desire to have "mathematics as a numerical language" in this sense was clearly expressed by Bishop in his article Bishop (1970) (with just that title). There are now many implementations of these ideas, for instance Nuprl, Coq Coq Development Team (2009), Agda Agda, Isabelle and Minlog, to mention only a few.

What are the requirements on a constructive logic that should guide us in our design?

- It should be as close as possible to the mathematical arguments we want to use. Variables should carry (functional) types, with free algebras (e.g., natural numbers) as base types. Over these, inductive definitions and the corresponding introduction and elimination axioms should be allowed.

- The constants of the language should denote computable functionals in the Scott-Ershov sense, and hence the higher-order quantifiers should range over their (mathematically correct) domain, the partial continuous functionals.
- The language of the logic should be strong (in the sense of being expressive), but the existence axioms used should be weak.
- Type parameters (ML style) should be allowed, but quantification over types should be disallowed in order to keep the theory predicative. Similarly, predicate variables should be allowed as placeholders for properties, but quantification over them should be disallowed, again to ensure predicativity.

On the technical side, since we need to actually construct formal proofs, we want to have some machine support in building them. In particular, to simplify equational reasoning, the system should identify terms with the same "normal form", and we should be able to add rewrite rules used to generate normal forms. Decidable predicates should be implemented via boolean valued functions, so that the rewrite mechanism applies to them as well.

Compared with the literature, the novel aspect of the present work is the development of elementary constructive analysis in such a way that witnesses have as low a type level as possible. This clearly is important for the complexity of the extracted programs. Here are some examples.

(1) A continuous function on the reals is determined by its values on the rationals, and hence can be represented by a type-one (rather than type-two) object.

(2) In the proof that the range of a continuous function on a compact intervall has a supremum, Brouwer's notion of a *totally bounded* set of reals (which has type-level two) is replaced by the notion of being *order located above* (which has type-level one).

(3) The Cauchy-Euler construction of approximate solutions to ordinary differential equations can be seen as a type-level one process.

CHAPTER 1

Real numbers

1.1. Approximation of square roots

To motivate real numbers, we show that there is a Cauchy sequence of rational numbers that does not converge to a rational number. First we show

LEMMA 1.1.1 (Irrationality of the square root of 2). *There is no rational number b with $b^2 = 2$.*

PROOF. Assume $b = \frac{n}{m} \in \mathbb{Q}$ such that $n^2 = 2m^2$. The number of prime factors 2 in n^2 ist even; however, it is odd in $2m^2$. This contradicts the uniqueness of prime factorization of natural numbers. □

THEOREM 1.1.2 (Approximation of square roots). *Let $a > 0$ and $a_0 > 0$ be given. Define the sequence a_n recursively by*

$$a_{n+1} := \frac{1}{2}\Big(a_n + \frac{a}{a_n}\Big).$$

Then

(a) $(a_n)_{n\in\mathbb{N}}$ *is a Cauchy sequence.*
(b) *If* $\lim_{n\to\infty} a_n = c$, *then* $c^2 = a$.

PROOF. By induction on n one can see easily that $a_n > 0$ for all $n \in \mathbb{N}$. Moreover,

$$a_{n+1}^2 \geq a \quad \text{for all } n; \tag{1}$$

this follows from

$$a_{n+1}^2 - a = \frac{1}{4}\Big(a_n^2 + 2a + \frac{a^2}{a_n^2}\Big) - a = \frac{1}{4}\Big(a_n^2 - 2a + \frac{a^2}{a_n^2}\Big) = \frac{1}{4}\Big(a_n - \frac{a}{a_n}\Big)^2 \geq 0.$$

Next

$$a_{n+2} \leq a_{n+1} \quad \text{for all } n, \tag{2}$$

since

$$a_{n+1} - a_{n+2} = a_{n+1} - \frac{1}{2}\Big(a_{n+1} + \frac{a}{a_{n+1}}\Big) = \frac{1}{2a_{n+1}}\Big(a_{n+1}^2 - a\Big) \geq 0.$$

Let

$$b_n := \frac{a}{a_n}.$$

Then $b_{n+1}^2 \le a$ for all n, since by (1) we have $\frac{1}{a_{n+1}^2} \le \frac{1}{a}$, hence also

$$b_{n+1}^2 = \frac{a^2}{a_{n+1}^2} \le \frac{a^2}{a} = a.$$

From (2) we obtain $b_{n+1} \le b_{n+2}$ for all n. Next we have

$$b_{n+1} \le a_{m+1} \quad \text{for all } n, m \in \mathbb{N}. \tag{3}$$

To see this, observe that – say for $n \ge m$ – we have $b_{n+1} \le a_{n+1}$ (this follows from (1) by multiplying with $1/a_{n+1}$), and $a_{n+1} \le a_{m+1}$ by (2).

We now show

$$a_{n+1} - b_{n+1} \le \frac{1}{2^n}(a_1 - b_1), \tag{4}$$

by induction on n. Basis: for $n = 0$ both sides are equal. Step:

$$\begin{aligned} a_{n+2} - b_{n+2} \le a_{n+2} - b_{n+1} &= \frac{1}{2}(a_{n+1} + b_{n+1}) - b_{n+1} \\ &= \frac{1}{2}(a_{n+1} - b_{n+1}) \le \frac{1}{2^{n+1}}(a_1 - b_1) \quad \text{by IH.} \end{aligned}$$

$(a_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, since for $n \le m$ by (2), (3) and (4)

$$|a_{n+1} - a_{m+1}| = a_{n+1} - a_{m+1} \le a_{n+1} - b_{n+1} \le \frac{1}{2^n}(a_1 - b_1).$$

Now assume $\lim a_n = c$. Then also $\lim b_n = c$, for

$$|c - b_{n+1}| \le |c - a_{n+1}| + |a_{n+1} - b_{n+1}|$$

and both summands can be made arbitrarily small for large n, by (4). Hence

$$c^2 = (\lim b_n)^2 = \lim b_n^2 \le a \le \lim a_n^2 = (\lim a_n)^2 = c^2$$

because of $b_{n+1}^2 \le a \le a_{n+1}^2$, and therefore $c^2 = a$. □

1.2. Cauchy sequences, equality

We shall view a real as a Cauchy sequence of rationals with a separately given modulus.

Definition 1.2.1. A real number x is a pair $((a_n)_{n\in\mathbb{N}}, M)$ with $a_n \in \mathbb{Q}$ and $M\colon \mathbb{Z}^+ \to \mathbb{N}$ such that $(a_n)_n$ is a *Cauchy sequence* with modulus M, that is

$$|a_n - a_m| \le \frac{1}{2^p} \quad \text{for } n, m \ge M(p)$$

and M is weakly increasing (that is $M(p) \le M(q)$ for $p \le q$). M is called *Cauchy modulus* of x.

We shall loosely speak of a real $(a_n)_n$ if the Cauchy modulus M is clear from the context or inessential. Every rational a is tacitly understood as the real represented by the constant sequence $a_n = a$ with the constant modulus $M(p) = 0$.

Definition 1.2.2. Two reals $x := ((a_n)_n, M)$, $y := ((b_n)_n, N)$ are called *equivalent* (or *equal* and written $x = y$, if the context makes clear what is meant), if

$$|a_{M(p+1)} - b_{N(p+1)}| \leq \frac{1}{2^p} \quad \text{for all } p \in \mathbb{Z}^+.$$

We want to show that this is an equivalence relation. Reflexivity and symmetry are clear. For transitivity we use the following lemma:

Lemma 1.2.3 (RealEqChar). *For reals $x := ((a_n)_n, M)$, $y := ((b_n)_n, N)$ the following are equivalent:*
(a) $x = y$;
(b) $\forall_p \exists_{n_0} \forall_{n \geq n_0} (|a_n - b_n| \leq \frac{1}{2^p})$.

Proof. (a) implies (b). For $n \geq M(p+2), N(p+2)$ we have

$$\begin{aligned} |a_n - b_n| &\leq |a_n - a_{M(p+2)}| + |a_{M(p+2)} - b_{N(p+2)}| + |b_{N(p+2)} - b_n| \\ &\leq \frac{1}{2^{p+2}} + \frac{1}{2^{p+1}} + \frac{1}{2^{p+2}}. \end{aligned}$$

(b) implies (a). Let $q \in \mathbb{Z}^+$, and $n \geq n_0, M(p+1), N(p+1)$ with n_0 provided for q by (b). Then

$$\begin{aligned} |a_{M(p+1)} - b_{N(p+1)}| &\leq |a_{M(p+1)} - a_n| + |a_n - b_n| + |b_n - b_{N(p+1)}| \\ &\leq \frac{1}{2^{p+1}} + \frac{1}{2^q} + \frac{1}{2^{p+1}}. \end{aligned}$$

The claim follows, because this holds for every $q \in \mathbb{Z}^+$. □

Remark 1.2.4 (RealSeqEqToEq). An immediate consequence is that any two reals with the same Cauchy sequence (but possibly different moduli) are equal.

Lemma 1.2.5 (RealEqTrans). *Equality between reals is transitive.*

Proof. Let $(a_n)_n$, $(b_n)_n$, $(c_n)_n$ be the Cauchy sequences for x, y, z. Assume $x = y$, $y = z$ and pick n_1, n_2 for $p + 1$ according to the lemma above. Then $|a_n - c_n| \leq |a_n - b_n| + |b_n - c_n| \leq \frac{1}{2^{p+1}} + \frac{1}{2^{p+1}}$ for $n \geq n_1, n_2$. □

1.3. The Archimedian property

For every function on the reals we certainly want compatibility with equality. This however is not always the case; here is an important example.

Lemma 1.3.1 (RealBound). *For every real $x := ((a_n)_n, M)$ we can find p_x such that $|a_n| \leq 2^{p_x}$ for all n.*

Proof. Let $n_0 := M(1)$ and p_x be such that $\max\{\, |a_n| \mid n \leq n_0 \,\} + \frac{1}{2} \leq 2^{p_x}$. Then $|a_n| \leq 2^{p_x}$ for all n. □

Clearly this assignment of p_x to x is not compatible with equality.

1.4. Nonnegative and positive reals

A real $x := ((a_n)_n, M)$ is called *nonnegative* (written $x \in \mathbb{R}^{0+}$) if

$$-\frac{1}{2^p} \leq a_{M(p)} \quad \text{for all } p \in \mathbb{Z}^+.$$

It is *p-positive* (written $x \in_p \mathbb{R}^+$, or $x \in \mathbb{R}^+$ if p is not needed) if

$$\frac{1}{2^p} \leq a_{M(p+1)}.$$

We want to show that both properties are compatible with equality. First we prove a useful characterization of nonnegative reals.

Lemma 1.4.1 (RealNNegChar). *For a real $x := ((a_n)_n, M)$ the following are equivalent:*

(a) $x \in \mathbb{R}^{0+}$;
(b) $\forall_p \exists_{n_0} \forall_{n \geq n_0} (-\frac{1}{2^p} \leq a_n)$.

Proof. (a) implies (b). For $n \geq M(p+1)$ we have

$$\begin{aligned} -\frac{1}{2^p} &\leq -\frac{1}{2^{p+1}} + a_{M(p+1)} \\ &= -\frac{1}{2^{p+1}} + (a_{M(p+1)} - a_n) + a_n \\ &\leq -\frac{1}{2^{p+1}} + \frac{1}{2^{p+1}} + a_n. \end{aligned}$$

(b) implies (a). Let $q \in \mathbb{Z}^+$ and $n \geq n_0, M(p)$ with n_0 provided by (b) (for q). Then

$$\begin{aligned} -\frac{1}{2^p} - \frac{1}{2^q} &\leq -\frac{1}{2^p} + a_n \\ &= -\frac{1}{2^p} + (a_n - a_{M(p)}) + a_{M(p)} \\ &\leq -\frac{1}{2^p} + \frac{1}{2^p} + a_{M(p)}. \end{aligned}$$

The claim follows, because this holds for every q. □

Lemma 1.4.2 (RealNNegCompat). *If $x \in \mathbb{R}^{0+}$ and $x = y$, then $y \in \mathbb{R}^{0+}$.*

PROOF. Let $x := ((a_n)_n, M)$ and $y := ((b_n)_n, N)$. Assume $x \in \mathbb{R}^{0+}$ and $x = y$, and let p be given. Pick n_0 according to the lemma above and n_1 according to the characterization of equality of reals in Lemma 1.2.3 (RealEqChar) (both for $p+1$). Then for $n \geq n_0, n_1$

$$-\frac{1}{2^p} \leq -\frac{1}{2^{p+1}} + a_n \leq (b_n - a_n) + a_n.$$

Hence $y \in \mathbb{R}^{0+}$ by definition. □

LEMMA 1.4.3 (RealPosChar). *For a real $x := ((a_n)_n, M)$ with $x \in_p \mathbb{R}^+$ we have*

$$\frac{1}{2^{p+1}} \leq a_n \quad \textit{for } M(p+1) \leq n.$$

Conversely, from $\forall_{n \geq n_0}(\frac{1}{2^q} \leq a_n)$ we can infer $x \in_{q+1} \mathbb{R}^+$.

PROOF. Assume $x \in_p \mathbb{R}^+$, that is $\frac{1}{2^p} \leq a_{M(p+1)}$. Then

$$\frac{1}{2^{p+1}} \leq -\frac{1}{2^{p+1}} + a_{M(p+1)} = -\frac{1}{2^{p+1}} + (a_{M(p+1)} - a_n) + a_n \leq a_n$$

for $M(p+1) \leq n$. Conversely,

$$\begin{aligned}
\frac{1}{2^{q+1}} &< -\frac{1}{2^{q+2}} + \frac{1}{2^q} \\
&\leq -\frac{1}{2^{q+2}} + a_n && \text{for } n_0 \leq n \\
&\leq (a_{M(q+2)} - a_n) + a_n && \text{for } M(q+2) \leq n.
\end{aligned}$$

Hence $x \in_{q+1} \mathbb{R}^+$. □

Positivity is compatible with equality, but only up to a shift of p:

LEMMA 1.4.4 (RealPosCompat). *If $x \in_p \mathbb{R}^+$ and $x = y$, then $y \in_{p+2} \mathbb{R}^+$.*

PROOF. Let $x := ((a_n)_n, M)$ and $y := ((b_n)_n, N)$. Assume $x = y$ and $x \in_p \mathbb{R}^+$, that is $\frac{1}{2^p} \leq a_{M(p+1)}$. The goal is $\frac{1}{2^{p+2}} \leq b_{N(p+3)}$. We have

$$\frac{1}{2^{p+2}} = \frac{1}{2^{p+1}} - \frac{1}{2^{p+2}} \leq a_{M(p+3)} + (b_{N(p+3)} - a_{M(p+3)})$$

using Lemma 1.4.3 (RealPosChar) with the monotonicity of M, and the definition of $x = y$. □

1.5. Arithmetical functions

Given real numbers $x := ((a_n)_n, M)$ and $y := ((b_n)_n, N)$, we define $x+y$, $-x$, $|x|$, $x \cdot y$, and $\frac{1}{x}$ (the latter only provided that $|x| \in_q \mathbb{R}^+$) as represented by the respective sequence (c_n) of rationals with modulus L:

	c_n	$L(p)$
$x+y$	$a_n + b_n$	$\max\bigl(M(p+1), N(p+1)\bigr)$
$-x$	$-a_n$	$M(p)$
$\lvert x\rvert$	$\lvert a_n\rvert$	$M(p)$
$x \cdot y$	$a_n \cdot b_n$	$\max\bigl(M(p+1+p_y), N(p+1+p_x)\bigr)$
$\frac{1}{x}$ for $\lvert x\rvert \in_q \mathbb{R}^+$	$\begin{cases} \frac{1}{a_n} & \text{if } a_n \neq 0 \\ 0 & \text{if } a_n = 0 \end{cases}$	$M(2(q+1)+p)$

where 2^{p_x} is the upper bound provided by Lemma 1.3.1 (RealBound).

Lemma 1.5.1. *For reals x, y also $x+y$, $-x$, $|x|$, $x \cdot y$ and (provided that $|x| \in_q \mathbb{R}^+$) also $1/x$ are reals.*

Proof. We restrict ourselves to the cases $x \cdot y$ and $1/x$.

$$\begin{aligned} |a_n b_n - a_m b_m| &= |a_n(b_n - b_m) + (a_n - a_m) b_m| \\ &\leq |b_n - b_m| \cdot |a_n| + |a_n - a_m| \cdot |b_m| \\ &\leq |b_n - b_m| \cdot 2^{p_x} + |a_n - a_m| \cdot 2^{p_y} \leq \frac{1}{2^p} \end{aligned}$$

for $n, m \geq \max\bigl(M(p+1+p_y), N(p+1+p_x)\bigr)$.

For $1/x$ assume $|x| \in_q \mathbb{R}^+$. Then by the (proof of our) characterization of positivity in Lemma 1.4.3 (RealPosChar), $\frac{1}{2^{q+1}} \leq |a_n|$ for $n \geq M(q+1)$. Hence

$$\begin{aligned} \Bigl|\frac{1}{a_n} - \frac{1}{a_m}\Bigr| &= \frac{|a_m - a_n|}{|a_n a_m|} \\ &\leq 2^{2(q+1)} |a_m - a_n| \quad \text{for } n, m \geq M(q+1) \\ &\leq \frac{1}{2^p} \quad \text{for } n, m \geq M(2(q+1)+p). \end{aligned}$$

The claim now follows from the assumption that M is weakly increasing. □

LEMMA 1.5.2. *For reals x, y, z*

$$\begin{aligned} x + (y + z) &= (x + y) + z & x \cdot (y \cdot z) &= (x \cdot y) \cdot z \\ x + 0 &= x & x \cdot 1 &= x \\ x + (-x) &= 0 & 0 < |x| \to x \cdot \frac{1}{x} &= 1 \\ x + y &= y + x & x \cdot y &= y \cdot x \end{aligned}$$

$$x \cdot (y + z) = x \cdot y + x \cdot z$$

PROOF. For $0 < |x| \to x \cdot \frac{1}{x} = 1$ the Cauchy sequences are finally the same, which suffices. In all other cases the Cauchy sequences are identical. □

LEMMA 1.5.3. *The functions $x + y$, $-x$, $|x|$, $x \cdot y$ and (provided that $|x| \in_q \mathbb{R}^+$) also $1/x$ are compatible with equality.*

PROOF. Routine. For instance in case $x + y$ because of the commutativity of $+$ it suffices to prove $x = y \to x + z = y + z$. But this follows immediately from Lemma 1.2.3 (RealEqChar): the n_0 for the conclusion can be the same as for the premise. □

LEMMA 1.5.4. *For reals x, y from $x \cdot y = 1$ we can infer $0 < |x|$.*

PROOF. Pick p such that $|b_n| \leq 2^p$ for all n. Pick n_0 such that $n_0 \leq n$ implies $\frac{1}{2} \leq a_n \cdot b_n$. Then $\frac{1}{2} \leq |a_n| \cdot 2^p$ for $n_0 \leq n$, and hence $\frac{1}{2^{p+1}} \leq |a_n|$. □

LEMMA 1.5.5. *For reals x, y,*

(a) $x, y \in \mathbb{R}^{0+} \to x + y, x \cdot y \in \mathbb{R}^{0+}$,
(b) $x, y \in \mathbb{R}^{+} \to x + y, x \cdot y \in \mathbb{R}^{+}$,
(c) $x \in \mathbb{R}^{0+} \to -x \in \mathbb{R}^{0+} \to x = 0$.

PROOF. (a), (b). Routine. (c). Let p be given. Pick n_0 such that $-\frac{1}{2^p} \leq a_n$ and $-\frac{1}{2^p} \leq -a_n$ for $n \geq n_0$. Then $|a_n| \leq \frac{1}{2^p}$. □

1.6. Comparison of reals

We write $x \leq y$ for $y - x \in \mathbb{R}^{0+}$ and $x < y$ for $y - x \in \mathbb{R}^{+}$. Unwinding the definitions yields that $x \leq y$ is to say that for every p, $a_{L(p)} \leq b_{L(p)} + \frac{1}{2^p}$ with $L(p) := \max(M(p), N(p))$, or equivalently (using Lemma 1.4.1 (RealNNegChar)) that for every p there exists n_0 such that $a_n \leq b_n + \frac{1}{2^p}$ for all $n \geq n_0$. Furthermore, $x < y$ is a shorthand for the presence of p with $a_{L(p+1)} + \frac{1}{2^p} \leq b_{L(p+1)}$ with L the maximum of M and N, or equivalently (using Lemma 1.4.3 (RealPosChar)) for the presence of p, q with $a_n + \frac{1}{2^p} \leq b_n$ for all $n \geq q$; we then write $x <_p y$ (or $x <_{p,q} y$) whenever we want to call these witnesses.

LEMMA 1.6.1 (RealApprox). $\forall_{x,p}\exists_a(|a-x| \le \frac{1}{2^p})$.

PROOF. Let $x = ((a_n), M)$. Given p, pick $a_{M(p)}$. We show $|a_{M(p)} - x| \le \frac{1}{2^p}$, that is $|a_{M(p)} - a_{M(q)}| \le \frac{1}{2^p} + \frac{1}{2^q}$ for every q. But this follows from

$$|a_{M(p)} - a_{M(q)}| \le |a_{M(p)} - a_{M(p+q)}| + |a_{M(p+q)} - a_{M(q)}| \le \frac{1}{2^p} + \frac{1}{2^q}. \qquad \square$$

LEMMA 1.6.2. *For reals x, y, z,*

$$\begin{array}{ll} x \le x & \\ x \le y \to y \le x \to x = y & x \not< x \\ x \le y \to y \le z \to x \le z & x < y \to y < z \to x < z \\ x \le y \to x + z \le y + z & x < y \to x + z < y + z \\ x \le y \to 0 \le z \to x \cdot z \le y \cdot z & x < y \to 0 < z \to x \cdot z < y \cdot z \end{array}$$

PROOF. From Section 1.5. □

Here we have left out information on witnesses p for the statements proving a $<$-formula. Such estimates can easily be given explicitely. Here are two examples.

LEMMA 1.6.3 (RealPosPlus). $0 \le x \to 0 <_p y \to 0 <_{p+3} x + y$.

PROOF. From $0 \le x$ we have $\forall_q\exists_{n_0}\forall_{n\ge n_0}(-\frac{1}{2^q} \le a_n)$. From $0 <_p y$ we have some n_1 such that $\forall_{n\ge n_1}(\frac{1}{2^{p+1}} \le b_n)$. Pick n_0 for $p+2$. Then $n_0, n_1 \le n$ implies $0 \le a_n + \frac{1}{2^{p+2}}$ and $\frac{1}{2^{p+2}} \le b_n - \frac{1}{2^{p+2}}$, hence $\frac{1}{2^{p+2}} \le a_n + b_n$. Now Lemma 1.4.3 (RealPosChar) gives $0 <_{p+3} x + y$. □

LEMMA 1.6.4. $x \le y \to y <_p z \to x <_{p+5} z$.

PROOF. This follows from Lemma 1.6.3 (RealPosPlus). □

As is to be expected in view of the existential and universal character of the predicates $<$ and $\le$ on the reals, we have:

LEMMA 1.6.5 (LeIsNotGt). $x \le y \leftrightarrow y \not< x$.

PROOF. $\to$. Assume $x \le y$ and $y < x$. By Lemma 1.6.4 we obtain $x < x$, a contradiction.

$\leftarrow$. It clearly suffices to show $0 \not< z \to z \le 0$, for a real z given by $(c_n)_n$. Assume $0 \not< z$. We must show $\forall_p\exists_{n_0}\forall_{n\ge n_0}(c_n \le \frac{1}{2^p})$. Let p be given. By assumption $0 \not< z$, hence $\neg\exists_q(\frac{1}{2^q} \le c_{M(q+1)})$. For $q := p+1$ this implies $c_{M(p+2)} < \frac{1}{2^{p+1}}$, hence $c_n \le c_{M(p+2)} + \frac{1}{2^{p+2}} < \frac{1}{2^p}$ for $M(p+2) \le n$. □

Constructively, we cannot compare two reals, but we can compare every real with a nontrivial interval.

LEMMA 1.6.6 (ApproxSplit). *Let x, y, z be given and assume $x < y$. Then either $z \leq y$ or $x \leq z$.*

PROOF. Let $x := ((a_n)_n, M)$, $y := ((b_n)_n, N)$, $z := ((c_n)_n, L)$. Assume $x <_p y$, that is (by definition) $\frac{1}{2^p} \leq b_n - a_n$ for $n := \max(M(p+2), N(p+2))$. Let $m := \max(n, L(p+2))$.

Case $c_m \leq \frac{a_n+b_n}{2}$. We show $z \leq y$. It suffices to prove $c_l \leq b_l$ for $l \geq m$. This follows from

$$c_l \leq c_m + \frac{1}{2^{p+2}} \leq \frac{a_n + b_n}{2} + \frac{b_n - a_n}{4} = b_n - \frac{b_n - a_n}{4} \leq b_n - \frac{1}{2^{p+2}} \leq b_l.$$

Case $c_m \not\leq \frac{a_n+b_n}{2}$. We show $x \leq z$. This follows from $a_l \leq c_l$ for $l \geq m$:

$$a_l \leq a_n + \frac{1}{2^{p+2}} \leq a_n + \frac{b_n - a_n}{4} \leq \frac{a_n + b_n}{2} - \frac{b_n - a_n}{4} \leq c_m - \frac{1}{2^{p+2}} \leq c_l. \qquad \square$$

Notice that the boolean object determining whether $z \leq y$ or $x \leq z$ depends on the representation of x, y and z. In particular this assignment is *not* compatible with our equality relation.

One might think that the non-available comparison of two reals could be circumvented by using a maximum function. Indeed, such a function can easily be defined (component-wise), and it has the expected properties $x, y \leq \max(x, y)$ and $x, y \leq z \to \max(x, y) \leq z$. But what is missing is the knowledge that $\max(x, y)$ equals one of its arguments, i.e., we do not have $\max(x, y) = x \vee \max(x, y) = y$.

However, in many cases it is sufficient to pick the up to ε largest real out of finitely many given ones. This is indeed possible. We give the proof for two reals; it can be easily generalized.

LEMMA 1.6.7 (Maximum of two reals). *Let $x := ((a_n)_n, M)$ and $y := ((b_n)_n, N)$ be reals, and $p \in \mathbb{Z}^+$. Then either $x \leq y + \frac{1}{2^p}$ or else $y \leq x + \frac{1}{2^p}$.*

PROOF. Let $m := \max\bigl(M(p+1), N(p+1)\bigr)$.

Case $a_m \leq b_m$. Then for $m \leq n$

$$a_n \leq a_m + \frac{1}{2^{p+1}} \leq b_m + \frac{1}{2^{p+1}} \leq b_n + \frac{1}{2^p}.$$

This holds for all $n \geq m$, therefore $x \leq y + \frac{1}{2^p}$.

Case $b_m < a_m$. Then for $m \leq n$

$$b_n \leq b_m + \frac{1}{2^{p+1}} < a_m + \frac{1}{2^{p+1}} \leq a_n + \frac{1}{2^p}.$$

This holds for all $n \geq m$, therefore $y \leq x + \frac{1}{2^p}$. $\square$

1.7. Cleaning of reals

After some computations involving real numbers the rational numbers occurring in the Cauchy sequences may become rather complex. Hence under computational aspects it is necessary to *clean up* a real, as follows.

LEMMA 1.7.1. *For every real $x = ((a_n)_n, M)$ we can construct an equivalent real $y = ((b_n)_n, N)$ where the rationals b_n are of the form $k_n/2^n$ with integers k_n, and with modulus $N(p) = p + 2$.*

PROOF. Let $k_n := \lfloor a_{M(n)} \cdot 2^n \rfloor$ and $b_n := \frac{k_n}{2^n}$, hence

$$\frac{k_n}{2^n} \le a_{M(n)} < \frac{k_n}{2^n} + \frac{1}{2^n} \quad \text{with } k_n \in \mathbb{Z}.$$

Then for $n \le m$

$$\begin{aligned} |b_n - b_m| &= |\frac{k_n}{2^n} - \frac{k_m}{2^m}| \\ &\le |\frac{k_n}{2^n} - a_{M(n)}| + |a_{M(n)} - a_{M(m)}| + |a_{M(m)} - \frac{k_m}{2^m}| \\ &\le \frac{1}{2^n} + \frac{1}{2^n} + \frac{1}{2^m} \\ &< \frac{4}{2^n}, \end{aligned}$$

hence $|b_n - b_m| \le \frac{1}{2^p}$ for $m \ge n \ge p + 2 =: N(p)$, so $(b_n)_n$ is a Cauchy sequence with modulus N.

To prove that x is equivalent to $y := ((b_n)_n, N)$, observe

$$|a_n - b_n| \le |a_n - a_{M(n)}| + |a_{M(n)} - \frac{k_n}{2^n}| \le \frac{1}{2^{p+1}} + \frac{1}{2^n} \le \frac{1}{2^p}$$

for $n \ge \max(p + 1, M(p + 1))$, and therefore $x = y$. □

CHAPTER 2

Sequences and series of real numbers

2.1. Completeness

Definition 2.1.1. A sequence $(x_n)_{n\in\mathbb{N}}$ of reals is a *Cauchy sequence* with modulus $M\colon \mathbb{Z}^+ \to \mathbb{N}$ whenever $|x_n - x_m| \leq \frac{1}{2^p}$ for $n, m \geq M(p)$, and *converges* with modulus $M\colon \mathbb{Z}^+ \to \mathbb{N}$ to a real y, its *limit*, whenever $|x_n - y| \leq \frac{1}{2^p}$ for $n \geq M(p)$.

Clearly the limit of a convergent sequence of reals is uniquely determined.

Lemma 2.1.2 (RatCauchyConvMod). *Every modulated Cauchy sequence of rationals converges with the same modulus to the real number it represents.*

Proof. Let $x := ((a_n)_n, M)$ be a real. We must show $|a_n - x| \leq \frac{1}{2^p}$ for $n \geq M(p)$. Fix $n \geq M(p)$. It suffices to show $|a_n - a_m| \leq \frac{1}{2^p}$ for $m \geq M(p)$. But this holds by assumption. □

By the triangle inequality, every convergent sequence of reals with modulus M is a Cauchy sequence with modulus $p \mapsto M(p+1)$. We now prove the reverse implication.

Theorem 2.1.3 (RealCompl). *For every Cauchy sequence of reals we can find a real to which it converges.*

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence of reals with modulus M; say x_n is $((a_{nl})_l, N_n)$. Note first that, for each $n \in \mathbb{N}$ and every p, by Lemma 2.1.2 (RatCauchyConvMod) we have $|x_n - a_{nl}| \leq \frac{1}{2^p}$ for all $l \geq N_n(p)$. Next, set

$$b_n := a_{nN_n(n)}$$

for every $n \in \mathbb{N}$, so that

$$|x_n - b_n| \leq \frac{1}{2^n} \quad \text{for all } n \in \mathbb{N}$$

by the particular case $l = N_n(n)$ of the foregoing consideration. Then

$$|b_m - b_n| \leq |b_m - x_m| + |x_m - x_n| + |x_n - b_n| \leq \frac{1}{2^m} + \frac{1}{2^{q+1}} + \frac{1}{2^n} \leq \frac{1}{2^q}$$

for all $m, n \geq \max(M(q+1), q+2)$, which is to say that $y := (b_n)_n$ is a Cauchy sequence with modulus $L(q) := \max(M(q+1), q+2)$. Moreover, again by Lemma 2.1.2 (RatCauchyConvMod)

$$|x_n - y| \leq |x_n - b_n| + |b_n - y| \leq \frac{1}{2^n} + \frac{1}{2^{q+1}} \leq \frac{1}{2^q}$$

for all $n \geq L(q+1)$. In other words: (x_n) converges to y with modulus $q \mapsto L(q+1)$. □

One can even say that (x_n) converges to y with the same modulus that (x_n) has as a Cauchy sequence. More generally, Lemma 2.1.2 (RatCauchyConvMod) holds for Cauchy sequences of reals as well.

Lemma 2.1.4 (RealCauchyConvMod). *Every modulated Cauchy sequence of reals converges with the same modulus to its limit.*

Proof. Let $(x_n)_n$ be a Cauchy sequence of reals with modulus M, that is

$$|x_n - x_m| \leq \frac{1}{2^p} \quad \text{for } n, m \geq M(p).$$

Let y be the limit of $(x_n)_n$, that is

$$|x_n - y| \leq \frac{1}{2^q} \quad \text{for } n \geq L(q).$$

We shall prove

$$|x_n - y| \leq \frac{1}{2^p} \quad \text{for } n \geq M(p).$$

Fix $n \geq M(p)$, and let $q \in \mathbb{Z}^+$. Then

$$\begin{aligned} |x_n - y| &\leq |x_n - x_m| + |x_m - y| \quad \text{for } m \geq M(p), L(q) \\ &\leq \frac{1}{2^p} + \frac{1}{2^q}. \end{aligned}$$

The claim follows, because this holds for every q. □

It will be useful to have a criterion for convergence of a sequence of reals, in terms of their approximations.

Lemma 2.1.5. *For reals x_n, x represented by $(a_{nl})_l, (b_l)_l$, we can infer that $(x_n)_n$ converges to x, i.e.,*

$$\forall_p \exists_{n_0} \forall_{n \geq n_0} (|x_n - x| \leq \frac{1}{2^p})$$

from

$$\forall_p \exists_{n_0} \forall_{n, l \geq n_0} (|a_{nl} - b_l| \leq \frac{1}{2^p}).$$

PROOF. Given p, we have to find n_0 such that $|x_n - x| \leq \frac{1}{2^p}$ for $n \geq n_0$. By Lemma 1.4.1 (RealNNegChar) it suffices to have $|a_{nl} - b_l| \leq \frac{1}{2^p} + \frac{1}{2^l}$ for $l \geq r$ with r depending on l. But by assumption we even have $|a_{nl} - b_l| \leq \frac{1}{2^p}$ for $l \geq n_0$. □

2.2. Limits and inequalities

We show that limits interact nicely with non-strict inequalities.

LEMMA 2.2.1 (RealNNegLim). *Let $(x_n)_{n\in\mathbb{N}}$ be a convergent sequence of reals and x its limit. Then $0 \leq x_n$ for all n implies $0 \leq x$.*

PROOF. By assumption $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence of reals, say with modulus M. Let x_n be $((a_{nl})_l, N_n)$. Assume $0 \leq x_n$ for all n, that is

$$-\frac{1}{2^q} \leq a_{nN_n(q)} \quad \text{for all } n \in \mathbb{N},\ q \in \mathbb{Z}^+.$$

By the theorem above, $b_n := a_{nN_n(n)}$ is a Cauchy sequence with modulus $L(p) := \max(M(p+1), p+2)$ representing x. We must show $0 \leq x$, that is

$$-\frac{1}{2^p} \leq b_{L(p)} \quad \text{for all } p \in \mathbb{Z}^+.$$

For $k' := L(p)$ and $n := L(p)$ we obtain

$$-\frac{1}{2^p} \leq -\frac{1}{2^{L(p)}} \leq a_{L(p)N_{L(p)}(L(p))} = b_{L(p)} \quad \text{for all } p \in \mathbb{Z}^+.$$

Note that $p < L(p)$ by definition of L. □

2.3. Series

Series are special sequences. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of reals, and define

$$s_n := \sum_{m=0}^{n} x_m.$$

We call s_n a *partial sum* of the sequence (x_n). The sequence

$$(s_n)_{n\in\mathbb{N}} = \Big(\sum_{m=0}^{n} x_m\Big)_{n\in\mathbb{N}} =: \sum_{m=0}^{\infty} x_m$$

is called the *series* determined by the sequence $(x_n)_{n\in\mathbb{N}}$. We say that the series $\sum_{m=0}^{\infty} x_m$ converges if and only if the sequence (s_n) converges. Its limit is somewhat sloppily denoted by $\sum_{m=0}^{\infty} x_m$ as well.

Example 2.3.1. Consider the series $\sum_{m=1}^{\infty} \frac{1}{m(m+1)}$. Its partial sums are

$$s_n := \sum_{m=1}^{n} \frac{1}{m(m+1)} = \frac{n}{n+1};$$

this can be proved by induction on n, as follows. For $n = 0$ the claim clearly holds, and in the induction step $n \mapsto n+1$ we have

$$\begin{aligned}\sum_{m=1}^{n+1} \frac{1}{m(m+1)} &= \sum_{m=1}^{n} \frac{1}{m(m+1)} + \frac{1}{(n+1)(n+2)} \\ &= \frac{n}{n+1} + \frac{1}{(n+1)(n+2)} \\ &= \frac{n(n+2)+1}{(n+1)(n+2)} \\ &= \frac{(n+1)^2}{(n+1)(n+2)} \\ &= \frac{n+1}{n+2}.\end{aligned}$$

Because of $\lim_{n\to\infty} \frac{n}{n+1} = 1$ we obtain $\sum_{m=1}^{\infty} \frac{1}{m(m+1)} = 1$.

Theorem 2.3.2 (Infinite geometric series). *For $|x| < 1$ we have*

$$\sum_{m=0}^{\infty} x^m = \frac{1}{1-x}.$$

Proof. Let $|x| < 1$. The n-th partial sum is

$$\sum_{m=0}^{n} x^m = \frac{1-x^{n+1}}{1-x},$$

which can be proved easily by induction. Hence

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

since $\lim x^{n+1} = 0$ for $|x| < 1$. □

For instance, $\sum_{i=-k}^{\infty} a_i \frac{1}{2^i}$ with $a_i \in \{-1, 0, 1\}$ converges, because

$$\Bigl|\sum_{i=m+1}^{l} a_i \frac{1}{2^i}\Bigr| \le \sum_{i=m+1}^{l} \frac{1}{2^i} < \sum_{i=m+1}^{\infty} \frac{1}{2^i} = \frac{1}{2^m}.$$

We show that every real x can be written in this form.

2.4. Representation of reals

Let $b \geq 2$ be a natural number. A series

$$\sum_{i=-k}^{\infty} a_i b^{-i},$$

where each a_i (for $-k \leq i$) is a natural number with $0 \leq a_i < b$, is called *b-adic expansion.*

Theorem 2.4.1. *Every b-adic expansion converges.*

Proof. Let

$$\sum_{i=-k}^{\infty} a_i b^{-i},$$

be a b-adic expansion. We consider the sequence

$$s_j := \sum_{i=-k}^{j} a_i b^{-i}.$$

Because of the completeness of $\mathbb{R}$ it suffices to show that $(s_j)_{j \geq -k}$ is a Cauchy sequence. This can be seen as follows. Let $-k \leq j \leq l$. Then

$$s_l - s_j = \sum_{i=j+1}^{l} a_i b^{-i} \leq \sum_{i=j+1}^{l} (b-1) b^{-i} = \frac{b-1}{b^{j+1}} \sum_{i=0}^{l-j-1} b^{-i} < \frac{b-1}{b^{j+1}} \cdot \frac{1}{1-\frac{1}{b}} = b^{-j}$$

which implies the claim. □

Theorem 2.4.2 (b-adic expansion). *Every real $x \in \mathbb{R}^{0+}$ can be represented by a b-adic expansion*

$$x = \sum_{i=-k}^{\infty} a_i b^{-i}.$$

Proof. By Lemma 1.3.1 (RealBound) we can find k with $x < b^{k+1}$. Hence we can find (even uniquely) a_{-k} with $0 \leq a_{-k} < b$ such that

$$a_{-k} b^k \leq x < (a_{-k} + 1) b^k$$

and therefore

$$0 \leq x - \sum_{i=-k}^{-k} a_i b^{-i} < b^k.$$

Assume we already have $a_{-k}, a_{-k+1}, \dots, a_j$ such that

$$0 \leq x - \sum_{i=-k}^{j} a_i b^{-i} < b^{-j}.$$

Then we can find (even uniquely) a_{j+1} with $0 \leq a_{j+1} < b$ such that

$$a_{j+1}b^{-j-1} \leq x - \sum_{i=-k}^{j} a_i b^{-i} < (a_{j+1}+1)b^{-j-1},$$

hence

$$0 \leq x - \sum_{i=-k}^{j+1} a_i b^{-i} < b^{-j-1}.$$

Therefore the corresponding b-adic series converges to x, i.e.

$$x = \sum_{i=-k}^{\infty} a_i b^{-i}. \qquad \square$$

Theorem 2.4.3 (Signed digit representation of reals). *Every real x can be represented in the form*

$$\sum_{i=-k}^{\infty} a_i 2^{-i} \quad \textit{with } a_i \in \{-1, 0, 1\}. \tag{5}$$

Proof. By Lemma 1.3.1 (RealBound) we can find k such that $-2^{k+1} \leq x \leq 2^{k+1}$. We recursively construct a_{-k}, a_{-k+1}, ..., a_j, ... such that

$$-2^{-j} \leq x - \sum_{i=-k}^{j} a_i 2^{-i} \leq 2^{-j} \quad \text{for } j \geq -k-1.$$

For $j = -k-1$ this holds by the choice of k. Now assume the claim holds for j; we need to construct a_{j+1} such that it holds for $j+1$ as well. Let $y := x - \sum_{i=-k}^{j} a_i 2^{-i}$, hence $-2^{-j} \leq y \leq 2^{-j}$. By comparing y first with $-2^{-j-1} < 0$ and then $0 < 2^{-j-1}$ we can define a_{j+1} such that

$$a_{j+1} = \begin{cases} -1 & \text{if } y \leq 0 \\ 0 & \text{if } -2^{-j-1} \leq y \leq 2^{-j-1} \\ 1 & \text{if } 0 \leq y. \end{cases}$$

Then in each of the three cases

$$(a_{j+1}-1)2^{-j-1} \leq y \leq (a_{j+1}+1)2^{-j-1},$$

hence

$$-2^{-j-1} \leq y - a_{j+1}2^{-j-1} \leq 2^{-j-1},$$

which was to be shown. $\square$

2.5. Theorem of Bolzano-Weierstraß

The Theorem of Bolzano-Weierstraß is an important instance of an existential theorem that only holds w.r.t. the classical existential quantifier, i.e., whose proof does not provide a construction of what is claimed to exist. For its formulation we need the notion of a subsequence. Let $(n_m)_{m\in\mathbb{N}}$ be a sequence of natural numbers satisfying

$$n_0 < n_1 < \cdots < n_m < n_{m+1} < \dots.$$

Then for an arbitrary sequence $(a_n)_{n\in\mathbb{N}}$ we call $(a_{n_m})_{m\in\mathbb{N}}$ the *subsequence* of (a_n) determined by $(n_m)_{m\in\mathbb{N}}$.

Theorem 2.5.1 (Bolzano-Weierstraß). *For every bounded sequence of reals there must exist a convergent subsequence.*

Proof. Let (x_n) be a bounded sequence of reals, and $a, b \in \mathbb{Q}$ such that

$$a \le x_n \le b \qquad \text{for all } n \in \mathbb{N}.$$

We define recursively $a_m, b_m \in \mathbb{Q}$ and $n_m \in \mathbb{N}$ such that

(i) $a = a_0 \le a_1 \le \cdots \le a_m < b_m \le \cdots \le b_1 \le b_0 = b$ and also $b_m = a_m + \frac{1}{2^m}(b-a)$,
(ii) $a_m \le x_{n_m} \le b_m$ and $n_0 < n_1 < \cdots < n_m$,
(iii) there must be infinitely many $n \ge n_m$ with $a_m \le x_n \le b_m$.

Such a subsequence will be a Cauchy sequence. To see this assume $l \ge m$. In case $x_{n_l} \le x_{n_m}$ we then have

$$0 \le x_{n_m} - x_{n_l} \le b_m - a_l \le b_m - a_m = \frac{1}{2^m}(b-a),$$

and in case $x_{n_m} \le x_{n_l}$

$$0 \le x_{n_l} - x_{n_m} \le b_l - a_m \le b_m - a_m = \frac{1}{2^m}(b-a).$$

For $m = 0$ let $a_0 = a$, $b_0 = b$ and $n_0 = 0$. Now assume that the recursive procedure has been done up to m. Divide the interval $[a_m, b_m]$ into two halfes $[a_m, c]$ and $[c, b_m]$ with $c := \frac{1}{2}(a_m + b_m)$. Since by (iii) there must be infinitely many x_n with $n \ge n_m$ in the interval $[a_m, b_m]$, also in one of the two halfes there must be infinitely many such x_n.

Case 1. For infininitely many $n \ge n_m$ we have $a_m \le x_n \le c$. Then let $a_{m+1} := a_m$, $b_{m+1} := c$ and choose n_{m+1} as the first $n > n_m$ with $a_m \le x_n \le c$.

Case 2. For infininitely many $n \ge n_m$ we have $c \le x_n \le b_m$. Then let $a_{m+1} :=:= c$, $b_{m+1} := b_m$ and choose n_{m+1} as the first $n > n_m$ with $c \le x_n \le b_m$. □

Remark 2.5.2. In the step of the recursive procedure we have used the axiom of dependent choice (DC), in the form

$$A(0,c_0) \to \forall_{m,c}(A(m,c) \to \tilde{\exists}_d A(m+1,d)) \to \tilde{\exists}_f \forall_m A(m,f(m))$$

where $A(m,c)$ expresses that there must be infinitely many $n \geq n_m$ such that $a_m \leq x_n \leq b_m$:

$$\forall_{n \geq n_m}(a_m \leq x_n \leq b_m \to \tilde{\exists}_{n'>n}(a_m \leq x_{n'} \leq b_m)).$$

Notice that both the statement of the Bolzano-Weierstraß theorem and DC are Harrop formulas, i.e., without computational content.

A sequence $(x_n)_{n\in\mathbb{N}}$ of reals is called

(i) *monotone increasing*, if $x_n \leq x_{n+1}$ for all $n \in \mathbb{N}$,
(ii) *strongly monotone increasing*, if $x_n < x_{n+1}$ for all $n \in \mathbb{N}$,
(iii) *monotone decreasing*, if $x_n \geq x_{n+1}$ for all $n \in \mathbb{N}$,
(iv) *strongly monotone decreasing*, if $x_n > x_{n+1}$ for all $n \in \mathbb{N}$,
(v) *monotone*, if it is monotone increasing or monotone decreasing.

Theorem 2.5.3. *Every bounded monotone sequence of reals must have*

(a) *a modulus of convergence, and*
(b) *a limit to which it converges with this modulus.*

Proof. Assume (x_n) is a monotone increasing bounded sequence.

(a) By the Theorem 2.5.1 (Bolzano-Weierstraß) there must be a convergent subsequence (x_{n_m}). We show that because on the monotonicity of (x_n) also the full sequence (x_n) must be a Cauchy sequence. Let $\varepsilon > 0$. Since (x_{n_m}) is a Cauchy sequence, we have $n_0 \in \mathbb{N}$ with

$$x_{n_m} - x_{n_l} \leq \varepsilon$$

for all $l \geq m \geq n_0$. Then for all $p \geq q \geq n_{n_0}$ we have (because of $n_p \geq p$ and $n_{n_0} \geq n_0$)

$$0 \leq x_p - x_q \leq x_{n_p} - x_{n_{n_0}} \leq \varepsilon.$$

(b) By Theorem 2.1.3 (RealCompl) we can find a real y to which (x_n) converges, and by Lemma 2.1.4 (RealCauchyConvMod) we know that the modulus of convergence is the one from (a). □

Also this theorem states Harrop formulas; its proof does not provide constructions of what is claimed to exist.

2.6. Convergence tests

We now consider some of the standard convergence tests for series.

THEOREM 2.6.1 (Cauchy convergence test). *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of reals. The series $\sum_{n=0}^{\infty} x_n$ converges if and only if for every $p \in \mathbb{Z}^+$ there is an $N \in \mathbb{N}$ such that for all $n \geq m \geq N$*

$$\Big|\sum_{\nu=m}^{n} x_\nu\Big| \leq \frac{1}{2^p}.$$

PROOF. The condition expresses that the sequence of partial sums is a Cauchy sequence. □

It follows that the convergence of series does not depend on a possible change of finitely many of its members. However, the limit of the series may well change.

THEOREM 2.6.2. *A necessary (but not sufficient) condition for the convergence of a series $\sum_{n=0}^{\infty} x_n$ is $\lim_{n\to\infty} x_n = 0$.*

PROOF. Assume $\sum_{n=0}^{\infty} x_n$ is convergent. We must show $\lim_{n\to\infty} x_n = 0$, which means

$$\forall_p \exists_{n_0} \forall_{n\geq n_0} (|x_n| \leq \frac{1}{2^p}).$$

So let $p \in \mathbb{Z}^+$. Then there is an $n_0 \in \mathbb{N}$ such that for all $n \geq m \geq n_0$

$$\Big|\sum_{\nu=m}^{n} x_\nu\Big| \leq \frac{1}{2^p}.$$

In particular then $|x_n| \leq \frac{1}{2^p}$ for $n \geq N$. □

EXAMPLE 2.6.3. The *harmonic series* $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges to $+\infty$. This can be seen by grouping its members together:

$$1 + \frac{1}{2} + \underbrace{\Big(\frac{1}{3} + \frac{1}{4}\Big)}_{\geq \frac{2}{4} = \frac{1}{2}} + \underbrace{\Big(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\Big)}_{\geq \frac{4}{8} = \frac{1}{2}} + \dots .$$

More precisely, one first shows that for all $m \in \mathbb{N}$

$$\sum_{\nu=2^m+1}^{2^{m+1}} \frac{1}{\nu} \geq 2^m \cdot \frac{1}{2^{m+1}} = \frac{1}{2}.$$

This implies

$$\sum_{\nu=1}^{2^{n+1}} \frac{1}{\nu} = 1 + \sum_{m=0}^{n} \sum_{\nu=2^m+1}^{2^{m+1}} \frac{1}{\nu} \geq 1 + \sum_{m=0}^{n} \frac{1}{2},$$

which implies the claim. The harmonic series is an example that the condition $\lim_{n\to\infty} x_n = 0$ does *not* ensure convergence of the series $\sum_{n=0}^{\infty} x_n$.

THEOREM 2.6.4 (Leibniz test for alternating series). *Let $(x_n)_{n\in\mathbb{N}}$ be a decreasing sequence of non-negative reals with $\lim_{n\to\infty} x_n = 0$. Then the series*

$$\sum_{n=0}^{\infty}(-1)^n x_n$$

converges.

PROOF. Because of $\lim_{n\to\infty} x_n = 0$ it suffices to show

$$\forall_{m,n}(0 \le (-1)^n \sum_{\nu=n}^{n+m}(-1)^\nu x_\nu \le x_n).$$

The proof is by induction on m. For $m = 0$ the claim is $0 \le (-1)^{2n}x_n = x_n$, and in the step $m \mapsto m+1$ we have

$$\begin{aligned}(-1)^n \sum_{\nu=n}^{n+m+1}(-1)^\nu x_\nu &= (-1)^n((-1)^n x_n + \sum_{\nu=n+1}^{n+m+1}(-1)^\nu x_\nu)\\ &= x_n - (-1)^{n+1}\sum_{\nu=n+1}^{n+1+m}(-1)^\nu x_\nu,\end{aligned}$$

and by induction hypothesis $0 \le \sum_{\nu=n+1}^{n+1+m}(-1)^\nu x_\nu \le x_{n+1}$. □

For example, the series $\sum_{n=1}^{\infty}\frac{(-1)^n}{n}$ converges by the Leibniz test.

DEFINITION 2.6.5. A series $\sum_{n=0}^{\infty} x_n$ is *absolutely convergent* if $\sum_{n=0}^{\infty}|x_n|$ converges.

Clearly every absolutely convergent series is convergent. The converse does not hold generally, by the example above.

THEOREM 2.6.6 (Comparison test). *Let $\sum_{n=0}^{\infty} y_n$ be a convergent series with non-negative y_n. If $|x_n| \le y_n$ for all $n \in \mathbb{N}$, then $\sum_{n=0}^{\infty} x_n$ is absolutely convergent.*

PROOF. We have to show that $\sum_{n=0}^{\infty}|x_n|$ converges. Let $\nu \in \mathbb{N}$. Since $\sum_{n=0}^{\infty} y_n$ converges, we have an $N \in \mathbb{N}$ such that for all $n \ge m \ge N$

$$\sum_{\nu=m}^{n} y_\nu \le \frac{1}{2^p}.$$

But then also

$$\sum_{\nu=m}^{n}|x_\nu| \le \sum_{\nu=m}^{n} y_\nu \le \frac{1}{2^p}.$$ □

EXAMPLE 2.6.7. The series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges for every $p \geq 2$. To see this, recall that $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ converges, hence also $\sum_{n=1}^{\infty} \frac{2}{n(n+1)}$. Because of $p \geq 2$ we have for all $n \geq 1$

$$\frac{1}{n^p} \leq \frac{1}{n^2} \leq \frac{2}{n(n+1)}.$$

Hence by the comparison test $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges as well.

THEOREM 2.6.8 (Ratio test). *Assume*

$$|x_{n+1}| \leq q|x_n| \quad \textit{for all } n \geq n_0$$

with $0 \leq q < 1$. *Then the series* $\sum_{n=0}^{\infty} x_n$ *is absolutely convergent.*

PROOF. Since the convergence of series does not depend on a possible change of finitely many of its members, we may assume $n_0 = 0$. By assumption we have for all n

$$|x_n| \leq q^n |x_0|;$$

this can be seen easily by induction. The geometric series $\sum_{n=0}^{\infty} q^n$ converges (because of $0 \leq q < 1$), hence also $\sum_{n=0}^{\infty} q^n |x_0|$. From the comparison test we can conclude the absolute convergence of $\sum_{n=0}^{\infty} x_n$. □

EXAMPLE 2.6.9. The series $\sum_{n=1}^{\infty} \frac{n^2}{2^n}$ converges. To see this, observe that for all $n \geq 3$

$$\frac{(n+1)^2 \cdot 2^n}{2^{n+1} \cdot n^2} = \frac{1}{2}\Big(1 + \frac{1}{n}\Big)^2 \leq \frac{1}{2} \cdot \frac{16}{9} = \frac{8}{9} < 1.$$

Hence the series converges by the ratio test.

2.7. Reordering

Let $\sum_{n=0}^{\infty} x_n$ be a series. If $\tau \colon \mathbb{N} \to \mathbb{N}$ is a bijective map, then the series $\sum_{n=0}^{\infty} x_{\tau(n)}$ is a *reordering* of $\sum_{n=0}^{\infty} x_n$.

THEOREM 2.7.1 (Reordering theorem). *Let* $\sum_{n=0}^{\infty} x_n$ *be absolutely convergent with limit* x. *Then every reordering of it converges to* x *as well.*

PROOF. (See Forster (2004)). We must show

$$\lim_{m \to \infty} \sum_{n=0}^{m} x_{\tau(n)} = x.$$

Let $p \in \mathbb{Z}^+$. Because of the absolute convergence of $\sum_{n=0}^{\infty} x_n$ we have an n_0 such that

$$\sum_{n=n_0}^{\infty} |x_n| \leq \frac{1}{2^{p+1}}.$$

Hence

$$\Big|x - \sum_{n=0}^{n_0-1} x_n\Big| = \Big|\sum_{n=n_0}^{\infty} x_n\Big| \leq \sum_{n=n_0}^{\infty} |x_n| \leq \frac{1}{2^{p+1}}.$$

Now choose N such that $\{\tau(0), \tau(1), \dots, \tau(N)\} \supseteq \{0, 1, \dots, n_0 - 1\}$. Then for all $m \geq N$

$$\begin{aligned}\Big|\sum_{n=0}^{m} x_{\tau(n)} - x\Big| &\leq \Big|\sum_{n=0}^{m} x_{\tau(n)} - \sum_{n=0}^{n_0-1} x_n\Big| + \Big|\sum_{n=0}^{n_0-1} x_n - x\Big| \\ &\leq \sum_{n=n_0}^{\infty} |x_n| + \frac{1}{2^{p+1}} \leq \frac{1}{2^p}.\end{aligned}$$

□

2.8. The exponential series

THEOREM 2.8.1. *The* exponential series

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

is absolutely convergent, for every real x.

PROOF.

$$\Big|\frac{x^{n+1}}{(n+1)!}\Big| \leq \frac{1}{2}\Big|\frac{x^n}{n!}\Big|$$

is equivalent to $2|x| \leq n + 1$. Hence the series converges absolutely by the ratio test. □

The Euler number e is defined as

$$e := \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!}.$$

THEOREM 2.8.2 (Estimate of the rest).

$$\Big|\sum_{n=N+1}^{\infty} \frac{x^n}{n!}\Big| \leq 2\frac{|x|^{N+1}}{(N+1)!} \quad \textit{for } |x| \leq 1 + \tfrac{N}{2}.$$

PROOF.

$$\begin{aligned}&\Big|\sum_{n=N+1}^{\infty} \frac{x^n}{n!}\Big| \\ &\leq \sum_{n=N+1}^{\infty} \frac{|x^n|}{n!}\end{aligned}$$

$$= \frac{|x|^{N+1}}{(N+1)!}\Big(1 + \frac{|x|}{N+2} + \dots + \frac{|x|^m}{(N+2)\dots(N+m+1)} + \dots\Big).$$

For $\frac{|x|}{N+2} \le \frac{1}{2}$ or $|x| \le 1 + \frac{N}{2}$ we can estimate this series against the geometric series, since

$$\frac{|x|^m}{(N+2)\dots(N+m+1)} \le \Big(\frac{|x|}{N+2}\Big)^m \le \frac{1}{2^m}.$$

Hence for $|x| \le 1 + \frac{N}{2}$

$$\Big|\sum_{n=N+1}^{\infty} \frac{x^n}{n!}\Big| \le \frac{|x|^{N+1}}{(N+1)!}\sum_{n=0}^{\infty}\frac{1}{2^n} = 2\frac{|x|^{N+1}}{(N+1)!}.$$ □

Theorem 2.8.3 (Cauchy product). *Assume $\sum_{n=0}^{\infty} x_n$ and $\sum_{n=0}^{\infty} y_n$ are absolutely convergent, and define*

$$z_n := \sum_{m=0}^{n} x_{n-m}y_m.$$

Then $\sum_{n=0}^{\infty} z_n$ is absolutely convergent as well, and

$$\sum_{n=0}^{\infty} z_n = \Big(\sum_{n=0}^{\infty} x_n\Big)\cdot\Big(\sum_{n=0}^{\infty} y_n\Big).$$

Proof. (See Forster (2004)). Define

$$Z_n := \sum_{m=0}^{n} z_m.$$

We first show

$$\lim_{n\to\infty} Z_n = \sum_{m=0}^{\infty} z_m = \Big(\sum_{n=0}^{\infty} x_n\Big)\cdot\Big(\sum_{n=0}^{\infty} y_n\Big).$$

For

$$Z_n^* := \Big(\sum_{m=0}^{n} x_m\Big)\cdot\Big(\sum_{m=0}^{n} y_m\Big),$$

we clearly have

$$\lim_{n\to\infty} Z_n^* = \Big(\sum_{n=0}^{\infty} x_n\Big)\cdot\Big(\sum_{n=0}^{\infty} y_n\Big).$$

Hence it suffices to show

$$\lim_{n\to\infty}(Z_n^* - Z_n) = 0.$$

To prove this, consider

$$P_n^* := \Big(\sum_{m=0}^{n} |x_m|\Big) \cdot \Big(\sum_{m=0}^{n} |y_m|\Big).$$

Since by assumption both $\sum_{n=0}^{\infty} x_n$ and $\sum_{n=0}^{\infty} y_n$ are absolutely convergent, $(P_n^*)_{n\in\mathbb{N}}$ converges.

Now let $p \in \mathbb{Z}^+$. From the convergence of $(P_n^*)_{n\in\mathbb{N}}$ we obtain an N such that for all $n \geq m \geq N$

$$P_n^* - P_m^* = \sum_{\substack{i,j\leq n\\ m<\max(i,j)}} |x_i||x_j| \leq \frac{1}{2^p}.$$

Hence for $n \geq 2N$

$$|Z_n^* - Z_n| = \Big|\sum_{\substack{i,j\leq n\\ n<i+j}} x_i x_j\Big| \leq \sum_{\substack{i,j\leq n\\ n<i+j}} |x_i||x_j| \leq \sum_{\substack{i,j\leq n\\ N<\max(i,j)}} |x_i||x_j| = P_n^* - P_N^* \leq \frac{1}{2^p}.$$

It remains to show that $\sum_{n=0}^{\infty} z_n$ is absolutely convergent. This follows from the comparison test and the previous arguments, applied to the series $\sum_{n=0}^{\infty} |x_n|$ und $\sum_{n=0}^{\infty} |y_n|$ instead of $\sum_{n=0}^{\infty} x_n$ and $\sum_{n=0}^{\infty} y_n$. For then

$$\sum_{n=0}^{\infty}\sum_{m=0}^{n} |x_{n-m}||y_m|$$

converges to $(\sum_{n=0}^{\infty} |x_n|) \cdot (\sum_{n=0}^{\infty} |y_n|)$. Because of

$$|z_n| = \Big|\sum_{m=0}^{n} x_{n-m}y_m\Big| \leq \sum_{m=0}^{n} |x_{n-m}||y_m|$$

the comparison test implies the absolute convergence of $\sum_{n=0}^{\infty} z_n$. □

If instead of the absolute convergence of $\sum_{n=0}^{\infty} x_n$ and $\sum_{n=0}^{\infty} y_n$ we only assume ordinary convergence, $\sum_{n=0}^{\infty} z_n$ in general will not converge.

THEOREM 2.8.4 (Functional equation for the exponential function).

$$\exp(x+y) = \exp(x)\exp(y) \quad \textit{for all } x, y \in \mathbb{R}.$$

PROOF. Applying the Cauchy product to the absolutely convergent series

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{and} \quad \exp(y) = \sum_{n=0}^{\infty} \frac{y^n}{n!}$$

gives

$$\exp(x)\exp(y) = \sum_{n=0}^{\infty}\sum_{m=0}^{n} \frac{x^{n-m}}{(n-m)!}\frac{y^m}{m!}$$

$$
\begin{aligned}
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{m=0}^{n} \binom{n}{m} x^{n-m} y^m \\
&= \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} \quad \text{by the Binomial theorem} \\
&= \exp(x+y).
\end{aligned}
$$

□

Corollary 2.8.5. (a) $0 \leq \exp(x)$ *for all* $x \in \mathbb{R}$.
(b) $0 < \exp(x)$ *for all* $x \in \mathbb{R}$.
(c) $\exp(-x) = \exp(x)^{-1}$ *for all* $x \in \mathbb{R}$.
(d) $\exp(k) = e^k$ *for every integer* $k \in \mathbb{Z}$.

Proof. First notice

$$\exp(x)\exp(-x) = \exp(x-x) = \exp(0) = 1.$$

(a). Since the goal is stable, we can distinguish cases (i) $0 \leq x$ and (ii) $\neg(0 \leq x)$ (see Appendix A). In case (i) we clearly have $\exp(x) \geq 1$. In case (ii) we have $x \leq 0$. Assume $\exp(x) \leq 0$. Then $1 = \exp(x)\exp(-x) \leq 0$ since $x \leq 0$ and therefore $0 \leq \exp(-x)$. This contradiction proves the claim.

(b). Pick $p \in \mathbb{Z}^+$ with $\exp(-x) \leq p$. Then $1 = \exp(x)\exp(-x) \leq \exp(x)p$ by (a). Hence the claim.

(c) is now immediate, and for (d) we use induction on n. Clearly $\exp(0) = 1 = e^0$; for $n \mapsto n+1$

$$\exp(n+1) = \exp(n)\exp(1) = e^n \cdot e = e^{n+1} \quad \text{by induction hypothesis,}$$

and for $k < 0$

$$\exp(k) = \frac{1}{\exp(-k)} = \frac{1}{e^{-k}} = e^k.$$

□

2.9. The exponential function for complex numbers

Later we shall define the sine and cosine functions by means of the complex exponential function, using the Euler equation

$$e^{ix} = \cos x + i \sin x.$$

As a preparation we introduce the complex numbers and prove their fundamental properties.

On the set $\mathbb{R} \times \mathbb{R}$ we define addition and multiplication by

$$
\begin{aligned}
(x_1, y_1) + (x_2, y_2) &:= (x_1 + x_2, y_1 + y_2), \\
(x_1, y_1) \cdot (x_2, y_2) &:= (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2).
\end{aligned}
$$

One can check easily that all the field axioms are satisfied if one defines $(0,0)$ as zero and $(1,0)$ als one. This field is called the field $\mathbb{C}$ of *complex* *numbers*. Because of

$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0),$$
$$(x_1, 0) \cdot (x_2, 0) = (x_1 x_2, 0)$$

a real number x can be identified with the complex number $(x, 0)$; in this sense we have $\mathbb{R} \subseteq \mathbb{C}$.

Defining

$$i := (0, 1),$$

we obtain

$$i^2 = (0, 1)(0, 1) = (-1, 0) = -1;$$

therefore in the field of complex numbers there is an element whose square is the negative of the unit element. Every complex number $z = (x, y)$ can – using the above identification – be written in the form

$$z = x + iy.$$

x is called the *real part* $\Re(z)$ and y the *imaginary part* $\Im(z)$ of z. Clearly two complex numbers are equal if and only if they have the same real and imaginary parts.

Every complex number $z = x+iy$ can be viewed as point in the Gaußian plane. The real part x is the projection of z to the x-axis and the imaginary part y the projection to the y-axis.

For every complex number $z = x + iy$ we define the *conjugated complex number* $\overline{z}$ durch $\overline{z} := x - iy$. In the Gaußian plane the conjugated complex number is obtained by mirroring at the x-axis. One can check easily that for all $z, z_1, z_2 \in \mathbb{C}$

$$\overline{\overline{z}} = z, \quad \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}.$$

Moreover for every $z \in \mathbb{C}$ we clearly have

$$\Re(z) = \frac{1}{2}(z + \overline{z}) \quad \text{and} \quad \Im(z) = \frac{1}{2i}(z - \overline{z}).$$

The *modulus* $|z|$ of a complex number z is defined by means of conjugated complex numbers; this will be useful for some of our later calculations. Let $z = x + iy$ with $x, y \in \mathbb{R}$. Then

$$z\overline{z} = (x + iy)(x - iy) = x^2 + y^2 \geq 0,$$

and we can define

$$|z| := \sqrt{z\overline{z}}.$$

Because of $|z| = \sqrt{x^2 + y^2}$ we can view $|z|$ as the distance of the point z in the Gaußian plane from the orign. Observe that for $z \in \mathbb{R}$ the modulus as defined for real numbers coincides with the modulus for complex numbers as we just defined it. Also we clearly have $|z| = |\overline{z}|$.

THEOREM 2.9.1. *For all $z, z_1, z_2 \in \mathbb{C}$ we have*
(a) $|z| \geq 0$, *and* $|z| = 0$ *iff* $z = 0$.
(b) $|z_1 z_2| = |z_1||z_2|$,
(c) $|z_1 + z_2| \leq |z_1| + |z_2|$ *(triangle inequality).*

PROOF. (a) is clear.
(b) $|z_1 z_2|^2 = z_1 z_2 \overline{z_1 z_2} = z_1 z_2 \overline{z_1}\,\overline{z_2} = z_1 \overline{z_1} z_2 \overline{z_2} = |z_1|^2 |z_2|^2$.
(c)

$$\begin{aligned}
|z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1} + \overline{z_2}) \\
&= z_1\overline{z_1} + z_1\overline{z_2} + z_2\overline{z_1} + z_2\overline{z_2} \\
&= |z_1|^2 + z_1\overline{z_2} + \overline{z_1\overline{z_2}} + |z_2|^2 \\
&= |z_1|^2 + 2\Re(z_1\overline{z_2}) + |z_2|^2 \\
&\leq (|z_1| + |z_2|)^2,
\end{aligned}$$

because of

$$\Re(z_1\overline{z_2}) \leq |z_1\overline{z_2}| = |z_1||\overline{z_2}| = |z_1||z_2|. \qquad \square$$

REMARK 2.9.2. A field with a modulus function satisfying the three properties of the theorem above is called a *valued field*. $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ are valued fields.

The notions and results above concerning the convergence of sequences and series can be carried over routinely from reals to complex numbers.

DEFINITION 2.9.3. A sequence $(c_n)_{n\in\mathbb{N}}$ of complex numbers is a *Cauchy sequence* with modulus $M\colon \mathbb{Z}^+ \to \mathbb{N}$ whenever $|c_m - c_n| \leq \frac{1}{2^p}$ for $m, n \geq M(p)$, and *converges* with modulus $M\colon \mathbb{Z}^+ \to \mathbb{N}$ to a complex number z, its *limit*, whenever $|c_n - z| \leq \frac{1}{2^p}$ for $n \geq M(p)$.

One can see easily that a sequence $(c_n)_{n\in\mathbb{N}}$ of complex numbers is a Cauchy sequence if and only if the two sequences of reals $(\Re(c_n))_{n\in\mathbb{N}}$ and $(\Im(c_n))_{n\in\mathbb{N}}$ are, and that it converges if and only if the two sequences of reals $(\Re(c_n))_{n\in\mathbb{N}}$ and $(\Im(c_n))_{n\in\mathbb{N}}$ converge. In this case we have

$$\lim_{n\to\infty} c_n = \lim_{n\to\infty} \Re(c_n) + i \lim_{n\to\infty} \Im(c_n).$$

THEOREM 2.9.4. *In $\mathbb{C}$ every Cauchy sequence converges.*

PROOF. The two sequences $(\Re(c_n))_{n\in\mathbb{N}}$ and $(\Im(c_n))_{n\in\mathbb{N}}$ are Cauchy sequences, hence converge in the reals. This implies the claim. $\square$

The treatment of the exponential series

$$\exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

can be carried over without any difficulty to the complex numbers. This also applies to the estimate of the rest, and the functional equation. As a consequence, we have $\exp(z) \neq 0$ for all $z \in \mathbb{C}$, because of $\exp(z)\exp(-z) = \exp(z-z) = \exp(0) = 1$. Notice also that

$$\exp(\overline{z}) = \overline{\exp(z)} \quad (z \in \mathbb{C}); \tag{6}$$

this follows from

$$\exp(\overline{z}) = \sum_{n=0}^{\infty} \frac{(\overline{z})^n}{n!} = \lim_{n\to\infty} \sum_{k=0}^{n} \overline{\left(\frac{z^k}{k!}\right)} = \lim_{n\to\infty} \overline{\left(\sum_{k=0}^{n} \frac{z^k}{k!}\right)} = \overline{\exp(z)}.$$

CHAPTER 3

Sets of real numbers

3.1. Intervals

For $x, y \in \mathbb{R}$ the *finite intervals* are defined by

$$\begin{aligned}
[x, y] &:= \{ z \in \mathbb{R} \mid x \leq z \leq y \}, \\
(x, y) &:= \{ z \in \mathbb{R} \mid x < z < y \}, \\
(x, y] &:= \{ z \in \mathbb{R} \mid x < z \leq y \}, \\
[x, y) &:= \{ z \in \mathbb{R} \mid x \leq z < y \}.
\end{aligned}$$

The interval $[x, y]$ is *closed*, (x, y) is *open*, $(x, y]$ is *half-open on the left*, $[x, y)$ is *half-open on the right*. We also allow the *infinite intervals*

$$\begin{aligned}
[x, \infty) &:= \{ z \in \mathbb{R} \mid x \leq z \}, \\
(x, \infty) &:= \{ z \in \mathbb{R} \mid x < z \}, \\
(\infty, y] &:= \{ z \in \mathbb{R} \mid z \leq y \}, \\
(\infty, y) &:= \{ z \in \mathbb{R} \mid z < y \}.
\end{aligned}$$

An inhabited, closed finite interval is called a *compact interval*. We use I, J to denote open (possibly infinite) intervals with rational end points.

3.2. Non-countability

Recall that every rational a is tacitly understood as the real represented by the constant sequence $a_n = a$ with the constant modulus $M(p) = 0$.

Lemma 3.2.1 ($\mathbb{Q}$ is dense in $\mathbb{R}$). *For any two reals $x < y$ there is a rational a such that $x < a < y$.*

Proof. Let $z := (x + y)/2$ be given by $(c_n)_n$. Then for some p we have $x <_p z <_p y$. Let $a := c_{M(p+1)}$, with M the Cauchy modulus of z. □

Notice that a depends on the representations of x and y.

Theorem 3.2.2 (Cantor). *Let a sequence (x_n) of reals be given. Then we can find a real y with $0 \leq y \leq 1$ that is apart from every x_n, in the sense that $x_n < y \vee y < x_n$.*

Proof. We construct sequences $(a_n)_n$, $(b_n)_n$ of rationals such that for all n

$$0 = a_0 \leq a_1 \leq \cdots \leq a_n < b_n \leq \cdots \leq b_1 \leq b_0 = 1, \tag{7}$$

$$x_n < a_{n+1} \vee b_{n+1} < x_n, \tag{8}$$

$$b_n - a_n \leq \frac{1}{2^n}. \tag{9}$$

Let $a_0, \dots, a_n$ and $b_0, \dots, b_n$ be already constructed such that (7)-(9) hold (as far as they are defined). Now compare the real x_n with $a_n < b_n$.

Case 1. $x_n < b_n$. Let $b_{n+1} := b_n$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, we can find a rational a_{n+1} such that

$$\max\bigl(x_n, a_n, b_n - \frac{1}{2^{n+1}}\bigr) < a_{n+1} < b_{n+1} = b_n.$$

Case 2. $a_n < x_n$. Let $a_{n+1} := a_n$, and find a rational b_{n+1} such that

$$a_n = a_{n+1} < b_{n+1} < \min\bigl(x_n, b_n, a_n + \frac{1}{2^{n+1}}\bigr).$$

Clearly (7)-(9) continue to hold for $n + 1$ (as far as defined). Now $y := (a_n)_n$ is a Cauchy sequence, since for $m \geq n$ we have $|a_m - a_n| = a_m - a_n \leq b_n - a_n \leq \frac{1}{2^n}$. Similarly $z := (b_n)_n$ is a Cauchy sequence. $y = z$ follows from (9), and from (8) together with (7) we obtain $x_n < y \vee z < x_n$. □

3.3. Supremum and infimum

Let S be a set of reals. A real y is an *upper bound* of S if $x \leq y$ for all $x \in S$. A real y is a *supremum* of S if y is an upper bound of S, and in addition for every rational $a < y$ there is real $x \in S$ such that $a \leq x$.

Every set S can have at most one supremum. To see this, assume that y, z are suprema of S. It is enough to show $y \leq z$, and for this it suffices to show $z \not< y$. So assume $z < y$. Then $z < a < y$ for some rational a, hence $a \leq x$ for some $x \in S$, contradicting the assumption that z is an upper bound of S. If the supremum of S exists, it is denoted by $\sup S$.

Definition 3.3.1. A set S of reals is *order located above* if for every $a < b$, either $x \leq b$ for all $x \in S$ or else $a \leq x$ for some $x \in S$.

Theorem 3.3.2 (Least-upper-bound principle). *Assume that S is an inhabited set of reals that is bounded above. Then S has a supremum if and only if it is order located above.*

Proof. If $\sup S$ exists and $a < b$, then either $\sup S < b$ or else $a < \sup S$. In the former case $x \leq b$ for all $x \in S$, and in the latter case clearly $a \leq x$ for some $x \in S$. Hence S is order located above.

For the converse it is useful to consider

$$\Pi_S(a,b): \qquad \text{both } y \le b \text{ for all } y \in S \text{ and } a \le x \text{ for some } x \in S$$

as a property of any pair a, b of rational numbers with $a < b$. By assumption we have $a, b \in \mathbb{Q}$ with $a < b$ such that $\Pi_S(a,b)$. We construct two sequences $(c_n)_n$ and $(d_n)_n$ of rationals such that for all n

$$a = c_0 \le c_1 \le \dots \le c_n < d_n \le \dots \le d_1 \le d_0 = b, \tag{10}$$

$$\Pi_S(c_n, d_n), \tag{11}$$

$$d_n - c_n \le \left(\frac{2}{3}\right)^n (b-a). \tag{12}$$

Let $c_0, \dots, c_n$ and $d_0, \dots, d_n$ be already constructed such that (10)-(12) hold. Let $c = c_n + \frac{1}{3}(d_n - c_n)$ and $d = c_n + \frac{2}{3}(d_n - c_n)$. Since S is order located above, either $s \le d$ for all $s \in S$ or else $c \le r$ for some $r \in S$. In the first case let $c_{n+1} := c_n$ and $d_{n+1} := d$, and in the second case let $c_{n+1} := c$ and $d_{n+1} := d_n$. Then clearly $\Pi_S(c_{n+1}, d_{n+1})$, (10) and (12) continue to hold for $n+1$, and the real number $x = y$ given by the modulated Cauchy sequences of rationals $(c_n)_n$ and $(d_n)_n$ is the least upper bound of S. □

A real y is a *lower bound* of S if $y \le x$ for all $x \in S$. A real y is an *infimum* of S if y is a lower bound of S, and in addition for every rational $a > y$ there is real $x \in S$ such that $a \ge x$. Clearly every set S can have at most one infimum. If the infimum of S exists, it is denoted by $\inf S$.

Definition 3.3.3. A set S or reals is *order located below* if for every $a < b$, either $x \le b$ for some $x \in S$ or else $a \le x$ for all $x \in S$.

Similarly to the least-upper-bound principle one proves

Theorem 3.3.4 (Greatest-lower-bound principle). *Assume that S is an inhabited set of reals that is bounded from below. Then S has an infimum if and only if it is order located below.*

The proofs given provide a reasonably fast algorithm to construct the supremum (or infimum), which makes use of the assumed order locatedness. If however we are only interested in the weak (or "classical") existence, then this assumption is not necessary.

Theorem 3.3.5 (Classical least-upper-bound principle). *Every non-empty bounded above (or below) set S of reals must have a supremum (or infimum).*

Proof. Let S be non-empty and y be an upper bound of S. Let $y_0 = y$ and x_0 an arbitrary element of S. We define two sequences $(x_n)_n$ and $(y_n)_n$ of reals such that for all n we have

(a) $x_0 \le x_1 \le \dots \le x_n \le y_n \le \dots \le y_1 \le y_0 = y$,

(b) $x_n \in S$ and y_n is an upper bound of S,
(c) $y_n - x_n \leq \frac{1}{2^n}(y_0 - x_0)$.

For $n = 0$ (a)-(c) are clearly satisfied. Assume that we already have (a)-(c) for n. Consider $z := \frac{1}{2}(x_n + y_n)$. We distinguish two cases.

Case 1. z is an upper bound of S. Let $y_{n+1} := z$ and $x_{n+1} := x_n$.

Case 2. z is not an upper bound of S. Then there must exist an $x \in S$ with $z < x$. Let x_{n+1} be such an x and $y_{n+1} := y_n$.

In both cases the validity of (a)-(c) is clear. Therefore (x_n) is a monotonically increasing bounded sequence of reals and hence by Theorem 2.5.3 must have a limit. Similarly (y_n) is a monotonically decreasing bounded sequence of reals and hence again by Theorem 2.5.3 must have a limit. By (c) both limits are equal, say $= z$. We still need to show that z satisfies the properties of a supremum of S.

z is an upper bound of S: let $x \in S$. Then $x \leq y_n$ for all n, hence also $x \leq \lim_n y_n$.

Assume $a < z$. Then $a \leq x_n$ for some n, and $x_n \in S$. □

CHAPTER 4

Continuous functions

4.1. Basic definitions

We consider real-valued functions defined on open, closed or half-open intervals $I \subseteq \mathbb{R}$. Let c^∞, d^∞ range over $\mathbb{R} \cup \{\pm\infty\}$.

Definition 4.1.1. A *uniformly continuous function* $f \colon I \to \mathbb{R}$ is given by

$$h \colon (I \cap \mathbb{Q}) \to \mathbb{N} \to \mathbb{Q} \qquad \text{(called approximating map)},$$

together with further data:

(a) A map $\alpha \colon \mathbb{Z}^+ \to \mathbb{N}$ such that for $a \in I$ each $(h(a,n))_n$ is a Cauchy sequence with (uniform) modulus α.

(b) A modulus $\omega \colon \mathbb{Z}^+ \to \mathbb{Z}^+$ of (uniform) continuity, such that $\omega(p)$ satisfies for all $a, b \in I$

$$|a - b| \leq \frac{1}{2^{\omega(p)-1}} \to |h(a,n) - h(b,n)| \leq \frac{1}{2^p} \quad \text{for } n \geq \alpha(p).$$

(c) Lower and upper bounds $\mu, \nu \in \mathbb{Q}$ for all $h(a,n)$ with $a \in I$.

We require $\alpha(p) \leq \alpha(q)$ and $\omega(p) \leq \omega(q)$ for all $p \leq q$.

Definition 4.1.2. A *continuous function* $g \colon (c^\infty, d^\infty) \to \mathbb{R}$ is given by an approximating map $h \colon ((c^\infty, d^\infty) \cap \mathbb{Q}) \to \mathbb{N} \to \mathbb{Q}$ and a family of uniformly continuous functions $(h{\restriction}[c,d], \alpha_{c.d}, \omega_{c.d}, \mu_{c.d}, \nu_{c.d})$ for $c^\infty \leq c < d \leq d^\infty$. We require for $c^\infty < c' \leq c < d \leq d' < d^\infty$ the monotonicity properties

$$\alpha_{c,d}(p) \leq \alpha_{c',d'}(p), \quad \omega_{c,d}(p) \leq \omega_{c',d'}(p), \quad \mu_{c',d'} \leq \mu_{c,d}, \quad \nu_{c,d} \leq \nu_{c',d'}.$$

Example 4.1.3 (Squaring). $\mathsf{sq} \colon (-\infty, \infty) \to \mathbb{R}$ is a continuous function given by

(a) the approximating map $h(a,n) := a^2$ and modulus $\alpha_{c,d}(p) := 0$;

(b) the modulus $\omega_{c,d}(p) := p + q + 1$ of uniform continuity, where q is such that $|a + b| \leq 2^q$ for $c \leq a < b \leq d$, because

$$|a - b| \leq \frac{1}{2^{p+q}} \to |a^2 - b^2| = |(a-b)(a+b)| \leq \frac{1}{2^q};$$

(c) the lower bound $\mu_{c,d} := c^2$ and upper bound $\nu_{c,d} := d^2$.

Similarly all polynomials with rational coefficients on finite intervals can be viewed as continuous functions in our sense.

Example 4.1.4 (Inverse). $\mathsf{inv}\colon (0, \infty) \to \mathbb{R}$ inverting its argument is a continuous function given by the approximating map $h(a, n) := \frac{1}{a}$. For every compact interval $[\frac{1}{2^q}, d]$ we have

(a) the Cauchy modulus $\alpha_{c,d}(p) := 0$;

(b) the modulus $\omega_{c,d}(p) := p + 2q + 1$ of uniform continuity, for

$$|a - b| \le \frac{1}{2^{p+2q}} \to \left|\frac{1}{a} - \frac{1}{b}\right| = \left|\frac{b-a}{ab}\right| \le \frac{1}{2^p},$$

because $ab \ge \frac{1}{2^{2q}}$;

(c) the lower bound $\mu_{c,d} := 1/d$ and upper bound $\nu_{c,d} := 2^q$.

Example 4.1.5 (Square root). The *square root function* differs from the previous ones in that the values on rational numbers will not be rationals any more. Given $a > 0$ and – for definiteness – $a_0 := 1$, recall from Theorem 1.1.2 of Chapter 1 that we can approximate $\sqrt{a}$ by

$$a_{n+1} := \frac{1}{2}\Big(a_n + \frac{a}{a_n}\Big).$$

One can verify easily that $\min(a, 1) \le a_n \le \max(a, 1)$, for all n. Hence the square root function on (c, d) $(0 < c < d)$ is a continuous function given by

(a) the approximating map $h(a, n) := a_n$;

(b) a modulus α, which can easily be computed from the fact (established in the proof of Theorem 1.1.2) that $|a_{n+1} - a_{m+1}| \le (a_1 - a'_1)/2^n$ for $n \le m$, with $a_1 = (1 + a)/2$ and $a'_1 = a/a_1$;

(c) the modulus of uniform continuity can be obtained from

$$\big|\sqrt{a} - \sqrt{b}\big| \le \frac{1}{\sqrt{a} + \sqrt{b}}|a - b|,$$

because $\sqrt{a} + \sqrt{b} \ge 2\min(c, 1)$;

(d) the lower bound $\mu_{c,d} := \min(c, 1)$ and upper bound $\nu_{c,d} := \max(d, 1)$.

In more detail, the argument for the modulus of uniform continuity runs as follows. Let $\sqrt{a} + \sqrt{b} \ge 2\min(c, 1) \ge \frac{1}{2^q}$. Assume $|a - b| \le \frac{1}{2^{p+q+1}}$. Then

$$\begin{aligned} |a_{n+1} - b_{n+1}| &\le |a_{n+1} - \sqrt{a}| + |\sqrt{a} - \sqrt{b}| + |\sqrt{b} - b_{n+1}| \\ &\le |a_{n+1} - a'_{n+1}| + \frac{1}{\sqrt{a} + \sqrt{b}}|a - b| + |b'_{n+1} - b_{n+1}| \\ &\le \frac{1}{2^{p+2}} + \frac{1}{2^{p+1}} + \frac{1}{2^{p+2}} = \frac{1}{2^p} \end{aligned}$$

provided n is such that $|a_{n+1} - a'_{n+1}| \le \frac{1}{2^n}(a_1 - a'_1) \le \frac{1}{2^{p+2}}$, and similarly for b. This can be achieved by choosing the Cauchy modulus α large enough.

Example 4.1.6 (Exponential). Our final example is the *exponential function* $\exp\colon (-\infty,\infty) \to \mathbb{R}$ given by

(a) the approximating map

$$h(a,n) := \sum_{k=0}^{n} \frac{a^k}{k!};$$

(b) a uniform Cauchy modulus $\alpha_{c,d}$, which can easily be computed from Theorem 2.8.2 (Estimate of the rest) of Chapter 2:

$$\Big|\sum_{k=0}^{n} \frac{a^k}{k!} - \sum_{k=0}^{m} \frac{a^k}{k!}\Big| = \Big|\sum_{k=n+1}^{m} \frac{a^k}{k!}\Big| \le \sum_{k=n+1}^{\infty} \frac{|a|^k}{k!} \le 2\frac{|a|^{n+1}}{(n+1)!}$$

for $|a| \le 1 + \frac{n}{2}$ and $n \le m$;

(c) the modulus $\omega_{c,d}$ of uniform continuity, which can be obtained from

$$\begin{aligned}\Big|\sum_{k=0}^{n} \frac{a^k}{k!} - \sum_{k=0}^{n} \frac{b^k}{k!}\Big| &= \Big|\sum_{k=1}^{n} \frac{a^k - b^k}{k!}\Big| = |a-b| \sum_{k=1}^{n} \frac{1}{k!}\Big|\sum_{l=0}^{k-1} a^{k-1-l} b^l\Big| \\ &\le |a-b| \sum_{k=1}^{n} \frac{kM^{k-1}}{k!} = |a-b| \sum_{k=0}^{n-1} \frac{M^k}{k!} < |a-b| \exp(M),\end{aligned}$$

where $M = \max(|c|,|d|)$;

(d) the lower bound $\mu_{c,d} := \sum_{k=0}^{\lceil M \rceil} \frac{M^k}{k!}$ where again $M := \max(|c|,|d|)$, and upper bound

$$\nu_{c,d} := \sum_{k=0}^{L} \frac{d^k}{k!} + 2\frac{|d|^{L+1}}{(L+1)!} \quad \text{with } L := 2\lceil |d| \rceil;$$

this can easily be verified, again using Theorem 2.8.2 of Chapter 2.

Definition 4.1.7 (Localization $g{\restriction}p$ of a continuous function). Let a continuous function g be given by $c^\infty, d^\infty, h, \alpha^\infty, \omega^\infty, \mu^\infty, \nu^\infty$. For $p \in \mathbb{Z}^+$ we define $[c,d]$ by

$$[c,d] := \begin{cases} [-2^p, 2^p] & \text{if } c^\infty = -\infty \text{ and } d^\infty = +\infty \\ [c^\infty + \frac{1}{2^p}, 2^p] & \text{if } c^\infty \in \mathbb{R} \text{ and } d^\infty = +\infty \\ [-2^p, d^\infty - \frac{1}{2^p}] & \text{if } c^\infty = -\infty \text{ and } d^\infty \in \mathbb{R} \\ [c^\infty + \frac{1}{2^p}, d^\infty - \frac{1}{2^p}] & \text{if } c^\infty, d^\infty \in \mathbb{R} \end{cases}$$

provided p is large enough to make $[c,d]$ a proper interval. The *localization* $g{\restriction}p$ is defined to consist of c,d as above, $h{\restriction}[c,d]$ and

$$\alpha(p) := \alpha^\infty_{c,d}(p), \quad \omega(p) := \omega^\infty_{c,d}(p), \quad \mu(p) := \mu^\infty_{c,d}(p), \quad \nu(p) := \nu^\infty_{c,d}(p).$$

Since the approximating map operates on rationals only, we need to define separately what it means to apply a continuous function in our sense to a real. It suffices to do this for uniformly continuous functions. For continuous ones g we in addition need a witness p for elementhood of the argument x in g's open interval (c^∞, d^∞), i.e., (c, d) (depending on p, as above) such that $x \in (c, d)$. Then we can define the restriction $g{\restriction}p$ as a uniformly continuous function, and use application $(g{\restriction}p)(x)$. Notice that $(g{\restriction}p)(x)$ does not depend on p, since the approximating map h of g is independent of interval bounds c, d, and by Remark 1.2.4 of Chapter 1 two real numbers are equal if their Cauchy sequences coincide from one point onwards.

Definition 4.1.8 (Application). Let $f\colon [c, d] \to \mathbb{R}$ be a uniformly continuous function given by $c, d, h, \alpha, \omega, \mu, \nu$. Let further $x = ((a_n)_n, M)$ be an arbitrary real. The *application* $f(x)$ of f to x is defined to be the Cauchy sequence $(h(\pi_{c,d}(a_n), n))_n$ with modulus

$$\lambda_p \max(\alpha(p + 1), M(\omega(p + 1) - 1)).$$

Here the projection $\pi_{c,d}$ is defined by

$$\pi_{c,d}(a) := \begin{cases} c & \text{if } a < c, \\ a & \text{if } c \leq a \leq d, \\ d & \text{if } d < a. \end{cases}$$

Lemma 4.1.9 (ContReal). *This is a modulus.*

Proof. We write a' for $\pi_{c,d}(a)$. Under the assumptions of the definition we have

$$\begin{aligned} |h(a'_n, n) - h(a'_m, m)| &\leq |h(a'_n, n) - h(a'_n, m)| + |h(a'_n, m) - h(a'_m, m)| \\ &\leq \frac{1}{2^{p+1}} + \frac{1}{2^{p+1}} \end{aligned}$$

if $n, m \geq \alpha(p+1)$ (this gives the first estimate) and $n, m \geq M(\omega(p+1)-1)$ (this gives the second estimate). To see the latter observe that because of $a'_n, a'_m \in [c, d]$ and $m \geq \alpha(p + 1)$ it suffices to prove

$$|a'_n - a'_m| \leq \frac{1}{2^{\omega(p+1)-1}}.$$

Because of $n, m \geq M(\omega(p + 1) - 1)$ we have

$$|a_n - a_m| \leq \frac{1}{2^{\omega(p+1)-1}}.$$

The claim now follows from $|a' - b'| \leq |a - b|$, which is easy to see. □

LEMMA 4.1.10 (ContAppCompat). *Let $f\colon [c,d] \to \mathbb{R}$ be a local continuous function given by $c, d, h, \alpha, \omega, \mu, \nu$, and $x, y \in [c,d]$. Then*

$$x = y \to f(x) = f(y).$$

PROOF. To prove $f(x) = f(y)$ we use Lemma 1.2.3 (RealEqChar) of Chapter 1. We again write a' for $\pi_{c,d}(a)$. Given p, it suffices to prove that

$$|h(a'_n, n) - h(b'_n, n)| \le \frac{1}{2^p} \qquad \text{holds finally.}$$

We apply Lemma 1.2.3 again (for $\omega(p) - 1$), this time to make use of our assumption $x = y$. It gives us an n_0 such that for $n \ge n_0$ we have

$$|a_n - b_n| \le \frac{1}{2^{\omega(p)-1}}$$

and hence also (as in the proof of Lemma 4.1.9 above)

$$|a'_n - b'_n| \le \frac{1}{2^{\omega(p)-1}}.$$

If in adddition we take $n \ge \omega(p)$ the claim follows. □

Next we show that indeed a continuous function f has ω as a modulus of uniform continuity.

LEMMA 4.1.11 (ContMod). *Let $f\colon [c,d] \to \mathbb{R}$ be a local continuous function given by $c, d, h, \alpha, \omega, \mu, \nu$, and $x, y \in [c,d]$. Then*

$$|x - y| \le \frac{1}{2^{\omega(p)}} \to |f(x) - f(y)| \le \frac{1}{2^p}.$$

PROOF. Assume $|a_n - b_n| \le \frac{1}{2^{\omega(p)-1}}$ for $n \ge n_0$. Then also $|a'_n - b'_n| \le \frac{1}{2^{\omega(p)-1}}$ for $a' := \pi_{c,d}(a)$. Hence for $n \ge n_0, \alpha(c,d,p)$

$$|h(a'_n, n) - h(b'_n, n)| \le \frac{1}{2^p},$$

that is $|f(x) - f(y)| \le \frac{1}{2^p}$. □

We define the *composition* of two continuous functions.

DEFINITION 4.1.12 (Composition). Assume that $f\colon [c,d] \to \mathbb{R}$ is given by $c, d, h, \alpha, \omega, \mu, \nu$ and $f'\colon [c', d'] \to \mathbb{R}$ by $c', d', h', \alpha', \omega', \mu', \nu'$. Assume further $c' \le \mu \le \nu \le d'$. Then $f' \circ f\colon [c,d] \to \mathbb{R}$ is defined by

$$\begin{aligned}
h^\circ(a,n) &:= h'(h(a,n), n) \\
\alpha^\circ(p) &:= \max(\alpha'(p+1), \alpha(\omega'(p+1) - 1)) \\
\omega^\circ(p) &:= \omega(\omega'(p) - 1) \\
\mu^\circ &:= \mu' \\
\nu^\circ &:= \nu'.
\end{aligned}$$

We need to show that this indeed defines a continuous function.

LEMMA 4.1.13 (ContComposeCorr). *Under the assumptions of the definition above we have*

(a) *Each $h^\circ(\pi_{c,d}(a), n)$ is a Cauchy sequence with (uniform) modulus $\alpha^\circ(p)$;*
(b) *$\omega^\circ(p)$ satisfies for all $a, b \in [c, d]$*

$$|a-b| \le \frac{1}{2^{\omega^\circ(p)-1}} \to |h(a,n) - h(b,n)| \le \frac{1}{2^p} \quad \textit{for } n \ge \alpha^\circ(p);$$

(c) *$[\mu^\circ, \nu^\circ]$ contains all $h^\circ(a, n)$ for $a \in [c, d]$.*

PROOF. Write a' for $\pi_{c,d}(a)$. (a). $(h(a', n)))_n$ is a real number with modulus α. By Lemma 4.1.9 application of h' to this real gives the Cauchy sequence $(h'(\pi_{c',d'}(h(a', n)), n))_n$ with $\alpha^\circ(p) = \max(\alpha'(p+1), \alpha(\omega(p+1)-1))$ its Cauchy modulus.

(b). Assume

$$|a-b| \le \frac{1}{2^{\omega^\circ(p)-1}} = \frac{1}{2^{\omega(\omega'(p)-1)-1}}$$

Then also $|a' - b'|$ satisfies this inequality and we obtain

$$|h(a', n) - h(b', n)| \le \frac{1}{2^{\omega'(p)-1}} \quad \text{if } n \ge \alpha(\omega'(p) - 1)$$

$$|h'(h(a', n), n) - h'(h(b', n), n)| \le \frac{1}{2^p} \quad \text{if } n \ge \alpha'(p).$$

Both conditions follow from $n \ge \alpha^\circ(p)$ by the monotonicity properties.

(c). Obvious. □

REMARK 4.1.14. Under the assumptions of the definition above we clearly have $(f' \circ f)(x) = f'(f(x))$ for all $x \in I$, since both reals have the same Cauchy sequence.

4.2. Properties of continuous functions

We show that continuous functions commute with limits.

LEMMA 4.2.1 (ContLim). *Let $(x_n)_n$ be a sequence of reals which converges to y. Assume $x_n, y \in I$ and let $f \colon I \to \mathbb{R}$ be continuous. Then $(f(x_n))_n$ converges to $f(y)$.*

PROOF. For a given p, pick n_0 such that for all n

$$n_0 \le n \to |x_n - y| \le \frac{1}{2^{\omega_f(p)}}.$$

Then by Lemma 4.1.11 (ContMod)

$$n_0 \le n \to |f(x_n) - f(y)| \le \frac{1}{2^p}.$$

Hence $(f(x_n))_n$ converges to $f(y)$. □

Lemma 4.2.2 (ContRat). *Assume that $f, g\colon I \to \mathbb{R}$ are continuous and coincide on all rationals $a \in I$. Then $f = g$.*

Proof. Let $x = ((a_n)_n, M)$. By Lemma 4.2.1 (ContdLim) the sequence $(f(a_n))_n$ converges to $f(x)$ and $(g(a_n))_n$ to $g(x)$. Now $f(a_n) = g(a_n)$ implies $f(x) = g(x)$. □

The supremum of the range of a continuous function on a finite interval can be shown to exist constructively.[1] We prove that the range is order located above, which entails (by the least-upper-bound principle) that it has a supremum.

Lemma 4.2.3. *Let $f\colon I \to \mathbb{R}$ be continuous and a finite subinterval J of I given by $c, d \in I$ with $c < d$. Then the range of f on J is order located above.*

Proof. Assume $a < b$. We show $\forall_{x \in J}(f(x) \le b) \vee \exists_{x \in J}(a \le f(x))$.

Let h be the approximating map for f, and α_f and ω_f be the moduli among the data for f. For the given c, d let $\alpha(p) := \alpha_f(c, d, p)$ and $\omega(p) := \omega_f(c, d, p)$. Fix p such that $\frac{1}{2^p} \le \frac{1}{3}(b - a)$. Take a partition $a_0, \dots, a_l$ of $[c, d]$ of mesh $\le \frac{1}{2^{\omega(p)-2}}$. Then for every a with $c < a < d$ there is an i such that $|a - a_i| \le \frac{1}{2^{\omega(p)-1}}$. Let $n_p := \alpha(p)$ and consider all finitely many

$$h(a_i, n_p) \quad \text{for } i = 0, \dots, l.$$

Let $h(a_j, n_p)$ be the maximum of all those.

Case $h(a_j, n_p) \le a + \frac{1}{3}(b - a)$. We show that $f(x) \le b$ for all x. Let $x = ((b_n)_n, M)$. Then for $n \ge n_p$

$$\begin{aligned} h(b_n, n) &\le h(b_n, n_p) + \frac{1}{2^p} \\ &\le h(a_i, n_p) + \frac{1}{2^{p-1}} \quad \text{for } i \text{ such that } |b_n - a_i| \le \tfrac{1}{2^{\omega(p)-1}} \\ &\le h(a_j, n_p) + \frac{1}{2^{p-1}} \\ &\le a + \frac{1}{3}(b - a) + \frac{2}{3}(b - a) = b. \end{aligned}$$

Case $a + \frac{1}{3}(b - a) < h(a_j, n_p)$. We show $a \le f(x)$ for $x := a_j$. Then $f(x)$ is given by the Cauchy sequence $(h(a_j, n))_n$. We have for $n \ge n_p$

$$h(a_j, n) \ge h(a_j, n_p) - \frac{1}{2^p} \ge a + \frac{1}{3}(b - a) - \frac{1}{2^p} \ge a.$$

Hence $a \le f(x)$. □

[1] This is proved in Bishop and Bridges (1985), using the notion of a “totally bounded” set. However, the latter is a type-level 2 concept, which we wish to avoid.

Corollary 4.2.4. *Let $f\colon I \to \mathbb{R}$ be continuous and a finite subinterval J of I given by $c, d \in I$ with $c < d$. Then the range of f on J has a supremum, denoted $\|f\|_J$.*

Proof. The range of f is bounded above, and by the last lemma it is order located above. Hence by Theorem 3.3.2 of Chapter 3 (Least-upper-bound principle) it has a supremum. □

Theorem 4.2.5. *Let $f\colon I \to \mathbb{R}$ be continuous and a finite subinterval J of I given by $c, d \in I$ with $c < d$. Then there must be a real $c < x < d$ such that $f(x) = \|f\|_J$.*

Proof. Consider the range of f. Rest to do. □

4.3. Intermediate value theorem

We next supply the standard constructive versions of the *intermediate value theorem*.

Theorem 4.3.1 (Approximate intermediate value theorem). *Let $f\colon I \to \mathbb{R}$ be continuous and $a < b$ rational numbers in I such that $f(a) \le 0 \le f(b)$. Then for every p we can find c with $a < c < b$ such that $|f(c)| \le \frac{1}{2^p}$.*

Proof. In the sequel we repeatedly invoke Lemma 1.6.6 (ApproxSplit) of Chapter 1. Given p, let $\varepsilon := \frac{1}{2^p}$. We compare $f(a)$ and $f(b)$ with $-\varepsilon < -\frac{\varepsilon}{2}$ and $\frac{\varepsilon}{2} < \varepsilon$, respectively. If $-\varepsilon < f(a)$ or $f(b) < \varepsilon$, then $|f(c)| < \varepsilon$ for $c = a$ or $c = b$; whence we may assume that

$$f(a) < -\frac{\varepsilon}{2} \quad \text{and} \quad \frac{\varepsilon}{2} < f(b).$$

Now pick q so that, for all $x, y \in [a, b]$, if $|x - y| \le \frac{1}{2^q}$, then $|f(x) - f(y)| \le \varepsilon$, and divide $[a, b]$ into $a = a_0 < a_1 < \cdots < a_m = b$ such that $|a_{i-1} - a_i| \le \frac{1}{2^q}$. Compare every $f(a_i)$ with $-\frac{\varepsilon}{2} < \frac{\varepsilon}{2}$. By assumption $f(a_0) < -\frac{\varepsilon}{2}$ and $\frac{\varepsilon}{2} < f(a_m)$; whence we can find j minimal such that

$$f(a_j) < \frac{\varepsilon}{2} \quad \text{and} \quad -\frac{\varepsilon}{2} < f(a_{j+1}).$$

Finally, compare $f(a_j)$ with $-\varepsilon < -\frac{\varepsilon}{2}$ and $f(a_{j+1})$ with $\frac{\varepsilon}{2} < \varepsilon$. If $-\varepsilon < f(a_j)$, we have $|f(a_j)| < \varepsilon$. If $f(a_{j+1}) < \varepsilon$, we have $|f(a_{j+1})| < \varepsilon$. If both $f(a_j) < -\frac{\varepsilon}{2}$ and $\frac{\varepsilon}{2} < f(a_{j+1})$, then we would have $|f(a_{j+1}) - f(a_j)| > \varepsilon$, contradicting $|a_{j+1} - a_j| \le \frac{1}{2^q}$. □

Alternative Proof. We give a different proof, which more directly makes use of the fact that our continuous functions come with witnessing data.

We may assume $f(a) < -\frac{1}{2^{p+1}}$ and $\frac{1}{2^{p+1}} < f(b)$ (see above). Divide $[a,b]$ into $a = a_0 < a_1 < \cdots < a_m = b$ such that $|a_{i-1} - a_i| \le \frac{1}{2^{\omega_f(p+1)}}$. Consider all finitely many

$$h(a_i, n_0) \quad \text{for } i = 1, \dots, m,$$

with $n_0 := \alpha_f(p+1)$. Pick j such that $h(a_{j-1}, n_0) \le 0 \le h(a_j, n_0)$; this can be done because $f(a) < -\frac{1}{2^{p+1}}$ and $\frac{1}{2^{p+1}} < f(b)$. We show $|f(a_j)| \le \frac{1}{2^p}$; for this it clearly suffices to show $|h(a_j, n)| \le \frac{1}{2^p}$ for $n \ge n_0$. Now

$$|h(a_j, n)| \le \bigl|h(a_j, n) - h(a_j, n_0)\bigr| + \bigl|h(a_j, n_0)\bigr| \le \frac{1}{2^{p+1}} + \frac{1}{2^{p+1}},$$

where the first estimate holds by the choice of n_0, and the second one follows from the choice of a_j and $|h(a_{i-1}, n) - h(a_i, n)| \le \frac{1}{2^{p+1}}$. □

For later use we prove a somewhat stronger form of the intermediate value theorem, where we pick the "last" approximate zero of the given function.

Theorem 4.3.2 (LastApproxZero). *Let $f\colon I \to \mathbb{R}$ be continuous and $a < b$ rational numbers in I such that $f(a) \le 0 \le f(b)$. Then for every p we can find c with $a \le c \le b$ such that $f(c) \le \frac{1}{2^p}$ and $0 \le f(z)$ for all $z \in [c,b]$.*

Proof. Let $4\varepsilon = \frac{1}{2^p}$, i.e., $\varepsilon := \frac{1}{2^{p+2}}$. Divide $[a,b]$ into $a = a_0 < a_1 < \cdots < a_m = b$ such that $|a_{i-1} - a_i| \le \frac{1}{2^{\omega_f(p+2)}}$. Consider all finitely many

$$h(a_i, n_0) \quad \text{for } i = 1, \dots, m,$$

with $n_0 := \alpha_f(p+2)$. Pick j such that $h(a_{j-1}, n_0) \le 2\varepsilon$ and $2\varepsilon \le h(a_i, n_0)$ for $j \le i \le m$; we may take the largest such j. Then for $n \ge n_0$

$$|h(a_{j-1}, n)| \le \bigl|h(a_{j-1}, n) - h(a_{j-1}, n_0)\bigr| + \bigl|h(a_{j-1}, n_0)\bigr| \le \varepsilon + 2\varepsilon,$$

where the first estimate holds by the choice of n_0, and the second one by the choice of j. Similarly for $j \le i \le m$ and $n \ge n_0$

$$h(a_i, n) \ge h(a_i, n_0) - \bigl|h(a_i, n) - h(a_i, n_0)\bigr| \ge 2\varepsilon - \varepsilon,$$

where the first estimate holds by the choice of j since $j \le i \le m$, and the second one by the choice of n_0. Let $c := a_{j-1}$. Then $f(c) \le 3\varepsilon < 4\varepsilon = \frac{1}{2^p}$, and for $z \in [c,b]$ we have an i such that $|z - a_i| \le \frac{1}{2^{\omega_f(p+2)}}$, hence $|f(z) - f(a_i)| \le \frac{1}{2^{p+2}} = \varepsilon$. Since $\varepsilon \le f(a_i)$ we obtain $0 \le f(z)$. □

A problem with all three of these proofs is that the algorithms they provide are rather bad: in each case one has to partition the interval into as many pieces as the modulus of the continuous function requires for the given error bound, and then for each of these (many) pieces perform certain operations. This problem seems to be unavoidable, since our continuous function may be rather flat. However, we can do somewhat better if we assume a uniform *modulus of increase* (or lower bound on the slope) of f, that is, some $q \in \mathbb{Z}^+$ such that for all $c, d \in \mathbb{Q}$ and all $p \in \mathbb{Z}^+$

$$\frac{1}{2^p} \leq d - c \to \frac{1}{2^{p+q}} \leq f(d) - f(c).$$

We begin with an auxiliary lemma, which from a "correct" interval $c < d$ (that is, $f(c) \leq 0 \leq f(d)$ and $\frac{1}{2^p} \leq d - c$) constructs a new one $c_1 < d_1$ with $d_1 - c_1 = \frac{2}{3}(d - c)$.

Lemma 4.3.3 (IVTAux). *Let $f\colon I \to \mathbb{R}$ be continuous, with a uniform modulus q of increase. Let $a < b$ be rational numbers in I such that $a \leq c < d \leq b$, say $\frac{1}{2^p} < d - c$, and $f(c) \leq 0 \leq f(d)$. Then we can construct c_1, d_1 with $d_1 - c_1 = \frac{2}{3}(d - c)$, such that again $a \leq c \leq c_1 < d_1 \leq d \leq b$ and $f(c_1) \leq 0 \leq f(d_1)$.*

Proof. Let $c_0 = \frac{2c+d}{3}$ and $d_0 = \frac{c+2d}{3}$. From $\frac{1}{2^p} < d - c$ we obtain $\frac{1}{p+2} \leq d_0 - c_0$, so $f(c_0) <_{p+2+q} f(d_0)$. Now compare 0 with this proper interval, using ApproxSplit. In the first case we have $0 \leq f(d_0)$; then let $c_1 = c$ and $d_1 = d_0$. In the second case we have $f(c_0) \leq 0$; then let $c_1 = c_0$ and $d_1 = d$. □

Theorem 4.3.4 (IVT). *Let $f\colon I \to \mathbb{R}$ be continuous, with a uniform modulus of increase. Let $a < b$ be rational numbers in I such that $f(a) \leq 0 \leq f(b)$. Then we can find $x \in [a, b]$ such that $f(x) = 0$.*

Proof. Iterating the construction in Lemma 4.3.3 (IVTAux), we construct two sequences $(c_n)_n$ and $(d_n)_n$ of rationals such that for all n

$$\begin{aligned}
&a = c_0 \leq c_1 \leq \cdots \leq c_n < d_n \leq \cdots \leq d_1 \leq d_0 = b,\\
&f(c_n) \leq 0 \leq f(d_n),\\
&d_n - c_n = \big(\frac{2}{3}\big)^n (b - a).
\end{aligned}$$

Let x, y be given by the Cauchy sequences $(c_n)_n$ and $(d_n)_n$ with the obvious modulus. As f is continuous, $f(x) = 0 = f(y)$ for the real number $x = y$. □

Remark 4.3.5. The proposition can also be proved for locally nonconstant functions. A continuous function $f\colon I \to \mathbb{R}$ is *locally nonconstant* if whenever $a < b$ are in I and c is an arbitrary real, then $f(x) \neq c$ for some real $x \in (a, b)$. Note that there is also a rational with that property. Strictly monotone functions are clearly locally nonconstant, and so are nonconstant real polynomials.

From the Intermediate Value Theorem we obtain

THEOREM 4.3.6 (Inv). *Let $f\colon I \to \mathbb{R}$ be continuous, with a uniform modulus of increase. Let $a < b$ be rationals in I such that $f(a) < f(b)$. We can find a continuous $g\colon (f(a), f(b)) \to \mathbb{R}$ such that $f(g(y)) = y$ for all $y \in (f(a), f(b))$ and $g(f(x)) = x$ for all $x \in [a, b]$ such that $f(a) \le f(x) \le f(b)$.*

PROOF. By assumption we have some $q \in \mathbb{Z}^+$ such that for all $c, d \in [a, b]$ and all $p \in \mathbb{Z}^+$

$$\frac{1}{2^p} \le d - c \to \frac{1}{2^{p+q}} \le f(c) - f(d).$$

We construct a continuous $g\colon (f(a), f(b)) \to \mathbb{R}$.

Let $u \in (a', b') \subseteq (f(a), f(b))$ be rational. Using $f(a) - u \le a' - u \le 0$ and $0 \le b' - u \le f(b) - u$, Theorem 4.3.4 (IVT) gives us an x such that $f(x) - u = 0$, as a Cauchy sequence (c_n). Let $h_g(u, n) := c_n$. Define the modulus α_g such that for $n \ge \alpha_g(p)$, $(\frac{2}{3})^n(b - a) \le \frac{1}{2^{\omega_f(p+q+2)}}$. For the uniform modulus ω_g of continuity assume $a' \le u < v \le b'$ and $p \in \mathbb{Z}^+$. We claim that with $\omega_g(p) := p + q + 2$ (q from the hypothesis on the slope) we can prove the required property

$$|u - v| \le \frac{1}{2^{\omega_g(p)+1}} \to |h_g(u, n) - h_g(v, n)| \le \frac{1}{2^p} \quad (n \ge \alpha_g(p)).$$

Let $a' \le u < v \le b'$ and $n \ge \alpha_g(p)$. For $c_n^{(u)} := h_g(u, n)$ and $c_n^{(v)} := h_g(v, n)$ assume that $|c_n^{(u)} - c_n^{(v)}| > \frac{1}{2^p}$; we must show $|u - v| > \frac{1}{2^{\omega_g(p)+1}}$.

By the proof of the Intermediate Value Theorem we have

$$d_n^{(u)} - c_n^{(u)} \le \big(\frac{2}{3}\big)^n (b - a) \le \frac{1}{2^{\omega_f(p+q+2)}} \quad \text{for } n \ge \alpha_g(p).$$

Using $f(c_n^{(u)}) - u \le 0 \le f(d_n^{(u)}) - u$, the fact that a continuous function f has ω_f as a modulus of uniform continuity gives us

$$|f(c_n^{(u)}) - u| \le |(f(d_n^{(u)}) - u) - (f(c_n^{(u)}) - u)| = |f(d_n^{(u)}) - f(c_n^{(u)})| \le \frac{1}{2^{p+q+2}}$$

and similarly $|f(c_n^{(v)}) - v| \le \frac{1}{2^{p+q+2}}$. Hence, using $|f(c_n^{(u)}) - f(c_n^{(v)})| \ge \frac{1}{2^{p+q}}$ (which follows from $|c_n^{(u)} - c_n^{(v)}| > \frac{1}{2^p}$ by the hypothesis on the slope),

$$|u - v| \ge |f(c_n^{(u)}) - f(c_n^{(v)})| - |f(c_n^{(u)}) - u| - |f(c_n^{(v)}) - v| \ge \frac{1}{2^{p+q+1}}.$$

To prove $f(g(u)) = u$ it suffices to show

$$|f(g(u)) - u| \le |f(g(u)) - f(h_g(u, n))| + |f(h_g(u, n)) - u| \le \frac{1}{2^p}$$

for sufficiently big n. The first term is $\leq \frac{1}{2^{p+1}}$ for $n \geq \alpha_g(\omega_f(p+1))$, since then

$$\begin{aligned} |g(u) - h_g(u,n)| &\leq \frac{1}{2^{\omega_f(p+1)}} && \text{by RealCauchyConvMod} \\ |f(g(u)) - f(h_g(u,n))| &\leq \frac{1}{2^{p+1}} && \text{by ConvMod} \end{aligned}$$

The second term is $\leq \frac{1}{2^{p+1}}$ for $n \geq M(\omega_f(p))$, where $M(p) := \log_{2/3} \frac{1}{2^{p(b-a)}}$. To see this, recall that for the Cauchy sequences $(c_n)_n$ and $(d_n)_n$ we have

$$\begin{aligned} f(c_n) &\leq 0 \leq f(d_n), \\ d_n - c_n &= \bigl(\frac{2}{3}\bigr)^n (b-a). \end{aligned}$$

Hence for $n \geq M(\omega_f(p+1))$ we have

$$\begin{aligned} d_n - c_n &\leq \frac{1}{2^{\omega_f(p+1)}} \\ f(d_n) - f(c_n) &\leq \frac{1}{2^{p-1}} \quad \text{by ConvMod} \end{aligned}$$

and therefore

$$|f(h_g(u,n)) - u| = |f(c_n) - u| = u - f(c_n) \leq f(d_n) - f(c_n) \leq \frac{1}{2^{p-1}}.$$

Since continuous functions are determined by their values on the rationals, we have $f(g(y)) = y$ for $y \in [a', b']$.

For every $x \in [a,b]$ with $a' \leq f(x) \leq b'$, from $g(f(x)) < x$ we obtain the contradiction $f(x) = f(g(f(x))) < f(x)$ by the hypothesis on the slope, and similarly for $>$. Using $u \not< v \leftrightarrow v \leq u$ we obtain $g(f(x)) = x$. □

As an example, consider the squaring function $f\colon [1,2] \to [1,4]$, given by the approximating map $h_f(a,n) := a^2$, constant Cauchy modulus $\alpha_f(p) := 1$, and modulus $\omega_f(p) := p+1$ of uniform continuity. A modulus of increase is $l := 1$, because for all $c, d \in [1,2]$

$$\frac{1}{2^p} \leq d - c \to \frac{1}{2^{p+1}} \leq d^2 - c^2.$$

Then $h_g(u,n) := c_n^{(u)}$, as constructed in the IVT for $x^2 - u$, iterating IVTAux. The Cauchy modulus α_g is such that $(\frac{2}{3})^n \leq \frac{1}{2^{p-3}}$ for $n \geq \alpha_g(p)$, and the modulus of uniform continuity is $\omega_f(p) := p+2$.

Theorem 4.3.7 (Attainment of the supremum, classical). *Let $f\colon [a,b] \to \mathbb{R}$ be continuous. Then there must be $x, y \in [a,b]$ such that $f(x)$ is the supremum and $f(y)$ the infimum of the range $\{\, f(x) \mid a \leq x \leq b \,\}$ of f.*

PROOF. We only treat the supremum case. Consider the set

$$S := \{\, f(x) \mid a \leq x \leq b \,\}.$$

If this set is bounded above, then by Theorem 3.3.5 of Chapter 3 it must have a supremum. Otherwise we consider $+\infty$ as its supremum. In each case we have a sequence (x_n) of real numbers in $[a, b]$ with

$$\lim f(x_n) = z := \sup S,$$

where $z \in \mathbb{R}$ or $z = +\infty$. By Theorem 2.5.1 (Bolzano-Weierstraß) every bounded sequence Folge (x_n) must have a convergent subsequence (x_{n_k}). For its limit we have

$$\lim_{k\to\infty} x_{n_k} =: x \in [a, b].$$

From the assumed continuity of f we obtain

$$z = \lim_{k\to\infty} f(x_{n_k}) = f(x),$$

hence in particular $z \in \mathbb{R}$. □

4.4. Continuity for functions of more than one variable

Without loss of generality we restrict ourselves to functions of two real variables.

DEFINITION 4.4.1. A *continuous function* $f\colon I_1 \times I_2 \to \mathbb{R}$ for compact intervals I_1, I_2 with rational end points is given by

(a) an approximating map $h_f\colon (I_1 \cap \mathbb{Q}) \times (I_2 \cap \mathbb{Q}) \times \mathbb{N} \to \mathbb{Q}$ and a map $\alpha_f\colon \mathbb{Z}^+ \to \mathbb{N}$ such that $(h_f(a, b, n))_n$ is a Cauchy sequence with (uniform) modulus α_f;

(b) a modulus $\omega_f\colon \mathbb{Z}^+ \to \mathbb{N}$ of (uniform) continuity, which satisfies

$$|a - a'|, |b - b'| \leq \frac{1}{2^{\omega_f(p)-1}} \to |h_f(a, b, n) - h_f(a', b', n)| \leq \frac{1}{2^p}$$

for $n \geq \alpha_f(p)$;

(c) a lower bound N_f and an upper bound M_f for all $h_f(a, b, n)$.

α_f and ω_f are required to be weakly increasing. A function $f\colon D \to \mathbb{R}$ on an arbitrary domain $D \subseteq \mathbb{R}^2$ is continuous if it is continuous on every $I_1 \times I_2 \subseteq D$, where I_1, I_2 are compact intervals with rational end points.

An example is the exponential function of a complex variable. Continuity of a function $f\colon D \to \mathbb{C}$ for some domain $D \subseteq \mathbb{C}$ is treated as continuity of the two real valued functions $\Re(f(z))$ and $\Im(f(z))$, and the latter as binary real valued functions, e.g., $\Re(f(x + iy))$. The example above of the continuity of the real exponential function can easily be modified to yield the continuity the exponential function of a complex variable, in the sense described.

CHAPTER 5

Differentiation

5.1. Derivatives

Definition 5.1.1. Let $f, g\colon I \to \mathbb{R}$ be continuous. g is called *derivative* of f with modulus $\delta_f\colon \mathbb{Z}^+ \to \mathbb{N}$ of differentiability if for $x, y \in I$ with $x < y$,

$$y \le x + \frac{1}{2^{\delta_f(p)}} \to \bigl|f(y) - f(x) - g(x)(y-x)\bigr| \le \frac{1}{2^p}(y-x).$$

f is said to be *differentiable* on I and g is called a *derivative* of f on I.

To say that g is a derivative of f we write

$$g = f', \quad g = Df, \quad \text{or} \quad g(x) = \frac{df(x)}{dx}.$$

If f has two derivatives, then clearly they are equal functions.

For example, a constant function has derivative 0, and the identity function has the constant 1 function as derivative.

We show that a bound on the derivative of f serves as a Lipschitz constant of f:

Lemma 5.1.2 (BoundSlope). *Let $f\colon I \to \mathbb{R}$ be continuous with derivative f'. Assume that f' is bounded by M on I. Then for $x, y \in I$ with $x < y$,*

$$\bigl|f(y) - f(x)\bigr| \le M(y-x).$$

Proof. Given $p \in \mathbb{Z}^+$, it suffices to prove

$$\bigl|f(y) - f(x)\bigr| \le M(y-x) + \frac{1}{2^p}.$$

Choose q such that $\frac{1}{2^q}(y-x) \le \frac{1}{2^p}$, and let $x = x_0 < x_1 < \cdots < x_n = y$ such that $x_{\nu+1} \le x_\nu + \frac{1}{2^{\delta_f(q)}}$, where δ_f is the modulus of differentiability of f. Then

$$\begin{aligned}
&\bigl|f(y) - f(x)\bigr| \\
&= \Bigl|\sum_{\nu<n} \bigl(f(x_{\nu+1}) - f(x_\nu)\bigr)\Bigr| \\
&= \Bigl|\sum_{\nu<n} \bigl(f'(x_\nu)(x_{\nu+1} - x_\nu)\bigr) + \sum_{\nu<n} \bigl(f(x_{\nu+1}) - f(x_\nu) - f'(x_\nu)(x_{\nu+1} - x_\nu)\bigr)\Bigr|
\end{aligned}$$

$$\leq M(y-x) + \frac{1}{2^q}(y-x). \qquad \square$$

COROLLARY 5.1.3 (DerivZero). *Let $f\colon I \to \mathbb{R}$ be continuous with derivative $f' = 0$. Then f is a constant.*

PROOF. Lemma 5.1.2 yields $f(x) = f(y)$ for $x, y \in I$, $x < y$. So $f(a)$ is constant for all rationals $a \in I$, hence also for all $x \in I$. □

LEMMA 5.1.4. *Let $f, g\colon I \to \mathbb{R}$ be continuous with derivatives f', g' of moduli δ_f, δ_g. Then*

$$(f+g)' := f' + g'$$

is a derivative of $f + g$ with modulus

$$\delta_{f+g}(p) := \max\bigl(\delta_f(p+1), \delta_g(p+1)\bigr).$$

PROOF. Let $x < y \leq x + \frac{1}{2^q}$. Then

$$\begin{aligned}
&\bigl|f(y) + g(y) - f(x) - g(x) - (f'(x) - g'(x))(y-x)\bigr| \\
&\leq \bigl|f(y) - f(x) - f'(x)(y-x)\bigr| + \Bigl|g(y) - g(x) - g'(x)(y-x)\Bigr| \\
&\leq \frac{1}{2^{p+1}}(y-x) + \frac{1}{2^{p+1}}(y-x).
\end{aligned}$$

for $q \geq \delta_f(p+1), \delta_g(p+1)$. □

LEMMA 5.1.5. *Let $f, g\colon I \to \mathbb{R}$ be continuous with derivatives f', g' of moduli δ_f, δ_g. Then*

$$(fg)' := f'g + fg'$$

is a derivative of fg with modulus

$$\delta_{fg}(s) := \max\bigl(\omega_g(r+s+1), \delta_f(s+q+2), \delta_g(s+p+2)\bigr),$$

where $2^r, 2^p, 2^q$ are upper bounds for f', f, g in I, respectively.

PROOF. Let $x < y \leq x + \frac{1}{2^m}$. Then, using Lemma 5.1.2

$$\begin{aligned}
&\bigl|f(y)g(y) - f(x)g(x) - f'(x)g(x)(y-x) - f(x)g'(x)(y-x)\bigr| \\
&= \bigl|\bigl(f(y) - f(x)\bigr)g(y) + f(x)\bigl(g(y) - g(x)\bigr) - \\
&\qquad f'(x)g(x)(y-x) - f(x)g'(x)(y-x)\bigr| \\
&= \bigl|\bigl(f(y) - f(x)\bigr)\bigl(g(y) - g(x)\bigr) + \bigl(f(y) - f(x) - f'(x)(y-x)\bigr)g(x) + \\
&\qquad f(x)\bigl(g(y) - g(x) - g'(x)(y-x)\bigr)\bigr| \\
&\leq 2^r(y-x)\frac{1}{2^{r+s+1}} + \frac{1}{2^{s+q+2}}(y-x)\bigl|g(x)\bigr| + \bigl|f(x)\bigr|(y-x)\frac{1}{2^{s+p+2}} \\
&\leq \frac{1}{2^s}(y-x)
\end{aligned}$$

for $m \geq \omega_g(r+s+1), \delta_f(s+q+2), \delta_g(s+p+2)$. □

LEMMA 5.1.6. *Let $g\colon I \to \mathbb{R}$ be continuous with derivative g' of modulus δ_g, and $|g'(x)| \le 2^q$ for all $x \in I$. Assume $\frac{1}{2^p} \le |g(x)|$ for all $x \in I$. Then*

$$\Big(\frac{1}{g}\Big)' := -\frac{g'}{g^2}$$

is a derivative of $\frac{1}{g}$ with modulus

$$\delta_{\frac{1}{g}}(r) := \max\big(\delta_g(r+2p+1), \omega_{\frac{1}{g}}(p+q+r+1)\big).$$

PROOF. Let $x < y \le x + \frac{1}{2^m}$. Then

$$\begin{aligned}
&\Big|\frac{1}{g(y)} - \frac{1}{g(x)} - \frac{g'(x)}{g(x)^2}(y-x)\Big| \\
&= \Big|\frac{1}{g(y)g(x)}\big(g(y) - g(x) - g'(x)(y-x)\big) + \frac{g'(x)}{g(x)}(y-x)\Big(\frac{1}{g(y)} - \frac{1}{g(x)}\Big)\Big| \\
&\le 2^p \cdot 2^p \cdot \frac{1}{2^{r+2p+1}} \cdot (y-x) + 2^p \cdot 2^q \cdot (y-x) \cdot \frac{1}{2^{p+q+r+1}} \\
&\le \frac{1}{2^r}(y-x)
\end{aligned}$$

for $m \ge \delta_g(r+2p+1), \omega_{\frac{1}{g}}(p+q+r+1)$. □

The well-known *quotient rule* can now be derived easily: under the appropriate assumptions we have

$$\Big(\frac{f}{g}\Big)' = f\frac{-g'}{g^2} + \frac{1}{g}f' = \frac{f'g - fg'}{g^2}.$$

THEOREM 5.1.7 (Chain Rule). *Let $f\colon I \to J$ and $g\colon J \to \mathbb{R}$ be continuous with derivatives f', g' of moduli δ_f, δ_g. Then*

$$(g \circ f)' = (g' \circ f) \cdot f'$$

is a derivative of $g \circ f$ with modulus

$$\delta_{g\circ f}(p) = \max\big(\delta_g(p+1+r), \delta_f(p+1+q)\big),$$

where $2^r, 2^q$ are upper bounds for f', g' in I, J, respectively.

Proof. Let $x < y \le x + \frac{1}{2^m}$. Then, using Lemma 5.1.2

$$\begin{aligned}
&\bigl|g(f(y)) - g(f(x)) - g'(f(x))f'(x)(y-x)\bigr| \\
&\le \bigl|g(f(y)) - g(f(x)) - g'(f(x))\bigl(f(y)-f(x)\bigr)\bigr| + \\
&\qquad \bigl|g'(f(x))\bigr| \cdot \bigl|f(y) - f(x) - f'(x)(y-x)\bigr| \\
&\le \frac{1}{2^{p+1+r}}\bigl|f(y)-f(x)\bigr| + 2^q\bigl|f(y) - f(x) - f'(x)(y-x)\bigr| \\
&\le \frac{1}{2^{p+1}}(y-x) + 2^q\bigl|f(y) - f(x) - f'(x)(y-x)\bigr| \\
&\le \frac{1}{2^{p+1}}(y-x) + 2^q\frac{1}{2^{p+1+q}}(y-x) \\
&= \frac{1}{2^p}(y-x)
\end{aligned}$$

for $m \ge \delta_g(p+1+r), \delta_f(p+1+q)$. □

We now show the well-known theorem of Rolle and as a consequence the mean value theorem.

Theorem 5.1.8 (Rolle). *Let $f\colon [a,b] \to \mathbb{R}$ be continuous with derivative f', and assume $f(a) = f(b)$. Then for every $p \in \mathbb{Z}^+$ we can find $c \in [a,b]$ such that $|f'(c)| \le \frac{1}{2^p}$.*

Proof. Let δ_f be the modulus of differentiability of f, and let $a = a_0 < a_1 < \dots < a_n = b$ such that $a_{\nu+1} \le a_\nu + \frac{1}{2^{\delta_f(p+2)}}$. Compare all $|f'(a_\nu)|$ with $\frac{1}{2^{p+1}} < \frac{1}{2^p}$. If we have found one $< \frac{1}{2^p}$, we are done. Otherwise we argue as in Lemma 5.1.2:

$$\begin{aligned}
&f(b) - f(a) \\
&= \sum_{\nu<n}\bigl(f(a_{\nu+1}) - f(a_\nu)\bigr) \\
&= \sum_{\nu<n}\bigl(f'(a_\nu)(a_{\nu+1}-a_\nu)\bigr) + \sum_{\nu<n}\bigl(f(a_{\nu+1}) - f(a_\nu) - f'(a_\nu)(a_{\nu+1}-a_\nu)\bigr) \\
&\ge \frac{1}{2^{p+1}}(b-a) - \frac{1}{2^{p+2}}(b-a) > 0.
\end{aligned}$$

This contradiction proves the claim. □

Theorem 5.1.9 (Mean value theorem). *Let $f\colon [a,b] \to \mathbb{R}$ be continuous with derivative f'. Then for every $p \in \mathbb{Z}^+$ we can find $c \in [a,b]$ such that*

$$\bigl|f(b) - f(a) - f'(c)(b-a)\bigr| \le \frac{1}{2^p}(b-a).$$

Proof. Let $\frac{1}{2^q} \le \frac{1}{2^p}(b-a)$ and define a continuous $h\colon [a,b] \to \mathbb{R}$ by

$$h(x) := (x-a)\bigl(f(b)-f(a)\bigr) - f(x)(b-a).$$

Then $h(a) = h(b) = -f(a)(b-a)$. Hence by Rolle's theorem we can find c in $[a,b]$ such that

$$\big|h'(c)\big| = \big|f(b) - f(a) - f'(c)(b-a)\big| \leq \frac{1}{2^q}. \qquad \square$$

5.2. Local extrema, convexity

DEFINITION 5.2.1. Let $f\colon I \to \mathbb{R}$ be continuous and $a < x < b$ for some $a, b \in I$. We call x *local maximum* if we have $p \in \mathbb{Z}^+$ such that

$$\forall_{\xi \in I}(|\xi - x| \leq \frac{1}{2^p} \to f(\xi) \leq f(x)).$$

The notion of a *local minimum* is defined similarly. By an *extremum* we mean either a maximum or a minimum.

THEOREM 5.2.2. *Let $f\colon I \to \mathbb{R}$ be continuous with derivative f', and $a < x < b$ for some $a, b \in I$. If x is a local extremum of f, then $f'(x) = 0$.*

PROOF. Assume x is a local maximum of f, and let $p \in \mathbb{Z}^+$ such that

$$\forall_{\xi \in I}(|\xi - x| \leq \frac{1}{2^p} \to f(\xi) \leq f(x)).$$

Since f' is the derivative of f we know

$$f'_+(x) = \lim_{\xi \searrow x} \frac{f(\xi) - f(x)}{\xi - x} \leq 0$$

and

$$f'_-(x) = \lim_{\xi \nearrow x} \frac{f(\xi) - f(x)}{\xi - x} \geq 0$$

and we have $f'_+(x) = f'_-(x) = f'(x)$, hence $f'(x) = 0$. In the case of a local minimum the proof is similar. □

Notice that the converse does not hold: a counterexample is the function $f(x) = x^3$.

THEOREM 5.2.3 (Rolle, classical). *Let $a < b$ and $f\colon [a,b] \to \mathbb{R}$ continuous with derivative f'. If $f(a) = f(b)$, then there must be an $x \in (a,b)$ such that $f'(x) = 0$.*

PROOF. In case f is constant the claim is clear. If f is not constant there must be an $x_0 \in (a,b)$ such that $f(a) < f(x_0)$ or $f(a) > f(x_0)$. Assume $f(a) < f(x_0)$. By Theorem 4.3.7 of Chapter 4 there must be some $x \in [a,b]$ with $f(x) = \sup f[a,b]$. Because of $f(a) < f(x_0)$ and $f(a) = f(b)$ we have $x \in (a,b)$. By Theorem 5.2.2 we obtain $f'(x) = 0$. □

Corollary 5.2.4 (Mean value theorem, classical). *Let $a < b$ and assume that $f\colon [a,b] \to \mathbb{R}$ is continuous with derivative f'. Then there must be an $x \in (a,b)$ such that*

$$\frac{f(b)-f(a)}{b-a} = f'(x).$$

Proof. Consider the function $g\colon [a,b] \to \mathbb{R}$ defined by

$$g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a).$$

g is continuous and has a derivative. Moreover $g(a) = g(b) = f(a)$. By Rolle's theorem 5.2.3 there must be an $x \in (a,b)$ such that $g'(x) = 0$. With

$$g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$$

the claim follows. □

Theorem 5.2.5 (Infimum of strictly convex functions). *Let $f, f'\colon [a,b] \to \mathbb{R}$ $(a < b)$ be continuous and f' derivative of f. Assume that f is* strictly convex *with witness q, in the sense that $f'(a) < 0 < f'(b)$ and*

$$\frac{1}{2^p} < d - c \to \frac{1}{2^{p+q}} < f'(d) - f'(c).$$

Then we can find $x \in (a,b)$ such that $f(x) = \inf_{y\in[a,b]} f(y)$.

Proof. To obtain x, apply the intermediate value theorem to f'. For $\forall_{y\in[a,b]}(f(x) \le f(y))$ (this is "non-computational", i.e., a Harrop formula) one can use the standard arguments in classical analysis (i.e., the mean value theorem 5.2.4 above). □

CHAPTER 6

Integration

To begin with, we define the integral of a continuous function on a compact interval with rational end points only. The reason for this restriction is that we need to establish $\int_a^x f(t)\,dt$ as a continuous function of x. Later we shall extend the definition of the integral to compact intervals whose end points are apart.

6.1. Riemannian sums

Definition 6.1.1. Let a, b be rationals with $a < b$. A list $P = a_0, \dots, a_n$ of rationals is a *partition* of the interval $[a, b]$, if $a = a_0 \le a_1 \le \dots \le a_n = b$. $\max\{\, a_{i+1} - a_i \mid i < n \,\}$ is the *mesh* of P. A partition $Q = a'_0, \dots, a'_m$ of $[a, b]$ is a *refinement* of P, if

$$\forall_{i \le n} \exists_{j \le m} a'_j = a_i.$$

If $f\colon [a, b] \to \mathbb{R}$ is a continuous function given by h_f, α_f and ω_f, and $P = a_0, a_1, \dots, a_n$ a partition of $[a, b]$, then an arbitrary sum of the form

$$\sum_{i=0}^{n-1} h_f(e_i, n) \cdot (a_{i+1} - a_i)$$

with $e_i \in [a_i, a_{i+1}]$ is denoted by $S(f, P)$. In particular for $a_i = a + \frac{i}{n}(b - a)$

$$S(f, n) := S(f, a, b, n) := \frac{b - a}{n} \sum_{i=0}^{n-1} h_f(a_i, n)$$

is one of the numbers $S(f, P)$.

Theorem 6.1.2. *Assume that $f\colon [a, b] \to \mathbb{R}$ is continuous with modulus ω_f of (uniform) continuity. Then*

$$\bigl(S(f, n)\bigr)_{n \in \mathbb{N}}$$

is a Cauchy sequence of rationals with modulus

$$M(p) = \max(2^{\omega_f(p+q+1)}(b - a), \alpha_f(p + q + 2)),$$

where q is such that $b-a \leq 2^q$; we denote this real by

$$\int_a^b f(x)\,dx.$$

Moreover, if P is a partition of mesh $\leq 2^{-\omega_f(l)}$, then

$$\Big|S(f,P) - \int_a^b f(x)\,dx\Big| \leq \frac{1}{2^l}(b-a).$$

Proof. Let k, l be given and $P = a_0, \dots, a_n$, $Q = b_0, \dots, b_m$ partitions of $[a,b]$ with mesh $\leq 2^{-\omega_f(k+1)}$ or $\leq 2^{-\omega_f(l+1)}$, respectively. Let $R = c_0, \dots, c_r$ be the common refinement of P and Q, obtained by arranging $a_0, \dots, a_n, b_0, \dots, b_m$ into a monotone sequence (here we make use of the assumption that a_i, b_i are rational numbers). Let $d_j \in [c_j, c_{j+1}]$ for $j < r$. For every $i < n$ denote by $\sum_i$ the summation over all indices j such that $a_i \leq c_j < a_{i+1}$. Then

$$\begin{aligned}
&|S(f,P) - S(f,R)| \\
&= \Big|\sum_{i=0}^{n-1} h_f(e_i,n)\cdot(a_{i+1}-a_i) - \sum_{i=0}^{r-1} h_f(d_i,r)\cdot(c_{i+1}-c_i)\Big| \\
&= \Big|\sum_{i=0}^{n-1} h_f(e_i,n)\sum_i(c_{j+1}-c_j) - \sum_{i=0}^{n-1}\sum_i h_f(d_j,r)\cdot(c_{j+1}-c_j)\Big| \\
&\leq \sum_{i=0}^{n-1}\sum_i |h_f(e_i,n) - h_f(d_j,r)|(c_{j+1}-c_j) \\
&\leq \sum_{i=0}^{n-1}\sum_i \big(|h_f(e_i,n) - h_f(e_i,r)| + |h_f(e_i,r) - h_f(d_j,r)|\big)(c_{j+1}-c_j) \\
&\leq \sum_{i=0}^{n-1}\sum_i \frac{1}{2^k}(c_{j+1}-c_j) \quad \text{for } n \geq \alpha_f(k+1) \\
&= \frac{1}{2^k}(b-a)
\end{aligned}$$

Similarly, for $n \geq \alpha_f(l+1)$

$$|S(f,Q) - S(f,R)| \leq \frac{1}{2^l}(b-a),$$

hence

$$|S(f,P) - S(f,Q)| \leq (\frac{1}{2^k} + \frac{1}{2^l})(b-a).$$

In particular

$$|S(f,m) - S(f,n)| \leq 2^{-k+1}(b-a)$$

for $m, n \geq 2^{\omega_f(k)}(b-a), \alpha_f(k+1)$. Hence $(S(f,n))_n$ is a Cauchy sequence.

Moreover we have for $n \geq 2^{\omega_f(l+1)}(b-a), \alpha_f(l+1)$

$$|S(f,P) - S(f,n)| \leq (\frac{1}{2^k} + \frac{1}{2^l})(b-a).$$

Now let $n \to \infty$ and $k \to \infty$. Then we obtain

$$|S(f,P) - \int_a^b f(x)\,dx| \leq \frac{1}{2^l}(b-a),$$

as was to be shown. □

Remark 6.1.3. We will also need to consider $S(f,n)$ in case $b < a$. Then we can use the same definition, and by the same argument we see that $(S(f,n))_n$ is a Cauchy sequence; its limit is denoted by $\int_a^b f(t)\,dt$. One can see easily that $\int_a^b f(t)\,dt = -\int_b^a f(t)\,dt$.

Immediately from the definition we obtain:

Corollary 6.1.4. *Assume that* $f\colon [a,b] \to \mathbb{R}$ *is continuous and* $c \in [a,b]$. *Then* $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$.

Corollary 6.1.5. *Assume that* $f, g\colon [a,b] \to \mathbb{R}$ *are continuous.*

(a) *If* $f \leq g$, *then* $\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$.
(b) $\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx$.
(c) $\int_a^b \bigl(f(x) + g(x)\bigr)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$.
(d) $\int_a^b \bigl(c \cdot f(x)\bigr)\,dx = c \cdot \int_a^b f(x)\,dx$.
(e) $\int_a^b c\,dx = c \cdot (b-a)$.

6.2. Integration and differentiation

Up to now we have considered the integral with respect to a fixed integration interval. Now we view the upper bound of this interval as variable and study the function obtained in this way. It is called the "undetermined integral".

Given $a_0 < c < b_0$ and a continuous $f\colon [a_0, b_0] \to \mathbb{R}$, we first need to establish $F(x) := \int_c^x f(t)\,dt$ as a continuous function. This means that we have to come up with h_F, α_F and ω_F; as lower bound we can take $N_F := (b_0 - a_0)N_f$ and as upper bound $M_F := (b_0 - a_0)M_f$. Let

$$h_F(a,n) := S(f,c,a,n).$$

By the theorem above we know that $(h_F(a,n))_n$ is a Cauchy sequence with modulus $p \mapsto 2^{\omega_f(p+1)}$. It remains to provide a modulus ω_F of (uniform) continuity. To this end, we may assume $c < a < b$. Divide the intervals $[c,a]$

and $[c, b]$ in n pieces each, and let $a_i := c + \frac{i}{n}(a - c)$ and $b_i := c + \frac{i}{n}(b - c)$. Then

$$\begin{aligned}
&\bigl|h_F(a,n) - h_F(b,n)\bigr| \\
&= \bigl|S(f,c,a,n) - S(f,c,b,n)\bigr| \\
&= \frac{1}{n}\bigl|(a-c)\sum_{i=0}^{n-1} h_f(a_i,n) - (b-c)\sum_{i=0}^{n-1} h_f(b_i,n)\bigr| \\
&\le \frac{1}{n}\sum_{i=0}^{n-1}\bigl((a-c)\bigl|h_f(a_i,n) - h_f(b_i,n)\bigr| + |a-b|\cdot|h_f(b_i,n)|\bigr) \\
&\le \frac{1}{n}\sum_{i=0}^{n-1}\bigl((a-c)\cdot\frac{1}{2^{p+1}} + |a-b|\cdot 2^q\bigr) \\
&\le \frac{1}{2^q}
\end{aligned}$$

provided $|a_i - b_i| \le 2^{-\omega_f(p+1)+1}$ and $|a - b| \le 2^{-p-q-1}$, where q is such that $h_f(b_i, n) \le 2^q$. So let

$$\alpha_F(p) := \max\bigl(\alpha_f(0), 2^{\omega_f(p+1)}\bigr), \quad \omega_F(p) := \max\bigl(p + q, \omega_f(p + 1)\bigr).$$

Proposition 6.2.1. *Let $f\colon [a, b] \to \mathbb{R}$ be continuous with modulus ω_f of (uniform) continuity. Fix $c \in [a, b]$ and let*

$$F(x) := \int_c^x f(t)\,dt.$$

be the continuous function just described. Then this function $F\colon [a, b] \to \mathbb{R}$ has f as derivative, with modulus ω_f. Morover, if G is any differentiable function on $[a, b]$ with $G' = f$, then the difference $F - G$ is a constant.

Proof.

$$\begin{aligned}
\bigl|F(y) - F(x) - f(x)(y-x)\bigr| &= \Bigl|\int_c^y f(t)\,dt - \int_c^x f(t)\,dt - f(x)(y-x)\Bigr| \\
&= \Bigl|\int_x^y f(t)\,dt - \int_x^y f(x)\,dt\Bigr| \\
&\le \int_x^y \bigl|f(t) - f(x)\bigr|\,dt \\
&\le \int_x^y \frac{1}{2^p}\,dt = \frac{1}{2^p}(y-x)
\end{aligned}$$

for $y \le x + 2^{-\omega_f(k)}$; this was to be shown. Now let G be any differentiable function on $[a, b]$ with $G' = f$. Then $(F - G)' = F' - G' = f - f = 0$, hence $F - G$ is a constant, by Corollary 5.1.3. □

THEOREM 6.2.2 (Fundamental theorem of calculus). *Let $f\colon I \to \mathbb{R}$ be continuous and $F\colon I \to \mathbb{R}$ such that $F' = f$. Then for all $a, b \in I$*

$$\int_a^b f(t)\,dt = F(b) - F(a).$$

PROOF. For $x \in I$ define

$$F_0(x) := \int_a^x f(t)\,dt.$$

By the proposition we have $F_0' = f$. Clearly

$$F_0(a) = 0 \qquad \text{and} \qquad F_0(b) = \int_a^b f(t)\,dt.$$

Hence for any $F\colon I \to \mathbb{R}$ such that $F' = f$, by Corollary 5.1.3 the function $F - F_0$ is a constant. Therefore

$$F(b) - F(a) = F_0(b) - F_0(a) = F_0(b) = \int_a^b f(t)\,dt. \qquad \square$$

It is common to use the notation

$$F(x)\Big|_a^b \quad \text{or} \quad \Big[F(x)\Big]_a^b \quad \text{for} \quad F(b) - F(a).$$

The formula from the fundamental theorem of calculus can then be written as

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b \quad \text{or} \quad \int_a^b f(x)\,dx = \Big[F(x)\Big]_a^b.$$

Let $f\colon I \to \mathbb{R}$ be continuous. For arbitrary reals $x, y \in I$ we define

$$\int_x^y f(t)\,dt := F(y) - F(x), \tag{13}$$

where F is the function from the proposition (which has f as derivative). Clearly this definition does not depend on the choice of the constant c implicit in the function F.

6.3. Substitution rule, partial integration

THEOREM 6.3.1 (Substitution rule). *Let $f\colon I \to \mathbb{R}$ be continuous and $\varphi\colon [a, b] \to \mathbb{R}$ differentiable such that $\varphi([a, b]) \subseteq I$. Then*

$$\int_a^b f(\varphi(t))\varphi'(t)\,dt = \int_{\varphi(a)}^{\varphi(b)} f(x)\,dx.$$

Remark 6.3.2. With the symbolic notation

$$d\varphi(t) := \varphi'(t)\,dt$$

the above formula can be written as

$$\int_a^b f(\varphi(t))\,d\varphi(t) = \int_{\varphi(a)}^{\varphi(b)} f(x)\,dx.$$

This is easy to remember, for one only has to replace x by $\varphi(t)$. The integration bounds can be inferred as well: if t ranges from a to b, then x $(= \varphi(t))$ ranges from $\varphi(a)$ to $\varphi(b)$.

Proof. Let $F\colon I \to \mathbb{R}$ be such that $F' = f$. For $F \circ \varphi\colon [a,b] \to \mathbb{R}$ we have by the chain rule

$$(F \circ \varphi)'(t) = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t).$$

Hence by the fundamental theorem of calculus

$$\int_a^b f(\varphi(t))\varphi'(t)\,dt = F(\varphi(b)) - F(\varphi(a)) = \int_{\varphi(a)}^{\varphi(b)} f(x)\,dx. \qquad \square$$

Theorem 6.3.3 (Partial integration). *Let $f, g\colon [a,b] \to \mathbb{R}$ be differentiable functions. Then*

$$\int_a^b f(x)g'(x)\,dx = f(x)g(x)\Big|_a^b - \int_a^b g(x)f'(x)\,dx.$$

Remark 6.3.4. A short notation for this formula is

$$\int f\,dg = fg - \int g\,df.$$

Proof. For $F := fg$ we have by the product rule

$$F'(x) = f'(x)g(x) + f(x)g'(x),$$

whence by the fundamental theorem of calculus

$$\int_a^b f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx = F(x)\Big|_a^b = f(x)g(x)\Big|_a^b. \qquad \square$$

6.4. Intermediate value theorem of integral calculus

Theorem 6.4.1 (Intermediate value theorem of integral calculus). *Let $f, \varphi\colon I \to \mathbb{R}$ be continuous and f locally nonconstant. Assume that we have rationals $a \le c < d \le b$ in I such that*

$$f(c) \le f(t) \le f(d) \quad (t \in [a,b]).$$

Assume further $\varphi \geq 0$ and $0 < \int_a^b \varphi(t)\,dt$. Then we can find $x \in [c,d]$ such that

$$\int_a^b f(t)\varphi(t)\,dt = f(x)\int_a^b \varphi(t)\,dt.$$

PROOF. By assumption

$$f(c)\int_a^b \varphi(t)\,dt \leq \int_a^b f(t)\varphi(t)\,dt \leq f(d)\int_a^b \varphi(t)\,dt,$$

whence we have $y \in [f(c), f(d)]$ such that

$$\int_a^b f(t)\varphi(t)\,dt = y\int_a^b \varphi(t)\,dt.$$

By the intermediate value theorem we obtain an $x \in [c,d]$ such that $f(x) = y$, as required. □

6.5. Inverse of the exponential function

We use the machinery developed in this section to define the inverse of the exponential function. To motivate the definition, suppose we already have a differentiable function $\ln\colon (0,\infty) \to \mathbb{R}$ such that $\exp(\ln(x)) = x$ for $x > 0$. Then the chain rule entails

$$\frac{d}{dx}\exp(\ln(x)) = \exp(\ln(x)) \cdot \frac{d}{dx}\ln(x) = 1,$$

hence

$$\frac{d}{dx}\ln(x) = \frac{1}{x}.$$

Because of $\exp(\ln(1)) = 1$ we must also have $\ln(1) = 0$.

Therefore we define

$$\ln(x) := \int_1^x \frac{dt}{t} \quad (x > 0). \tag{14}$$

Because of $\exp(x) > 0$ the composite function $\ln \circ \exp$ is continuous on $\mathbb{R}$. Its derivative is

$$\frac{d}{dx}\ln(\exp(x)) = \frac{1}{\exp(x)} \cdot \exp(x) = 1 = \frac{d}{dx}x.$$

By Corollary 5.1.3 the function $\ln(\exp(x)) - x$ is a constant, and because of $\ln(\exp(0)) = \ln(1) = 0$ this constant must be 0. Hence

$$\ln(\exp(x)) = x \quad (x \in \mathbb{R}). \tag{15}$$

Now fix $x > 0$ and let $y := \exp(\ln(x))$. Then

$$\ln(y) = \ln(\exp(\ln(x))) = \ln(x)$$

by (15), hence

$$0 = \ln(y) - \ln(x) = \int_x^y \frac{dt}{t}.$$

Assuming $x < y$ clearly leads to a contradiction, hence $x \geq y$. Similarly we obtain $y \geq x$ and therefore $x = y$. Hence

$$\exp(\ln(x)) = x \quad (x > 0). \tag{16}$$

To prove the familiar functional equation for the logarithm, fix $y > 0$ and consider $\ln(xy) - \ln(y)$ $(x > 0)$. Then

$$\frac{d}{dx}\big(\ln(xy) - \ln(y)\big) = \frac{1}{xy} \cdot y - 0 = \frac{1}{x} = \frac{d}{dx}\ln(x).$$

By Corollary 5.1.3 $x \mapsto \ln(xy) - \ln(y) - \ln(x)$ is a constant, which must be 0 since this expression vanishes at $x = 1$. Hence

$$\ln(xy) = \ln(y) + \ln(x) \quad (x, y > 0). \tag{17}$$

Now we can define general exponentiation by

$$x^y := \exp(y \cdot \ln(x)) \quad (x > 0)$$

and derive easily all its usual properties.

CHAPTER 7

Sequences of functions

7.1. Taylor series

We now study more systematically the development of functions in power series.

THEOREM 7.1.1 (Taylor formula). *Let $f\colon I \to \mathbb{R}$ be $(n+1)$-times differentiable. Then for all $a, x \in I$*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n+1}(x)$$

with

$$R_{n+1}(x) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt.$$

PROOF. By induction on n. Basis $n = 0$. By the fundamental theorem of calculus

$$f(x) = f(a) + \int_a^x f'(t)\,dt.$$

Step $n \to n+1$. By induction hypothesis

$$\begin{aligned} R_n(x) &= \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f^{(n)}(t)\,dt \\ &= -\int_a^x f^{(n)}(t)\Big(\frac{d}{dt}\frac{(x-t)^n}{n!}\Big)\,dt \\ &= -f^{(n)}(t)\frac{(x-t)^n}{n!}\Big|_{t=a}^{t=x} + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,dt \\ &= \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt. \qquad \square \end{aligned}$$

COROLLARY 7.1.2. *Let $f\colon I \to \mathbb{R}$ be $(n+1)$-times differentiable with $f^{(n+1)}(x) = 0$ for all $x \in I$. Then f is a polynomial of degree $\le n$.*

PROOF. In this case we have $R_{n+1}(x) = 0$. □

THEOREM 7.1.3 (Lagrange). *Let $f\colon I \to \mathbb{R}$ be $(n+1)$-times differentiable and $a, x \in I$. Assume that $f^{(n+1)}$ is locally nonconstant and that we have* *rationals c, d with $a \leq c < d \leq x$ such that*

$$f^{(n+1)}(c) \leq f^{(n+1)}(t) \leq f^{(n+1)}(d) \quad (t \in [a, x]).$$

Then we can find $\xi \in [a, b]$ such that

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

Proof. By the intermediate value theorem of integral calculus we can construct $\xi \in [c, d]$ such that

$$\begin{aligned} R_{n+1}(x) &= \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt \\ &= f^{(n+1)}(\xi)\int_a^x \frac{(x-t)^n}{n!}\,dt \\ &= -f^{(n+1)}(\xi)\frac{(x-t)^{n+1}}{(n+1)!}\Big|_{t=a}^{t=x} \\ &= \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}. \end{aligned}$$ □

7.2. Uniform convergence

We define the notion of uniform convergence of a sequence of continuous functions $f_n : I \to \mathbb{R}$ to a continuous function $f : I \to \mathbb{R}$. The definition is in terms of witnesses for the given continuous functions f_n, in order to ensure that from a proof of uniform convergence we can extract the right data.

Definition 7.2.1. Let $f_n, f : I \to \mathbb{R}$ be continuous, with approximating maps h_n, h and Cauchy moduli α_n, α. The sequence $(f_n)_{n\in\mathbb{N}}$ is *uniformly convergent* to f if

$$\forall_p \exists_q \forall_{n\geq q} \forall_{a\in I}\big(\big|h_n(a, \alpha_n(p+2)) - h(a, \alpha(p+2))\big| \leq \frac{1}{2^p}\big).$$

The next lemma gives a useful characterization of uniform convergence.

Lemma 7.2.2 (UnifConvChar). *Let $f_n, f : I \to \mathbb{R}$ be continuous, with approximating maps h_n, h. Then the following are equivalent.*

(a) *The sequence $(f_n)_{n\in\mathbb{N}}$ is uniformly convergent to f.*
(b) $\forall_p \exists_{q_1} \forall_{n\geq q_1} \exists_{q_2} \forall_{k\geq q_2} \forall_{a\in I}(|h_n(a, k) - h(a, k)| \leq \frac{1}{2^p})$.

Proof. (a) $\Rightarrow$ (b). Given p, pick q_1 by (a) for $p+1$. Given $n \geq q_1$,

$$\begin{aligned}\bigl|h_n(a,k) - h(a,k)\bigr| &\leq \bigl|h_n(a,k) - h_n(a,\alpha_n(p+3))\bigr| + \\ &\quad\ \bigl|h_n(a,\alpha_n(p+3)) - h(a,\alpha(p+3))\bigr| + \\ &\quad\ \bigl|h(a,\alpha(p+3)) - h(a,k)\bigr| \\ &\leq 2^{-p-3} + \frac{1}{2^{p+1}} + 2^{-p-3}\end{aligned}$$

if $k \geq q_2 := \alpha_n(p+3), \alpha(p+3)$. Then the first and last term are $\leq 2^{-p-3}$, and the middle term is $\leq \frac{1}{2^{p+1}}$ by the choice of q_1.

(b) $\Rightarrow$ (a).

$$\begin{aligned}\bigl|h_n(a,\alpha_n(p+2)) - h(a,\alpha(p+2))\bigr| &\leq \bigl|h_n(a,\alpha_n(p+2)) - h_n(a,k)\bigr| + \\ &\quad\ \bigl|h_n(a,k) - h(a,k)\bigr| + \\ &\quad\ \bigl|h(a,k) - h(a,\alpha(p+2))\bigr| \\ &\leq \frac{1}{2^{p+2}} + \frac{1}{2^{p+1}} + \frac{1}{2^{p+2}}\end{aligned}$$

if $k \geq \alpha_n(p+2), \alpha(p+2)$ (for the first and last term) and in addition $n, k \geq p$ with p provided for $p+1$ by (b). □

We now show that a uniformly convergent sequence indeed is uniformly convergent in the usual sense.

Lemma 7.2.3. *Let $f_n, f\colon I \to \mathbb{R}$ be continuous. Assume that the sequence $(f_n)_{n\in\mathbb{N}}$ is uniformly convergent to f. Then*

$$\forall_p \exists_q \forall_{n\geq q} \forall_x \bigl(\bigl|f_n(x) - f(x)\bigr| \leq \frac{1}{2^p}\bigr).$$

Proof. Let $x = ((a_k)_k, M)$, and let h_n, h be approximating maps for f_n, f, respectively. By Lemma 7.2.2 (UnifConvChar)

$$\forall_p \exists_{q_1} \forall_{n\geq q_1} \exists_{q_2} \forall_{k\geq q_2} \forall_{a\in I} (|h_n(a,k) - h(a,k)| \leq \frac{1}{2^p}),$$

whence the claim. □

The next lemma gives a useful criterion as to when and how we can construct the limit function. It will be used below.

Lemma 7.2.4 (UnifConvLim). *Let $f_n\colon I \to \mathbb{R}$ be continuous functions, given by approximating functions h_n and moduli α of Cauchyness and ω of (uniform) continuity, where the latter two are* independent *of n. Assume we have a weakly increasing modulus $\delta\colon \mathbb{N} \to \mathbb{N}$ of* uniform convergence *satisfying*

$$\bigl|h_n(a,k) - h_m(a,k)\bigr| \leq \frac{1}{2^p}$$

for $n, m \geq \delta(p)$ and $k \geq \alpha(p)$, and all $a \in I$. Then $(f_n)_{n\in\mathbb{N}}$ uniformly converges to the continuous function $f\colon I \to \mathbb{R}$ given by

$$h_f(a,k) := h_k(a,k), \quad \alpha_f(p) := \max\bigl(\delta(p+1), \alpha(p+1)\bigr), \quad \omega_f := \omega.$$

Proof. It is easy to see that this function $f\colon I \to \mathbb{R}$ given by h_f, α_f and ω_f is indeed continuous: α_f is a Cauchy modulus, because

$$\begin{aligned}|h_k(a,k) - h_l(a,l)| &\leq |h_k(a,k) - h_l(a,k)| + |h_l(a,k) - h_l(a,l)| \\ &\leq \frac{1}{2^{p+1}} + \frac{1}{2^{p+1}}\end{aligned}$$

for $k, l \geq \alpha_f(p)$, and ω is a modulus of (uniform) continuity, because

$$|a-b| \leq 2^{-\omega(p)+1} \to |h_k(a,k) - h_k(b,k)| \leq \frac{1}{2^p}$$

for $k \geq \alpha_f(p)$. Moreover, for a given p we may pick $p := \max(\delta(p), \alpha(p))$. Then for $n, k \geq p$ and all a, $|h_n(a,k) - h_k(a,k)| \leq \frac{1}{2^p}$. By Lemma 7.2.2 (UnifConvChar), this implies that $(f_i)_{i\in\mathbb{N}}$ uniformly converges to f. □

7.3. Integration, differentiation and limits

We show that for a uniformly convergent sequence of continuous functions, integration and limits can be exchanged.

Theorem 7.3.1 (IntLimit). *Let $f_n, f\colon [a,b] \to \mathbb{R}$ be continuous, and assume that for f_n the moduli of Cauchyness and of (uniform) continuity are independent of n. Assume that the sequence $(f_n)_{n\in\mathbb{N}}$ is uniformly convergent to f. Then*

$$\lim_{n\to\infty} \int_a^b f_n(t)\,dt = \int_a^b f(t)\,dt.$$

Proof. Let $a_{nk} := S(f_n, k)$, $a_k := S(f, k)$,

$$\int_a^b f_n(t)\,dt = (a_{nk})_k =: x_n \quad \text{and} \quad \int_a^b f(t)\,dt = (a_k)_k =: x.$$

We show that $(x_n)_n$ converges to x, that is $|x_n - x| \leq \frac{1}{2^p}$ for $n \geq M(p)$. Observe that for any k,

$$|x_n - x| \leq |x_n - a_{nk}| + |a_{nk} - a_k| + |a_k - x|.$$

Recall that by definition

$$\begin{aligned}&\int_a^b f_n(t)\,dt = (S(f_n, n), M_n) \quad \text{with} \\ &M_n(p) = \max(2^{\omega_{f_n}(p+q+1)}(b-a), \alpha_{f_n}(p+q+2)),\end{aligned}$$

where q is such that $b - a \leq 2^q$. In our case, the moduli α_{f_n} of Cauchyness and ω_{f_n} of (uniform) continuity are independent of n, say α and ω. So instead of M_n we can take $M(p) := \max(2^{\omega_f(p+q+1)}(b-a), \alpha_f(p+q+2))$.

We now estimate each of the three parts of $|x_n - a_{nk}| + |a_{nk} - a_k| + |a_k - x|$ separately.

First, $|x_n - a_{nk}| \leq \frac{1}{2^{p+2}}$ for $k \geq M(p+2)$; here we need Lemma 2.1.2 (RatCauchyConvMod).

Second, for a given l such that $b - a \leq 2^l$, by Lemma 7.2.2 (UnifConvChar) we can pick q_1 such that for all $n \geq q_1$ we can pick q_2 such that for all $k \geq q_2$ we have $|h_n(a_i, k) - h(a_i, k)| \leq 2^{-p-1-l}$. Hence

$$\begin{aligned} |a_{nk} - a_k| &\leq \frac{b-a}{k} \sum_{i=0}^{k-1} \bigl|h_n(a_i,k) - h(a_i,k)\bigr| \\ &\leq \frac{b-a}{k} k 2^{-p-1-l} \leq \frac{1}{2^{p+1}}. \end{aligned}$$

Third, $|a_k - x| \leq \frac{1}{2^{p+2}}$ for $k \geq \alpha'(p+2)$ with α' the Cauchy modulus of $(a_k)_k$; here again we need Lemma 2.1.2 (RatCauchyConvMod).

Finally, for $n \geq q_1$ and $k \geq \max(M(p+2), \alpha'(p+2), q_2)$ (with q_2 depending on n) we have all three estimates simultaneously and hence

$$\begin{aligned} |x_n - x| &\leq |x_n - a_{nk}| + |a_{nk} - a_k| + |a_k - x| \\ &\leq \frac{1}{2^{p+2}} + \frac{1}{2^{p+1}} + \frac{1}{2^{p+2}} = \frac{1}{2^p}. \end{aligned}$$

Therefore it suffices to take $n \geq q_1$. □

The final theorem gives a sufficient criterium as to when differentiation and limits can be exchanged.

THEOREM 7.3.2 (DiffLimit). *Let the continuous functions $f_n \colon [a,b] \to \mathbb{R}$ be uniformly convergent to a continuous $f \colon [a,b] \to \mathbb{R}$. Assume that each f_n is differentiable with derivative f_n', and assume that for f_n' the moduli of Cauchyness and of (uniform) continuity are independent of n. Moreover assume that the sequence $(f_n')_{n\in\mathbb{N}}$ is uniformly convergent to a continuous $f^* \colon [a,b] \to \mathbb{R}$. Then f is differentiable with derivative f^*.*

PROOF. By the fundamental theorem of calculus

$$f_n(c) = f_n(a) + \int_a^c f_n'(t)\,dt \quad (c \in [a,b]).$$

By the theorem above, $\int_a^c f_n'(t)\,dt$ converges for $n \to \infty$ to $\int_a^c f^*(t)\,dt$. whence

$$f(c) = f(a) + \int_a^c f^*(t)\,dt \quad (c \in [a,b]).$$

Now let $x = (c_k)_k$ be a real in $[a, b]$. Then

$$f(x) = f(a) + \int_a^x f^*(t)\,dt \quad (x \in [a, b]).$$

By Section 6.2, f is differentiable with derivative f^*. □

CHAPTER 8

Trigonometric functions

8.1. Euler's formula

For all $x \in \mathbb{R}$ let

$$\cos x := \Re(e^{ix}), \quad \sin x := \Im(e^{ix}),$$

hence

$$e^{ix} = \cos x + i \sin x \quad (\textit{Euler's formula}).$$

Notice that for all $x \in \mathbb{R}$ we have $|e^{ix}| = 1$, because

$$|e^{ix}|^2 = e^{ix}\overline{e^{ix}} = e^{ix}e^{-ix} = e^0 = 1.$$

Therefore e^{ix} is a point on the unit circle of the Gaußian plane and $\cos x$, $\sin x$ are the projections of this point to the x- and y-axis. Immediately from the definitions we have

$$\begin{aligned}
&\cos x = \frac{1}{2}(e^{ix} + e^{-ix}) \quad \text{and} \quad \sin x = \frac{1}{2i}(e^{ix} - e^{-ix}),\\
&\cos(-x) = \cos x \quad \text{and} \quad \sin(-x) = -\sin x,\\
&\cos^2 x + \sin^2 x = 1.
\end{aligned}$$

THEOREM 8.1.1. *The functions* $\cos x$ *and* $\sin x$ *are continuous on* $\mathbb{R}$.

PROOF. Omitted. □

8.2. Addition theorems

THEOREM 8.2.1 (Addition Theorems). *For all* $x, y \in \mathbb{R}$ *we have*

$$\begin{aligned}
\cos(x+y) &= \cos x \cos y - \sin x \sin y,\\
\sin(x+y) &= \sin x \cos y + \cos x \sin y.
\end{aligned}$$

PROOF. From the functional equation of the exponential function

$$e^{i(x+y)} = e^{ix}e^{iy}$$

we obtain by Euler's formula

$$\begin{aligned}
&\cos(x+y) + i\sin(x+y)\\
&= (\cos x + i \sin x)(\cos y + i \sin y)\\
&= (\cos x \cos y - \sin x \sin y) + i(\sin x \cos y + \cos x \sin y).
\end{aligned}$$

Comparing the real and imaginary parts gives the claim. □

COROLLARY 8.2.2. *For all $x, y \in \mathbb{R}$ we have*

$$\begin{aligned} \sin x - \sin y &= 2\cos\frac{x+y}{2}\sin\frac{x-y}{2}, \\ \cos x - \cos y &= -2\sin\frac{x+y}{2}\sin\frac{x-y}{2}. \end{aligned}$$

PROOF. Let $u := \frac{x+y}{2}$ and $v := \frac{x-y}{2}$; that $x = u+v$ and $y = u-v$. The addition theorem for sin entails

$$\begin{aligned} \sin x - \sin y &= \sin(u+v) - \sin(u-v) \\ &= \sin u\cos v + \cos u \sin v - \sin u \cos(-v) - \cos u \sin(-v) \\ &= \sin u \cos v + \cos u \sin v - \sin u \cos v + \cos u \sin v \\ &= 2\cos u \sin v. \end{aligned}$$

The second equation is proved similarly. □

THEOREM 8.2.3. *For all $x \in \mathbb{R}$ we have*

$$\begin{aligned} \cos x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - + \dots, \\ \sin x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - + \dots. \end{aligned}$$

Both series converge absolutely.

PROOF. Absolute convergence follows from the absolute convergence of the exponential series. Using

$$i^n = \begin{cases} 1, & \text{if } n = 4m; \\ i, & \text{if } n = 4m+1; \\ -1, & \text{if } n = 4m+2; \\ -i, & \text{if } n = 4m+3 \end{cases} \qquad (m \in \mathbb{Z}).$$

we obtain for all $x \in \mathbb{R}$

$$\begin{aligned} e^{ix} &= \sum_{n=0}^{\infty} i^n \frac{x^n}{n!} \\ &= 1 + ix - \frac{x^2}{2!} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} - \frac{x^6}{6!} - i\frac{x^7}{7!} + \frac{x^8}{8!} + \dots \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \end{aligned}$$

Comparing the real and imaginary parts gives the claim. □

8.3. Estimate of the rest

Theorem 8.3.1 (Estimate of the rest). *For all $x \in \mathbb{R}$ we have*

$$\cos x = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} + r_{2n+2}(x),$$

$$\sin x = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + r_{2n+3}(x),$$

where

$$|r_{2n+2}(x)| \leq \frac{|x|^{2n+2}}{(2n+2)!} \quad \textit{for } |x| \leq 2n+3,$$

$$|r_{2n+3}(x)| \leq \frac{|x|^{2n+3}}{(2n+3)!} \quad \textit{for } |x| \leq 2n+4.$$

Remark 8.3.2. These estimates are valid for *all* $x \in \mathbb{R}$; this can be proved by means of the Taylor formula.

Proof. For all $x \in \mathbb{R}$ we have

$$r_{2n+2}(x) = \pm \frac{x^{2n+2}}{(2n+2)!} \big(1 - \frac{x^2}{(2n+3)(2n+4)} + \dots \pm \frac{x^{2k}}{(2n+3)\cdots(2n+2k+2)} \mp \dots \big).$$

For $k \geq 1$ let

$$a_k := \frac{x^{2k}}{(2n+3)(2n+4)\cdots(2n+2k+2)}.$$

Then for all $|x| \leq 2n+3$

$$1 \geq a_1 \geq a_2 \geq \dots \geq a_k \geq 0.$$

As in the proof of the Leibniz test we obtain

$$a_1 - a_2 + a_3 - + \dots \mp a_k \geq 0 \quad \text{and} \quad 1 - a_1 + a_2 - a_3 + - \dots \pm a_k \geq 0,$$

hence

$$|r_{2n+2}(x)| = \frac{|x|^{2n+2}}{(2n+2)!} (1 - a_1 + a_2 - a_3 + - \dots \pm a_k \mp \dots) \leq \frac{|x|^{2n+2}}{(2n+2)!}.$$

The second estimate is proved similarly. □

Corollary 8.3.3.

$$\lim_{\substack{x \to 0 \\ x \neq 0}} \frac{\sin x}{x} = 1.$$

Proof. We use the 3rd order rest, i.e.,

$$\sin x = x + r_3(x)$$

with

$$|r_3(x)| \leq \frac{|x|^3}{3!} \quad \text{for } |x| \leq 4.$$

This gives for $0 < |x| \leq 3$

$$\left|\frac{\sin x}{x} - 1\right| = \frac{|r_3(x)|}{|x|} \leq \frac{|x|^2}{6}$$

and hence the claim. □

8.4. Definition of pi

Theorem 8.4.1. *The function* cos *has exactly one zero in the interval* $[0, 2]$.

For the proof we need three auxiliary lemmata.

Lemma 8.4.2. $\cos 2 \leq -\frac{1}{3}$.

Proof. We use the 4th order rest, i.e.,

$$\cos x = 1 - \frac{x^2}{2} + r_4(x)$$

with

$$|r_4(x)| \leq \frac{|x|^4}{4!} \quad \text{for } |x| \leq 5.$$

Hence

$$\cos 2 = 1 - 2 + r_4(2)$$

with

$$|r_4(2)| \leq \frac{16}{24} = \frac{2}{3}$$

and hence the claim $\cos 2 \leq -\frac{1}{3}$. □

Lemma 8.4.3. $\sin x > 0$ *for all* $x \in]0, 2]$.

Proof. Because of

$$\sin x = x + r_3(x)$$

with

$$|r_3(x)| \leq \frac{|x|^3}{3!} \quad \text{for } |x| \leq 4$$

we have for $x \in]0, 2]$

$$\sin x \geq x - \frac{x^3}{6} = \frac{x}{6}(6 - x^2) > 0.$$

This proves the claim. □

LEMMA 8.4.4. *The function* cos *is strictly decreasing in the interval* $[0, 2]$.

PROOF. Let $0 \le x < x' \le 2$. From the Corollary of the addition theorem we have

$$\cos x' - \cos x = -2 \sin \frac{x' + x}{2} \sin \frac{x' - x}{2}.$$

$0 < \frac{x'-x}{2} \le 1$ implies $\sin \frac{x'-x}{2} > 0$ by the second auxiliary lemma, and $0 < \frac{x'+x}{2} \le 2$ implies $\sin \frac{x'+x}{2} > 0$ again by the second auxiliary lemma, hence $\cos x' - \cos x < 0$. □

PROOF OF THE THEOREM. $\cos 0 = 1$, $\cos 2 \le -\frac{1}{3}$ and the fact that cos is strictly decreasing in the interval $[0, 2]$ imply the claim. □

We now define the real number $\pi/2$ as the (uniquely determined) zero of the function cos in the interval $[0, 2]$.

THEOREM 8.4.5 (Special values of the exponential function).

$$e^{i\frac{\pi}{2}} = i, \quad e^{i\pi} = -1, \quad e^{i\frac{3\pi}{2}} = -i, \quad e^{2\pi i} = 1.$$

PROOF. Because of $\cos^2 x + \sin^2 x = 1$ and the definition of $\frac{\pi}{2}$ we have $\sin \frac{\pi}{2} = \pm 1$, hence by the second auxiliary lemma above $\sin \frac{\pi}{2} = 1$. This implies

$$e^{i\frac{\pi}{2}} = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} = i.$$

Hence by the functional equation

$$\begin{aligned} e^{i\pi} &= e^{i\frac{\pi}{2}} e^{i\frac{\pi}{2}} = i^2 = -1 \\ e^{i\frac{3\pi}{2}} &= e^{i\pi} e^{i\frac{\pi}{2}} = -i \\ e^{2\pi i} &= e^{i\pi} e^{i\pi} = 1. \end{aligned}$$

□

Therefore

x	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
$\sin x$	0	1	0	-1	0
$\cos x$	1	0	-1	0	1

COROLLARY 8.4.6. *For all* $x \in \mathbb{R}$ *we have*

(a) $\cos(x + 2\pi) = \cos x$, $\sin(x + 2\pi) = \sin x$
(b) $\cos(x + \pi) = -\cos x$, $\sin(x + \pi) = -\sin x$
(c) $\cos x = \sin(\frac{\pi}{2} - x)$, $\sin x = \cos(\frac{\pi}{2} - x)$

PROOF. (a) $e^{i(x+2\pi)} = e^{ix} e^{2\pi i} = e^{ix}$.
(b) $e^{i(x+\pi)} = e^{ix} e^{i\pi} = -e^{ix}$.
(c) $e^{ix} = e^{ix - i\frac{\pi}{2} + i\frac{\pi}{2}} = e^{i(x-\frac{\pi}{2})} e^{i\frac{\pi}{2}} = i e^{i(x-\frac{\pi}{2})}$. Hence

$$\cos x = -\sin\bigl(x - \frac{\pi}{2}\bigr) = \sin(\frac{\pi}{2} - x),$$

$$\sin x = \cos(\frac{\pi}{2} - x) = \cos\bigl(x - \frac{\pi}{2}\bigr). \qquad \square$$

COROLLARY 8.4.7 (Zeros of sine and cosine). *In the interval* $[0, 2\pi[$ *the function* cos *has exactly the zeros* $\frac{\pi}{2}$ *and* $\frac{3\pi}{2}$*, and* sin *has exactly the zeros* 0 *and* π*.*

PROOF. 1. $\cos\frac{\pi}{2} = 0$ by definition, hence also $\cos\frac{3\pi}{2} = -\cos\frac{\pi}{2} = 0$. Moreover, cos is strictly decreasing in $[0, \frac{\pi}{2}]$, and $\cos(\frac{\pi}{2} + x) = -\cos(\frac{\pi}{2} - x)$. Hence cos is strictly decreasing in $[\frac{\pi}{2}, \pi]$ as well. Therefore $\frac{\pi}{2}$ is the unique zero of cos in $[0, \pi]$. Furthermore $\cos(\pi + x) = \cos(-\pi - x) = \cos(\pi - x)$; hence cos has exactly one zero in $[\pi, 2\pi]$, namely $\frac{3\pi}{2}$.

2. Because of $\sin x = \cos(\frac{\pi}{2} - x) = \cos(x - \frac{\pi}{2})$ the claim follows from the fist part. □

COROLLARY 8.4.8. *In the interval* $[0, 2\pi[$*, the function* cos *assumes the value* 1 *exactly in the point* 0*.*

PROOF. We have just shown that cos is strictly decreasing in $[0, \pi]$, and that $\cos(\pi + x) = \cos(\pi - x)$. Because of $\cos 0 = 1$ the claim follows. □

We can now define the tangens function for $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ by

$$\tan x := \frac{\sin x}{\cos x}.$$

8.5. The inverse functions arcsin, arccos, arctan

The inverse functions arccos for cos, arcsin for sin and arctan for tan may now be defined similarly to how we defined the logarithm as the inverse of the exponential function, i.e., by means of integrals. We carry this out for the sine function. To motivate the definition of the inverse arcsin of the sine function, suppose we already have a differentiable function $\arcsin\colon (-1, 1) \to (-\pi/2, \pi/2)$ such that $\sin(\arcsin x) = x$ for $-1 < x < 1$. Then the chain rule entails

$$\frac{d}{dx}\sin(\arcsin x) = \cos(\arcsin x) \cdot \frac{d}{dx}\arcsin x = 1,$$

hence

$$\frac{d}{dx}\arcsin x = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - \sin^2(\arcsin x)}} = \frac{1}{\sqrt{1 - x^2}}.$$

Because of $\sin(\arcsin(0)) = 0$ we must also have $\arcsin(0) = 0$.

Therefore we define

$$\arcsin x := \int_0^x \frac{dt}{\sqrt{1 - t^2}} \quad (-1 < x < 1). \tag{18}$$

Because of $-1 < \sin x < 1$ the composite function $\arcsin \circ \sin$ is continuous on $(-\pi/2, \pi/2)$. Its derivative is

$$\frac{d}{dx}\arcsin(\sin x) = \frac{1}{\sqrt{1-\sin^2 x}} \cdot \cos x = 1 = \frac{d}{dx}x.$$

By Corollary 5.1.3 the function $\arcsin(\sin x) - x$ is a constant, and because of $\arcsin(\sin(0)) = \arcsin(0) = 0$ this constant must be 0. Hence

$$\arcsin(\sin x) = x \quad (-\pi/2 < x < \pi/2). \tag{19}$$

Now fix $x > 0$ and let $y := \sin(\arcsin x)$. Then

$$\arcsin y = \arcsin(\sin(\arcsin x)) = \arcsin x$$

by (19), hence

$$0 = \arcsin y - \arcsin x = \int_x^y \frac{dt}{\sqrt{1-t^2}}.$$

Assuming $x < y$ clearly leads to a contradiction, hence $x \geq y$. Similarly we obtain $y \geq x$ and therefore $x = y$. Hence

$$\sin(\arcsin x) = x \quad (-1 < x < 1). \tag{20}$$

Similarly we can introduce $\arccos\colon (-1,1) \to (0,\pi)$ by

$$\arccos x := \frac{\pi}{2} - \int_0^\pi \frac{dt}{\sqrt{1-t^2}}. \tag{21}$$

For the inverse $\arctan\colon \mathbb{R} \to (-1,1)$ of the tangens function $\tan x := \frac{\sin x}{\cos x}$ first recall that by the quotient rule

$$\frac{d}{dx}\tan x = \frac{1}{\cos^2 x}.$$

To motivate the definition of the inverse arctan of the tangens function, suppose we already have a differentiable function $\arctan\colon \mathbb{R} \to (-1,1)$ such that $\tan(\arctan x) = x$ for $x \in \mathbb{R}$. Then the chain rule entails

$$\frac{d}{dx}\tan(\arctan x) = \frac{1}{\cos^2(\arctan x)} \cdot \frac{d}{dx}\arctan x = 1,$$

hence

$$\frac{d}{dx}\arctan x = \cos^2(\arctan x).$$

Now let $y := \arctan x$, hence $x = \tan y$. Then

$$x^2 = \tan^2 y = \frac{\sin^2 y}{\cos^2 y} = \frac{1-\cos^2 y}{\cos^2 y} = \frac{1}{\cos^2 y} - 1,$$

hence

$$cos^2 y = \frac{1}{1+x^2}$$

and therefore

$$\frac{d}{dx}\arctan x = \frac{1}{1+x^2}.$$

Because of $\tan(0) = 0$ we must also have $\arctan(0) = 0$. So we define

$$\arctan x := \int_0^x \frac{dt}{1+t^2} \quad (x \in \mathbb{R}). \tag{22}$$

Clearly the composite function $\arctan \circ \tan$ is continuous on $(-\pi/2, \pi/2)$. Its derivative is

$$\frac{d}{dx}\arctan(\tan x) = \frac{1}{1+\tan^2 x} \cdot \frac{1}{\cos^2 x} = 1 = \frac{d}{dx}x.$$

By Corollary 5.1.3 the function $\arctan(\tan x) - x$ is a constant, and because of $\arctan(\tan(0)) = \arctan(0) = 0$ this constant must be 0. Hence

$$\arctan(\tan x) = x \quad (-\pi/2 < x < \pi/2). \tag{23}$$

Now fix $x > 0$ and let $y := \tan(\arctan x)$. Then

$$\arctan y = \arctan(\tan(\arctan x)) = \arctan x$$

by (23), hence

$$0 = \arctan y - \arctan x = \int_x^y \frac{dt}{1+t^2}.$$

Assuming $x < y$ clearly leads to a contradiction, hence $x \geq y$. Similarly we obtain $y \geq x$ and therefore $x = y$. Hence

$$\tan(\arctan x) = x \quad (x \in \mathbb{R}). \tag{24}$$

8.6. Polar coordinates

THEOREM 8.6.1 (Polar coordinates). *Every complex number $z \neq 0$ can be written uniquely in the form*

$$z = re^{i\varphi} \quad \textit{with } r = |z| \textit{ and } \varphi \in [0, 2\pi).$$

PROOF. Let $\xi := \frac{z}{|z|}$, $x := \Re(\xi)$ and $y := \Im(\xi)$. Because of $|\xi| = 1$ we have $x^2 + y^2 = 1$, hence $x, y \in [-1.1]$. Let $\alpha \in [0, \pi]$ we the unique real such that $\cos\alpha = x$. Because of $y^2 = 1 - x^2 = 1 - \cos^2\alpha = \sin^2\alpha$ we have

$$y = \pm\sin\alpha.$$

Let

$$\varphi := \begin{cases} \alpha, & \text{if } y = \sin\alpha; \\ 2\pi - \alpha, & \text{if } y = -\sin\alpha. \end{cases}$$

Then in any case $\sin\varphi = y$ (we may assume here $|y| \geq \frac{1}{3}$, otherwise we work with x instead), and hence

$$e^{i\varphi} = \cos\varphi + i\sin\varphi = x + iy = \xi = \frac{z}{|z|}.$$

For uniqueness, assume $e^{i\varphi_1} = e^{i\varphi_2}$ with $0 \leq \varphi_1 \leq \varphi_2 < 2\pi$. Then $e^{i(\varphi_2 - \varphi_1)} = 1$ with $\varphi_2 - \varphi_1 \in [0, 2\pi)$, hence $\varphi_2 - \varphi_1 = 0$. □

Remark 8.6.2. The product of two complex numbers can now simply be written as

$$r_1 e^{i\varphi_1} \cdot r_2 e^{i\varphi_2} = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}.$$

Corollary 8.6.3 (nth root of unity). *Let n be a natural number ≥ 2. The equation $z^n = 1$ has exactly n complex roots, namely*

$$e^{i\frac{2k\pi}{n}} \quad \textit{for } k = 0, \dots, n-1.$$

Proof. First notice that

$$\left(e^{i\frac{2k\pi}{n}}\right)^n = e^{2k\pi i} = 1.$$

Now let $z \in \mathbb{C}$ with $z^n = 1$. Then $|z| = 1$, hence by the theorem z can be written uniquely in the form $e^{i\varphi}$ with $\varphi \in [0, 2\pi)$. By asumption $(e^{i\varphi})^n = e^{in\varphi} = 1$, hence $n\varphi = 2k\pi$ for some $k \in \mathbb{Z}$, hence $\varphi = \frac{2k\pi}{n}$. Because of $\varphi \in [0, 2\pi)$ we obtain $0 \leq k < n$. □

CHAPTER 9

Metric spaces

We now generalize our treatment of the reals to metric spaces.

9.1. Cauchy sequences, equality, completeness

Definition 9.1.1. A *metric* on a countable set Q of approximations is a map $\rho\colon Q \to Q \to \mathbb{R}$ such that for all $u, v, w \in Q$

(a) $\rho(u, v) = 0$ iff $u = v$,
(b) $\rho(u, v) = \rho(v, u)$ (symmetry), and
(c) $\rho(u, w) \leq \rho(u, v) + \rho(v, w)$ (triangle inequality).

A *metric aproximation space* is a pair (Q, ρ) consisting of a countable set Q and a metric ρ on Q. The real $\rho(u, v)$ is called *distance* of u and v w.r.t. ρ.

Remark 9.1.2. The axioms entail $\rho(u, v) \geq 0$ for all $u, v \in Q$. This follows from the triangle inequality applied to u, v, u:

$$0 = \rho(u, u) \leq \rho(u, v) + \rho(v, u) = 2\rho(u, v).$$

Lemma 9.1.3 (MetrUB, MetrLB). *Let (Q, ρ) be a metric space. Then*

$$|\rho(u, w) - \rho(v, w)| \leq \rho(u, v) \leq \rho(u, w) + \rho(v, w).$$

Proof. From the triangle inequality and symmetry we obtain both

$$\begin{aligned} \rho(u, w) - \rho(v, w) &\leq \rho(u, v), \\ \rho(v, w) - \rho(u, w) &\leq \rho(u, v) \end{aligned}$$

and hence the first inequality. The second one follows immediately from the triangle inequality and symmetry. □

The *completion* of a metric approximation space (Q, ρ) consists of all pairs $((u_n)_n, M)$ such that $(u_n)_n$ is a Cauchy sequence with modulus M, that is,

$$\rho(u_n, u_m) \leq \frac{1}{2^p} \quad \text{for } n, m \geq M(p).$$

Let X be the set of all such pairs, called *points*. We extend ρ to points $x = ((u_n)_n, M)$ and $y = ((v_n)_n, N)$ of X by

$$\rho(x, y) := ((\rho(u_n, v_n))_n, L) \quad \text{with } L(p) := \max(M(p + 1), N(p + 1)).$$

It is easy to see that the right hand side is a real number, since

$$\begin{aligned}&|\rho(u_n,v_n)-\rho(u_m,v_m)|\\&\leq|\rho(u_n,v_n)-\rho(u,v_m)|+|\rho(u_n,v_m)-\rho(u_m,v_m)|\\&\leq\frac{1}{2^{p+1}}+\frac{1}{2^{p+1}}\quad\text{for } n,m\geq\max(M(p+1),N(p+1)).\end{aligned}$$

Note that then also $\rho(x,u)$ is defined, by viewing u as the constant sequence with modulus 1.

We call (X,ρ,Q) (with the extended ρ above) a *metric space*[1], and the elements $u,v,w\in Q$ its *approximations*.

Definition 9.1.4 (Equality). Two modulated Cauchy sequences $x := ((u_n)_n, M)$ and $y := ((v_n)_n, N)$ are *equal* if

$$\rho(u_{M(p+1)},v_{N(p+1)})\leq\frac{1}{2^p}\quad\text{for all } p\in\mathbb{Z}^+.$$

To see that this is an equivalence relation we can proceed as in Lemma 1.2.3 of Chapter 1 (RealEqChar) and show that $x=y$ is equivalent to

$$\forall_p\exists_{n_0}\forall_{n\geq n_0}(\rho(u_n,v_n)\leq\frac{1}{2^p}).$$

Definition 9.1.5. A sequence $(x_n)_{n\in\mathbb{N}}$ of points in a metric space (X,ρ,Q) is a *Cauchy sequence* with modulus $M\colon\mathbb{Z}^+\to\mathbb{N}$ whenever

$$\rho(x_n,x_m)\leq\frac{1}{2^p}\quad\text{for } n,m\geq M(p),$$

and *converges* with modulus $M\colon\mathbb{Z}^+\to\mathbb{N}$ to a point y, its *limit*, whenever $\rho(x_n,y)\leq\frac{1}{2^p}$ for $n\geq M(p)$.

Lemma 9.1.6 (ApproxCauchyConvMod). *Any modulated Cauchy sequence of approximations in a metric space (X,ρ,Q) converges with the same modulus to the point it represents.*

Proof. Let $x := ((u_n)_n, M)$ be a point. We must show $\rho(u_n,x)\leq\frac{1}{2^p}$ for $n\geq M(p)$. Fix $n\geq M(p)$. It suffices to show $\rho(u_n,u_m)\leq\frac{1}{2^p}$ for $m\geq M(p)$. But this holds by assumption. □

By the triangle inequality, every convergent sequence of points in a metric space with modulus M is a Cauchy sequence with modulus $p\mapsto M(p+1)$. As in Theorem 2.1.3 of Chapter 1 (RealCompl) we can prove the reverse implication, this time using Lemma 9.1.6 (ApproxCauchyConvMod).

Theorem 9.1.7 (MetrCompl). *For every Cauchy sequence of points in a metric space we can find a point to which it converges.*

[1] Separated metric space (or separable metric space if Q is left implicit) might be a more appropiate name.

9.2. Located sets

Definition 9.2.1 (Located set). Let (X, ρ, Q) be a metric space. A subset V of Q is *located*[2] if for every approximation $u \in Q$ and all rationals $c, d \in \mathbb{Q}$ with $c < d$

$$\forall_{v \in V}(c \leq \rho(u, v)) \vee \exists_{v \in V}(\rho(u, v) \leq d).$$

Theorem 9.2.2 (Distance from located sets). *Let (X, ρ, Q) be a metric space and V a subset of Q. Assume that V is* sufficiently inhabited*, i.e., $\forall_u \exists_{v \in V}(0 < \rho(u, v))$. Then the following are equivalent.*

(a) *V is located.*

(b) *$\rho(u, V) := \inf\{\, \rho(u, v) \mid v \in V \,\}$ exists for all $u \in Q$.*

Proof. (b) → (a). Let $u \in Q$ and $c, d \in \mathbb{Q}$ with $c < d$. Then either $c \leq \rho(u, V)$ or else $\rho(u, V) \leq \frac{c+d}{2}$. In the first case we have $c \leq \rho(u, v)$ for all $v \in V$, and in the second case we have $\rho(u, v) \leq d$ for some $v \in V$.

(a) → (b). Let $u \in Q$. Consider

$$\Pi_{V,u}(c, d) := \forall_{v \in V}(c \leq \rho(u, v)) \wedge \exists_{v \in V}(\rho(u, v) \leq d)$$

as a property of pairs c, d of rational numbers with $c < d$. Pick $v \in V$ and d such that $0 < \rho(u, v) \leq d$. Then $\Pi_{V,u}(0, d)$. We construct two sequences $(c_n)_n$ and $(d_n)_n$ of rationals such that for all n

$$0 = c_0 \leq c_1 \leq \cdots \leq c_n < d_n \leq \cdots \leq d_1 \leq d_0 = d \tag{25}$$

$$\Pi_{V,u}(c_n, d_n), \tag{26}$$

$$d_n - c_n \leq \left(\frac{2}{3}\right)^n d. \tag{27}$$

Let $c_0, \dots, c_n$ and $d_0, \dots, d_n$ be already constructed such that (25)-(27) hold. Let $c' = c_n + \frac{1}{3}(d_n - c_n)$ and $d' = c_n + \frac{2}{3}(d_n - c_n)$. By the locatedness of V

$$\forall_{v \in V}(c' \leq \rho(u, v)) \vee \exists_{v \in V}(\rho(u, v) \leq d').$$

In the first case let $c_{n+1} := c_n$ and $d_{n+1} := d'$, and in the second case let $c_{n+1} := c'$ and $d_{n+1} := d_n$. Then clearly (25)-(27) continue to hold for $n + 1$, and the real number given by the two modulated Cauchy sequences of rationals $(c_n)_n$ and $(d_n)_n$ is the infimum of $\{\, \rho(u, v) \mid v \in V \,\}$. □

The next lemma employs a technique first used by Bishop (1967, p.177, Lemma 7), which is known under the name λ-technique. The formulation below is adapted from Bishop and Bridges (1985, p.92).

[2] In Bishop (1967); Bishop and Bridges (1985) locatedness is defined by the existence of infima, as in part (b) of Theorem 9.2.2. Our notion of locatedness is sometimes called "order located below" in the literature.

Lemma 9.2.3. *Let (X, ρ, Q) be a metric space and V a located subset of Q. Fix $u \in Q$ and assume $0 < \rho(u, v_0)$ for some $v_0 \in V$. Then there is a point $((v_n)_n, M) =: y \in X$ with all v_n in V such that for any p*

$$\frac{1}{2^p} < \rho(u, y) \quad \textit{implies} \quad \forall_{v \in V}(\frac{1}{2^{p+1}} \leq \rho(u, v)).$$

Proof. For simplicity assume $\rho(u, v_0) = 1$. By simultaneous recursion we define sequences $(v_n)_n$ of approximations in V and $(\lambda_n)_n$ of decreasing booleans (in the sense that $\lambda_n = \mathtt{tt}$ implies $\lambda_m = \mathtt{tt}$ for all $m \geq n$) such that

$$\lambda_n = \mathtt{ff} \to \rho(u, v_n) \leq \frac{1}{2^n},$$
$$\lambda_n = \mathtt{tt} \to \forall_{v \in V}(\frac{1}{2^{n+1}} \leq \rho(u, v)).$$

Let $\lambda_0 := \mathtt{ff}$. In the step we are given v_n, λ_n and must define v_{n+1}, λ_{n+1}. *Case* $\lambda_n = \mathtt{ff}$. By the locatedness of V w.r.t. $\frac{1}{2^{n+2}} < \frac{1}{2^{n+1}}$ we know

$$\forall_{v \in V}(\frac{1}{2^{n+2}} \leq \rho(u, v)) \vee \exists_{v \in V}(\rho(u, v) \leq \frac{1}{2^{n+1}})$$

In the first case let $v_{n+1} := v_n$ and $\lambda_{n+1} := \mathtt{tt}$, and in the second case let v_{n+1} be the element provided and $\lambda_{n+1} := \mathtt{ff}$. *Case* $\lambda_n = \mathtt{tt}$. Let $v_{n+1} := v_n$ and $\lambda_{n+1} := \mathtt{tt}$. – We show that $(v_n)_n$ is a Cauchy sequence with modulus $M(p) := p + 1$, i.e.

$$\rho(v_n, v_m) \leq \frac{1}{2^p} \quad \text{for } p + 1 \leq n \leq m.$$

Assume $n \leq m$. *Case* $\lambda_m = \mathtt{ff}$ (hence also $\lambda_n = \mathtt{ff}$). Then

$$\rho(v_n, v_m) \leq \rho(u, v_n) + \rho(u, v_m) \leq \frac{2}{2^n} \leq \frac{1}{2^p} \quad \text{for } p + 1 \leq n.$$

Case $\lambda_n = \mathtt{ff}$, $\lambda_m = \mathtt{tt}$. Take l with $n \leq l < m$ and $\lambda_l = \mathtt{ff}$, $\lambda_{l+1} = \mathtt{tt}$. Then

$$\rho(v_n, v_m) = \rho(v_n, v_l) \leq \frac{1}{2^p} \quad \text{for } p + 1 \leq n,$$

as in the previous case. *Case* $\lambda_n = \mathtt{tt}$ (hence also $\lambda_m = \mathtt{tt}$). Then $v_n = v_m$, hence $\rho(v_n, v_m) = 0$. – Let $y := ((v_n)_n, M)$ and assume $\frac{1}{2^p} < \rho(u, y)$. Recall $\rho(u, y) = ((\rho(u, v_n))_n, M)$. Hence $\frac{1}{2^p} < \rho(u, v_p)$ and therefore $\lambda_p = \mathtt{tt}$, which implies $\forall_{v \in V}(\frac{1}{2^{p+1}} \leq \rho(u, v))$. □

9.3. Continuous functions

Definition 9.3.1 (Continuous function). Let (X, ρ, Q), (Y, σ, R) be metric spaces. A *continuous function* $f\colon (X, \rho, Q) \to (Y, \sigma, R)$ is given by an *approximating map*

$$h\colon Q \to \mathbb{N} \to R$$

together with further data dependent on w, r (center and radius of a ball):

(a) a map $\alpha \colon Q \to \mathbb{Z}^+ \to \mathbb{Z}^+ \to \mathbb{N}$ such that $\alpha_{w,r}(p)$ is a (uniform) *Cauchy modulus* of the *Cauchy sequence* $(h(u,n))_n$ for $\rho(u,w) \le \frac{1}{2^r}$, that is

$$\sigma(h(u,n), h(u,m)) \le \frac{1}{2^p} \quad \text{for } n, m \ge \alpha_{w,r}(p);$$

(b) a modulus $\omega \colon Q \to \mathbb{Z}^+ \to \mathbb{Z}^+ \to \mathbb{Z}^+$ of (uniform) continuity, such that $\omega_{w,r}(p)$ satisfies for $n \ge \alpha_{w,r}(p)$ and $\rho(u,w), \rho(v,w) \le \frac{1}{2^r}$

$$\rho(u,v) \le \frac{2}{2^{\omega_{w,r}(p)}} \to \sigma(h(u,n), h(v,n)) \le \frac{1}{2^p};$$

(c) maps $\gamma \colon Q \to \mathbb{Z}^+ \to R$, $\delta \colon Q \to \mathbb{Z}^+ \to \mathbb{Z}^+$ such that $\gamma(w,r)$ and $\delta(w,r)$ are center and radius of a ball containing all $h(u,n)$ (for $\rho(u,w) \le \frac{1}{2^r}$):

$$\rho(u,w) \le \frac{1}{2^r} \to \sigma(h(u,n), \gamma(w,r)) \le \frac{1}{2^{\delta(w,r)}}.$$

$\alpha, \omega, \gamma, \delta$ are supposed to have some monotonicity properties: if the ball $B(w,r)$ is contained in the ball $B(w',r')$, i.e.,

$$\rho(w,w') + \frac{1}{2^r} \le \frac{1}{2^{r'}},$$

then we require

$$\alpha(w,r,p) \le \alpha(w',r',p), \qquad \omega(w,r,p) \le \omega(w',r',p)$$

and that $B(\gamma(w,r), \delta(w,r))$ is contained in $B(\gamma(w',r'), \delta(w',r'))$, i.e.,

$$\sigma(\gamma(w,r), \gamma(w',r')) + \frac{1}{2^{\delta(w,r)}} \le \frac{1}{2^{\delta(w',r')}}.$$

Notice that a continuous function is given by objects of type level ≤ 1 only, since it suffices to define its values on Q.

Since the approximating map operates on approximations only, we need to define separately what it means to apply a continuous function in our sense to an arbitrary element of the metric space.

Definition 9.3.2. Let a continuous function $f \colon (X, \rho, Q) \to (Y, \sigma, R)$ be given (by $h, \alpha, \omega, \gamma, \delta$), and also $x = ((u_n)_n, M)$ and w, r with $\rho(u_n, w) < \frac{1}{2^r}$. The *application* $f(x)$ of f to x is defined to be the Cauchy sequence $(h(u_n, n))_n$ with modulus

$$\max(\alpha_{w,r}(p+2), M(\omega_{w,r}(p+1) - 1)).$$

Lemma 9.3.3 (MetrContAppCorr). *This is a modulus.*

Proof. Under the assumptions of the definition we have (omitting w, r)

$$\begin{aligned} &\sigma(h(u_n,n), h(u_m,m)) \\ &\le \sigma(h(u_n,n), h(u_n,l)) + \sigma(h(u_n,l), h(u_m,l)) + \sigma(h(u_m,l), h(u_m,m)) \end{aligned}$$

$$\leq \frac{1}{2^{p+2}} + \frac{1}{2^{p+1}} + \frac{1}{2^{p+2}}$$

if $n, m \geq M(\omega(p+1) - 1)$ and $l \geq \alpha(p+1)$ (for the middle term), and moreover $n, m, l \geq \alpha(p+2)$ (for the first and last term). □

Notice that $f(x)$ is independent from the assumed w, r with $x \in B(w, r)$, in the sense that $f_{c,d}(x) = f_{c',d'}(x)$. This holds since the Cauchy sequences of these two elements are the same (only their moduli may be different).

We show that application is compatible with equality.

Lemma 9.3.4 (MetrContAppCompat). *Let $f \colon (X, \rho, Q) \to (Y, \sigma, R)$ be a continuous function. Then*

$$x = y \to f(x) = f(y).$$

Proof. We may assume that $x = ((u_n)_n, M)$ and $y := ((v_n)_n, N)$ are such that $\rho(u_n, w), \rho(v_n, w) \leq \frac{1}{2^p}$ for all n. Assume $x = y$. To prove $f(x) = f(y)$ it suffices to prove $\sigma(f(x), f(y)) = 0$. This follows from

$$\sigma(h(u_n, n), h(v_n, n)) \leq \frac{1}{2^p}$$

if $n \geq \alpha(w, r, p)$ and in addition $n \geq l$ with l provided by $\omega(w, r, p) - 1$. □

Next we show that indeed a continuous function f has ω as a modulus of uniform continuity.

Lemma 9.3.5 (MetrContMod). *Let $f \colon (X, \rho, Q) \to (Y, \sigma, R)$ be a continuous function and $x, y \in B(w, r) \subseteq X$. Then*

$$\rho(x, y) \leq \frac{1}{2^{\omega(p)}} \to \sigma(f(x), f(y)) \leq \frac{1}{2^p}.$$

Proof. Again we may assume that $x = ((u_n)_n, M)$ and $y := ((v_n)_n, N)$ are such that $\rho(u_n, w), \rho(v_n, w) \leq \frac{1}{2^p}$ for all n. Assume $\rho(u_n, v_n) \leq \frac{2}{2^{\omega(w,r,p)}}$ for $n \geq l$. Then for $n \geq l, \alpha(w, r, p)$

$$\sigma(h(u_n, n), h(v_n, n)) \leq \frac{1}{2^p},$$

that is $\sigma(f(x), f(y)) \leq \frac{1}{2^p}$. □

We show that continuous functions commute with limits.

Lemma 9.3.6 (MetrContLim). *Let $(x_n)_n$ be a sequence of elements in a metric space (X, ρ, Q) which converges to y. Assume $x_n, y \in B(w, r)$ and let $f \colon (X, \rho, Q) \to (Y, \sigma, R)$ be continuous. Then $(f(x_n))_n$ converges to $f(y)$.*

Proof. For a given p, pick l such that for all n

$$l \leq n \to \rho(x_n, y) \leq \frac{1}{2^{\omega_f(p)}}.$$

Then by Lemma 9.3.5 (MetrContMod)

$$l \leq n \to \sigma(f(x_n), f(y)) \leq \frac{1}{2^p}.$$

Hence $(f(x_n))_n$ converges to $f(y)$. □

Lemma 9.3.7 (MetrContRat). *Assume that $f, g\colon (X, \rho, Q) \to (Y, \sigma, R)$ are continuous and coincide on all approximations $u \in Q$. Then $f = g$.*

Proof. Let $x = ((u_n)_n, M)$. By Lemma 9.3.6 (MetrContLim) the sequence $(f(u_n))_n$ converges to $f(x)$ and $(g(u_n))_n$ to $g(x)$. Now $f(u_n) = g(u_n)$ implies $f(x) = g(x)$. □

We define the *composition* of two continuous functions.

Definition 9.3.8. Let $f\colon (X, \rho, Q) \to (Y, \sigma, R)$ and $g\colon (Y, \sigma, R) \to (Z, \tau, S)$ be continuous functions, with f given by $h_f, \alpha_f, \omega_f, \gamma_f, \delta_f$ and g given by $h_g, \alpha_g, \omega_g, \gamma_g, \delta_g$. Assume that for $B(w, r) \subseteq Q$ we know that $B(\gamma_{cd}, \delta_{cd}) \subseteq R$ (with $\gamma_{cd} := \gamma_f(c, d)$ and $\delta_{cd} := \delta_f(c, d)$). The *composition* $g \circ f\colon (X, \rho, Q) \to (Z, \tau, S)$ is defined by

$$\begin{aligned}
h_{g\circ f}(a, n) &:= h_g(h_f(a, n), n)\\
\alpha_{g\circ f}(c, d, k) &:= \max(\alpha_g(\gamma_{cd}, \delta_{cd}, k+2), \alpha_f(c, d, \omega_g(\gamma_{cd}, \delta_{cd}, k+1) - 1))\\
\omega_{g\circ f}(c, d, k) &:= \omega_f(c, d, \omega_g(\gamma_{cd}, \delta_{cd}, k) - 1) + 1\\
\gamma_{g\circ f}(c, d) &:= \gamma_g(\gamma_{cd}, \delta_{cd})\\
\delta_{g\circ f}(c, d) &:= \delta_g(\gamma_{cd}, \delta_{cd})
\end{aligned}$$

We need to show that this indeed defines a continuous function.

Lemma 9.3.9 (MetrContComposeCorr). *Under the assumptions of the definition above and for any $c \in \mathbb{Q}$ and $d \in \mathbb{Q}^+$ with $B(c, d) \cap \mathbb{Q} \subseteq X$ we have*

(a) *Each $h_{g\circ f}(a, n)$ (for $a \in B(c, d) \cap \mathbb{Q}$) is a Cauchy sequence with (uniform) modulus $\alpha_{g\circ f}(c, d, k)$*

(b) *$\omega_{g\circ f}(c, d, k)$ satisfies for all $a, b \in B(c, d) \cap \mathbb{Q}$*

$$|a - b| \leq 2^{-\omega_{g\circ f}(c,d,k)+1} \to |h(a, n) - h(b, n)| \leq \frac{1}{2^p} \quad \text{for } n \geq \alpha_{g\circ f}(c, d, k);$$

(c) *$\gamma_{g\circ f}(c, d)$ and $\delta_{g\circ f}(c, d)$ are center and radius of a ball containing all $h_{g\circ f}(a, n)$ ($a \in B(c, d) \cap \mathbb{Q}$).*

PROOF. (a). $(h_f(a,n))_n$ is a real with modulus α_f. For $a \in B(c,d) \cap \mathbb{Q}$ we have $\gamma_{cd} < h_f(a,n) < \delta_{cd}$. By Lemma 4.1.9 (ContReal), application of g to this real gives the Cauchy sequence $(h_g(h_f(a,n),n))_n$ with Cauchy modulus $\alpha_{g \circ f}(k)$.

(b). Let $a, b \in B(c,d) \cap \mathbb{Q}$ and $|a - b| \leq 2^{\omega_f(c,d,\omega_g(\gamma_{cd},\delta_{cd},k)-1)+1}$. Then $|h_f(a,n) - h_f(b,n)| \leq 2^{-\omega_g(\gamma_{cd},\delta_{cd},k)+1}$ for $n \geq \alpha_f(c,d,\omega_g(\gamma_{cd},\delta_{cd},k)-1)$, and therefore $|h_g(h_f(a,n),n) - h_g(h_f(b,n),n)| \leq \frac{1}{2^p}$ if also $n \geq \alpha_g(\gamma_{cd},\delta_{cd},k)$. Both conditions on n hold for $n \geq \alpha_{g \circ f}(c,d,k)$, since α_g, α_f and ω_g are weakly increasing.

(c). Because of $h_f(a,n) \in B(\gamma_{cd},\delta_{cd})$ we have

$$h_g(h_f(a,n),n) \in B(\gamma_g(\gamma_{cd},\delta_{cd}), \gamma_g(\gamma_{cd},\mu_{cd})). \qquad \square$$

REMARK 9.3.10. Under the assumptions of the definition above we have $(g \circ f)(x) = g(f(x))$ for all x, since both points have the same Cauchy sequence.

9.4. Totally bounded sets

DEFINITION 9.4.1. Let (X, ρ, Q) be a metric space. A subset V of Q is *totally bounded*[3] if for every positive integer p there is a finite list $(u_n)_{n<m}$ of elements of V such that for every $u \in V$ there is an $n < m$ with $\rho(u, u_n) \leq \frac{1}{2^p}$. We call the list $(u_n)_{n<m}$ an $\frac{1}{2^p}$*-net* for V.

From a computational point of view, a problem with the notion of total boundedness is that for a given accuracy $\frac{1}{2^p}$ it requires to work with huge lists of approximations. In contrast, the notion of locatedness allows usage of the trisection method (as in Theorem 9.2.2) and hence is much preferable. It is therefore a relief that inside a totally bounded set (for instance $[c,d]^n$ in $\mathbb{R}^n$) the notions of total boundedness and locatedness coincide. – In the next two propositions we fix a metric space (X, ρ, Q).

PROPOSITION 9.4.2. *Every totally bounded set V is located.*

PROOF. Given $u \in Q$ and rationals $c, d \in \mathbb{Q}$ with $c < d$ we must show

$$\forall_{v \in V}(c \leq \rho(u,v)) \vee \exists_{v \in V}(\rho(u,v) \leq d).$$

Pick p with $c + \frac{1}{2^p} < d$, and $(v_n)_{n<m}$ in V such that

$$\forall_{v \in V} \exists_{n<m} (\rho(v, v_n) \leq \frac{1}{2^p}).$$

[3] The notion of total boundedness is due to Brouwer. It replaces the classical notion of compactness in the present constructive setting.

Case $\forall_{n<m}(c + \frac{1}{2^p} \leq \rho(u, v_n))$. Let $v \in V$. Pick n with $\rho(v, v_n) \leq \frac{1}{2^p}$. Then

$$c + \rho(v, v_n) \leq c + \frac{1}{2^p} \leq \rho(u, v_n) \leq \rho(u, v) + \rho(v, v_n).$$

Thus $c \leq \rho(u, v)$. Since $v \in V$ was arbitrary, the l.h.s. of the alternative holds. *Case* $\exists_{n<m}(\rho(u, v_n) \leq d)$. Then the r.h.s. of the alternative holds. □

PROPOSITION 9.4.3. *Every located subset V of a totally bounded set U is totally bounded.*

PROOF. Fix p. Since V is located we have for every $u \in U$

$$\forall_{v \in V}(\frac{1}{2^{p+1}} \leq \rho(u, v)) \vee \exists_{v \in V}(\rho(u, v) \leq \frac{3}{2^{p+2}}).$$

From the $\frac{1}{2^{p+2}}$-net *us* for the totally bounded U select the sublist $(u_n)_{n<m}$ of all those approximations which satify the r.h.s. of the alternative. Thus for every u_n we have a v_n such that $\rho(u_n, v_n) \leq \frac{3}{2^{p+2}}$. Now consider an arbitrary $v \in V$. Since U is totally bounded we have some u in *us* with $\rho(v, u) \leq \frac{1}{2^{p+2}}$. Thus this u is some u_n. Consider the associated $v_n \in V$. Then

$$\rho(v, v_n) \leq \rho(v, u_n) + \rho(u_n, v_n) \leq \frac{1}{2^{p+2}} + \frac{3}{2^{p+2}} = \frac{1}{2^p}.$$ □

CHAPTER 10

Normed spaces

Important examples of metric spaces are normed (linear) spaces. We only consider linear spaces over $\mathbb{R}$.

10.1. Groups

DEFINITION 10.1.1. A *group* $(G, \circ)$ (or just G) is given by a map $\circ\colon G \to G \to G$ with the following properties.

(1) $(u \circ v) \circ w = u \circ (v \circ w)$ for all $u, v, w \in G$ (*associativity*).
(2) There is $e \in G$ (called *neutral element* of G) such that
  (a) $e \circ u = u$ for all $u \in G$;
  (b) for every $u \in G$ there is an $u' \in G$ such that $u' \circ u = e$ (u' is called *inverse element* for u).

G is called *abelean*, if in addition $u \circ v = v \circ u$ holds for all $u, v \in G$ (*commutativity*).

LEMMA 10.1.2 (LeftInvRightInv). *Let $e \in G$ be a neutral element and $u, u' \in G$ such that $u' \circ u = e$. Then we also have $u \circ u' = e$.*

PROOF. Choose u'' with $u'' \circ u' = e$. Then

$$u \circ u' = e \circ u \circ u' = u'' \circ \underbrace{u' \circ u}_{e} \circ u' = u'' \circ u' = e. \qquad \square$$

LEMMA 10.1.3 (LeftNeutralRightNeutral). *Let $e \in G$ be a neutral element. Then $u \circ e = u$ for all $u \in G$.*

PROOF. $u \circ e = u \circ u' \circ u = e \circ u = u$ by LeftInvRightInv. □

LEMMA 10.1.4 (NeutralUnique). *There is exactly one neutral element $e \in G$.*

PROOF. Let e, e^* be neutral elements of G. Then $e^* = e \circ e^* = e$ by LeftNeutralRightNeutral. □

LEMMA 10.1.5 (InvUnique). *For every $u \in G$ there is exactly one inverse element $u' \in G$.*

Proof. Let u', u^* be inverse elements for $u \in G$. Then

$$u^* = u^* \circ e = u^* \circ u \circ u' = e \circ u = u'.$$

Here we have used the three previous lemmata. □

The uniquely determined inverse element for u is denoted u^{-1}.

Lemma 10.1.6 (ZeroCrit). $u + u = 0 \to u = 0$.

Proof. $u = 0 + u = (-u + u) + u = -u + (u + u) = -u + u = 0$. □

10.2. Linear spaces

We define linear spaces over the field $\mathbb{R}$ of reals.

Definition 10.2.1. A *linear space* $(X, +, \cdot)$ (or just X) is given by two maps $+\colon X \to X \to X$ (addition) and $\cdot\colon \mathbb{R} \to X \to X$ (scalar multiplication) with the following properties.

(1) $(X, +)$ is an abelean group.
(2) For all $u, v \in X$ and all $x, y \in \mathbb{R}$ we have

$$\begin{aligned} (x+y)u &= (xu) + (yu), \\ x(u+v) &= xu + xv, \\ x(yu) &= (xy)u, \\ 1u &= u. \end{aligned}$$

The elements of X are called *vectors*, and the elements of $\mathbb{R}$ *scalars*.

As usual we write $u - v$ for $u + (-v)$ and nu for $u + u + \cdots + u$ (with n occurrences of u).

Lemma 10.2.2 (TimesZeroLeft). $0u = u$.

Proof. $0u = (0 + 0)u = 0u + 0u$. □

Lemma 10.2.3 (TimesZeroRight). $x0 = 0$.

Proof. $x0 = x(0 + 0) = x0 + x0$. □

Lemma 10.2.4 (RealInvTimes). *If $yx = 1$ and $xu = 0$, then $u = 0$.*

Proof. $u = 1u = (yx)u = y(xu) = y0 = 0$. □

Lemma 10.2.5 (TimesMinusOne). $(-1)u = -u$.

Proof. $u + (-1)u = 1u + (-1)u = (1 + (-1))u = 0u = 0$. □

Definition 10.2.6. Let X and Y be linear spaces. A map $f\colon X \to Y$ is *linear* if for all $u, v \in X$ and all $x \in \mathbb{R}$

$$\begin{aligned} f(u+v) &= f(u) + f(v), \\ f(xu) &= xf(u). \end{aligned}$$

Lemma 10.2.7 (LinZero). *$f(0) = 0$.*

Proof. $f(0) = f(0u) = 0f(u) = 0$. □

Lemma 10.2.8 (LinTimes). *$f(-u) = -f(u)$.*

Proof. $f(-u) + f(u) = f(-u + u) = f(0) = 0$. □

10.3. Normed spaces

Definition 10.3.1. A *norm* on a linear space X is a map $\| \cdot \| \colon X \to \mathbb{R}$ such that for all $x \in \mathbb{R}$ and $u, v \in X$

(a) $\|xu\| = |x| \cdot \|u\|$,
(b) $\|u\| = 0 \to u = 0$, and
(c) $\|u + v\| \leq \|u\| + \|v\|$ (triangle inequality).

Lemma 10.3.2 (NormZero). *$\|0\| = 0$.*

Proof. $\|0\| = \|0 \cdot 0\| = |0| \cdot \|0\| = 0$. □

Lemma 10.3.3 (NormInv). *$\| - u\| = \|u\|$.*

Proof. $\| - u\| = \|(-1) \cdot u\| = | - 1| \cdot \|u\| = \|u\|$. □

From a norm on X one can define a metric by $\rho(u, v) := \|u - v\|$.

Lemma 10.3.4 (NormMetric). *$\|u - w\| \leq \|u - v\| + \|v - w\|$.*

Proof. $\|u - w\| = \|(u - v) + (v - w)\| \leq \|u - v\| + \|v - w\|$. □

A complete normed space is called *Banach space*.

A linear map $f \colon X \to Y$ between normed spaces is *strongly extensional* when $0 < \|u\|$ implies $0 < \|f(u)\|$. It has been shown in Bridges and Ishihara (1990) that every linear map from a Banach space X to a normed space Y is strongly extensional. This was obtained as a corollary of a more general result. Here we give a direct proof, due to Hannes Diener.

Lemma 10.3.5. *Every linear map from a Banach space X to a normed space Y is strongly extensional.*

Proof. Let $u \in X$ be such that $0 < \|f(u)\|$. We show $0 < \|u\|$. Fix an increasing binary sequence (λ_n) such that

$$\|u\| < \frac{1}{2^n(n+1)} \quad \text{if } \lambda_n = 0,$$
$$\frac{1}{2^{n+1}(n+1)} < \|u\| \quad \text{if } \lambda_n = 1.$$

Define a sequence (v_n) in X by

$$v_n := \begin{cases} 0 & \text{if } \lambda_n = 0, \\ mu & \text{if } \lambda_n = 1 \text{ and m is the minimal such number.} \end{cases}$$

Then (v_n) is a Cauchy sequence.

Let v be its limit in X. Pick n_0 such that $\|f(v)\| \leq n_0 \|f(u)\|$. Assume that there is some $n > n_0$ with $\lambda_n = 1$. Then $v = v_m = mu$ where m is minimal such that $\lambda_m = 1$. Hence

$$n\|f(u)\| > \|f(v)\| = m\|f(u)\| > n\|f(u)\|$$

a contradiction. Hence all λ_n would be zero and therefore $u = 0$. But then $f(u) = 0$ by the linearity of f, contradicting our assumption $0 < \|f(u)\|$. $\square$

LEMMA 10.3.6 (Ishihara's trick). *Let f be a linear map from a Banach space X into a normed space Y, and let (u_n) be a sequence in X converging to 0. Then for $0 < a < b$ either $a \leq \|fu_n\|$ for some n or $\|fu_n\| \leq b$ for all n.*

PROOF. Let M be a modulus of convergence of (u_n) to 0; we can assume $M0 = 0$. Call m a *hit* on n if $M_n \leq m < M_{n+1}$ and $a \leq \|fu_m\|$. Our first goal is to define a function $h\colon \mathbb{N} \to \mathbb{N}$ such that

(i) $h_n = 0$ if for all $n' \leq n$ there is no hit;
(ii) $h_n = m + 2$ if at n for the first time we have a hit, with m;
(iii) $h_n = 1$ if there is an $n' < n$ with a hit.

We will need the bounded least number operator $\mu_n g$ defined recursively as follows. Here g is a variable of type $\mathbb{N} \to \mathbb{B}$.

$$\mu_0 g := 0,$$

$$\mu_{Sn} g := \begin{cases} 0 & \text{if } g0 \\ S\mu_n(g \circ S) & \text{otherwise.} \end{cases}$$

From $\mu_n g$ we define

$$\mu_{n_0}^n g := \begin{cases} (\mu_{n-n_0} \lambda_m g(m + n_0)) + n_0 & \text{if } n_0 \leq n \\ 0 & \text{otherwise.} \end{cases}$$

To define h we will make use of a function g of type $\mathbb{N} \to \mathbb{B}$ (to be defined from `cApproxSplit`) such that

$$\begin{cases} a \leq \|fu_m\| & \text{if } gm \\ \|fu_m\| \leq b & \text{otherwise.} \end{cases}$$

Then we can define $h_n := H(g, M, n)$ where

$$H(g, M, n) := \begin{cases} 0 & \text{if } M_n \leq \mu_{M_n} g \text{ and } M_{n+1} \leq \mu_{M_n}^{M_{n+1}} g \\ \mu_{M_n}^{M_{n+1}} g + 2 & \text{if } M_n \leq \mu_{M_n} g \text{ and } \mu_{M_n}^{M_{n+1}} g < M_{n+1} \\ 1 & \text{if } \mu_{M_n} g < M_n. \end{cases}$$

To avoid multiple computations we split this definition into parts.

$$\begin{aligned} H(g, M, n) &:= \operatorname{HitPast}(g, M, M_n, n) \\ \operatorname{HitPast}(g, M, n_0, n) &:= \begin{cases} 1 & \text{if } \mu_{n_0} g < n_0 \\ \operatorname{HitHere}(g, n_0, M_{n+1}) & \text{otherwise} \end{cases} \\ \operatorname{HitHere}(g, n_0, n_1) &:= \operatorname{Hit}(\mu_{n_0}^{n_1} g, n_1) \\ \operatorname{Hit}(m, n) &:= \begin{cases} m + 2 & \text{if } m < n \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Clearly h has the properties listed above.

The next goal is to define from h a sequence (v_n) in X such that

(i) $v_n = 0$ if $h_n = 0$;
(ii) $v_n = nu_m$ if $h_n = m + 2$;
(iii) $v_n = v_{n-1}$ if $h_n = 1$.

Let ξ be the type of elements of X, and us a variable of type $\mathbb{N} \to \xi$. Define $v_n := V_\xi(g, M, us, n)$ where (writing u_m for $us(m)$)

$$V_\xi(g, M, us, n) := \begin{cases} 0 & \text{if } H(g, M, n) = 0 \\ nu_m & \text{if } H(g, M, n) = m + 2 \\ 0 \quad \text{(arbitrary)} & \text{if } H(g, M, n) = 1 \text{ and } n = 0 \\ V_\xi(g, M, us, n-1) & \text{if } H(g, M, n) = 1 \text{ and } n > 0. \end{cases}$$

Again we split the definition to avoid recomputations.

$$V_\xi(g, M, us, n) := \operatorname{Seq}_\xi(H(g, M), H(g, M, n), us, n)$$

$$\begin{aligned} \operatorname{Seq}_\xi(h, 0, us, n) &:= 0 \\ \operatorname{Seq}_\xi(h, m + 2, us, n) &:= nu_m \\ \operatorname{Seq}_\xi(h, 1, us, 0) &:= 0 \quad \text{(arbitrary)} \\ \operatorname{Seq}_\xi(h, 1, us, n + 1) &:= \operatorname{Seq}_\xi(h, h_n, us, n). \end{aligned}$$

One can show that (v_n) has the properties listed above.

Next we show that (v_n) is a Cauchy sequence with modulus $N(k) := 2k + 1$, which satisfies

$$\frac{N(k) + 1}{2^{N(k)}} \leq \frac{1}{2^k}.$$

Since our goal is stable, we may employ arbitrary case distinctions.

Case 1. There is no hit. Then h_n is always 0, hence (v_n) is identically zero and therefore a Cauchy sequence with any modulus.

Case 2. Assume that there is a hit. Let n be the one, say with value m. Given k, let $N(k) < n_1 < n_2$. We show $\|v_{n_1} - v_{n_2}\| \le 1/2^k$. If $n \le n_1$ or $n_2 < n$, then $v_{n_1} = v_{n_2}$ and we are done. Assume $n_1 < n \le n_2$. Then $v_{n_1} = 0$, hence $\|v_{n_1} - v_{n_2}\| = \|v_{n_2}\| = \|v_n\|$. By definition $v_n = nu_m$, hence

$$\begin{aligned}\|v_n\| &= n\|u_m\| \\ &\le (n+1)/2^n \quad \text{since } M_n \le m \text{ and } M \text{ is a modulus for } (u_n) \\ &\le (N(k)+1)/2^{N(k)} \quad \text{since } n \mapsto (n+1)/2^n \text{ is monotone} \\ &\le 1/2^k.\end{aligned}$$

By the assumed completeness of X we have a limit v of (v_n). Pick n_0 such that $\|fv\| \le n_0 a$. Assume that there is a first hit at some $n > n_0$, with value m. Then $v = v_n = nu_m$ and

$$na \le n\|fu_m\| = \|n(fu_m)\| = \|f(nu_m)\| = \|fv\| \le n_0 a < na,$$

a contradiction. Hence beyond this n_0 we cannot have a first hit.

If $\forall_{n \le n_0} h_n = 0$ then there is no hit at all and we have $\|fu_n\| \le b$ for all n. Otherwise there is a hit before n_0 and we have $a \le \|fu_n\|$ for some n. □

The computational content machine extracted from this proof is

```
[f,us,M,a,a0,k]
 [let g
   ([n]negb(cAC([n0]cApproxSplitBooleRat
                     a a0 lnorm(f(us n0))k)n))
   [case (H g M
          (cRealPosRatBound
           lnorm(f((cXCompl xi)
                    ((V xi)g M us)
                    ([k0]abs(IntS(2*k0)max 0))))
           a))
    (Zero -> False)
    (Succ n -> True)]]
```

Here H and V are the functionals defined above. `cAC` is the computational content of the axiom of choice

```
(pp "AC")
all m ex boole (Pvar nat boole)^ m boole ->
ex g all m (Pvar nat boole)^ m(g m)
```

and hence the identity. `cApproxSplitBooleRat` and `cRealPosRatBound` are the computational content of lemmata

```
all a,b,x,k(Real x -> 1/2**k<=b-a ->
 ex boole((boole -> x<<=b) andu ((boole -> F) -> a<<=x)))

all x,a(Real x -> 0<a -> ex n x<<=n*a)
```

In our formulation of Ishihara's trick we have used the "decorated" disjunction $\vee^{\mathrm{u}}$ (u for uniform) to express the final alternative. This means that the computational content of the lemma returns just a boolean, expressing which side of the disjunction holds, but not returning a witness for the existential quantifier in the left hand side, $\exists_n a \leq \|fu_n\|$. We can change this and use the "left" disjunction $\vee^{\mathrm{l}}$ instead. Then literally the same proof works. However, in the extracted term a subterm starting with cRealPosRatBound occurs twice. We take it out by introducing a second "let", via another use of the identity lemma at the point in the proof where the existence of this bound is proved. The extracted term then is

```
[f,us,M,a,a0,k]
 [let g
   ([n]negb(cAC([n0]cApproxSplitBooleRat
                    a a0 lnorm(f(us n0))k)n))
   [let n
    (cRealPosRatBound
    lnorm(f((cXCompl xi)
            ((V xi)g M us)
            ([k0]abs(IntS(2*k0)max 0))))
    a)
    [case (H g M n)
     (Zero -> (DummyR nat))
     (Succ n0 -> Inl right(cHFind g M n))]]]
```

Note that the required witness is obtained by an application of `cHFind`, the computational content of a lemma `HFind`:

```
(pp "HFind")
all g,M,n(M Zero=Zero -> (H g M n=Zero -> F) ->
 ex n0,m(n0<=n & H g M n0=m+2))
```

Lemma 10.3.7 (Ishihara's second trick). *Let f be a linear map from a Banach space X into a normed space Y, and let (u_n) be a sequence in X converging to 0. Then for $0 < a < b$*

$$\textit{either } \forall_n \exists_{m \geq n} a \leq \|fu_m\| \textit{ or } \exists_n \forall_{m \geq n} \|fu_m\| \leq b.$$

Proof. By the previous lemma applied to the subsequences $(u_{m+n})_n$ we have functions g of type $\mathbb{N} \to \mathbb{B}$ and $h' \colon \mathbb{N} \to \mathbb{N}$ such that for all m

$$\begin{cases} \forall_n \|fu_{m+n}\| \leq b & \text{if } gm \\ a \leq \|fu_{m+h'_m}\| & \text{otherwise.} \end{cases}$$

If $g0$ then $\forall_n \|fu_n\| \leq b$ and the claim holds. Now let $g0$ be false. We say the we have a *hit* at n with value m if $M_n \leq m < M_{n+1}$ and gm. As in the proof of the previous lemma we define a function $h \colon \mathbb{N} \to \mathbb{N}$ such that

(i) $h_n = 0$ if for all $n' \leq n$ there is no hit;
(ii) $h_n = m + 2$ if at n for the first time we have a hit, with value m;
(iii) $h_n = 1$ if there is an $n' < n$ with a hit.

From h define a sequence (v_n) in X by

(i) $v_n = 0$ if $h_n = 0$;
(ii) $v_n = nu_{m-1+h'_{m-1}}$ if $h_n = m + 2$;
(iii) $v_n = 0$ (arbitrary) if $h_n = 1$ and $n = 0$;
(iv) $v_n = v_{n-1}$ if $h_n = 1$ and $0 < n$.

Again as before one shows that (v_n) is a Cauchy sequence with modulus $N(k) := 2k + 1$. By the assumed completeness of X we have a limit v of (v_n). Pick n_0 such that $\|fv\| \leq n_0 a$. Assume that there is a first hit at some $n > n_0$, with value m. Because of gm we have $0 < m$. Then $g(m-1)$ is false, and $v = nu_{m-1+h'_{m-1}}$. Hence

$$na \leq n\|fu_{m-1+h'_{m-1}}\| = \|f(nu_{m-1+h'_{m-1}})\| = \|fv\| \leq n_0 a < na,$$

a contradiction. Hence beyond n_0 we cannot have a first hit.

If $\forall_{n \leq n_0} h_n = 0$ then there is no hit at all and hence $a \leq \|fu_{m+h'_m}\|$ for all m, i.e., $\forall_n \exists_{m \geq n} a \leq \|fu_m\|$. Otherwise there is a hit, say n with value m. Then gm, hence $\forall_n \|fu_{m+n}\| \leq b$ and therefore $\exists_n \forall_{m \geq n} \|fu_m\| \leq b$. □

CHAPTER 11

Ordinary differential equations

11.1. The Cauchy-Euler approximation method

We consider a differential equation

$$y' = f(x, y), \tag{28}$$

where $f\colon D \to \mathbb{R}$ is a continuous function on some subset D of $\mathbb{R}^2$. A solution of (28) on an interval I is a function $\varphi\colon I \to \mathbb{R}$ with a continuous derivative φ' such that for all $x \in I$

$$(x, \varphi(x)) \in D \text{ (hence } f(x, \varphi(x)) \text{ is defined) and } \varphi'(x) = f(x, \varphi(x)).$$

We want to construct approximate solutions to (28). Let $f\colon D \to \mathbb{R}$ be continuous, and consider an interval I. A function $\varphi\colon I \to \mathbb{R}$ is an *approximate solution up to the error* $\frac{1}{2^p}$ of (28) if

(a) φ is *admissible*, i.e., $(x, \varphi(x)) \in D$ for $x \in I$.
(b) φ is continuous.
(c) φ has a piecewise continuous derivative on I.
(d) $|\varphi'(x) - f(x, \varphi(x))| \le \frac{1}{2^p}$ for all $x \in I$ where $\varphi'(x)$ is defined.

Notice that we only required the differential equation (28) to be satisfied up to the error $\frac{1}{2^p}$. Later we shall see that under certain conditions which guarantee a unique exact solution, every approximate solution differs from the exact one by a constant multiple of its error.

Theorem 11.1.1 (Cauchy-Euler approximation). *Let $f\colon D \to \mathbb{R}$ be continuous, and $(a_0, b_0) \in D$ such that the rectangle R given by $|x - a_0| \le a$, $|y - b_0| \le b$ is in D. Assume $|f(x, y)| \le M$ for $(x, y) \in R$, and let $h := \min(a, b/M)$. Then for every $p \in \mathbb{Z}^+$ we can construct an approximate solution $\varphi\colon [a_0 - h, a_0 + h] \to \mathbb{R}$ of (28) up to the error $\frac{1}{2^p}$ such that $\varphi(a_0) = b_0$.*

Proof. By definition of h, the rectangle

$$S\colon\ |x - a_0| \le h,\ |y - b_0| \le Mh$$

is in D. Since f is continuous, it comes with a modulus of (uniform) continuity. Hence for our given p we have $q \in \mathbb{Z}^+$ such that

$$|f(\tilde{x}, \tilde{y}) - f(x, y)| \leq \frac{1}{2^{p+1}} \tag{29}$$

for $(\tilde{x}, \tilde{y}), (x, y)$ in D and $|\tilde{x} - x|, |\tilde{y} - y| \leq \frac{1}{2^q}$.

We now divide the interval $[a_0, a_0 + h]$ such that

(a) $a_0 < a_1 < \cdots < a_{n-1} < a_n = a_0 + h$
(b) $a_i - a_{i-1} \leq \min(\frac{1}{2^q}, \frac{1}{2^q}/M)$ for $i = 1, \dots, n$

and construct an approximate solution on $[a_0, a_0 + h]$; similarly this can be done on $[a_0 - h, a_0]$.

The idea is to start at (a_0, b_0) and draw a line with slope $f(a_0, b_0)$ until it intersects $x = a_1$, say at (a_1, b_1), then starting from (a_1, b_1) draw a line with slope $f(a_1, b_1)$ until it intersects $x = a_2$, say at (a_2, b_2), etc. Since we want an approximate solution which maps rationals to rationals, we approximate the slopes $f(a_{i-1}, b_{i-1})$ by rationals s_{i-1}.

More precisely, we recursively define for $i = 1, \dots, n$

$$\begin{aligned}
&\varphi(x) = b_{i-1} + (x - a_{i-1})s_{i-1} \quad \text{for } a_{i-1} \leq x \leq a_i,\\
&b_i := \varphi(a_i),\\
&-M \leq s_{i-1} \leq M \quad \text{such that } |s_{i-1} - f(a_{i-1}, b_{i-1})| \leq \tfrac{1}{2^{p+1}}.
\end{aligned}$$

Clearly φ is continuous, admissible and has piecewise derivatives

$$\varphi'(x) = s_{i-1} \quad \text{for } a_{i-1} < x < a_i.$$

Now for $a_{i-1} < x < a_i$ we have $|x - a_{i-1}| \leq \frac{1}{2^q}$ and

$$|\varphi(x) - b_{i-1}| \leq |x - a_{i-1}| \cdot |s_{i-1}| \leq \frac{\frac{1}{2^q}}{M} \cdot M = \frac{1}{2^q},$$

hence by (29)

$$\begin{aligned}
|\varphi'(x) - f(x, \varphi(x))| &= |s_{i-1} - f(x, \varphi(x))|\\
&\leq |s_{i-1} - f(a_{i-1}, b_{i-1})| + |f(a_{i-1}, b_{i-1}) - f(x, \varphi(x))|\\
&\leq \frac{1}{2^{p+1}} + \frac{1}{2^{p+1}} = \frac{1}{2^p}.
\end{aligned}$$

Hence φ is an approximate solution up to the error $\frac{1}{2^p}$. □

The approximate solutions we have constructed are *rational polygons*, i.e., piecewise differentiable continuous functions with rational corners and rational slopes.

Lemma 11.1.2 (Rational polygons). *Given a rational polygon φ on $[a, b]$ and $c \in [a, b]$. Then one of the following alternatives will hold.*

(a) *$0 \leq \varphi(x)$ for $a \leq x \leq b$, or*

(b) $\varphi(x) \leq 0$ *for* $a \leq x \leq b$, *or*

(c) *there is a* $d < c$ *such that* $\varphi(d) = 0$ *and either* $0 \leq \varphi(x)$ *for* $d \leq x \leq c$, *or else* $\varphi(x) \leq 0$ *for* $d \leq x \leq c$.

Proof. Let $a = a_0 < a_1 < \cdots < a_n = b$ be the exception points for φ. We can locate c in $a = a_0 < a_1 < \cdots < a_n = b$. Pick i maximal such that $a_{i-1} \leq c$. Compare $\varphi(a_0)$, $\varphi(a_1)$, $\varphi(a_{i-1})$ and $\varphi(c)$. If all have the same sign, we are done. Otherwise pick j maximal such that $\varphi(a_j)$ and $\varphi(a_{j+1})$ (or $\varphi(c)$, respectively) change sign. Then we are done as well. □

11.2. The fundamental inequality

We need an additional restriction in order to estimate the difference of approximate solutions. A function $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^2$ is said to satisfy a *Lipschitz condition* w.r.t. its second argument for the constant $L > 0$, if for every $(x, y_1), (x, y_2) \in D$

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|.$$

We begin by giving an easy estimate for the solutions of linear differential inequalities.

Lemma 11.2.1 (LinDiffIneq). *Let* $\sigma\colon [a, b] \to \mathbb{R}$ *be continuous with a piecewise continuous derivative* σ' *such that*

$$\sigma'(x) \leq L\sigma(x) + \varepsilon$$

for all $x \in [a, b]$ *where* σ' *is defined. Then*

$$\sigma(x) \leq e^{L(x-a)}\sigma(a) + \frac{\varepsilon}{L}\bigl(e^{L(x-a)} - 1\bigr) \quad \textit{for all } x \in [a, b].$$

Proof. Since $\sigma'(x) \leq L\sigma(x) + \varepsilon$ we have

$$\int_a^x e^{-Lt}\bigl(\sigma'(t) - L\sigma(t)\bigr)\,dt \leq \varepsilon \int_a^x e^{-Lt}\,dt.$$

The integrand on the left-hand side has finitely many discontinuities but a continuous indefinite integral, so

$$\begin{aligned}
&\bigl[e^{-Lt}\sigma(t)\bigr]_a^x \leq \varepsilon\bigl[-\frac{1}{L}e^{-Lt}\bigr]_a^x \\
&e^{-Lx}\sigma(x) - e^{-La}\sigma(a) \leq \frac{\varepsilon}{L}\Bigl(e^{-La} - e^{-Lx}\Bigr) \\
&\sigma(x) \leq e^{L(x-a)}\sigma(a) + \frac{\varepsilon}{L}\bigl(e^{L(x-a)} - 1\bigr),
\end{aligned}$$

which is the required inequality. □

Next we give an estimate on the differences of approximate solutions, provided the differential equation (28) satisfies a Lipschitz condition.

LEMMA 11.2.2 (LipDiffApprox). *Let $f\colon D \to \mathbb{R}$ be continuous, and satisfy a Lipschitz condition w.r.t. its second argument for the constant $L > 0$. Let*

$$\varphi, \psi\colon [a, b] \to \mathbb{R}$$

be approximate solutions up to the error $\frac{1}{2^p}, \frac{1}{2^q}$ of (28)*. Then*

$$\psi'(x) - \varphi'(x) \leq L|\psi(x) - \varphi(x)| + \varepsilon$$

with $\varepsilon := \frac{1}{2^p} + \frac{1}{2^q}$.

PROOF. For all points except finitely many we have

$$\bigl|\varphi'(x) - f(x, \varphi(x))\bigr| \leq \frac{1}{2^p} \quad \text{and} \quad \bigl|\psi'(x) - f(x, \psi(x))\bigr| \leq \frac{1}{2^q},$$

hence with $\varepsilon := \frac{1}{2^p} + \frac{1}{2^q}$

$$\psi'(x) - \varphi'(x) \leq \bigl|f(x, \psi(x)) - f(x, \varphi(x))\bigr| + \varepsilon \leq L\bigl|\psi(x) - \varphi(x)\bigr| + \varepsilon$$

by the Lipschitz condition. □

THEOREM 11.2.3 (Fundamental inequality). *Let $f\colon D \to \mathbb{R}$ be continuous, and satisfy a Lipschitz condition w.r.t. its second argument for the constant $L > 0$. Let*

$$\varphi, \psi\colon [a, b] \to \mathbb{R}$$

be approximate solutions up to the error $\frac{1}{2^p}, \frac{1}{2^q}$ of (28)*. Then for all $x \in [a, b]$*

$$\bigl|\psi(x) - \varphi(x)\bigr| \leq e^{L(x-a)}\bigl|\psi(a) - \varphi(a)\bigr| + \frac{\frac{1}{2^p} + \frac{1}{2^q}}{L}\bigl(e^{L(x-a)} - 1\bigr).$$

PROOF. Let $\sigma := \psi - \varphi$ and $x \in [a, b]$. We may assume $0 \leq \sigma(x)$. We distinguish cases as to whether

$$\forall_{y\in[a,x]}(0 \leq \sigma(y)) \mathbin{\tilde{\vee}} \tilde{\exists}_{y\in[a,x]}(\sigma(y) < 0).$$

By Appendix A this case distinction is possible in our constructive setting, since the goal is an inequality between real numbers and hence stable.

Case (a): $\forall_{y\in[a,x]}(0 \leq \sigma(y))$. Then by LipDiffApprox we have

$$\sigma'(x) \leq L\sigma(x) + \varepsilon \quad (\textit{not } |\sigma(x)|)$$

with $\varepsilon = \frac{1}{2^p} + \frac{1}{2^q}$. Now we can use LinDiffIneq.

Case (b): $\tilde{\exists}_{y\in[a,x]}(\sigma(y) < 0)$. Let $y \in [a, x]$ and assume $\sigma(y) < 0$. Let n be arbitrary. By Theorem 4.3.2 (LastApproxZero) we have $c \in [a, b]$ with

$$\sigma(c) \leq \frac{1}{2^n} \frac{1}{e^{L(x-c)}}$$

and $0 \leq \sigma(z)$ for all $z \in [c, b]$. Now we argue as in (a). By LipDiffApprox

$$\sigma'(x) \leq L\sigma(x) + \varepsilon$$

on $[c,x]$, with $\varepsilon = \frac{1}{2^p} + \frac{1}{2^q}$. By LinDiffIneq

$$\sigma(y) \leq e^{L(y-c)}\sigma(c) + \frac{\varepsilon}{L}\bigl(e^{L(y-c)} - 1\bigr)$$

for all $y \in [c,x]$, hence also for x. By the estimate for $\sigma(c)$ above we obtain

$$\sigma(x) \leq \frac{1}{2^n} + \frac{\varepsilon}{L}\bigl(e^{L(x-a)} - 1\bigr).$$

Since n is arbitrary, the required inequality is a consequence. □

11.3. Uniqueness

To prove uniqueness of solutions we again need the Lipschitz condition.

THEOREM 11.3.1 (Uniqueness). *Let $f\colon D \to \mathbb{R}$ be continuous, and satisfy a Lipschitz condition w.r.t. its second argument. Let*

$$\varphi, \psi\colon I \to \mathbb{R}$$

be two (exact) solutions of (28). *If $\varphi(a) = \psi(a)$ for some $a \in I$, then $\varphi(x) = \psi(x)$ for all $x \in I$.*

PROOF. We show $\varphi(x) = \psi(x)$ for $a \leq x \leq a + \frac{1}{2L}$, $x \in I$. Similarly this can be shown for $a - \frac{1}{2L} \leq x \leq a$; hence the claim follows.

Integrating the two equations

$$\varphi'(x) = f(x, \varphi(x)) \quad \text{and} \quad \psi'(x) = f(x, \psi(x))$$

we obtain from $\varphi(a) = \psi(a)$

$$\varphi(x) - \psi(x) = \int_a^x \bigl(f(t, \varphi(t)) - f(t, \psi(t))\bigr)\, dt$$

and hence

$$|\varphi(x) - \psi(x)| \leq \int_a^x \bigl|f(t, \varphi(t)) - f(t, \psi(t))\bigr|\, dt \leq L \int_a^x \bigl|\varphi(t) - \psi(t)\bigr|\, dt.$$

Let M be the supremum of the range of $|\varphi - \psi|$ on $[a, a + \frac{1}{2L}]$. Then for $a \leq x \leq a + \frac{1}{2L}$

$$|\varphi(x) - \psi(x)| \leq L(x-a)M \leq \frac{1}{2}M,$$

hence $M = 0$ and therefore $\varphi = \psi$ on $[a, a + \frac{1}{2L}]$. □

The example

$$y' = y^{1/3}, \quad y(0) = y_0. \tag{30}$$

shows that the Lipschitz condition is indeed necessary for uniqueness: for $y_0 = 0$ we have two solutions $\varphi(x) = 0$ and $\varphi(x) = (\frac{2}{3}x)^{3/2}$.

11.4. Construction of an exact solution

To prove the existence of an exact solution we again assume the Lipschitz condition.

THEOREM 11.4.1 (Exact solutions). *Let $f\colon D \to \mathbb{R}$ be continuous, and satisfy a Lipschitz condition w.r.t. its second argument. Let $(a_0, b_0) \in D$ such that the rectangle R given by $|x - a_0| \le a$, $|y - b_0| \le b$ is in D. Assume $|f(x, y)| \le M$ for $(x, y) \in R$, and let $h := \min(a, b/M)$. Then we can construct an exact solution $\varphi\colon [a_0 - h, a_0 + h] \to \mathbb{R}$ of* (28) *such that $\varphi(a_0) = b_0$.*

PROOF. By the Cauchy-Euler approximation theorem in Section 11.1 we have an approximate solution φ_n up to the error $\frac{1}{2^n}$, which is a rational polygon, over $I := [a_0 - h, a_0 + h]$. By Lemma 7.2.4 the sequence $(\varphi_n)_{n\in\mathbb{N}}$ uniformly converges over I to a continuous function φ. Hence the sequence $(f(x, \varphi_n(x)))_n$ of continuous functions uniformly converges over I to the continuous function $f(x, \varphi(x))$.[1] Therefore, by Theorem 7.3.1

$$\lim_{n\to\infty} \int_a^b f(t, \varphi_n(t))\, dt = \int_a^b f(t, \varphi(t))\, dt.$$

We now prove that φ is an exact solution. By the choice of φ_n

$$\left|\varphi_n'(x) - f(x, \varphi_n(x))\right| \le \frac{1}{2^n}$$

for all $x \in I$ where $\varphi_n'(x)$ is defined. Integrating this inequality from a_0 to x gives

$$\Big|\int_{a_0}^x \big[\varphi_n'(t) - f(t, \varphi_n(t))\big]\, dt\Big| \le \frac{1}{2^n}(x - a_0) \le \frac{1}{2^n} h.$$

Since φ_n is continuous, by the fundamental theorem of calculus

$$\Big|\varphi_n(x) - \varphi_n(a_0) - \int_{a_0}^x f(t, \varphi_n(t))\, dt\Big| \le \frac{1}{2^n} h.$$

Approaching the limit for $n \to \infty$ gives

$$\varphi(x) - \varphi(a_0) - \int_{a_0}^x f(t, \varphi(t))\, dt = 0.$$

Differentiation yields $\varphi'(x) = f(x, \varphi(x))$, and $\varphi_n(a_0) = b_0$ entails $\varphi(a_0) = b_0$. □

[1] This is best proved by a slightly more general setup, where metric spaces (e.g., $\mathbb{R}^2$) are considered. One shows that if φ_n, ψ_n are uniformly convergent to φ, ψ, respectively, then (φ_n, ψ_n) is uniformly convergent to (φ, ψ), and if φ_n is uniformly convergent to φ, then $f(\varphi_n(x))$ is uniformly convergent to $f(\varphi(x))$.

For the construction above of an exact solution we have made use of the Lipschitz condition. However, it is well known that classically one has Peano's existence theorem, which does not require a Lipschitz condition.

Following Aberth (1970) and Bridges (2003) we now want to argue that Peano's existence theorem entails that for every real x we can decide whether $x \geq 0$ or $x \leq 0$, hence we cannot expect to be able to prove it constructively.

Consider again the initial value problem (30). First note that for $0 < a < y < b$ the continuous function $f(x, y) := y^{1/3}$ satisfies a Lipschitz condition w.r.t. its second argument (and similarly for $b < y < a < 0$). To see this, by Lemma 5.1.2 it suffices to find a bound on the derivative of $y^{1/3}$. But this is easy, since $\frac{d}{dy} y^{1/3} = \frac{1}{3} y^{-2/3} < \frac{1}{3} a^{-2/3}$ for $0 < a < y < b$.

Therefore in case $y_0 > 0$ the solution $\varphi_+(x) := (\frac{2}{3}x + y_0^{2/3})^{3/2}$ is unique, and similarly in case $y_0 < 0$ the solution $\varphi_-(x) := -(\frac{2}{3}x + |y_0|^{2/3})^{3/2}$ is unique. Pick $|y_0|$ small enough such that $(\frac{2}{3} - |y_0|^{2/3})^{3/2} > \frac{1}{2}$.

Now suppose that (30) has a solution φ, for a given real y_0. Compare $\varphi(1)$ with $[-1/2, 1/2]$. If $\varphi(1) < 1/2$, then $y_0 \not> 0$, hence $y_0 \leq 0$. If $-1/2 < \varphi(1)$, then $y_0 \not< 0$, hence $y_0 \geq 0$.

We finally show that an approximate solution of (28) up to the error $\frac{1}{2^p}$ differs from the exact solution by a constant multiple of $\frac{1}{2^p}$.

Theorem 11.4.2. *Let $f\colon D \to \mathbb{R}$ be continuous, and satisfy a Lipschitz condition w.r.t. its second argument. Let $(a_0, b_0) \in D$ such that the rectangle R given by $|x - a_0| \leq a$, $|y - b_0| \leq b$ is in D. Assume $|f(x, y)| \leq M$ for $(x, y) \in R$, and let $h := \min(a, b/M)$. Assume further that we have an exact solution $\varphi\colon [a_0 - h, a_0 + h] \to \mathbb{R}$ of (28) such that $\varphi(a_0) = b_0$, that ψ is an approximate solution up to the error $\frac{1}{2^p}$ such that $\psi(a_0) = b_0$, and that $\varphi \leq \psi$ or $\psi \leq \varphi$. Then there is a constant N independent of k such that $|\varphi(x) - \psi(x)| \leq \frac{1}{2^p} N$ for $|x - a_0| \leq h$.*

Proof. By the fundamental inequality

$$|\varphi(x) - \psi(x)| \leq \frac{\frac{1}{2^p}}{L}(e^{Lh} - 1).$$

Hence we can define $N := (e^{Lh} - 1)/L$. □

CHAPTER 12

Notes

There are many approaches to exact real number computation in the literature. One of those - using Möbius (or linear fractional) transformations - has been put forward by Edalat; a good survey can be found in Edalat (2003) (see also Edalat and Heckmann (2001)). Exact real numbers based on the so-called redundant b-adic notation have been treated in Wiedmer (1980) and in Boehm and Cartwright (1990), and based on continued fractions by Gosper (1972), Vuillemin (1990), Nielsen and Kornerup (1995) and also by Geuvers and Niqui (2000).

Generally, one can see these approaches as either using Cauchy sequences with (fixed or separately given) modulus, or else Dedekind cuts. We prefer Cauchy sequences over Dedekind cuts, since the latter are given by sets, and hence we would need additional enumerating devices in order to compute approximations of a real number presented as a Dedekind cut. There is not much of a difference between fixed or separate moduli for Cauchy sequences: one can always transform one form into the other. However, in order to keep the standard series representations of particular reals (like e) we prefer to work with separate moduli.

Another treatment (including an implementation in Mathematica) has been given in the Master's thesis of Andersson (2001) (based on Palmgren (1996)). He treats trigonometric functions, and includes Picard's and Euler's methods to constructively prove the existence of solutions for differential equations.

Some authors (in particular the so-called Russian school) restrict attention to computable real numbers. We do not want to make this restriction, since it makes sense, also constructively, to speak about arbitrary sequences. This view of higher type computability is the basis of Scott/Ershov domain theory, and we would like to adopt it here.

However, the domain theoretic setting for dealing with exact real numbers (cf. Edalat and Pattinson (2003)) is usually done in such a way that continuous functions are viewed as objects of the function domain, and hence are objects of type level 2. This clearly is one type level higher than necessary, since a continuous function is determined by its values on the rational numbers already. In particular from the point of view of program extraction it seems crucial to place objects as basic as continuous functions at the lowest possible type level. Therefore we propose a special concept of continuous functions, as type 1 objects.

Some of the (rather standard) calculus material, for instance in the section on sequences and series of real numbers, is taken form Forster's text Forster (2004). The section on ordinary differential equations is based on Chapter 1 of Hurewicz (1958), adapted to our constructive setting. I have also made use of Weghorn (2012) and a note of Bridges (2003).

APPENDIX A

Classical arguments in constructive proofs

Our underlying logical system is minimal logic; recent expositions are in Troelstra and van Dalen (1988) and also in Troelstra and Schwichtenberg (2000). Minimal logic is an appropriate logical framework for constructive mathematics. Already in 1933 both Gentzen and Gödel independently observed that classical logic can be viewed as a subsystem of minimal logic. Here we review some consequences of this insight.

Weak versions of existence and disjunction. We distinguish between two kinds of "exists" and two kinds of "or": the "weak" or classical ones and the "strong" or non-classical ones, with constructive content. In the present context both kinds occur together and hence we must mark the distinction; we shall do this by writing a tilde above the weak disjunction and existence symbols thus

$$A \tilde{\vee} B := \neg A \to \neg B \to \mathbf{F}, \qquad \tilde{\exists}_x A := \neg \forall_x \neg A.$$

These weak variants of disjunction and the existential quantifier are no stronger than the proper ones (in fact, they are weaker):

$$A \vee B \to A \tilde{\vee} B, \qquad \exists_x A \to \tilde{\exists}_x A.$$

Since $\tilde{\exists}_x \tilde{\exists}_y A$ unfolds into a rather awkward formula we extend the $\tilde{\exists}$-terminology to lists of variables:

$$\tilde{\exists}_{x_1,\dots,x_n} A := \forall_{x_1,\dots,x_n}(A \to \mathbf{F}) \to \mathbf{F}.$$

Moreover let

$$\tilde{\exists}_{x_1,\dots,x_n}(A_1 \tilde{\wedge} \dots \tilde{\wedge} A_m) := \forall_{x_1,\dots,x_n}(A_1 \to \dots \to A_m \to \mathbf{F}) \to \mathbf{F}.$$

This allows to stay in the $\to, \forall$ part of the language. Notice that $\tilde{\wedge}$ only makes sense in this context, i.e., in connection with $\tilde{\exists}$.

Stable formulas. Many properties of finitary mathematical objects (like natural numbers, integers and rational numbers) are "decidable" in the sense that they are given by a total computable function f into the boolean objects ff, tt. Then $f t_1 \dots t_n$ for closed argument terms can be evaluated to a boolean and hence decided. Examples are the order relations $\leq$ and $<$

on natural numbers, integers and rational numbers. Decidable formulas are closed under the logical connectives $\to$, $\wedge$, $\vee$ and bounded quantification.

We define negation $\neg A$ by $A \to \mathbf{F}$, where $\mathbf{F}$ is the (arithmetical) falsity defined by $\mathsf{ff} \equiv \mathsf{tt}$ where $\equiv$ is Leibniz equality. A formula A is called *stable* if

$$\neg\neg A \to A \qquad \text{“principle of indirect proof”}$$

holds for A. Every decidable formula is stable, but not conversely: an important example is $\leq$ on real numbers. If B is stable, then so is $A \to B$, and stable formulas are closed under conjunction and universal quantification (see Schwichtenberg and Wainer (2012, p.14)).

Recall some general logical facts on weak existence where stability may or may not play a role. Let $\mathrm{FV}(A)$ denote the set of variables free in A.

Lemma A.0.1. *The following are derivable.*

$$(\tilde{\exists}_x A \to B) \to \forall_x (A \to B) \quad \textit{if } x \notin \mathrm{FV}(B), \tag{31}$$

$$(\neg\neg B \to B) \to \quad \forall_x (A \to B) \to \tilde{\exists}_x A \to B \quad \textit{if } x \notin \mathrm{FV}(B), \tag{32}$$

$$(\mathbf{F} \to B[x{:=}c]) \to (A \to \tilde{\exists}_x B) \to \tilde{\exists}_x (A \to B) \quad \textit{if } x \notin \mathrm{FV}(A), \tag{33}$$

$$\tilde{\exists}_x (A \to B) \to A \to \tilde{\exists}_x B \quad \textit{if } x \notin \mathrm{FV}(A). \tag{34}$$

The last two items can also be seen as simplifying a weakly existentially quantified implication whose premise does not contain the quantified variable. In case the conclusion does not contain the quantified variable we have

$$(\neg\neg B \to B) \to \quad \tilde{\exists}_x (A \to B) \to \forall_x A \to B \quad \textit{if } x \notin \mathrm{FV}(B), \tag{35}$$

$$\forall_x (\neg\neg A \to A) \to (\forall_x A \to B) \to \tilde{\exists}_x (A \to B) \quad \textit{if } x \notin \mathrm{FV}(B). \tag{36}$$

Proof. See Schwichtenberg and Wainer (2012, p.15). □

Therefore when working with stable formulas we have by (35) and (36)

$$\tilde{\exists}_x (A \to B) \leftrightarrow (\forall_x A \to B) \quad \text{if } A, B \text{ are stable and } x \notin \mathrm{FV}(B). \tag{37}$$

There is a similar lemma on weak disjunction:

Lemma A.0.2. *The following are derivable.*

$$(A \mathbin{\tilde{\vee}} B \to C) \to (A \to C) \wedge (B \to C), \tag{38}$$

$$(\neg\neg C \to C) \to (A \to C) \to (B \to C) \to A \mathbin{\tilde{\vee}} B \to C, \tag{39}$$

$$(\mathbf{F} \to B) \to \quad (A \to B \mathbin{\tilde{\vee}} C) \to (A \to B) \mathbin{\tilde{\vee}} (A \to C), \tag{40}$$

$$(A \to B) \mathbin{\tilde{\vee}} (A \to C) \to A \to B \mathbin{\tilde{\vee}} C, \tag{41}$$

$$(\neg\neg C \to C) \to (A \to C) \mathbin{\tilde{\vee}} (B \to C) \to A \to B \to C, \tag{42}$$

$$(\mathbf{F} \to C) \to \quad (A \to B \to C) \to (A \to C) \mathbin{\tilde{\vee}} (B \to C). \tag{43}$$

PROOF. See Schwichtenberg and Wainer (2012, p.17). □

The weak existential quantifier $\tilde{\exists}$ and weak disjunction $\tilde{\vee}$ satisfy the same introduction axioms as the strong ones: this follows from the derivability of $\exists_x A \to \tilde{\exists}_x A$ and $A \vee B \to A \mathbin{\tilde{\vee}} B$. They also satisfy the same elimination axioms, provided one restricts the conclusion to stable formulas. For $\tilde{\exists}$ this has been proved in (32), and for $\tilde{\vee}$ in (39).

Therefore when proving a stable goal in minimal logic more proof techniques are available than in the general case. For instance, case distinction on an arbitrary formula A is possible by (39), since $A \mathbin{\tilde{\vee}} \neg A$ is (easily) derivable. Another important example is

LEMMA A.0.3. *The following is derivable.*

$$\forall_x \neg A \mathbin{\tilde{\vee}} \tilde{\exists}_x A.$$

PROOF. Unfolding $\tilde{\vee}$ and $\tilde{\exists}$ gives

$$(\forall_x (A \to \mathbf{F}) \to \mathbf{F}) \to (\underbrace{(\forall_x (A \to \mathbf{F}) \to \mathbf{F})}_{\tilde{\exists}_x A} \to \mathbf{F}) \to \mathbf{F}.$$ □

It is often helpful to use this lemma in a slightly more general form, for instancce

$$\forall_{x,y}(A \to B \to \mathbf{F}) \mathbin{\tilde{\vee}} \tilde{\exists}_{x,y}(A \mathbin{\tilde{\wedge}} B).$$

The proof is again immediate, since the right hand side $\tilde{\exists}_{x,y}(A \mathbin{\tilde{\wedge}} B)$ unfolds into the negated left hand side.

APPENDIX B

Detailed proof of Ishihara's trick

We give a detailed proof of Ishihara's trick, which was used as a guide for the formalization.

Let M be a modulus of convergence of (u_n) to 0; we can assume $M0 = 0$. Call m a *hit* on n if $M_n \leq m < M_{n+1}$ and $a \leq \|fu_m\|$. Our first goal is to define a function $h\colon \mathbb{N} \to \mathbb{N}$ such that

(i) $h_n = 0$ if for all $n' \leq n$ there is no hit;
(ii) $h_n = m + 2$ if at n for the first time we have a hit, with m;
(iii) $h_n = 1$ if there is an $n' < n$ with a hit.

We will need the bounded least number operator $\mu_n g$ defined recursively as follows. Here g is a variable of type $\mathbb{N} \to \mathbb{B}$.

$$\begin{aligned} &\mu_0 g := 0, \\ &\mu_{Sn} g := \begin{cases} 0 & \text{if } g0 \\ S\mu_n(g \circ S) & \text{otherwise.} \end{cases} \end{aligned}$$

Then we obtain

$$\begin{aligned} &\texttt{NatLeastBound:} && \mu_n g \leq n, \\ &\texttt{NatLeastLeIntro:} && gm \to \mu_n g \leq m, \\ &\texttt{NatLeastLtElim:} && \mu_n g < n \to g(\mu_n g). \end{aligned}$$

From $\mu_n g$ we define

$$\mu^n_{n_0} g := \begin{cases} (\mu_{n-n_0} \lambda_m g(m + n_0)) + n_0 & \text{if } n_0 \leq n \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $\mu^n_0 g = \mu_n g$. Generally we have

$$\begin{aligned} &\texttt{NatLeastUpLBound:} && n_0 \leq n \to n_0 \leq \mu^n_{n_0} g, \\ &\texttt{NatLeastUpBound:} && \mu^n_{n_0} g \leq n, \\ &\texttt{NatLeastUpLeIntro:} && n_0 \leq m \to gm \to \mu^n_{n_o} g \leq m, \\ &\texttt{NatLeastUpLtElim:} && n_0 \leq \mu^n_{n_0} g < n \to g(\mu^n_{n_0} g). \end{aligned}$$

To define h we will make use of a function g of type $\mathbb{N} \to \mathbb{B}$ (to be defined from `cApproxSplit`) such that

$$\begin{cases} a \leq \|fu_m\| & \text{if } gm \\ \|fu_m\| \leq b & \text{otherwise.} \end{cases}$$

Then we can define $h_n := H(g, M, n)$ where

$$H(g,M,n) := \begin{cases} 0 & \text{if } M_n \leq \mu_{M_n} g \text{ and } M_{n+1} \leq \mu_{M_n}^{M_{n+1}} g \\ \mu_{M_n}^{M_{n+1}} g + 2 & \text{if } M_n \leq \mu_{M_n} g \text{ and } \mu_{M_n}^{M_{n+1}} g < M_{n+1} \\ 1 & \text{if } \mu_{M_n} g < M_n. \end{cases}$$

To avoid multiple computations we split this definition into parts.

$$\begin{aligned} H(g,M,n) &:= \mathrm{HitPast}(g,M,M_n,n) \\ \mathrm{HitPast}(g,M,n_0,n) &:= \begin{cases} 1 & \text{if } \mu_{n_0} g < n_0 \\ \mathrm{HitHere}(g,n_0,M_{n+1}) & \text{otherwise} \end{cases} \\ \mathrm{HitHere}(g,n_0,n_1) &:= \mathrm{Hit}(\mu_{n_0}^{n_1} g, n_1) \\ \mathrm{Hit}(m,n) &:= \begin{cases} m+2 & \text{if } m < n \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

We show that h has the properties listed above.

Lemma B.0.1 (HProp01). *$\forall_{g,M,n}(H(g,M,n) = 0 \to M_n \leq \mu_{M_n} g)$.*

Lemma B.0.2 (HProp01Cor). *$\forall_{g,M,n,m}(H(g,M,n) = 0 \to m < M_n \to gm \to \mathbf{F})$.*

Lemma B.0.3 (HProp02). *$\forall_{g,M,n}(H(g,M,n) = 0 \to M_{n+1} \leq \mu_{M_n}^{M_{n+1}} g)$.*

Lemma B.0.4 (HProp02Cor).

$\forall_{g,M,n,m}(H(g,M,n) = 0 \to M_n \leq M_{n+1} \to m < M_{n+1} \to gm \to \mathbf{F})$.

Lemma B.0.5 (HProp0Cor). *$\forall_{g,M,n,m}(H(g,M,n)=0 \to m<M_{n+1} \to gm \to \mathbf{F})$.*

Lemma B.0.6 (HProp22). *$\forall_{g,M,n,m}(H(g,M,n) = m+2 \to \mu_{M_n}^{M_{n+1}} g < M_{n+1})$.*

Lemma B.0.7 (HProp2Cor).

$$\forall_{g,M,n,m}(H(g,M,n) = m+2 \to M_n \leq M_{n+1} \to g(\mu_{M_n}^{M_{n+1}} g)).$$

Lemma B.0.8 (HProp2Val). *$\forall_{g,M,n,m}(H(g,M,n) = m+2 \to \mu_{M_n}^{M_{n+1}} g = m)$.*

Lemma B.0.9 (HProp2gVal).

$$\forall_{g,M,n,m}(H(g,M,n) = m+2 \to M_n \leq M_{n+1} \to gm).$$

Lemma B.0.10 (HProp1). $\forall_{g,M,n}(H(g,M,n+1) = 1 \to \mu_{M_n} g < M_n)$.

Lemma B.0.11 (H0DownMon).

$$\forall_{g,M,n,n_1}(H(g,M,n) = 0 \to \forall_{n,m}(n \leq m \to M_n \leq M_m) \to \\ n_1 \leq n \to H(g,M,n_1) = 0).$$

Lemma B.0.12 (H1Down). $\forall_{g,M,n}(H(g,M,n+1) = 1 \to H(g,M,n) \neq 0)$.

Lemma B.0.13 (HFind).

$$\forall_{g,M,n}(M_0 = 0 \to H(g,M,n) \neq 0 \to \exists_{n_1,m}(n_1 < n \wedge H(g,M,n_1) = m{+}2)).$$

Further properties of H.

Lemma B.0.14 (H2Succ).

$$\forall_{g,M,n,m}(H(g,M,n) = m+2 \to \forall_{n,m}(n \leq m \to M_n \leq M_m) \to \\ H(g,M,n+1) = 1).$$

Lemma B.0.15 (H1Succ).

$$\forall_{g,M,n}(H(g,M,n) = 1 \to \forall_{n,m}(n \leq m \to M_n \leq M_m) \to \\ H(g,M,n+1) = 1).$$

Lemma B.0.16 (H2Up).

$$\forall_{g,M,n,m}(H(g,M,n) = m+2 \to \forall_{n,m}(n \leq m \to M_n \leq M_m) \to \\ \forall_{n_1} H(g,M,n+n_1+1) = 1).$$

Lemma B.0.17 (H2Down).

$$\forall_{g,M,n,m}(H(g,M,n+1) = m+2 \to \forall_{n,m}(n \leq m \to M_n \leq M_m) \to \\ H(g,M,n) = 0).$$

The next goal is to define from h a sequence (v_n) in X such that

(i) $v_n = 0$ if $h_n = 0$;
(ii) $v_n = nu_m$ if $h_n = m+2$;
(iii) $v_n = v_{n-1}$ if $h_n = 1$.

Let ξ be the type of elements of X, and us a variable of type $\mathbb{N} \to \xi$. Define $v_n := V_\xi(g,M,us,n)$ where (writing u_m for $us(m)$)

$$V_\xi(g,M,us,n) := \begin{cases} 0 & \text{if } H(g,M,n) = 0 \\ nu_m & \text{if } H(g,M,n) = m+2 \\ 0 \quad \text{(arbitrary)} & \text{if } H(g,M,n) = 1 \text{ and } n = 0 \\ V_\xi(g,M,us,n-1) & \text{if } H(g,M,n) = 1 \text{ and } n > 0. \end{cases}$$

Again we split the definition to avoid recomputations.

$$V_\xi(g, M, us, n) := \mathrm{Seq}_\xi(H(g, M), H(g, M, n), us, n)$$

$$\begin{aligned}
\mathrm{Seq}_\xi(h, 0, us, n) &:= 0\\
\mathrm{Seq}_\xi(h, m+2, us, n) &:= nu_m\\
\mathrm{Seq}_\xi(h, 1, us, 0) &:= 0 \quad \text{(arbitrary)}\\
\mathrm{Seq}_\xi(h, 1, us, n+1) &:= \mathrm{Seq}_\xi(h, h_n, us, n).
\end{aligned}$$

We show that (v_n) has the properties listed above.

Lemma B.0.18 (VIfH0). $\forall_{g,M,us,n}(H(g, M, n) = 0 \to V_\xi(g, M, us, n) = 0)$.

Lemma B.0.19 (VIfH2).

$$\forall_{g,M,us,n,m}(H(g, M, n) = m + 2 \to V_\xi(g, M, us, n) = nu_m).$$

Lemma B.0.20 (VIfH1).

$$\forall_{g,M,us,n}(H(g, M, n) = 1 \to V_\xi(g, M, us, n) = V_\xi(g, M, us, n \dot{-} 1)).$$

Lemma B.0.21 (VIfH2Up).

$$\begin{aligned}
\forall_{g,M,n,m,us}(H(g, M, n) = m + 2 \to \forall_{n,m}(n \le m \to M_n \le M_m) \to\\
\forall_{n_1} V_\xi(g, M, us, n + n_1) = nu_m).
\end{aligned}$$

Next we show that (v_n) is a Cauchy sequence with modulus $N(k) := 2k + 1$. Note that

$$\frac{N(k) + 1}{2^{N(k)}} \le \frac{1}{2^k}.$$

Since our goal is stable, we may employ arbitrary case distinctions.

Case 1. There is no hit. Then h_n is always 0, hence (v_n) is identically zero by `VIfH0` and therefore a Cauchy sequence with any modulus.

Case 2. Assume that there is a hit. Let n be one, say with value m. Given k, let $N(k) < n_1 < n_2$. We show $\|v_{n_1} - v_{n_2}\| \le 1/2^k$. If $n \le n_1$ or $n_2 < n$, then $v_{n_1} = v_{n_2}$ by `VIfH2Up` and we are done. Assume $n_1 < n \le n_2$. Then $v_{n_1} = 0$, hence $\|v_{n_1} - v_{n_2}\| = \|v_{n_2}\| = \|v_n\|$. By definition $v_n = nu_m$, hence

$$\begin{aligned}
\|v_n\| &= n\|u_m\|\\
&\le (n+1)/2^n \quad \text{since } M_n \le m \text{ and } M \text{ is a modulus for } (u_n)\\
&\le (N(k)+1)/2^{N(k)} \quad \text{since } n \mapsto (n+1)/2^n \text{ is monotone}\\
&\le 1/2^k \quad \text{by the note above.}
\end{aligned}$$

Recall that X is a Banach space, i.e., a *complete* linear space. We express this an axiom stating the assumed property of a limit operator:

Axiom (Compl).

$$\forall_{vs,N}(\forall_{k,n,m}(N(k) \le n \to n < m \to \|v_n - v_m\| \le 1/2^k) \to$$
$$\forall_{k,n}(N(k) \le n \to \|v_n - \lim(vs, N)\| \le 1/2^k)).$$

Note that we can assume here that the modulus N of convergence of (v_n) to v is the same as the Cauchy modulus of (v_n). As a consequence we have

Lemma B.0.22 (H2Compl).

$$\forall_{g,M,us,n,m,vs}(H(g, M, n+1) = m + 2 \to \forall_{n,m}(n \le m \to M_n \le M_m) \to$$
$$vs = V_\xi(g, M, us) \to$$
$$\forall_{k,n,m}(2k + 1 \le n \to n < m \to \|v_n - v_m\| \le 1/2^k) \to$$
$$nu_m = \lim(vs, \lambda_k(2k + 1))).$$

Now we can carry out the proof of Ishihara's trick. First we introduce abbreviations for the function g mentioned above, the special modulus N and also vs and v. Then we pick an n_0 such that `n0Prop`: $\|f(v)\| \le n_0 a$ (using `RealPosRatBound`). The main case distinction then is (a) $h_{n_0} = 0$ or (b) $\exists_n h_{n_0} = n + 1$.

In the first case (a) we prove the second alternative $\forall_m \|f(u_m)\| \le b$, as follows. First from `H0DownMon` we obtain $\forall_{n \le n_0} h_n = 0$. Next we prove $\forall_{n > n_0} h_n = 0$. Pick $n_1 > n_0$. Since the atom $h_{n_1} = 0$ is stable, is suffices to obtain a contradiction from the assumption $h_{n_1} \ne 0$. From `HFind` we obtain an $n < n_1$ and an m with $h_n = m + 2$. In case $n \le n_0$ we get the desired contradiction from $\forall_{n \le n_0} h_n = 0$. Hence assume $n_0 < n$. Then we obtain the desired contradiction as in Ishihara's Ishihara (1990), from `H2Compl` (i.e., $v = nu_m$) and

$$na \le n\|fu_m\| = \|n(fu_m)\| = \|f(nu_m)\| = \|fv\| \le n_0 a < na$$

using `HProp2gVal`, `gProp` and `n0Prop`. Hence we have $\forall_{n > n_0} h_n = 0$, and therefore $\forall_n h_n = 0$. To obtain the second alternative $\forall_m \|f(u_m)\| \le b$ by `gProp` it suffices to prove $gm \to F$. But this follows from `HProp01Cor` using $\forall_n h_n = 0$ and $n < M(n + 1)$.

In the second case (b) $\exists_n h_{n_0} = n + 1$ we prove the first alternative $\exists_m a \le \|f(u_m)\|$. To this end we use `HFind` to obtain $n_1 < n_0$ and m such that $h_{n_1} = m+2$. To show that this m has the required property $a \le \|f(u_m)\|$ it suffices to prove gm, because of `gProp`. But this follows from `HProp2gVal`. This concludes the proof.

Bibliography

Oliver Aberth. Computable analysis and differential equations. In Kino et al. (1970), pages 47–52.

Agda. Agda. http://wiki.portal.chalmers.se/agda/.

Patrik Andersson. Exact real arithmetic with automatic error estimates in a computer algebra system. Master's thesis, Mathematics department, Uppsala University, 2001.

Errett Bishop. *Foundations of Constructive Analysis.* McGraw-Hill, New York, 1967.

Errett Bishop. Mathematics as a numerical language. In Kino et al. (1970), pages 53–71.

Errett Bishop and Douglas Bridges. *Constructive Analysis*, volume 279 of *Grundlehren der mathematischen Wissenschaften.* Springer Verlag, Berlin, Heidelberg, New York, 1985.

Hans Boehm and Robert Cartwright. Exact real arithmetic. formulating real numbers as functions. In *Research Topics in Functional Programming*, chapter 3, pages 43–64. Addison Wesley, 1990.

Douglas Bridges. A Note on the Existence of Solutions for ODEs. Unpublished note, June 2003.

Douglas Bridges and Hajime Ishihara. Linear mappings are fairly well-behaved. *Archiv der Mathematik*, 54(6):558–662, 1990.

Coq Development Team. *The Coq Proof Assistant Reference Manual – Version 8.2.* Inria, 2009.

Abbas Edalat. Exact real number computation using linear fractional transformations. Technical report, Imperial College, Dept. of Computing, 2003.

Abbas Edalat and Reinhold Heckmann. Computing with real numbers: (i) LFT approach to real computations, (ii) domain-theoretic model of computational geometry. In G. Barthe et al., editors, *Proc APPSEM Summer School in Portugal*, LNCS. Springer Verlag, Berlin, Heidelberg, New York, 2001.

Abbas Edalat and Dirk Pattinson. Initial value problems in domain theory. to appear: Computation in Analysis, 2003.

Otto Forster. *Analysis 1.* Vieweg, 7th edition, 2004.

Herman Geuvers and Milad Niqui. Constructive reals in coq: Axioms and categoricity. In P. Callaghan, Z. Luo, J. McKinna, and R. Pollack, editors, *Proc. Types 2000*, volume 2277 of *LNCS*, pages 79–95. Springer Verlag, Berlin, Heidelberg, New York, 2000.

Witold Hurewicz. *Lectures on Ordinary Differential Equations.* MIT Press, Cambridge, Mass., 1958.

Hajime Ishihara. A constructive closed graph theorem. Technical report, Division of Mathematical and Information Sciences, Faculty of Integrated Arts and Sciences, Hiroshima University, 1990. ccgt.pdf.

Akiko Kino, John Myhill, and Richard E. Vesley, editors. *Intuitioninism and Proof Theory*, Studies in Logic and the Foundations of Mathematics, 1970. North-Holland, Amsterdam.

Asger Munk Nielsen and Peter Kornerup. MSB-first digit serial arithmetic. *Journal of Univ. Comp. Science*, 1(7):523–543, 1995.

Erik Palmgren. Constructive nonstandard analysis. In A. Petry, editor, *Méthods et analyse non standard*, volume 9, pages 69–97. Cahiers du Centre de Logique, 1996.

Helmut Schwichtenberg and Stanley S. Wainer. *Proofs and Computations.* Perspectives in Logic. Association for Symbolic Logic and Cambridge University Press, 2012.

Anne S. Troelstra and Helmut Schwichtenberg. *Basic Proof Theory.* Cambridge University Press, second edition, 2000.

Anne S. Troelstra and Dirk van Dalen. *Constructivism in Mathematics. An Introduction*, volume 121, 123 of *Studies in Logic and the Foundations of Mathematics.* North-Holland, Amsterdam, 1988.

Jean Vuillemin. Exact real computer arithmetic with continued fractions. *IEEE Transactions on Computers*, 39(8):1087–1105, 1990.

Thilo Weghorn. Verwendung klassischer Schlußweisen in der konstruktiven Analysis. Master's thesis, Mathematisches Institut der Universität München, 2012.

Edwin Wiedmer. Computing with infinite objects. *Theoretical Computer Science*, 10:133–155, 1980.

Index