Ludwig-Maximilians-Universität München

Fakultät für Mathematik

The constructive Stone-Weierstrass theorem

Bachelorarbeit
von
Sophie-Louise Schwenke

Betreuer: Dr. Iosif Petrakis
Abgabetermin: 20.08.2019

Ich erkläre hiermit, dass ich die Bachelor's Thesis selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

München, den 13.08.2019

Contents

1 Intoduction

One of the most well-known achievements of the American mathematician Errett Bishop is certainly Bishop's theorem, which generalizes the Stone-Weierstrass theorem. Published in 1961, "A generalization of the Stone-Weierstrass theorem" is one of Errett Bishop's earlier works. Since then, Errett Bishop has been particularly interested in constructive analysis. However, the theorem he gave his name to is not constructive.

This may seem paradoxical to some, considering that Bishop himself is known as a convinced constructivist and later in his life was seen as the leading mathematician in the field of constructive analysis. He earned this reputation foremost through his book "Foundations of constructive Analysis", published in 1967. Bishop's constructive analysis took Brouwer's constructive ideas and developed them further. It was on this basis that his book, a standard work on classical analysis with the aim of constructively proving large parts of classical analysis including the constructive Stone-Weierstrass theorem, was written, a project that many mathematicians had described as impossible at the time.

In this work we will constructively prove the Stone-Weierstrass theorem for compact metric spaces and at the same time study Bishop's constructive analysis. Our proof is based on Bishop's standard work "Foundation of Constructive Analysis" and its extension "Constructive Analysis". We will also use definitions, propositions, etc. contained therein, but we will provide more detailed proofs.

2 Calculus and the Real Numbers

2.1 The Real Number System

Bishop's definition of real numbers differs from that of classical analysis. Based on the rational numbers and their usual laws, we define the real numbers as a regular sequence of rational numbers. Accordingly, we also redefine the operations on real numbers.

Definition 1. *A sequence $\{x_n\}$ of rational numbers is regular if for all $m, n \in \mathbb{Z}^+$*

$$|x_m - x_n| \leq \frac{1}{m} + \frac{1}{n}$$

A real number is a regular sequence of rational numbers. Two real numbers $x := \{x_n\}$ and $y := \{y_n\}$ are equal if for all $n \in \mathbb{Z}^+$

$$|x_n - y_n| \leq \frac{2}{n}$$

The set of real numbers is denoted by $\mathbb{R}$.

Definition 2. *If K_x is the least integer which is greater than $|x_1| + 2$, then we call K_x the canonical bound for the real number x. It then obviously holds for all $n \in \mathbb{Z}^+$*

$$|x_n| < K_x.$$

Definition 3. *Let $x := \{x_n\}$ and $y := \{y_n\}$ be real numbers with respective canonical bounds K_x and K_y. Write $k := \max\{K_x, K_y\}$. Let α be any rational number. We define*

$$\begin{gathered}
(a)\ x + y := \{x_{2n} + y_{2n}\}_{n=1}^{\infty} \\
(b)\ xy := \{x_{2kn} y_{2kn}\}_{n=1}^{\infty} \\
(c)\ \max\{x, y\} := \{\max\{x_n, y_n\}\}_{n=1}^{\infty} \\
(d)\ -x := \{-x_n\}_{n=1}^{\infty} \\
(e)\ \alpha^* := \{\alpha, \alpha, \ldots\}.
\end{gathered}$$

As one can easily see, the sequences just defined are also real numbers. It can also be shown that they follow the same arithmetic laws which we already know from classical analysis.

Based on the definitions of positivity and negativity of real numbers, we define the order relations "$<$" and "$\leq$".

Definition 4. *A real number $x := \{x_n\}$ is said to be positive, or $x \in \mathbb{R}^+$, if*

$$x_n > \frac{1}{n}$$

for some n in $\mathbb{Z}^+$. A real number $x := \{x_n\}$ is said to be nonnegative, or $x \in \mathbb{R}_0^+$,if

$$x_n \geq -\frac{1}{n}$$

for all n in $\mathbb{Z}^+$.

Definition 5. *Let x and y be real numbers. We define $x > y$ (or $y < x$) if $x - y \in \mathbb{R}^+$ and $x \geq y$ (or $y \leq x$) if $x - y \in \mathbb{R}_0^+$. A real number x is negative if $x < 0^*$, that is, if $-x$ is positive.*

Definition 6. *For real numbers x and y we write $x \neq y$ if and only if $x < y$ or $x > y$.*

In Bishop's constructive analysis it is not possible to compare arbitrary real numbers as we know it from classical analysis. Whereas in classical analysis it is always valid that for two arbitrary real numbers $x, y \in \mathbb{R}$ at least one of the following properties holds

$$x < y \ \vee \ x = y \ \vee \ x > y.$$

This statement is not true in Bishop's constructive analysis. Although we can say in constructive mathematics that $x \leq y$ is equivalent to the negation of $x > y$, it is not true that $x > y$ is equivalent to the negation of $x \leq y$. This is because we can't constructively prove that from $x > y$ follows the negation of $x \leq y$. In order to still be able to compare arbitrary real numbers, we will need the following two results.

Proposition 1. *If $x_1, \ldots, x_n$ are real numbers with $x_1 + \ldots + x_n > 0$, then $x_i > 0$ for some $i \in \{1, \ldots, n\}$.*

Corollary 1. *If x, y, and z are real numbers with $y < z$, then either $x < z$ or $x > y$.*

2.2 Sequences and Series of Real Numbers

Definition 7. *A sequence $\{x_n\}$ of real numbers converges to a real number x_0 if for each k in $\mathbb{Z}^+$ there exists N_k in $\mathbb{Z}^+$ such that for all $n \geq N_k$*

$$|x_n - x_0| \leq k^{-1}.$$

To express that $\{x_n\}$ converges to x_0 we write $\lim_{n\to\infty} x_n = x_0$ or $x_n \to x_0$ as $n \to \infty$.

It should be noted that the definition of a convergent sequence of real numbers includes not only the sequence $\{x_n\}$ itself, but also its limit x_0 and the sequence $\{N_k\}$. The same is true for other definitions mentioned later.

Definition 8. *A sequence $\{x_n\}$ of real numbers is a Cauchy sequence if for each k in $\mathbb{Z}^+$ there exists M_k in $\mathbb{Z}^+$ such that for all $m, n \geq M_k$*

$$|x_m - x_n| \leq k^{-1}$$

Theorem 1. *A sequence $\{x_n\}$ of real numbers converges if and only if it is a Cauchy sequence.*

Proposition 2. *Assume that $x_n \to x_0$ as $n \to \infty$, and $y_n \to y_0$ as $n \to \infty$. Then*

$$(a)\ x_n + y_n \to x_0 + y_0 \ as\ n \to \infty$$
$$(b)\ x_n y_n \to x_0 y_0 \ as\ n \to \infty$$
$$(c)\ \max\{x_n, y_n\} \to \max\{x_0, y_0\} \ as\ n \to \infty$$
$$(d)\ x_0 = c \ whenever\ x_n = c \ for\ all\ n$$
$$(e)\ x_n^{-1} \to x_0^{-1} \ as\ n \to \infty \ whenever\ x_0 \neq 0 \ and\ x_n \neq 0 \ for\ all\ n$$
$$(f)\ x_0 \leq y_0 \ if\ x_n \leq y_n \ for\ all\ n$$

Definition 9. *For each sequence $\{x_n\}$ of real numbers the number*

$$s_n := \sum_{k=1}^{n} x_k$$

is called the nth partial sum of $\{x_n\}$, and $\{s_n\}$ is called the sequence of partial sums of the sequence $\{x_n\}$. A sum s_0 of $\{x_n\}$ is a limit of the sequence $\{s_n\}$ of partial sums, and we write

$$s_0 = \sum_{n=1}^{\infty} x_n$$

to indicate that s_0 is a sum of $\{x_n\}$. A sequence which is meant to be summed is called a series. A series is said to converge to its sum.

Even if Bishop does not, we give a proof of the infinite geometric series, because it will be of considerable use in the further explanations.

Theorem 2. *If $|x| < 1$, then $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$.*

Proof. Let $|x| < 1$. We first show that

$$\sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x} \tag{1}$$

by induction.
If $n = 0$, then $\sum_{k=0}^{n} x^k = 1 = \frac{1-x}{1-x}$. We assume that (1) is true for a fixed $n \in \mathbb{Z}_0^+$ and we show that (1) must be true for $n + 1$ as well.

$$\sum_{k=0}^{n+1} x^k = x^{n+1} + \sum_{k=0}^{n} x^k = x^{n+1} + \frac{1 - x^{n+1}}{1 - x} = \frac{1 - x^{n+2}}{1 - x}$$

Since for $|x| < 1$ we have $\lim_{n\to\infty} x^{n+1} = 0$, it follows that

$$\lim_{n\to\infty} \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x} \cdot (1 - \lim_{n\to\infty} x^{n+1}) = \frac{1}{1 - x}.$$

□

Example 1. *It is $\sum_{n=1}^{\infty} 2^{-n} = 1$. Since $0 < \frac{1}{2} < 1$, we get*

$$\sum_{n=1}^{\infty} 2^{-n} = \frac{1}{1 - \frac{1}{2}} - 1 = 1.$$

In classical analysis, we get as a result that a sequence of nonnegative terms converges when the partial sums are bounded. Although this is not the case in constructive analysis, we can refer to the following result as a substitute, which is also known under the name comparison test.

Proposition 3. *If $\sum_{n=1}^{\infty} y_n$ is a convergent series of nonnegative terms, and if $|x_n| \leq y_n$ for each n, then $\sum_{n=1}^{\infty} x_n$ converges.*

Proof. By Theorem 1 we get that the sequence of patial sums of $\{y_n\}$ is a Cauchy sequence. Thus let $k \in \mathbb{Z}^+$ and $M_k \in \mathbb{Z}^+$ such that for all $m, n \geq M_k$ it is

$$\left|\sum_{i=1}^{m} y_i - \sum_{i=1}^{n} y_i\right| = \sum_{i=\min\{m,n\}+1}^{\max\{m,n\}} y_i \ \leq \frac{1}{k},$$

then for $m, n \geq M_k$ follows

$$\left|\sum_{i=1}^{m} x_i - \sum_{i=1}^{n} x_i\right| \leq \sum_{i=\min\{m,n\}+1}^{\max\{m,n\}} |x_i| \ \leq \sum_{i=\min\{m,n\}+1}^{\max\{m,n\}} y_i \ \leq \frac{1}{k}.$$

We have shown that the sequence of partial sums of $\{x_n\}$ is a Cauchy sequence and therefore $\sum_{n=1}^{\infty} x_n$ converges. □

The next test that represents a test for convergence is called the ratio test.

Proposition 4. *Let $\sum_{n=1}^{\infty} x_n$ be a series, c a positive real number, and N a positive integer. Then $\sum_{n=1}^{\infty} x_n$ converges if $c < 1$ and*

$$|x_{n+1}| \leq c|x_n|$$

for all $n \geq N$.

Proof. We assume that $c < 1$ and that for all $n \geq N$ it is $x_{n+1}| \leq c|x_n|$. We show by induction, that for all $n \geq N$

$$|x_n| \leq c^{n-N}|x_N| \tag{2}$$

holds. If $n = N$, then (2) is obviously true. We now assume that (2) is true for fixed $n \in \mathbb{Z}^+$ with $n \geq N$ and we show that (2) must be true for $n+1$ as well.

$$|x_{n+1}| \leq c|x_n| \leq c \cdot c^{n-N}|x_n| = c^{(n+1)-N}|x_N|$$

Therefore (2) is proved. Since $|c| < 1$, it follows from Theorem 2 that the series

$$\sum_{n=1}^{\infty} c^{n-N}|x_N| = \frac{|x_N|}{c^N}\sum_{n=1}^{\infty} c^n$$

converges. We now get that the series $\sum_{n=1}^{\infty} x_n$ converges by applying the comparison test. □

The next test is based on Kummer's criterion.

Lemma 1. *Let $\{a_n\}$ and $\{x_n\}$ be sequences of positive numbers, c a positive number, and N a positive integer. Then $\sum_{n=1}^{\infty} x_n$ converges if $a_n x_n \to 0$ as $n \to \infty$ and for all $n \geq N$*

$$a_n x_n x_{n+1}^{-1} - a_{n+1} \geq c.$$

Lemma 2. *Let $\{y_n\}$ be a sequence of positive numbers, c a positive number, and N a positive integer such that for all $n \geq N$*

$$n(y_n y_{n+1}^{-1} - 1) \geq c.$$

Then $\lim_{n\to\infty} y_n = 0$.

The last convergence test executed in the following is also called Raabe's test in classical analysis.

Proposition 5. *Let $\sum_{n=1}^{\infty} x_n$ be a series of positive numbers such that $n(x_n x_{n+1}^{-1} - 1)$ converges to a limit L. Then $\sum_{n=1}^{\infty} x_n$ converges if $L > 1$.*

2.3 Continuous Functions

There is a proposition in classical analysis which states that a function that is pointwise continuous on a compact interval is also uniformly continuous. However, no constructive proof has yet been found for this proposition. In order to be compatible with classical analysis, Bishop does not work with the concept of pointwise continuity. Instead, he directly introduces the concept of uniform continuity.

Definition 10. *A real-valued function f defined on a compact interval I (i.e. I is a non-empty interval of the form $I = [a, b]$ with $a < b$ and a,b finite) is continuous on I if for each $\epsilon > 0$ there exists $w(\epsilon) > 0$ such that for all $x, y \in I$*

$$|x - y| \leq w(\epsilon) \Rightarrow |f(x) - f(y)| \leq \epsilon.$$

The operation $\epsilon \to w(\epsilon)$ is called a modulus of continuity for f. A real-valued function f on an arbitrary interval J is continuous on J if it is continuous on every compact subinterval I of J.

In this context, it should be noted that a modulus of continuity w is always an indispensable part of the definition of a continuous function on a compact interval.

Definition 11. *A real number c is called the least upper bound (respectively, greatest lower bound) of a subset A of $\mathbb{R}$ and written $c := l.u.b.\ A$ (respectively, $c := g.l.b.\ A$) if $x \leq c$ (respectively, $x \geq c$) for all x in A and if for each $\epsilon > 0$ there exists x in A with $c - x < \epsilon$ (respectively, $x - c < \epsilon$).*

Theorem 3. *Let the subset A of $\mathbb{R}$ have the property that for each $\epsilon > 0$ there exist finitely many points $y_1, \ldots, y_n$ in A such that for each x in A at least one of the numbers $|x - y_1|, \ldots, |x - y_n|$ is less than ϵ. (Such a set is called totally bounded.) Then l.u.b. A and g.l.b. A exist.*

Corollary 2. *If $f\colon [a, b] \to \mathbb{R}$ is a continuous function on a compact interval, then the quantities $c := l.u.b.\{f(x)\colon x \in [a, b]\}$ and $d := g.l.b.\{f(x)\colon x \in [a, b]\}$ (called, respectively, the supremum and the infimum of f on the interval $[a, b]$) exist.*

Proof. Let w be a modulus of continuity for f and let $\epsilon > 0$. We choose $a_0, \ldots, a_n \in \mathbb{R}$ such that

$$a = a_0 \leq a_1 \leq \ldots \leq a_n = b \quad and \quad a_{i+1} - a_i \leq w(\epsilon)$$

for all $i \in \{0, \ldots, n-1\}$. Then whenever it is $x \in [a, b]$ we have $|x - a_i| \leq w(\epsilon)$ for at least one i, and therefore $|f(x) - f(a_i)| \leq \epsilon$ for the corresponding i. So we have shown that $\{f(x)\colon x \in [a, b]\}$ is totally bounded. By Theorem 3, it follows that $\sup f$ and $\inf f$ exist. □

Theorem 4. *Let f and g be continuous real-valued functions defined on an interval I. Then the functions $f + g$, fg, and $\max\{f, g\}$ are continuous on I. If f is bounded away from 0 on every compact subinterval J of I, that is, if $|f(x)| \geq c$ for all x in J and some $c > 0$ (depending on J), then f^{-1} is continuous on I.*

Proof. By the definition of continuity, it is enough to consider the case in which I is compact. In preperation of the proof, let w_f and w_g be moduli of continuity for f and g on I. Let $\epsilon > 0$.

Let $x, y \in I$. If we set $|x-y| \leq w(\epsilon) := \min\left\{w_f\left(\frac{\epsilon}{2}\right), w_g\left(\frac{\epsilon}{2}\right)\right\}$, we get

$$\begin{aligned}|(f+g)(x)-(f+g)(y)| &\leq |f(x)-f(y)+g(x)-g(y)| \\ &\leq |f(x)-f(y)|+|g(x)-g(y)| \\ &\leq \frac{\epsilon}{2}+\frac{\epsilon}{2} \leq \epsilon\end{aligned}$$

It follows that $f+g$ is continuous on I with modulus of continuity w.

Because of the Corollary above, we can set $M := \sup\{|f(x)| : x \in I\}$. Then if $x, y \in I$ and $|x-y| \leq w(\epsilon) := w_f\left(\frac{\epsilon}{2M}\right)$, we have

$$\begin{aligned}|f^2(x)-f^2(y)| &= |(f(x)-f(y))\cdot(f(x)+f(y))| \\ &\leq |f(x)-f(y)|\cdot(|f(x)|+|f(y)|) \\ &\leq |f(x)-f(y)|\cdot 2M \leq \frac{\epsilon}{2M}\cdot 2M = \epsilon\end{aligned}$$

and therefore f^2 is continuous on I.

Let $\alpha \in \mathbb{R}$. Then if $x, y \in I$ and $|x-y| \leq w(\epsilon) := w_f\left(\frac{\epsilon}{|\alpha|}\right)$, we have

$$|(\alpha f)(x)-(\alpha f)(y)| = |\alpha||f(x)-f(y)| \leq |\alpha|\cdot|\frac{\epsilon}{|\alpha|} = \epsilon$$

and therefore αf is continuous on I.

Since $fg = \frac{1}{2}((f+g)^2 - f^2 - g^2)$, it follows that fg is continuous on I.

Let $x, y \in I$. If we set $|x-y| \leq w(\epsilon) := w_f(\epsilon)$, we get

$$\Big|(|f|)(x)-(|f|)(y)\Big| = \Big||f(x)|-|f(y)|\Big| \leq |f(x)-f(y)| \leq \epsilon$$

and therefore $|f|$ is continuous on I.

Since it is

$$\max\{f,g\}(x) = \frac{f(x)+g(x)}{2} + \frac{|f(x)-g(x)|}{2}$$

and

$$\min\{f,g\}(x) = \frac{f(x)+g(x)}{2} - \frac{|f(x)-g(x)|}{2},$$

it follows that both functions are continuous on I.

Let $|f(x)| \geq c$ for all x in I and some $c > 0$. Then whenever $x, y \in I$ and $|x-y| \leq w(\epsilon) := w_f(M^2 \cdot \epsilon)$, where $M := \inf\{|f(x)| : x \in I\}$, we have

$$\begin{aligned}|f^{-1}(x) - f^{-1}(y)| &= \left|\frac{1}{f(x)} - \frac{1}{f(y)}\right| = \left|\frac{f(y)-f(x)}{f(x)\cdot f(y)}\right| \\ &= \frac{|f(x)-f(y)|}{|f(x)|\cdot|f(y)|} \leq \frac{M^2\cdot\epsilon}{M^2} = \epsilon,\end{aligned}$$

and therefore f^{-1} is continuous on I.
It follows that $\frac{g}{f} = g \cdot f^{-1}$ is continuous on I. □

Theorem 5. *The composition of continuous functions is continuous, in the sense that if $f\colon I \to J$ and $g\colon J \to \mathbb{R}$ are continuous, then $g \circ f$ is continuous, provided that f maps every compact subinterval of I into a compact subinterval of J.*

Proof. Again, we only consider the case in which I and J are both compact. Let w_f be a modulus of continuity of f and w_g a modulus of continuity of g. Let $\epsilon > 0$. Then if $x, y \in I$ and $|x-y| \leq w(\epsilon) := w_f(w_g(\epsilon))$, we have

$$|f(x)-f(y)| \leq w_g(\epsilon) \Rightarrow |(g\circ f)(x) - (g \circ f)(y)| = |g(f(x)) - g(f(y))| \leq \epsilon.$$

□

Definition 12. *A sequence $\{f_n\}$ of continuous functions on a compact interval I converges on I to a continuous function f if for each $\epsilon > 0$ there exists N_ϵ in $\mathbb{Z}^+$ such that for all $x \in I$ and $n \geq N_\epsilon$*

$$|f_n(x) - f(x)| \leq \epsilon$$

A sequence $\{f_n\}$ of continuous functions on an arbitrary interval J converges on J to a continuous function f if it converges to f on every compact subinterval I of J. Notations to express the fact that $\{f_n\}$ converges to f are $\lim_{n\to\infty} f_n = f$ and $f_n \to f$ as $n \to \infty$.

Definition 13. *A sequence $\{f_n\}$ of continuous functions on a compact interval I is a Cauchy sequence on I if for each $\epsilon > 0$ there exists M_ϵ in $\mathbb{Z}^+$ such that for all $x \in I$ and $m, n \geq M_\epsilon$*

$$|f_m(x) - f_n(x)| \leq \epsilon$$

A sequence $\{f_n\}$ of continuous functions on an arbitrary interval J is a Cauchy sequence on J if it is a Cauchy sequence on every compact subinterval of J.

Theorem 6. *A sequence $\{f_n\}$ of continuous functions on an interval J converges on J if and only if it is a Cauchy sequence on J.*

Definition 14. *To each sequence $\{f_n\}$ of continuous functions on an interval I corresponds a sequence $\{g_n\}$ of partial sums, defined by*

$$g_n := \sum_{k=1}^{n} f_k$$

If $\{g_n\}$ converges to a continuous function g on I, then g is the sum of the series $\sum_{n=1}^{\infty} f_n$,

$$g := \sum_{n=1}^{\infty} f_n$$

and the series is said to converge to g on I.

Remark 1. *The comparison test and the ratio test carry over to series of functions. The comparison test states that if $\sum_{n=1}^{\infty} g_n$ is a convergent series of nonnegative continuous functions on an interval I, then the series $\sum_{n=1}^{\infty} f_n$ of continuous functions on I converges on I whenever $|f_n(x)| \leq g_n(x)$ for all n in $\mathbb{Z}^+$ and all x in I.*
The ratio test states that if $\sum_{n=1}^{\infty} f_n$ is a series of continuous functions on an interval J such that for each compact subinterval I of J there exists a constant c_I, $0 < c_I < 1$, and a positive integer N_I with for all $x \in I$ and for all $n \geq N_I$

$$|f_{n+1}(x)| \leq c_I |f_n(x)|,$$

then $\sum_{n=1}^{\infty} |f_n|$ converges on J.

Proposition 6. *Let the series*

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

have the property that there exists $r > 0$ and N in $\mathbb{Z}^+$ such that $|a_{n+1}| \leq r^{-1}|a_n|$ for all $n \geq N$. Then the series $\sum_{n=0}^{\infty} |a_n(x - x_0)^n|$ converges on the interval $(x_0 - r, x_0 + r)$.

Proof. Let $I \subset (x_o - r, x_0 + r)$ be an arbitrary compact interval. If we define $r_0 := |x - x_0|$, we get

$$\begin{aligned} |a_{n+1}(x - x_0)^{n+1}| &= |a_{n+1}||x - x_=|^{n+1} \\ &\leq r^{-1}|a_n||x - x_0|^{n+1} = \frac{r_0}{r} \cdot |a_n||x - x_0|^n. \end{aligned}$$

Since $|x - x_0| < r$, we have $r_0 r^{-1} < 1$ and thus the series converges by the ratio test. □

2.4 Differentation

In accordance to Bishop-continuity we define differentiability directly on compact intervals.

Definition 15. *Let f and g be continuous functions on a compact proper interval I (i.e. a and b are finite and $a < b$) such that for each $\epsilon > 0$ there exists $\delta(\epsilon) > 0$ such that for all $x, y \in I$ it is*

$$|y - x| \leq \delta(\epsilon) \Rightarrow |f(y) - f(x) - g(x)(y - x)| \leq \epsilon|y - x|.$$

Then f is said to be differentiable on I, g is called a derivative of f on I, and δ is called a modulus of differentiability for f on I. If f and g are continuous on the proper interval J, then g is a derivative of f on J if it is a derivative of f on every compact proper subinterval I of J, and f is said to be differentiable on J. To express that g is a derivative of f we write $g = f'$ or $g = Df$.

Theorem 7. *Let f_1 and f_2 be differentiable functions on an interval I. Then $f_1 + f_2$ and $f_1 \cdot f_2$ are differentiable on I. In case f_1 is bounded away from 0 on every compact subinterval of I, then f_1^{-1} is differentiable on I. The function $x \to x$ is differentiable on $\mathbb{R}$. For each c in $\mathbb{R}$ the function $x \to c$*

is differentiable on $\mathbb{R}$. The derivatives in question are given by the following equations.

$$(a)\ D(f_1 + f_2) = Df_1 + Df_2$$
$$(b)\ D(f_1 \cdot f_2) = f_1 Df_2 + f_2 Df_1$$
$$(c)\ Df_1^{-1} = -f_1^{-2} Df_1$$
$$(d)\ \frac{dx}{dx} = 1$$
$$(e)\ \frac{dc}{dx} = 0$$

Proof. By the definition of differentiability, it is enough to consider the case in which I is compact. To prove all this, let δ_1 and δ_2 be moduli of differentiability for f_1 and f_2 on I, let f_1' and f_2' be the corresponding derivatives and let w_1 be a modulus of continuity for f_1 on I.
(a) Let $x, y \in I$. If we set $|y - x| \leq \delta(\epsilon) := \min\{\delta_1(\frac{\epsilon}{2}), \delta_2(\frac{\epsilon}{2})\}$, we get

$$\begin{aligned}
&|f_1(y) + f_2(y) - (f_1(x) + f_2(x)) - (f_1'(x) + f_2'(x))(y - x)| \\
\leq\ &|f_1(y) - f_1(x) - f_1'(x)(y - x)| + |f_2(y) - f_2(x) - f_2'(x)(y - x)| \\
\leq\ &\frac{\epsilon}{2} \cdot |y - x| + \frac{\epsilon}{2} \cdot |y - x| = \epsilon \cdot |y - x|.
\end{aligned}$$

It follows that $f_1 + f_2$ is differentiable on I with derivative $f_1' + f_2'$ and modulus of differentiablity δ.
(b) Since f_1, f_2 and f_2' are continuous and I is compact, their suprema and infima on I exist. Therefore we can define $M := \max\{\max\{|f_1(x)| : x \in I\}, \max\{|f_2(x)| : x \in I\}, \max\{|f_2'(x)| : x \in I\}\}$. Then if $x, y \in I$ and

$|y-x| \le \delta(\epsilon) := \min\{\delta_1\left(\frac{\epsilon}{3M}\right), \delta_2\left(\frac{\epsilon}{3M}\right), w_1\left(\frac{\epsilon}{3M}\right)\}$, we have

$$\begin{aligned}
&|f_1(y)f_2(y) - f_1(x)f_2(x) - (f_1(x)f_2'(x) + f_2(x)f_1'(x))(y-x)| \\
&= |f_1(y)f_2(y) - f_1(y)f_2(x) - f_1(y)f_2'(x)(y-x) + f_1(y)f_2'(x)(y-x) \\
&\quad -f_1(x)f_2'(x)(y-x) + f_2(x)f_1(y) - f_2(x)f_1(x) - f_2(x)f_1'(x)(y-x)| \\
&\le |f_1(y)||f_2(y) - f_2(x) - f_2'(x)(y-x)| + |f_1(y) - f_1(x)||f_2'(x)||y-x| \\
&\quad +|f_2(x)||f_1(y) - f_1(x) - f_1'(x)(y-x)| \\
&\le |f_1(y)| \cdot \frac{\epsilon}{3M} \cdot |y-x| + |f_2'(x)| \cdot \frac{\epsilon}{3M} \cdot |y-x| + |f_2(x)| \cdot \frac{\epsilon}{3M} \cdot |y-x| \\
&\le M \cdot \frac{\epsilon}{3M} \cdot |y-x| + M \cdot \frac{\epsilon}{3M} \cdot |y-x| + M \cdot \frac{\epsilon}{3M} \cdot |y-x| \\
&= \epsilon \cdot |y-x|.
\end{aligned}$$

(c) Let $M := \max\{\max\{|f_1^{-1}(x)| : x \in I\}, \max\{|f_1'(x)| : x \in I\}\}$. Then whenever $x, y \in I$ and $|y-x| \le \delta(\epsilon) := \min\{\delta_1\left(\frac{\epsilon}{2M^2}\right), w_1\left(\frac{\epsilon}{2M^4}\right)\}$, we have

$$\begin{aligned}
&|f_1^{-1}(y) - f_1^{-1}(x) + f_1^{-2}f_1'(x)(y-x)| \\
&= |f_1^{-1}(x) - f_1^{-1}(y) - f_1^{-2}f_1'(x)(y-x)| \\
&= |f_1^{-1}(x)f_1^{-1}(y)f_1(y) - f_1^{-1}(x)f_1^{-1}(y)f_1(x) \\
&\quad -f_1^{-1}(x)f_1^{-1}(y)f_1(y)f_1^{-1}(x)f_1'(x)(y-x)| \\
&= |f_1^{-1}(x)f_1^{-1}(y)||f_1(y) - f_1(x) - f_1(y)f_1^{-1}(x)f_1'(x)(y-x)| \\
&\le M^2|f_1(y) - f_1(x) - f_1'(x)(y-x) \\
&\quad +f_1'(x)(y-x) - f_1'(x)f_1^{-1}(x)f_1(y)(y-x)| \\
&\le M^2|f_1(y) - f_1(x) - f_1'(x)(y-x)| \\
&\quad +M^2|f_1'(x)f_1^{-1}(x)f_1(y)(y-x) - f_1'(x)(y-x)| \\
&= M^2|f_1(y) - f_1(x) - f_1'(x)(y-x)| + M^2|f_1'(x)f_1^{-1}(x)||f_1(y) - f_1(x)||y-x| \\
&\le M^2 \cdot \frac{\epsilon}{2M^2} \cdot |y-x| + M^2M^2 \cdot \frac{\epsilon}{2M^4} \cdot |y-x| = \epsilon \cdot |y-x|.
\end{aligned}$$

(d) Let I be an arbitrary compact interval in $\mathbb{R}$. Then for all x,y in I and $|y-x| \le \delta(\epsilon) := \epsilon$, we have

$$|y - x - 1(y-x)| = 0 \le \epsilon \cdot |y-x|.$$

Since I was arbitrary, it follows the function $x \to x$ is differentiable on $\mathbb{R}$.
(e) Let I be again an arbitrary compact interval in $\mathbb{R}$. Then if x,y in I and $|y - x| \leq \delta(\epsilon) := \epsilon$, we get

$$|c - c - 0(y - x)| = 0 \leq \epsilon \cdot |y - x|.$$

$\square$

Corollary 3. *For all positive integers n,*

$$\frac{dx^n}{dx} = nx^{n-1}.$$

Theorem 8. *Let $f\colon I \to \mathbb{R}$ and $g\colon J \to \mathbb{R}$ be differentiable functions such that f maps every compact subinterval of I into some compact subinterval of J. Then $g \circ f$ is differentiable, and $(g \circ f)' = (g' \circ f)f'$.*

The following theorem is a counterpart to Rolle's theorem. In the proof it is particularly well recognizable, which characteristics one must consider when proving constructively. In this case, this stems mainly from the fact, that we can't compare arbitrary real numbers.

Theorem 9. *Let f be differentiable on the interval $[a, b]$, and let $f(a) = f(b)$. Then for each $\epsilon > 0$ there exists x in $[a, b]$ with*

$$|f'(x)| \leq \epsilon$$

Proof. Since f is differentiable, we know that f' is continuous on $[a, b]$. We take w to be a modulus of continuity of f' on $[a, b]$ and we take δ to be a modulus of differentiability of f on $[a, b]$. We choose $x_0, \ldots, x_n \in \mathbb{R}$ such that

$$a = x_0 < x_1 < \ldots < x_n = b \textit{ and } x_{i+1} - x_i \leq \min\{\delta\left(\frac{\epsilon}{2}\right), w\left(\frac{\epsilon}{2}\right)\}$$

for all i $(0 \leq i \leq n - 1)$. Then we have

$$\begin{aligned} f(x_{i+1}) - f(x_i) &= f'(x_i)(x_{i+1} - x_i) + f(x_{i+1}) - f(x_i) - f'(x_i)(x_{i+1} - x_i) \\ &\leq \left(f'(x_i) + \frac{\epsilon}{2}\right)|x_{i+1} - x_i| < (f'(x_i) + \epsilon)|x_{i+1} - x_i| \end{aligned}$$

for all i in $\{0, \ldots, n-1\}$ and therefore it is

$$0 = f(b) - f(a) = \sum_{i=0}^{n-1} f(x_{i+1}) - f(x_i) < \sum_{i=0}^{n-1} (f'(x_i) + \epsilon)|x_{i+1} - x_i|.$$

If we apply Proposition 1, we get that $f'(x_i) > -\epsilon$ for at least one value of i. Furthermore, we have

$$\begin{aligned} f(x_i) - f(x_{i+1}) &= f(x_i) - f(x_{i+1}) - f'(x_{i+1})(x_i - x_{i+1}) + f'(x_{i+1})(x_i - x_{i+1}) \\ &\leq \frac{\epsilon}{2} \cdot |x_i - x_{i+1}| + f'(x_{i+1})(x_i - x_{i+1}) \\ &< \epsilon \cdot |x_i - x_{i+1}| + f'(x_{i+1})(x_i - x_{i+1}) \\ &= \epsilon \cdot |x_i - x_{i+1}| - f'(x_{i+1})|x_i - x_{i+1}| \\ &= (\epsilon - f'(x_{i+1}))|x_i - x_{i+1}| \end{aligned}$$

for all i in $\{0, ..., n-1\}$ and therefore, we have

$$0 = f(a) - f(b) = \sum_{i=0}^{n-1} f(x_i) - f(x_{i+1}) < \sum_{i=0}^{n-1} (\epsilon - f'(x_{i+1}))|x_i - x_{i+1}|.$$

Applying Proposition 1 again, we get that $f'(x_i) < \epsilon$ for at least one value of i. By the Corollary to Proposition 1 and since $\frac{\epsilon}{2} < \epsilon$, it is either $|f'(x_i)| < \epsilon$ or $|f'(x_i)| > \frac{\epsilon}{2}$ for all i $(0 \leq i \leq n)$. If $|f'(x_i)| < \epsilon$ for some i, we are done. Therefore, let us consider the case in which $|f'(x_i)| > \frac{\epsilon}{2}$ for all i. Since $|x_{i+1} - x_i| \leq w(\frac{\epsilon}{2})$, the continuity of f implies $|f'(x_{i+1}) - f'(x_i)| \leq \frac{\epsilon}{2}$ and thus we obtain that the quantities $f'(x_{i+1})$ and $f'(x_i)$ are either both positive or both negative. This must then obviously apply to all i. As $f'(x_i) > -\epsilon$ for at least one value of i and $f'(x_i) < \epsilon$ for at least one value of i, we get that $0 < |f'(x_i)| < \epsilon$ for at least one value of i. □

Definition 16. *Let $f, f^{(1)}, f^{(2)}, \ldots, f^{(n)}$ be differentiable functions on an interval I with $Df = f^{(1)}, Df^{(1)} = f^{(2)}, \ldots, Df^{(n-1)} = f^{(n)}$ on I. Then $f^{(n)}$ is called the nth derivative of f on I and written $D^{(n)}f$, or simply $f^{(n)}$, and f is said to be n times differentiable on I. The function f itself may be written $f^{(0)}$ or $D^0 f$.*

Definition 17. *If f is an n times differentiable function on an Interval I, and a is a point in I, then we call*

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

the nth Taylor polynomial for f about a and for a given value b

$$R := f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k$$

is called the remainder term. If f is infinitely differentiable on an open interval $I = (a - x, a + x)$, then for $t \in I$ we call

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(t-a)^n$$

the Taylor series for f about a.

On the basis of the following theorem we see one more time that difficulties stem from the fact that in constructive analysis in contrast to classical analysis arbitrary real numbers can not be compared.

Theorem 10. *Let f be an (n+1) times differentiable function on an interval I. Let ϵ be a positive constant, and let a and b be points of I. Then there exists c, with* $\min\{a, b\} \le c \le \max\{a, b\}$, *such that*

$$\left| R - \frac{f^{(n+1)}(c)}{n!}(b-c)^n(b-a) \right| \le \epsilon,$$

where R represents the remainder term from the deinition above.

Proof. We start our proof by defining

$$M := 1 + \max\left\{ \left|\frac{f^{(1)}(a)}{1!}\right|, \left|\frac{f^{(2)}(a)}{2!}\right|, \cdots, \left|\frac{f^{(n)}(a)}{n!}\right| \right\}$$

and δ such that

$$0 < \delta < \min\left\{1, \frac{\epsilon}{2nM}, w\left(\frac{\epsilon}{2}\right)\right\},$$

where w is a modulus of continuity for f on $[\min\{a, b\}, \max\{a, b\}]$. By the Corollary of Proposition 1, it is either $0 < |a - b|$ or $|a - b| < \delta$. We can therefore distinguish between these two cases. First suppose $|a-b|<\delta$. If we set $c := b$, we get

$$\begin{aligned}|R| &= \Big|f(b)-f(a)-\sum_{k=1}^{n}\frac{f^{(k)}(a)}{k!}(b-a)^k\Big| \\ &\leq |f(b)-f(a)| + \sum_{k=1}^{n}\left|\frac{f^{(k)}(a)}{k!}\right|\cdot|b-a|^k \\ &\leq \frac{\epsilon}{2} + M\sum_{k=1}^{n}\delta^k < \frac{\epsilon}{2} + M\sum_{k=1}^{n}\frac{\epsilon}{2nM} = \epsilon.\end{aligned}$$

Now suppose that $0<|a-b|$. We consider the function

$$\begin{aligned}g(x) := f(b) - f(x) - \frac{f'(x)}{1!}(b-x) - \frac{f''(x)}{2!}(b-x)^2 - \dots \\ -\frac{f^{(n)}(x)}{n!}(b-x)^n - R(b-x)(b-a)^{-1}.\end{aligned}$$

It follows immediately that $g(a)=g(b)=0$. The function g is differentiable on I as a composition of differentiable functions on I and we get

$$\begin{aligned}g'(x) &= -f'(x)+f'(x)-f''(x)(b-x)+f''(x)(b-x)-\dots \\ &\quad+\frac{f^{(n)}(x)}{(n-1)!}(b-x)^{n-1}-\frac{f^{(n+1)}(x)}{n!}(b-x)^n+R(b-a)^{-1} \\ &= -\frac{f^{(n+1)}}{n!}(b-x)^n+R(b-a)^{-1}.\end{aligned}$$

By Rolle's theorem, we know that for each $\epsilon>0$ there exists c with $\min\{a,b\}\leq c\leq\max\{a,b\}$ and $|g'(c)|\leq\epsilon$. Since g is differentiable on $[\min\{a,b\}\ \max\{a,b\}]$ and $g(a)=g(b)$, we can apply Rolle's theorem

$$\begin{aligned}|g'(c)| = \Big|-\frac{f^{(n+1)}}{n!}(b-c)^n + R(b-a)^{-1}\Big| \leq \epsilon|b-a|^{-1} \\ \Leftrightarrow \Big|R - \frac{f^{(n+1)}}{n!}(b-c)^n(b-a)\Big| \leq \epsilon\end{aligned}$$

and have thus finished our proof. $\square$

3 Metric Spaces

3.1 Fundamental Definitions and Constructions

Definition 18. *A metric on a set X is a function $d\colon X \times X \to \mathbb{R}_0^+$ such that*
(a) $d(x,y) = 0$ if and only if $x = y$
(b) $d(x,y) = d(y,x)$
(c) $d(x,y) \leq d(x,z) + d(z,y)$
Elements x and y of X are unequal, $x \neq y$, if and only if $d(x,y) > 0$. A set X which is endowed with a metric is called a metric space. The most important example of a metric space is the set of real numbers, metrized by defining $d(x,y) := |x-y|$.

Definition 19. *Let $\{(X_n, d_n)\}$ be a sequence of metric spaces, each bounded by 1, i.e. $d_n(x_n, y_n) \leq 1$ for all $x_n, y_n \in X_n$. The product metric d on $X := \prod_{n=1}^{\infty} X_n$ is defined by*

$$d(\{x_n\}, \{y_n\}) := \prod_{n=1}^{\infty} 2^{-n} d_n(x_n, y_n)$$

for all $\{x_n\}, \{y_n\} \in X$.

Proposition 7. *Let (X,d) be a metric space. Let $h\colon \mathbb{R}_0^+ \to \mathbb{R}_0^+$ satisfy the conditions*
(i) $h(u) = 0$ if and only if $u = 0$
(ii) $h(u+v) \leq h(u) + h(v)$ for all $u, v \in \mathbb{R}_0^+$.
Then $d' := h \circ d$ is a metric on X.

Proof. We show that d' satisfies the metric properties from Definition 18.
(a) Because of (i) and the metric properties of d, it is

$$(h \circ d)(x,y) = h(d(x,y)) = 0 \Leftrightarrow d(x,y) = 0 \Leftrightarrow x = y.$$

(b) Since d is a metric, we have

$$(h \circ d)(x,y) = h(d(x,y)) = h(d(y,x)) = (h \circ d)(y,x).$$

(c) Let $c \in \mathbb{R}_0^+$ such that $d(x,y) + c = d(x,z) + d(z,y)$, then by (ii) we get

$$\begin{aligned}(h \circ d)(x,y) = h(d(x,y)) &\leq h(d(x,y)+c) - h(c) \leq h(d(x,y)+c) \\ &= h(d(x,z) + d(z,y)) \leq h(d(x,z)) + h(d(z,y)) \\ &= (h \circ d)(x,z) + (h \circ d)(z,y).\end{aligned}$$

□

Corollary 4. *If d is any metric on a set X, then the function d' defined by*

$$d'(x, y) := \min\{d(x, y), 1\}$$

for all $x, y \in X$ is a bounded metric on X.

Proof. Obviously d' is bounded on X. We show that d' is a metric on X. It is

$$\min\{d(x, y), 1\} = \frac{1}{2}(d(x, y) + 1 - |d(x, y) - 1|) = g(d(x, y)),$$

where $g\colon \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is a function given by $g(x) := \frac{1}{2}(x + 1 - |x - 1|)$. It now suffices to show that g satisifies the conditons (i) and (ii) from Proposition 6. We see immediately that $g(0) = 0$. Futhermore we have

$$0 = g(x) = \frac{1}{2}(x + 1 - |x - 1|) \Rightarrow |x - 1| = x + 1$$

Because $|x - 1| = \max\{x - 1, -(x - 1)\}$, it is either $|x - 1| = x - 1$ or $|x - 1| = -(x - 1)$. But since $x - 1 \neq x + 1$ and $-(x - 1) = x + 1 \Rightarrow x = 0$, it follows from $h(x) = 0$, that $x = 0$. Furthermore

$$\begin{aligned} h(u + v) &= \frac{1}{2}(u + v + 1 - |u + v - 1|) = \frac{1}{2}(u + v + 1 - |u - 1 + v - 1 + 1|) \\ &\leq \frac{1}{2}(u + v + 1 - |u - 1| - |v - 1| + 1) = h(u) + h(v). \end{aligned}$$

Thus d' is a metric on X. □

We will assume from now on that whenever we deal with a sequence of metric spaces each space is bounded by 1.

Definition 20. *A function $f\colon X_1 \to X_2$ from a metric space (X_1, d_1) to a metric space (X_2, d_2) is uniformly continuous if there exists $w\colon \mathbb{R}^+ \to \mathbb{R}^+$ such that for all $x, y \in X_1$ and for all $\epsilon \in \mathbb{R}^+$*

$$d_1(x, y) \leq w(\epsilon) \Rightarrow d_2(f(x), f(y)) \leq \epsilon.$$

The function w is called a modulus of continuity of f on X_1.

Example 2. *Let X be an arbitrary metric space and consider the function $x \to d(x, x_0)$. Let $x, y \in X$. As the metric properties imply that $d(x, x_0) \leq d(x, y) + d(y, x_0)$, we get*

$$|d(x, x_0) - d(y, x_0)| \leq d(x, y) \leq \epsilon,$$

if $d(x, y) \leq \epsilon$. Therefore the function $x \to d(x, x_0)$ is uniformely continuous with modulus of continuity $\epsilon \to \epsilon$.

Example 3. *Let (X, d) be the product of sequence of metric spaces $\{(X_n, d_n)\}_{n=1}^{\infty}$. Let $x := \{x_n\}, y := \{y_n\} \in X$ and for each $n \in Z^+$ let $\pi_n(x) := x_n$. If $\sum_{k=1}^{\infty} 2^{-k} d(x_k, y_k) \leq 2^{-n}\epsilon$, then we have $2^{-n} d(x_n, y_n) \leq 2^{-n}\epsilon$ and thus $d(x_n, y_n) \leq \epsilon$. So we have shown that π_n is uniformly continuous with modulus of continuity $\epsilon \to 2^{-n}\epsilon$.*

Definition 21. *A sequence $\{f_n\}$ of functions from a set X_1 to a metric space X_2 converges uniformly to a function $f\colon X_1 \to X_2$ if for each $\epsilon > 0$ there exists N_ϵ in $\mathbb{Z}^+$ such that for all $x \in X_1$ and $n \geq N_\epsilon$*

$$d(f_n(x), f(x)) \leq \epsilon.$$

Proposition 8. *The sum $f + g$ and product fg of uniformly continuous functions $f\colon X \to \mathbb{R}$ and $g\colon X \to \mathbb{R}$ are uniformly continuous, and f^{-1} is uniformly continuous if $|f(x)| \geq c$ for all x in X and some $c > 0$. The composition $f_2 \circ f_1$ of uniformly continuous functions $f_1\colon X_1 \to X_2$, $f_2\colon X_2 \to X_3$ is uniformly continuous. The limit $f\colon X_1 \to X_2$ of a uniformly convergent sequence $\{f_n\}$ of uniformly continuous functions from a metric space X_1 to a metric space X_2 is uniformly continuous.*

Definition 22. *A sequence $\{x_n\}$ of elements of a metric space X converges to an element x of X, written either $\lim\limits_{n\to\infty} x_n = x$ or $x_n \to x$ as $n \to \infty$ if*

$$\lim_{n\to\infty} d(x_n, x) = 0$$

Definition 23. *Let X be a metric space. A subset Y of X is closed if and only if it contains all points that are limits of sequences of points of Y. The closure of a subset Y of X consists of all y in Y and of all x in X that are limits of sequences of points of Y and can be written as*

$$\overline{Y} := \{x \in X\colon \forall\epsilon > 0\ \exists y \in Y\colon d(x, y) \leq \epsilon\}.$$

A subset Y of X is dense in X if its closure is X.

3.2 Compactness

Definition 24. *A Cauchy sequence $\{x_n\}$ of elements of a metric space X is a sequence such that for each $\epsilon > 0$ there exists a positive integer N_ϵ with*

$$d(x_m, x_n) \leq \epsilon$$

for all $m, n \geq N_\epsilon$. The metric space X is called complete if every Cauchy sequence converges.

Theorem 11. *The product (X, d) of a sequence $\{(X_n, d_n)\}$ of complete metric spaces is complete.*

Proof. As mentioned earlier we assume that the (X_n, d_n) are bounded by 1 for all n. Let $\{x^k\}$, where for each n it is $x^k := \{x_n^k\}_{n=1}^{\infty}$ with $x_n^k \in X_n$, be a Cauchy sequence of elements of X. We show that $\{x^k\}$ converges. Let $n \in \mathbb{Z}^+$ and $\epsilon > 0$. We begin by showing that $\{x_n^k\}_{k=1}^{\infty}$ is a Cauchy sequence. Since $\{x^k\} \in X$ is a Cauchy sequence, we can choose $N \in \mathbb{Z}^+$ such that for all $k, j \geq N$ get $d(x^k, x^j) \leq \frac{\epsilon}{2^n}$. Then we have

$$d_n(x_n^k, x_n^j) \leq 2^n \sum_{i=1}^{\infty} 2^{-i} d_i(x_i^k, x_i^j) = 2^n d(x^k, x^j) \leq 2^n \cdot \frac{\epsilon}{2^n} = \epsilon$$

for all $k, j \geq N$.
Since (X_n, d_n) is complete for each n, the Cauchy sequence $\{x_n^k\}_{k=1}^{\infty}$ converges to a point $x_n^0 \in X_n$. We are now ready to show that $\{x^k\}$ converges to the point $x^0 := \{x_n^0\}_{n=1}^{\infty} \in X$. Since $\lim\limits_{k \to \infty} d_n(x_n^k, x_n^0) = 0$, we get for k sufficiently large and arbitrary $N \in \mathbb{Z}^+$ that

$$\begin{aligned}
d(x^k, x^0) &= \sum_{n=1}^{N} 2^{-n} d_n(x_n^k, x_n^0) + \sum_{n=N+1}^{\infty} 2^{-n} d_n(x_n^k, x_n^0) \\
&\leq \sum_{n=1}^{N} 2^{-n} d_n(x_n^k, x_n^0) + \sum_{n=N+1}^{\infty} 2^{-n} \\
&= \sum_{n=1}^{N} 2^{-n} d_n(x_n^k, x_n^0) + \Big(1 - \sum_{n=1}^{N} 2^{-n}\Big) \\
&= \sum_{n=1}^{N} 2^{-n} d_n(x_n^k, x_n^0) + \left(1 - \left(\frac{1 - 2^{-(N+1)}}{2^{-1}} - 1\right)\right) \\
&= \sum_{n=1}^{N} 2^{-n} d_n(x_n^k, x_n^0) + 2^{-N} \leq 2^{-N} + 2^{-N} = 2^{-N+1},
\end{aligned}$$

where we used the geometric series and the fact that $\sum_{n=1}^{\infty} 2^{-n} = 1$. Since N was arbitrary we are done. $\square$

Definition 25. *A metric space X is totally bounded if for each $\epsilon > 0$ there exists a finite subset $\{x_1, \ldots, x_n\}$ (where n is a positive integer depending on ϵ) of X, called an ϵ approximation to X, such that for each x in X at least one of the numbers $d(x, x_1), \ldots, d(x, x_n)$ is less than ϵ.*

The definition in constructive analysis differs from that in classical analysis, according to which a set is compact if each open cover has a finite subcover. In contrast, we define compactness via the properties of total boundness and completeness.

Definition 26. *A compact metric space, or simply a compact space, is a metric space which is complete and totally bounded.*

Proposition 9. *The product of a sequence of compact spaces is compact.*

Proof. We have already shown that the product of a sequence of complete metric spaces is complete. It can easily be seen, that the product of a sequence of totally bounded metric spaces is totally bounded. □

Proposition 10. *The image $f(X)$ of a totally bounded metric space X under a uniformly continuous function $f\colon X \to Y$ is totally bounded.*

Proof. Let $\epsilon > 0$ and let w be a modulus of continuity for f. Since X is totally bounded, there exists a finite subset $\{x_1, \cdots, x_n\}$ of X, such that for each $x \in X$ at least one of the numbers $d(x, x_1), \dots, d(x, x_n)$ is less than $w(\frac{\epsilon}{2})$. Due to the continuity of f, it follows that

$$d(f(x), f(x_i)) \leq \frac{\epsilon}{2} < \epsilon$$

for at least one value of $i \in \{1, \dots, n\}$. Therefore $\{f(x_1), \dots, f(x_n)\}$ is an ϵ approximation to $f(X)$ and thus $f(X)$ is totally bounded. □

Corollary 5. *Let $f\colon X \to \mathbb{R}$ be a uniformly continuous function on a totally bounded metric space X. Then the least upper bound and greatest lower bound of the subset $f(X)$ of $\mathbb{R}$, called respectively the supremum or sup and the infimum or inf of f on X, and written* $\sup\{f(x)\colon x \in X\}$ *and* $\inf\{f(x)\colon x \in X\}$*, exist.*

Proof. We just showed that $f(X)$ is totally bounded. Since we know that on a totally bounded set the least upper bound and the greatest lower bound exist, the Corollary follows. □

Definition 27. *For each compact space X and each metric space Y, we shall use $C(X, Y)$ to denote the set of all (uniformly) continuous functions from*

X to Y. For $C(X,\mathbb{R})$ we write simply $C(X)$. The metric d on $C(X,Y)$ is defined by

$$d(f,g) := \sup\{d(f(x),g(x)) \colon x \in X\}$$

for $f,g \in C(X,Y)$. In case $Y = \mathbb{R}$, the metric d is related to the norm

$$\|f\| := \sup\{|f(x)| \colon x \in X\}$$

on $C(X)$ by the equality

$$d(f,g) := \|f - g\|$$

for $f,g \in C(X)$.

4 The constructive Stone-Weierstrass theorem

4.1 Groundwork for the Stone-Weierstrass theorem

We now turn to the Stone-Weierstrass theorem, but for this we need some groundwork first. Lemma 3 will be proved in the following, in contrast to the explanations of Bishop, very detailed.

Definition 28. *A polynomial (of degree N in n variables) is a function $p \colon \mathbb{R}^n \to \mathbb{R}$ of the form*

$$p(x_1,\ldots,x_n) = \sum_{0 \le i_1+\ldots+i_n \le N} a_{i_1\cdots i_n} x_1^{i_1} \cdots x_n^{i_n}$$

If $p(0) = 0$, the polynomial p is strict.

Lemma 3. *For each $\epsilon > 0$ there exists a strict polynomial $p \colon \mathbb{R} \to \mathbb{R}$ such that for all $x \in \mathbb{R}$ with $|x| \le 1$*

$$\Big| |x| - p(x) \Big| \le \epsilon$$

Proof. In order to prove our Lemma, let us first consider the function f defined by

$$f \colon (-1,1) \to \mathbb{R},\ t \to (1-t)^{\frac{1}{2}}$$

To determine the Taylor series from f about 0, we need the nth derivative of f. We therefore examine the following examples:

$$f^{(1)}(t) = (-1)^1 \cdot \frac{1}{2} \cdot (1-t)^{-\frac{1}{2}}$$
$$f^{(2)}(t) = (-1)^2 \cdot \frac{1}{2} \cdot -\frac{1}{2} \cdot (1-t)^{-\frac{3}{2}}$$
$$f^{(3)}(t) = (-1)^3 \cdot \frac{1}{2} \cdot -\frac{1}{2} \cdot -\frac{3}{2} \cdot (1-t)^{-\frac{5}{2}}$$

So we can assume that

$$f^{(n)}(t) = (-1)^n \cdot \left(\prod_{i=1}^{n} \frac{3-2i}{2} \right) \cdot (1-t)^{\frac{1}{2}-n} \tag{3}$$

$$= -(2)^n \cdot \left(\prod_{i=2}^{n} 2i-3 \right) \cdot (1-t)^{\frac{1}{2}-n}$$

We proof our assertion by induction. We have already shown the case n=1. We assume (1) is true for a fixed $n \in \mathbb{Z}^+$ and we show that (1) must be true for $n+1$ as well.

$$\begin{aligned}
f^{(n+1)}(t) = (f^{(n)}(t))' &= (-1)^n \cdot \prod_{i=1}^{n} \left(\frac{3-2i}{2} \right) \cdot ((1-t)^{\frac{1}{2}-n})' \\
&= (-1)^n \cdot \prod_{i=1}^{n} \left(\frac{3-2i}{2} \right) \cdot \left(\frac{1}{2} - n \right) \cdot (1-t)^{\frac{1}{2}-n-1} \cdot (-1) \\
&= (-1)^{n+1} \cdot \prod_{i=1}^{n} \left(\frac{3-2i}{2} \right) \cdot \left(\frac{3-2(n+1)}{2} \right) \cdot (1-t)^{\frac{1}{2}-(n+1)} \\
&= f^{(n+1)}(t)
\end{aligned}$$

Therefore the Taylor series for f about 0 is given by

$$\sum_{n=0}^{\infty} \frac{f(n)(0)}{n!}(t-0)^n = 1 - \sum_{n=1}^{\infty} (2^n n!)^{-1} \left(\prod_{i=2}^{n} 2i-3 \right) t^n. \tag{4}$$

We will show that this series converges to $(1-t)^{\frac{1}{2}}$. To this end, we first define

$$R_N := (1-t)^{\frac{1}{2}} - 1 + \sum_{n=1}^{N} (2^n n!)^{-1} \left(\prod_{i=2}^{n} 2i-3 \right) t^n,$$

where R_N is the associated remainder term to the Nth taylor polynomial. By Taylor's theorem, we know that for each $\epsilon > 0$ there exists c such that $\min\{0,t\} \leq c \leq \max\{0,t\}$ and

$$\left| R_N - \frac{f^{(N+1)}(c)}{N!}(t-c)^N t \right| \leq \epsilon.$$

Since the set $\{c \in \mathbb{R} : \min\{0,t\} \leq c \leq \max\{0,t\}\}$ is compact,

$$m := inf\left\{ \left| R_N - \frac{f^{(N+1)}(c)}{N!}(t-c)^N t \right| : \min\{0,t\} \leq c \leq \max\{0,t\} \right\}$$

exists. Since ϵ is an arbitrary positive number, it is $m = 0$. Therefore we know that there exists c in $[\min\{0,t\}, \max\{0,t\}]$ such that

$$R_N = \frac{f^{(N+1)}(c)}{N!}(t-c)^N t = -(2^{N+1}N!)^{-1}\left(\prod_{i=2}^{N+1} 2i-3\right)(1-c)^{\frac{1}{2}-N}(t-c)^N t.$$

Hence

$$\begin{aligned} |R_N| &= (2^{N+1}N!)^{-1}\left(\prod_{n=2}^{N+1} 2n-3\right)|t(1-c)^{\frac{1}{2}}| \cdot |(t-c)(1-c)^{-1}|^N \\ &\leq |t(1+|t|)|^{\frac{1}{2}}(2^{N+1}N!)^{-1}\left(\prod_{n=2}^{N+1} 2n-3\right) \cdot |(t-c)(1-c)^{-1}|^N. \end{aligned}$$

It is easily seen that for all $n \in \mathbb{Z}^+$, it is

$$(2^{n+1}n!)^{-1}\left(\prod_{n=2}^{n+1} 2n-3\right) \leq \left(2^n\Big((n-1)!\Big)\right)^{-1}\left(\prod_{n=2}^{n} 2n-3\right).$$

Thus by the ratio test, it follows that if $|x| < 1$ the series

$$\sum_{n=1}^{\infty}(2^{n+1}n!)^{-1}\left(\prod_{n=2}^{n+1} 2n-3\right)x^n$$

converges and with that we have (for $|x| < 1$)

$$\lim_{n\to\infty}(2^{n+1}n!)^{-1}\left(\prod_{n=2}^{n+1} 2n-3\right)x^n = 0.$$

For $t \in (-1, 1)$ the term $|t(1+|t|)|^{\frac{1}{2}}$ is is obviously bounded and futhermore we have

$$\left|\frac{t-c}{1-c}\right| < 1 \quad (c \in [\min\{0,t\}, \max\{0,t\}]).$$

Together we now get $\lim\limits_{N\to\infty} R_N = 0$ and have finally shown that the Taylor series from (2) converges to f on $(-1, 1)$. We now show that this is also the case on $[-1, 1]$. It holds for $|t| < 1$, that

$$\sum_{n=1}^{\infty}(2^n n!)^{-1}\left(\prod_{i=2}^{n} 2i-3\right)t^n \leq \sum_{n=1}^{\infty}(2^n n!)^{-1}\left(\prod_{i=2}^{n} 2i-3\right) \tag{5}$$

If we apply Raabe's test, we get that (3) converges, as

$$\lim_{n\to\infty} n\left(\frac{\prod_{k=2}^{n} 2k-3}{2^n \cdot n!} \cdot \frac{2^{n+1}\cdot(n+1)!}{\prod_{k=2}^{n+1} 2k-3} - 1\right) = \frac{3}{2} > 1.$$

Thus our Taylor series from (2) must also converge to a continuous function on $[-1, 1]$. For reasons of continuity, this function must be f again.
We have therefore shown, that for all $t \in [-1, 1]$ and all $\epsilon > 0$ for some N it is

$$|(1-t)^{\frac{1}{2}} - g(t)| \leq \frac{\epsilon}{2},$$

where

$$g(t) = 1 - \sum_{n=1}^{N}(2^n n!)^{-1}\left(\prod_{i=2}^{n} 2i-3\right)t^N.$$

It follows for all $\epsilon > 0$ and $|x| \leq 1$, that there exists N such that

$$\Big||x| - g(1-x^2)\Big| = \Big|(1-(1-x^2))^{\frac{1}{2}} - g(1-x^2)\Big| \leq \frac{\epsilon}{2}.$$

Because $g(1-x^2)$ is not a strict polynomial, we define

$$p(x) := g(1-x^2) - g(1).$$

Then p is strict, since $p(0) = g(1) - g(1) = 0$ and since $g(1)$ converges to 0, we have for all $\epsilon > 0$ and $|x| \leq 1$

$$\Big||x| - p(x)\Big| = \Big||x| - g(1-x^2) + g(1)\Big| \leq \Big||x| - g(1-x^2)\Big| + g(1) \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for some N. $\square$

Definition 29. *Let G be a family of real-valued functions on a set X. Then $U(G)$ is the family of all real-valued functions f on X of the form*

$$f := p \circ g = \sum_{0 \leq i_1 + \ldots + i_n \leq N} a_{i_1, \ldots, i_n} g_1(x)^{i_1} \cdot \cdots \cdot g_n(x)^{i_n},$$

where $p \colon \mathbb{R}^n \to \mathbb{R}$ is a strict polynomial and where $g \colon X \to \mathbb{R}^n$ has the form

$$g(x) := (g_1(x), \cdots, g_n(x))$$

for some $g_1 \cdots, g_n$ in G.

Theorem 12. *$U(G)$ is the smallest family of real-valued functions on X including G and closed with respect to the operations of addition, multiplication, and multiplication by real numbers.*

Proof. First we show that $U(G)$ includes G and that $U(G)$ is closed with respect to the operations of addition, multiplication, and multiplication by real numbers. To this end, first note that $G \subseteq U(G)$, since for every $g \in G$ we have $(p \circ g)(x) = g(x)$, if p is given by $p \colon \mathbb{R} \to \mathbb{R},\ p(x) := x$. It then also holds that $p(0) = 0$ and therefore p is a strict polynomial and it follows that every g in G is also in $U(G)$. Take $f_1, f_2 \in U(G)$. It is easily seen that the sum and product of f_1 and f_2 must be of the form $f = p \circ g$, where p is a polynomial and g is again as in the definition above. Furthermore, if p is the polynomial associated with the sum of $f_1 + f_2$ and p_1 and p_2 are the polynomials associated with f_1 and f_2, then $p(0) = 0$ must apply, since $p(0) = p_1(0) + p_2(0) = 0 + 0 = 0$ and analogously, if p is the polynomial associated with the product of $f_1 \cdot f_2$, then $p(0) = p_1(0) \cdot p_2(0) = 0 \cdot 0 = 0$ holds. It therefore follows that $f_1 + f_2$ and $f_1 \cdot f_2$ are in $U(G)$. Take $f \in U(G)$ and k to be a real number. Then

$$(kf)(x) = (k(p \circ g))(x) = k(p(g(x))) = (kp)g(x) = ((kp) \circ g)(x)$$

and since $(kp)(0) = k \cdot p(0) = k \cdot 0 = 0$, we have $kf \in U(G)$.
Now consider the function

$$f := p \circ g = \sum_{0 \leq i_1 + \cdots + i_n \leq N} a_{i_1, \ldots, i_n} g_1(x)^{i_1} \cdot \cdots \cdot g_n(x)^{i_n}.$$

Since the function f is the result of finite multiplication and addition of functions of G and their multiplication with real numbers, it follows that each family of real-valued functions on X, which includes G and is closed with respect to the operations of addition, multiplication, and multiplication by real numbers, must contain f. Therefore $U(G)$ is the smallest of such families. $\square$

Theorem 13. *If $U(G)$ is closed with respect to the operations of addition, multiplication, and multiplication by real numbers, then so is $\overline{U(G)}$.*

Proof. First we remember that on a compact set X for every continuous function $f\colon X \to \mathbb{R}$ the supremum $\sup\{f(x)\colon x \in X\}$ exists. Let $\epsilon > 0$. Let $f_1, f_2 \in \overline{U(G)}$, $g_1, g_2 \in U(G)$ such that $\sup\{|f_i(x) - g_i(x)|\colon x \in X\} < \frac{\epsilon}{2}$ for $i \in \{1, 2\}$ and define $g := g_1 + g_2 \in U(G)$ and $f := f_1 + f_2$. From the obvious properties of the supremum then follows

$$
\begin{aligned}
&\sup\{|f(x) - g(x)|\colon x \in X\}\\
&= \sup\{|f_1(x) + f_2(x) - g_1(x) - g_2(x)|\colon x \in X\}\\
&\leq \sup\{|f_1(x) - g_1(x)| + |f_2(x) - g_2(x)|\colon x \in X\}\\
&\leq \sup\{|f_1(x) - g_1(x)|\colon x \in X\} + \sup\{|f_2(x) - g_2(x)|\colon x \in X\} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}
$$

and therefore we have $f \in \overline{U(G)}$.

Let $f \in \overline{U(G)}$ and let $g \in U(G)$ such that $\sup\{|f(x) - g(x)|\colon x \in X\} < \frac{\epsilon}{2M}$, where $M := \max\{\sup\{|f(x)|\colon x \in X\}, \sup\{|g(x)|\colon x \in X\}\}$. Then it is $g^2 \in U(G)$ and we get

$$
\begin{aligned}
&\sup\{|f^2(x) - g^2(x)|\colon x \in X\}\\
&= \sup\{|f(x) - g(x)| \cdot |f(x) + g(x)|\colon x \in X\}\\
&\leq \sup\{|f(x) - g(x)|\colon x \in X\} \cdot \sup\{|f(x) + g(x)|\colon x \in X\}\\
&< \frac{\epsilon}{2M} \cdot (\sup\{|f(x)|\colon x \in X\} + \sup\{|g(x)|\colon x \in X\}) \leq \frac{\epsilon}{2M} \cdot 2M = \epsilon.
\end{aligned}
$$

and therefore $f^2 \in \overline{U(G)}$.

Let $f \in \overline{U(G)}$, $\alpha \in \mathbb{R}$ and let $g \in U(G)$ such that $\sup\{|f(x) - g(x)|\colon x \in X\} < \frac{\epsilon}{|\alpha|}$. Then it is $\alpha g \in U(G)$ and we have

$$
\begin{aligned}
&\sup\{|(\alpha f)(x) - (\alpha g)(x)|\colon x \in X\}\\
&= \sup\{|\alpha| \cdot |f(x) - g(x)|\colon x \in X\}\\
&= |\alpha| \cdot \sup\{|f(x) - g(x)|\colon x \in X\} < |\alpha| \cdot \frac{\epsilon}{|\alpha|} = \epsilon.
\end{aligned}
$$

and therfore it is $\alpha f \in \overline{U(G)}$.

Since $fg = \frac{1}{2}((f+g)^2 - f^2 - g^2)$, it follows $fg \in \overline{U(G)}$ whenever $f, g \in \overline{U(G)}$. □

Lemma 4. *Let G be any family of continuous functions on a compact space X. Let f and g be any functions belonging to the closure H of $U(G)$ in $C(X)$. Then $|f|$, $\max\{f,g\}$ and $\min\{f,g\}$ also belong to H.*

Proof. Because of the last two theorems, it is $H = U(H)$. We choose $c > 0$ such that

$$\|cf\| = \sup\{|cf(x)| \colon x \in X\} = c \cdot \sup\{|f(x)| \colon x \in X\} \leq 1.$$

This is possible, since X is compact and thus $\sup\{|f(x)| \colon x \in X\}$ exists. We can now insert cf into the statement of Lemma 3 and obtain for all $\epsilon > 0$ and all $x \in X$

$$\Big||cf(x)| - p(cf(x))\Big| \leq \epsilon,$$

where p is a strict polynomial function defined as in Lemma 3. Since H is closed and $p \circ (cf) \in H$, we get $|cf| \in H$. Now it follows from $|cf| \in H$, that $|\frac{1}{c}||cf| = |f| \in H$ and thus

$$\max\{f,g\} = \frac{1}{2}\Big(f + g + |f-g|\Big) \in H$$
$$\min\{f,g\} = \frac{1}{2}\Big(f + g - |f-g|\Big) \in H$$

□

Definition 30. *Let X be a compact space. A family $G \subset C(X)$ is separating if there exists a function $\delta \colon \mathbb{R}^+ \to \mathbb{R}^+$
such that (Sep1) whenever $\epsilon > 0$, $x, y \in X$, and $d(x,y) \geq \epsilon$ there exists g in G satisfying for all $z \in X$*

$$d(x,z) \leq \delta(\epsilon) \Rightarrow |g(z)| \leq \epsilon \;\; \textit{and}$$
$$d(y,z) \leq \delta(\epsilon) \Rightarrow |g(z) - 1| \leq \epsilon$$

and such that (Sep2) whenever $\epsilon > 0$ and $y \in X$ there exists g such that for all $z \in X$

$$d(y,z) \leq \delta(\epsilon) \Rightarrow |g(z) - 1| \leq \epsilon.$$

Lemma 5. *Let G be a separating family of continuous functions on a compact space X, and let H be the closure of $U(G)$ in $C(X)$. Let h be a function in H with $c := \inf\{|h(x)|\colon x \in X\} > 0$. Then $h^{-1} \in H$.*

Proof. We define

$$\lambda := \frac{h^2}{(\sup\{|h(x)|\colon x \in X\})^2}.$$

Then

$$\frac{c^2}{\|h\|^2} = \frac{(\inf\{|h(x)|\colon x \in X\})^2}{(\sup\{|h(x)|\colon x \in X\})^2} \leq \frac{h^2}{(\sup\{|h(x)|\colon x \in X\})^2} = \lambda \leq 1$$

and therefore

$$\|1-\lambda\| = \sup\{|1-\lambda(x)|\colon x \in X\} \leq 1 - c^2\|h\|^{-2} < 1.$$

We define the series

$$\sum_{n=0}^{\infty} \frac{h(1-\lambda)^n}{\|h\|^2}.$$

Using the geometric series, we get

$$\sum_{n=0}^{\infty} \frac{h(1-\lambda)^n}{\|h\|^2} = \frac{h}{\|h\|^2}\sum_{n=0}^{\infty}(1-\lambda)^n = \frac{h}{\|h\|^2} \cdot \frac{1}{1-(1-\lambda)} = h^{-1}.$$

Since every partial sum of the series is in H and H is closed, h^{-1} must also be in H. □

4.2 The Stone-Weierstrass theorem

Having presented the groundwork, we now turn to the core of the constructive Stone-Weierstrass theorem.

Theorem 14. *Let G be a seperating family of continuous functions on a compact space X. Then U(G) is dense in C(X).*

Proof. We want to show that $\overline{U(G)} = C(X)$. In order to show this, let H be the closure of U(G) in C(X). As we have seen before, it is $H = U(H)$. We begin by proving $1 \in H$. It suffices to construct a function $h \in H$ such that

$\inf\{|h(x)| : x \in X\} > 0$. Then by Lemma 5, we have $h^{-1} \in H$ and therefore $h \cdot h^{-1} = 1 \in H$. For this purpose, let $x_1, \ldots, x_n$ be a $\delta(\frac{1}{2})$ approximation to X, i.e. for each x in X we have $d(x, x_i) < \delta(\frac{1}{2})$ for at least one i, where δ is the operation of Definition 30. Let $z \in X$. Because of (Sep2) of Definition 30, we have that for each i ($1 \leq i \leq n$) with $d(x_i, z) \leq \delta(\frac{1}{2})$ there exists g_i in G satisfying

$$|g_i(z) - 1| \leq \frac{1}{2} \Leftrightarrow |1 - g_i(z)| \leq \frac{1}{2} \Rightarrow 1 - g_i(z) \leq \frac{1}{2} \Leftrightarrow g_i(z) \geq \frac{1}{2}$$

Because of Lemma 4, we know that $h := \max\{g_1, \ldots, g_n\} \in H$. Since $d(x_i, z) < \delta(\frac{1}{2})$ for at least one i, it follows $h(z) \geq \frac{1}{2} > 0$ and $1 \in H$ as claimed.

To finish our proof, our next step will be to show that for each y in X and each r with $0 < r \leq \frac{1}{4}$ there exists $\lambda \in H$ with $0 \leq \lambda \leq 1$ such that $\lambda(z) = 1$ whenever $d(y, z) \leq \delta(r)$ and $\lambda(z) = 0$ whenever $d(y, z) \geq 3r$. For this purpose, let y be an arbitrary point in X and let r be an arbitrary point satisfying $0 < r \leq \frac{1}{4}$. Let $\{x_1, \ldots, x_n\}$ be a $c := \min\{r, \delta(r)\}$ approximation to X, where δ is given by Definition 30 as before. Futhermore, let S be the set of all i ($1 \leq i \leq n$) with $d(x_i, y) > r$ and let T be the set of all i ($1 \leq i \leq n$) with $d(x_i, y) < 2r$. Then the set $\{1, \ldots, n\}$ is apparently the union of the finite sets S and T. Because of (Sep1) of Definition 30, we have that for each i in S there exists g_i in G such that in case $d(x_i, z) \leq \delta(r)$ we have $|g_i(z)| \leq r \Rightarrow g_i(z) \leq r \leq \frac{1}{4}$ and in case $d(y, z) \leq \delta(r)$ we have $|g_i(z) - 1| \leq r \Leftrightarrow |1 - g_i(z)| \leq r \Rightarrow 1 - g_i(z) \leq r \Leftrightarrow \frac{3}{4} \leq 1 - r \leq g_i(z)$. For each i in S we define

$$h_i := \min\{1, \max\{0, 2g_i - \frac{1}{2}\}\}.$$

Then the following holds: $h_i \in H$, $0 \leq h_i \leq 1$, $h_i(z) = 0$ if $d(x_i, z) \leq \delta(r)$ and $h_i(z) = 1$ if $d(y, z) \leq \delta(r)$. We further define

$$\lambda := \prod_{i \in S} h_i.$$

It holds: $\lambda \in H$, $0 \leq \lambda \leq 1$ and $\lambda(z) = 1$ if $d(y, z) \leq \delta(r)$. We now consider z in X with $d(y, z) \geq 3r$ and we choose i ($1 \leq i \leq n$) with $d(x_i, y) \leq c \leq r$, which is possible because $\{x_1, \cdots, x_n\}$ is a c approximation to X. Applying the metric property $d(y, z) \leq d(x_i, y) + d(x_i, z)$, we get

$$d(x_i, y) \geq d(y, z) - d(x_i, z) \geq 3r - r = 2r.$$

Hence, $i \in S$ and since $d(x_i, z) \leq c \leq \delta(r)$ it is $\lambda(z) = h_i(z) = 0$ and we have thus proven our desired result.

We are now sufficiently prepared to show that every f in C(X) is also in H. Let w be a modulus of continuity of f. Let ϵ be an arbitrary positive real number and we define $r := \min\{\frac{1}{4}, \frac{1}{4}w(\epsilon)\}$. Let $\{y_1, \ldots, y_m\}$ be a $\delta(r)$ approximation to X. As we have shown above, for each y_j there exists λ_j in H satisfying $0 \leq \lambda_j \leq 1$, $\lambda_j(z) = 1$ in case $d(y_j, z) \leq \delta(r)$ and $\lambda_j(z) = 0$ in case $d(y_j, z) \geq 3r$. We continue by defining

$$\lambda := \sum_{j=1}^{m} \lambda_j.$$

We thus have $\lambda \in H$ and $\lambda \geq 1$, since for each $z \in X$ we have $d(y_j, z) \leq \delta(r)$ for at least one j and therefore $\lambda(z) \geq \lambda_j(z) \geq 1$. By Lemma 5, it follows $\lambda^{-1} \in H$. Let g be given by

$$g := \sum_{j=1}^{m} f(y_j)\lambda^{-1}\lambda_j.$$

Since the $f(y_j)$ $(0 \leq j \leq m)$ are constant, it is $g \in H$. Let $z \in X$. It is $\lambda_j(z) = 0$ if $d(y_j, z) \geq 3r$ and because of the continuity we have $|f(z) - f(y_j)| \leq \epsilon$ if $d(y_j, z) \leq w(\epsilon)$. Thus if we recognize that $3r \leq \frac{3}{4}w(\epsilon) < w(\epsilon)$ we can conclude that

$$|f(z) - g(z)| = \Big|\sum_{j=1}^{m}(f(z) - f(y_j))\lambda^{-1}(z)\lambda_j(z)\Big|$$

$$\leq \sum_{j=1}^{m} |f(z) - f(y_j)|\lambda^{-1}(z)\lambda_j(z) \leq \sum_{j=1}^{m} \epsilon\lambda^{-1}(z)\lambda_j(z) = \epsilon.$$

Since ϵ was arbitrary and H is closed, we finally get $f \in H$ and have thus completed our proof. □

Corollary 6. *Let $\{(X_n, d_n)\}_{n=1}^{\infty}$ be compact spaces, with product (X, d). Let G be the set of all functions on X of the form*

$$g := f_n \circ \pi_n$$

where $\pi_n \colon X \to X_n$ is the projection of X onto X_n and $f_n \colon X_n \to \mathbb{R}$ is continuous. Then the set H of all h in $C(X)$ of the form $h = h_1 + \cdots + h_n$, where each h_i is a finite product of functions in G, is dense in $C(X)$.

Proof. Theorem 14 tells us that for each separating family G of continuous functions in a compact space X, we have U(G) is dense in C(X). In Example 3 we have already seen that every projection $\pi_n \colon X \to X_n$ is continuous and since every f_n is continuous by assumption, we get that every g in G is also continuous. Since we have already seen that the space (X, d) is compact, to apply Theorem 14 we only have to show that G is a separating family. To this end, let $\epsilon > 0$ and $x := \{x_n\}, y := \{y_n\}$ be points of X with $d(x, y) \geq \epsilon$. As the series $\sum_{n=1}^{\infty} 2^{-n} = 1$ converges, we can choose $N \in \mathbb{Z}^+$ satisfying $\sum_{n=N+1}^{\infty} 2^{-n} < \frac{\epsilon}{2}$. Using that for each n the metric d_n is bounded by 1, we estimate

$$\begin{aligned}\sum_{n=1}^{N} 2^{-n} d_n(x_n, y_n) &> \sum_{n=1}^{N} 2^{-n} d_n(x_n, y_n) + \sum_{n=N+1}^{\infty} 2^{-n} d_n(x_n, y_n) - \frac{\epsilon}{2} \\ &= d(x, y) - \frac{\epsilon}{2} \geq \frac{\epsilon}{2}.\end{aligned}$$

From this follows

$$\begin{aligned}\sum_{n=1}^{N} 2^{-n} d_n(x_n, y_n) > \left(\frac{\epsilon}{2}\right) \cdot \sum_{n=1}^{N} 2^{-n} &\Leftrightarrow \sum_{n=1}^{N} 2^{-n} \left(d_n(x_n, y_n) - \frac{\epsilon}{2}\right) > 0 \\ \Rightarrow d_n(x_n, y_n) - \frac{\epsilon}{2} > 0 &\Leftrightarrow d_n(x_n, y_n) > \frac{\epsilon}{2}\end{aligned}$$

for some n with $n \leq N$, where we used Proposition 1. We now define the function

$$f_n \colon X_n \to \mathbb{R},\ z_n \to \frac{d_n(x_n, z_n)}{d_n(x_n, y_n)},$$

which is continuous as we have seen in Example 2. Futhermore, we define $g := f_n \circ \pi_n$. Let $z := \{z_n\} \in X$. If we have $d(x, z) \leq \epsilon^2 2^{-N-1}$, then

$$\begin{aligned}|g(z)| = |f_n(z_n)| &= \left|\frac{d_n(x_n, z_n)}{d_n(x_n, y_n)}\right| \leq \frac{2}{\epsilon} \cdot d_n(x_n, z_n) \\ &\leq \frac{2}{\epsilon} \cdot 2^n \cdot d(x, z) \leq \frac{2}{\epsilon} \cdot 2^N \cdot d(x, z) \leq \frac{2}{\epsilon} \cdot 2^N \cdot \epsilon^2 2^{-N-1} = \epsilon\end{aligned}$$

and if we have $d(y,z) \le \epsilon^2 2^{-N-1}$, then

$$\begin{aligned}|g(z)-1| = |f_n(z_n)-f_n(y_n)| &= \left|\frac{d_n(x_n,z_n)-d_n(x_n,y_n)}{d_n(x_n,y_n)}\right| \\ &\le \frac{2}{\epsilon}\cdot |d_n(x_n,z_n)-d_n(x_n,y_n)| \le \frac{2}{\epsilon}\cdot d_n(y_n,z_n) \\ &\le \frac{2}{\epsilon}\cdot 2^N \cdot d(y,z) \le \epsilon,\end{aligned}$$

where we have used that $d_n(x_n,z_n) \le d_n(x_n,y_n)+d_n(y_n,z_n)$ holds, since d_n is a metric. So we showed that G satisfies (Sep1) of Definition 30 with $\delta(\epsilon) := \epsilon^2 2^{-N-1}$. To show that δ also satiesfies (Sep2), let $\epsilon > 0$ and $y := \{y_n\} \in X$. If we set $g\colon X \to \mathbb{R},\ g(z) = 1$, then for all $z \in X$ we get

$$d(y,z) \le \delta(\epsilon) \Rightarrow |g(z)-1| = 0 \le \epsilon$$

and thus G is a separating family. It therefore follows that $U(G)$ is dense $C(X)$.
H obviously includes G and is closed with respect to + and $\cdot$. Since every constant function from X to $\mathbb{R}$ is contained in G, it follows that H is also closed with respect to multiplication by real numbers. If we consider a function of the form

$$h = h_1 + \cdots + h_n \quad,\ h_i = g_1 \cdot \cdots \cdot g_n,$$

where $m, n \in \mathbb{Z}^+$,$i \in \{1,\ldots,n\}$ and $g_1, \cdots, g_n \in G$, then we immediately see that h must be an element of $U(G)$. Together $H = U(G)$ follows and thus H is dense in $C(X)$. $\square$

Corollary 7. *Let X be a compact space with* $\sup\{d(x,y)\colon x,y \in X\} > 0$*, and let G consist of all functions $x \to d(x,x_o)$, with $x_0 \in X$. Then $U(G)$ is dense in $C(X)$.*

Proof. We have already seen that $x \to d(x,x_0)$ is continuous for all x_0 in X. Thus every function in $U(G)$ is also continuous as a composition of continuous functions. We now show that $U(G)$ is separating, because then follows $U(U(G)) = U(G)$ is dense in $C(X)$.
To this end, let $\epsilon > 0$ and $x, y \in X$ with $d(x,y) \ge \epsilon$. We define the function g in $U(G)$ by

$$g(z) = \frac{d(x,z)}{d(x,y)}.$$

Then if $d(x,z) \leq \epsilon^2$, we have

$$|g(z)| = \left|\frac{d(x,z)}{d(x,y)}\right| \leq \frac{d(x,z)}{\epsilon} \leq \frac{\epsilon^2}{\epsilon} = \epsilon$$

and if $d(y,z) \leq \epsilon^2$, then we have

$$|g(z) - 1| = \left|\frac{d(x,z) - d(x,y)}{d(x,y)}\right| \leq \frac{|d(x,z) - d(x,y)|}{\epsilon} \leq \frac{d(y,z)}{\epsilon} \leq \epsilon^2,$$

where we have used that $d(x,z) \leq d(x,y) + d(y,z)$ holds, since d is a metric. So we have shown that (Sep1) of Definition 30 holds for $\delta(\epsilon) = \epsilon^2$.

Now let again $\epsilon > 0$, we still need to show that (Sep2) from Definition 30 also holds. To show this we will need some preparation. From Proposition 9 it follows that $M := \sup\{d(x,y) \colon x, y \in X\}$ exists, since X is compact and the identity function from X to X is uniformly continuous. Let $\{x_1, \dots, x_n\}$ be a finite $\frac{1}{7} \cdot M$ approximation to X. We suppose that there exists $x \in X$ such that $d(x, x_i) < \frac{2}{7} \cdot M$ for all $i \in \{1, \dots, n\}$. Futhermore let $y, z \in X$ and choose $j, k \in \{1, \dots, n\}$ such that $d(y, x_j) < \frac{1}{7} \cdot M$ and $d(z, x_k) < \frac{1}{7} \cdot M$. Then

$$d(y,z) \leq d(y, x_j) + d(x_j, x) + d(x, x_k) + d(x_k, z) < \frac{6}{7} \cdot M$$

holds, because of the metric properties. Now it follows that $M \leq \frac{6}{7} \cdot M$, which is a contradiction. We therefore can conclude that for every $x \in X$ it is $d(x, x_i) > \frac{1}{7} \cdot M$ for at least one i in $\{1, \dots, n\}$. If we define

$$g(z) := \frac{d(z, x_i)}{d(x, x_i)},$$

then $g \in U(G)$ and if $d(x,z) \leq \frac{1}{7} \cdot M\epsilon$, we get

$$|g(z) - 1| = \left|\frac{d(z, x_i) - d(x, x_i)}{d(x, x_i)}\right| \leq \frac{7 \cdot (|d(z, x_i) - d(x, x_i)|)}{M} \leq \frac{7 \cdot d(x,z)}{M} \leq \epsilon.$$

It follows that $U(G)$ is separating with $\delta(\epsilon) := \min\{\epsilon^2, \frac{1}{7} \cdot M\epsilon\}$. □

An interesting aspect of the third Corollary that follows is that the special case $n = 1$ and $X = [-1, 1]$ is the famous Weierstrass approximation theorem.

Corollary 8. *Every continuous function f on a compact set $X \subset \mathbb{R}^n$ can be arbitrarily closely approximated on X by polynomial functions $p\colon \mathbb{R}^n \to \mathbb{R}$.*

Proof. We prove the statement first for $n = 1$, so let $X \subset \mathbb{R}$ be compact. We recall Lemma 3. We already know that for all $x \in [-1, 1]$ the function $x \to |x|$ can be arbitrarily closely approximated by polynomial functions. Now consider a function of the form $g\colon X \to \mathbb{R},\ x \to |x - x_0|$ for $x_0 \in X$. Then $g(X)$ is compact as g is uniformly continuous on X and since polynomial functions are obviously closed with respect to multiplication by real numbers, it follows that the function g can be arbitrarily closely approximated on X by polynomial functions. Let G be the set of all functions of the form g. Applying Corollary 8 follows then $U(G)$ is dense in $C(X)$ and thus we can conclude that any function $f \in C(X)$ can be arbitrarily closely approximated on X by polynomial functions.

If we apply Corollary 7 to what has just been proved, the validity of the statement also follows in the case $X = [a, b]^n$ for some compact interval $[a, b]$.

Let now $X \subset \mathbb{R}^n$ be an arbitrary compact interval. Then a compact interval $[a, b]$ exists such that $X \subset [a, b]^n$. Since we know that each function $x \to d(x, x_0)$ with $x_0 \in [a, b]^n$ can be arbitrarily closely approximated on $[a, b]^n$ by polynomial functions, the statement follows also in the general case by another application of Corollary 8.

□

5 Bibliography

Bishop, Errett/ Bridges, Douglas: *Constructive Analysis*, Springer-Verlag, 1985

Bishop, Errett: *Foundations of Constructive Analysis*, McGraw-Hill Book Company, 1967

Bishop, Errett: *A generalization of the Stone-Weierstrass theorem*, Pacific J. Math. (Volume 11, Number 3, 777-783), 1961