Functional Analysis

Cordula Bernert

01.09.2013

Inhaltsverzeichnis

Literaturverzeichnis

[1] H. Heuser: Funktionalanalysis, BG Teubner 2006

[2] Appell, Väth: Elemente der Funktionalanalysis, Vieweg 2005

[3] D. Werner: Funktionalanalysis, Springer Lehrbuch, 6. Auflage 2007

[4] K. Burg, H. Haf, F. Wille, A. Meister: Partielle Differentialgleichungen und funktionalanalytische Grundlagen, Vieweg+Teubner,4.Auflage 2009

[5] K. Saxe: Beginning Functional Analysis, Springer 2000

[6] Joseph Muscat: Functional Analysis, Springer 2014

[7] A.N. Kolmogorov, S.V. Fomin: Elements of the Theory of Functions and Functional Analysis, Martino Publishing 2012

[8] G. Shilov: Elementary Functional Analysis, Dover publications 2013

1 Metric Spaces

A set of common mathematical objects with equal properties is called a space. The „easiest" one is the metric space. It belongs to the topological spaces.

Example 1.1 $\mathbb{R}^3 = \{\mathbf{x} = (x_1, x_2, x_3)^T \quad | \quad x_i \in \mathbb{R}; \quad i = 1, 2, 3\}$
** You are surely familiar with this 3-dimensional Euclidian space. It is a set of points with 3 components*
** The distance between the points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ is defined by:*

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$$

** d is something which can be measured (ruler!) and it is nonnegative. d is equal to zero if and only if you calculate $d(x, x)$. The distance between $\mathbf{x}$ and $\mathbf{y}$ is the same as the distance between $\mathbf{y}$ and $\mathbf{x}$. Also d satisfies the triangle inequality. $(*)$*
** The definition of the distance d gives us the possibility to describe the geometry of real objects and their position (length, width, height, radius, shape,...).*
** In mathematics the function of distance is called the metric or the metric function. Its definition is not unique. If d satisfies only the properties $(*)$ then d is a metric.*
** The metric d defines a topology. That means you can define a special system of open sets in $\mathbb{R}^3$ which is based on the metric function. On this system of open sets you can do analysis in $\mathbb{R}^3$, for example you can define the limit of a sequence of elements of $\mathbb{R}^3$, you can define the continuity of functions, derivatives, integrals,...*

Our goal is now to introduce the metrics for common sets of mathematical objects, for example in sets of functions, sequences, bodies, differential equations,... To get a metric space the metric function must satisfy the three conditions $(*)$. In metric spaces you can define a topology and do analysis.

1.1 Metric spaces

Definition 1.1 *A **metric space** $\mathbb{X}$ is defined to be a nonempty set $\mathbb{X}$ together with a real function d, $\mathbb{X} \times \mathbb{X} \to \mathbb{R}$, satisfying 3 conditions:*

1. $d(\mathbf{x}, \mathbf{y}) \geq 0, \quad \wedge \quad d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$ (nonnegativity, nondegeneracy)
2. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{X}$ (symmetry)
3. $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{X}$ (triangle inequality)

$d(\mathbf{x},\mathbf{y})$ is called the distance function or the **Metric in** $\mathbb{X}$.

Conclusion 1.1 *Common triangle inequality*

$$d(\mathbf{x}_1,\mathbf{x}_n) \le d(\mathbf{x}_1,\mathbf{x}_2) + d(\mathbf{x}_2,\mathbf{x}_3) + ... + d(\mathbf{x}_{n-1},\mathbf{x}_n) \quad \forall \mathbf{x}_1,\mathbf{x}_2,...\mathbf{x}_n \in \mathbb{X}$$

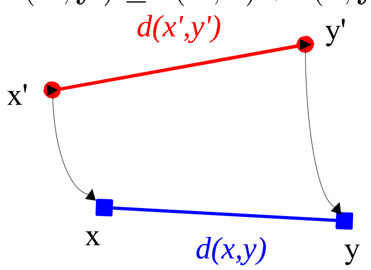
Conclusion 1.2 $|d(\mathbf{x},\mathbf{z}) - d(\mathbf{y},\mathbf{z})| \le d(\mathbf{x},\mathbf{y}) \quad \forall \mathbf{x},\mathbf{y},\mathbf{z} \in \mathbb{X}$

Proof.
$d(\mathbf{x},\mathbf{z}) \le d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z}) \quad \curvearrowright \quad d(\mathbf{x},\mathbf{z}) - d(\mathbf{y},\mathbf{z}) \le d(\mathbf{x},\mathbf{y})$
$d(\mathbf{y},\mathbf{z}) \le d(\mathbf{y},\mathbf{x}) + d(\mathbf{x},\mathbf{z}) \quad \curvearrowright \quad d(\mathbf{y},\mathbf{z}) - d(\mathbf{x},\mathbf{z}) \le d(\mathbf{y},\mathbf{x})$ ■

Conclusion 1.3 *Continuity of the metric*

$$|d(\mathbf{x},\mathbf{y}) - d(\mathbf{x}',\mathbf{y}')| \le d(\mathbf{x},\mathbf{x}') + d(\mathbf{y},\mathbf{y}') \quad \forall \mathbf{x},\mathbf{x}',\mathbf{y},\mathbf{y}' \in \mathbb{X}$$

Proof. .
$d(\mathbf{x},\mathbf{y}) \le d(\mathbf{x},\mathbf{x}') + d(\mathbf{x}',\mathbf{y}') + d(\mathbf{y}',\mathbf{y})$
$d(\mathbf{x}',\mathbf{y}') \le d(\mathbf{x}',\mathbf{x}) + d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{y}')$



■

What are some examples of metric spaces in addition to $\mathbb{R}^3$?

Example 1.2 $\mathbb{X} = \mathbb{R} : d(\mathbf{x},\mathbf{y}) = |\mathbf{x} - \mathbf{y}| \quad \forall \mathbf{x},\mathbf{y} \in \mathbb{R}$

Example 1.3 $\mathbb{X} = \mathbb{R}^2 : d(\mathbf{x},\mathbf{y}) = |x_1 - y_1| + |x_2 - y_2| \quad \forall \mathbf{x} = \binom{x_1}{x_2}, \mathbf{y} = \binom{y_1}{y_2} \in \mathbb{R}^2$

In these two examples the triangle inequality is based on the triangle inequality of the real numbers.

Example 1.4 $\mathbb{X} = \mathbb{R}^n$ *or* $\mathbb{X} = \mathbb{C}^n$

a) $d(\mathbf{x},\mathbf{y}) = \max_{1 \le i \le n} |\, x_i - y_i|$

b) $d(\mathbf{x},\mathbf{y}) = \left(\sum_{i=1}^{n} |\, x_i - y_i|^p \right)^{1/p}$ *for* $1 \le p < \infty$

Proof. *of the triangle inequalities:*
a)

$$\begin{aligned}|x_i - y_i| &\leq |x_i - z_i| + |z_i - y_i| \\ &\leq \max_{1\leq i\leq n} |x_i - z_i| + \max_{1\leq i\leq n} |z_i - y_i| \\ \max_{1\leq i\leq n} |x_i - y_i| &\leq \max_{1\leq i\leq n} |x_i - z_i| + \max_{1\leq i\leq n} |z_i - y_i|\end{aligned}$$

b) By the Minkowski inequality we get:

$$\begin{aligned}d(\mathbf{x}, \mathbf{y}) &= \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p} \\ &\leq \left(\sum_{i=1}^{n} |x_i - z_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |z_i - y_i|^p\right)^{1/p} \\ &= d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})\end{aligned}$$

■

Example 1.5 *Let* $\mathbb{X}$ *be the set of all sequences of real or complex numbers* $\mathbf{x} = \{x_i\}_{i=1}^{\infty}$, $\mathbf{y} = \{y_i\}_{i=1}^{\infty}$.

a) $d(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^{\infty} | x_i - y_i|^p\right)^{1/p}$ *for* $1 \leq p < \infty$

$\forall\ \mathbf{x}, \mathbf{y} \in \mathbb{X}$ *with* $\sum_{i=1}^{\infty} | x_i|^p < \infty$ *and* $\sum_{i=1}^{\infty} | y_i|^p < \infty$

b) $d(\mathbf{x}, \mathbf{y}) = \sup_i | x_i - y_i|$

$\forall\ \mathbf{x}, \mathbf{y} \in \mathbb{X}$ *with* $\sup_i | x_i| < \infty$ *and* $\sup_i | y_i| < \infty$

Proof. of the triangle inequalities:
a) By the Minkowski inequality we get:

$$\begin{aligned}d(\mathbf{x}, \mathbf{y}) &= \lim_{n\to\infty} \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p} \\ &\leq \lim_{n\to\infty} \left[\left(\sum_{i=1}^{n} |x_i - z_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |z_i - y_i|^p\right)^{1/p}\right] \\ &= d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})\end{aligned}$$

This is possible because we get on the same way

$$\begin{aligned} d(\mathbf{x},\mathbf{y}) &= \lim_{n\to\infty}\left(\sum_{i=1}^{n}|x_i-0+0-y_i|^p\right)^{1/p} \\ &\leq \lim_{n\to\infty}\left[\left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p}+\left(\sum_{i=1}^{n}|y_i|^p\right)^{1/p}\right] < \infty. \end{aligned}$$

b)

$$\begin{aligned} |x_i-y_i| &\leq |x_i-z_i|+|z_i-y_i| \\ &\leq \sup_i|x_i-z_i|+\sup_i|z_i-y_i| \\ \sup_i|x_i-y_i| &\leq \sup_i|x_i-z_i|+\sup_i|z_i-y_i| \end{aligned}$$

and analogous to a):

$$\begin{aligned} d(\mathbf{x},\mathbf{y}) &\leq d(\mathbf{x},0)+d(0,\mathbf{y}) \\ &\leq \sup_i|x_i|+\sup_i|y_i| < \infty. \end{aligned}$$

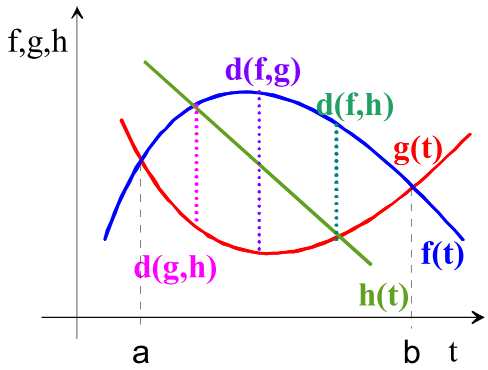
■

Example 1.6 $\mathbb{X}=C[a,b]$: *set of all continuous functions* $\mathbf{f}\,(t)$ *with real or complex values over* $[a,b]$

$$d(\mathbf{f},\mathbf{g})=\max_{a\leq t\leq b}|\,\mathbf{f}\,(t)-\mathbf{g}\,(t)\,| \quad \textit{with} \quad \mathbf{f},\mathbf{g}\in\ \mathbf{C}\,[a,b]$$

Proof. of the triangle inequality:

$$\begin{aligned} |\mathbf{f}(t)-\mathbf{g}(t)| &\leq |\mathbf{f}(t)-\mathbf{h}(t)|+|\mathbf{h}(t)-\mathbf{g}(t)| \quad \forall t\in[a,b] \\ &\leq \max_{t\in[a,b]}|\mathbf{f}(t)-\mathbf{h}(t)|+\max_{t\in[a,b]}|\mathbf{h}(t)-\mathbf{g}(t)| \\ \max_{t\in[a,b]}|\mathbf{f}-\mathbf{g}| &\leq \max_{t\in[a,b]}|\mathbf{f}-\mathbf{h}|+\max_{t\in[a,b]}|\mathbf{h}-\mathbf{g}| \\ d(\mathbf{f},\mathbf{g}) &\leq d(\mathbf{f},\mathbf{h})+d(\mathbf{h},\mathbf{g}) \quad \forall\mathbf{h}(t)\in\ \mathbf{C}\,[a,b] \end{aligned}$$



■

Example 1.7 *$\mathbb{X}-$ set of all continuous functions $\mathbf{f}(t)$ with real or complex values over any interval (a,b), such that: $\int_a^b |\mathbf{f}(t)|^p dt < \infty$.*

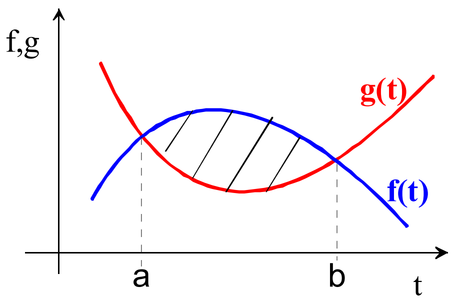
$$d(\mathbf{f},\mathbf{g}) = \left(\int_a^b |\mathbf{f}(t) - \mathbf{g}(t)|^p dt\right)^{1/p} \qquad \begin{array}{c} \forall \ \mathbf{f},\mathbf{g} \in \mathbb{X} \\ 1 \le p < \infty \end{array}$$

The triangle inequality can be proven using the MINKOWSKI inequality:

$$\left(\int_a^b |\mathbf{f}(t) - \mathbf{g}(t)|^p dt\right)^{1/p} \le \left(\int_a^b |\mathbf{f}(t) - \mathbf{h}(t)|^p dt\right)^{1/p} + \left(\int_a^b |\mathbf{h}(t) - \mathbf{g}(t)|^p dt\right)^{1/p}$$

Therefore $\mathbb{X}$ is a metric space. Interpretation for $p = 1$:

$$d(\mathbf{f},\mathbf{g}) = \int_a^b |\mathbf{f}(t) - \mathbf{g}(t)| dt$$



This is the absolute value of the area between the two functions in the interval $[a,b]$.

1.2 Fundamentals of Topology

Definition 1.2 *Let $(\mathbb{X}, d)$ be a metric space, $\mathbf{x}_0 \in \mathbb{X}$; The set*

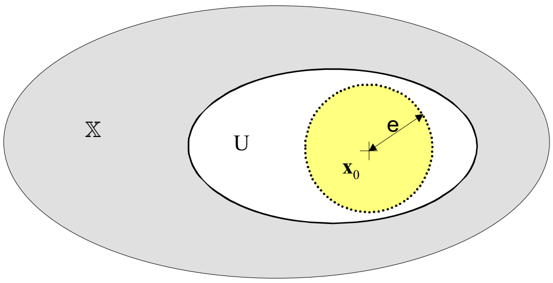
$$K_\varepsilon(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{X} \mid d(\mathbf{x}, \mathbf{x}_0) < \varepsilon\}$$

*is called an **open ball** centered at $\mathbf{x}_0$ with the radius ε, or is called the $\varepsilon-$ **neighbourhood of** $\mathbf{x}_0$.*

Definition 1.3 *The proper subset $A \subset \mathbb{X}$ is called **open**, if*

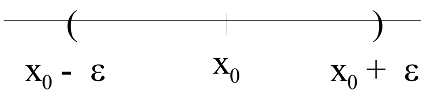
$$\forall \mathbf{x} \in A \quad \exists r > 0 \mid K_r(\mathbf{x}) \subseteq A.$$

Definition 1.4 *The proper subset $U \subset \mathbb{X}$ is called the **neighbourhood of** $\mathbf{x}_0$, if it contains an $\varepsilon-$neighbourhood of $\mathbf{x}_0$.*



Now we want to discuss some examples of open balls.

Example 1.8 $\mathbb{X} = \mathbb{R}; \quad d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$
Open balls are the the open intervalls $(\mathbf{x}_0 - \varepsilon; \mathbf{x}_0 + \varepsilon)$.



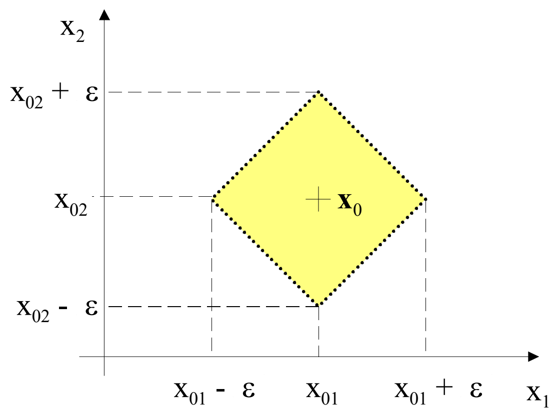
Example 1.9 $\mathbb{X} = \mathbb{R}^2,\ \mathbf{x}, \mathbf{y} \in \mathbb{R}^2,\ \mathbf{x} = (x_1, x_2)^T,\ \mathbf{y} = (y_1, y_2)^T$
a) $d(\mathbf{x}, \mathbf{y}) = |x_1 - y_1| + |x_2 - y_2|$

$$K_\varepsilon(\mathbf{x}_0) = \left\{\mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2 \mid |x_1 - x_{01}| + |x_2 - x_{02}| < \varepsilon\right\}$$

For an interpretation we use the equality $|x_1 - x_{01}| + |x_2 - x_{02}| = \varepsilon$. In a first case we get by applying the definition of the absolute value

$$\begin{aligned} x_1 - x_{01} + x_2 - x_{02} &= \varepsilon \\ x_2 &= -x_1 + (\varepsilon + x_{01} + x_{02}) \end{aligned}$$

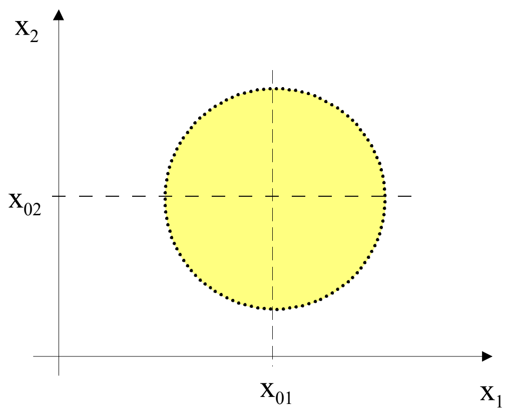
This is a linear equation in x_1 and x_2.



b) $d(\mathbf{x},\mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$

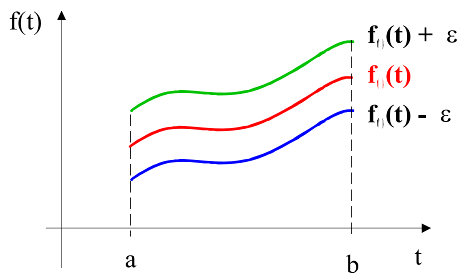
$$K_\varepsilon(\mathbf{x}_0) = \left\{\mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2 \mid (x_1 - x_{01})^2 + (x_2 - x_{02})^2 < \varepsilon\right\}$$

This is the interior of a circle centered at $\mathbf{x}_0$ with the radius $r = \varepsilon$.



Example 1.10 $\mathbb{X} = C[a,b]$; $\quad d(\mathbf{f},\mathbf{g}) = \max_{a\le t\le b} |\mathbf{f}(t) - \mathbf{g}(t)|$

$$K_\varepsilon(\mathbf{f}_0) = \left\{\mathbf{f}(t) \in C[a,b] \mid \max_{a\le t\le b} |\mathbf{f}(t) - \mathbf{f}_0(t)| < \varepsilon\right\} \quad \curvearrowright |\mathbf{f}(t) - \mathbf{f}_0(t)| < \varepsilon \quad \forall t \in [a,b]$$



Theorem 1.1 *For any (countable) collection* $\{A_\lambda\}$, $\lambda \in L$ *of open sets, the union* $\bigcup_{\lambda \in L} A_\lambda$ *is open.*

Proof. Let A be the union $A = \bigcup_{\lambda \in L} A_\lambda$ and $\mathbf{x} \in A$. $\curvearrowright$ $\exists \lambda_0 \mid \mathbf{x} \in A_{\lambda_0}$.
A_{λ_0} is an open set. $\curvearrowright$ $\exists r > 0 \mid K_r(\mathbf{x}) \subseteq A_{\lambda_0} \subset A$. ■

Theorem 1.2 *For any finite collection* $\{A_i\}_{i=1}^n$ *of open sets, the intersection* $\bigcap_{i=1}^n A_i$ *is open.*

Proof. Let A be the intersection $A = \bigcap_{i=1}^n A_i$ and $\mathbf{x} \in A$. $\curvearrowright$ $\mathbf{x} \in A_i$; $i = 1, 2, ..., n$
$\curvearrowright$ $\exists r_i > 0 \mid K_{r_i}(\mathbf{x}) \subseteq A_i$; $i = 1, 2, ..., n$
$r = \min_{1 \le i \le n} r_i > 0$ because of the finite number of sets.
$\curvearrowright$ $K_r(\mathbf{x}) \subset A_i$; $i = 1, 2, ..., n$ $\curvearrowright$ $K_r(\mathbf{x}) \subset A$ ■

Notation 1.1 *For any (countable) collection* $\{A_\lambda\}$, $\lambda \in L$ *of open sets, the intersection* $\bigcap_{\lambda \in L} A_\lambda$ *is not open in general.*

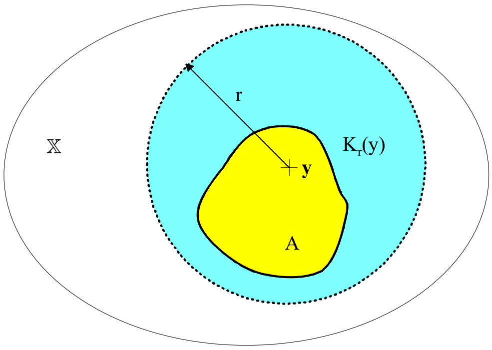
Example 1.11 $\mathbb{X} = R$; $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$
$A_n = \left(-\frac{1}{n}; 1 + \frac{1}{n}\right)$; $n = 1, 2, ...$ $\Rightarrow$ $\bigcap_{i=1}^{\infty} A_i = [0; 1]$ *is closed.*

Notation 1.2 *The metric space* $\mathbb{X}$ *and the empty set* $\varnothing$ *are open.*
(By definition the empty set $\varnothing$ *fulfills any condition.)*

Definition 1.5 x_0 *is an interior point of* A*, i.e.* $x_0 \in \overset{o}{A}$ $\iff$ $\exists \varepsilon > 0 \mid K_\varepsilon(x_0) \subset A$

Notation 1.3 *The set* $A \subset \mathbb{X}$ *is open* $\iff$ *any point of* A *is an interior point.*

Definition 1.6 *The proper subset $A \subset \mathbb{X}$ is called **bounded**, if A is completely contained within an open ball $K_r(y)$, $\mathbf{y} \in \mathbb{X}$, $0 < r < \infty$.*



Theorem 1.3 *For any finite collection $\{A_i\}_{i=1}^{n}$ of bounded sets, the union $\bigcup_{i=1}^{n} A_i$ is bounded.*

Proof. Let A be the union $A = \bigcup_{i=1}^{n} A_i$. $\rightarrow$ $\exists r_i \in \mathbb{R},\ \mathbf{y}_i \in \mathbb{X} \mid A_i \subseteq K_{r_i}(\mathbf{y}_i)$; $i = 1, 2, ..., n$
$a = \max_{2 \le i \le n} d(\mathbf{y}_i, \mathbf{y}_{i-1}) < \infty$. We define $\mathbf{y} = \mathbf{y}_1$ and choose any $\mathbf{x} \in A$.
Without loss of generality $\mathbf{x} \in A_n$ implies

$$\begin{aligned} d(\mathbf{x}, \mathbf{y}) &= d(\mathbf{x}, \mathbf{y}_1) \\ &\le d(\mathbf{x}, \mathbf{y}_n) + d(\mathbf{y}_n, \mathbf{y}_{n-1}) + ... + d(\mathbf{y}_2, \mathbf{y}_1) \\ &\le r_n \quad + \quad (n-1)a = r < \infty. \quad \curvearrowright \\ \mathbf{x} &\in K_r(\mathbf{y} = \mathbf{y}_1) \end{aligned}$$

■

Notation 1.4 *For any (countable) collection $\{A_\lambda\}$, $\lambda \in L$ of bounded sets, the union $\bigcup_{\lambda \in L} A_\lambda$ is not bounded in general:*

Example 1.12 $\mathbb{X} = \mathbb{R}$; $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$
$A_i = [i; i+1)$; $i = 0, 1, 2, ...$ $\rightarrow$ $\bigcup_i A_i = [0; \infty)$

Definition 1.7 *The point $\mathbf{x}_0 \in \mathbb{X}$ is a limit point of a set $A \subset \mathbb{X}$, if every open ball centered at $\mathbf{x}_0$ contains a point $\mathbf{x} \in A$; $\mathbf{x} \neq \mathbf{x}_0$. The set of all limit points of A is called the **derivated set** A^+.*

Definition 1.8 *The set* $\overline{A} = A \cup A^+$ *is called the* ***closure*** *or the* ***closed cover*** *of* A.

Definition 1.9 *The subset* $A \subset X$ *is called* ***closed****, if* $A^+ \subseteq A$.

Definition 1.10 *The set* B *is called* ***dense in*** A*, if* $B \subset A \quad \wedge \quad \overline{B} = A$.

Example 1.13 $\mathbb{X} = \mathbb{R} : d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$
a) $A = [0; 1]$:
$A = A^+ = \overline{A}$; *All points of* A *are limit points.*

b) $B = (0; 1)$:
$B \subset B^+ = \overline{B} = [0; 1] = A$;

c) $C = \{\frac{1}{n} \mid n \in \mathbb{N};\ n \neq 0\}$
The only one limit point is $\mathbf{x}_0 = 0$*:* $C^+ = \{0\}$
C *is not closed because of* $C^+ \nsubseteq C$.
C *is not open in* $\mathbb{R}$.

d) The set $\mathbb{Q}$ *of the rational numbers is dense in* $\mathbb{R}$.

Notation 1.5 *For any collections of closed sets the intersection is closed and for any finite collection of closed sets the union is closed, too.*

Notation 1.6 *The empty set and* $\mathbb{X}$ *are closed sets.*

Notation 1.7 *The subset* $A \subset \mathbb{X}$ *is closed* $\Longleftrightarrow B = \mathbb{X} \setminus A$ *is open.*

1.3 Convergence and Completeness

Let $\mathbb{X}$ be a metric space with the distance function $d(\mathbf{x}, \mathbf{y}),\ \mathbf{x}, \mathbf{y} \in \mathbb{X}$.

Definition 1.11 *A* ***sequence*** $\{\mathbf{x}_n\}_{n=1}^{\infty} \subset \mathbb{X}$ ***is called convergent****, if there exists an element* $\mathbf{x}_0 \in \mathbb{X}$ *fulfilling the condition* $\lim_{n\to\infty} d(\mathbf{x}_n, \mathbf{x}_0) = 0$. $\mathbf{x}_0$ *is called the* ***limit*** *of the sequence.*

We write: $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}_0$ or $\mathbf{x}_n \stackrel{n\to\infty}{\to} \mathbf{x}_0$ or in $\delta - \varepsilon-$notation:
$\forall \varepsilon > 0\ \exists n_0(\varepsilon) \mid d(\mathbf{x}_n, \mathbf{x}_0) < \varepsilon \quad \forall n > n_0$.

Theorem 1.4 *The limit of a sequence is unique.*

Proof. We suppose that $\mathbf{x}_0$, $\mathbf{y}_0$ are limits of the sequence $\{\mathbf{x}_n\}_{n=1}^{\infty}$ and $\mathbf{x}_0 \neq \mathbf{y}_0$. $\curvearrowright$

$$0 \leq d(\mathbf{x}_0, \mathbf{y}_0) \leq d(\mathbf{x}_0, \mathbf{x}_n) + d(\mathbf{x}_n, \mathbf{y}_0)$$

As n tends to infinity we get $d(\mathbf{x}_0, \mathbf{y}_0) = 0$. Contradiction ■

Example 1.14 $\mathbb{X} = \mathbb{R}^n, \quad d(\mathbf{x},\mathbf{y}) = \sqrt{(\sum_{i=1}^n (x_i - y_i)^2)}$:

$\lim_{k\to\infty} \mathbf{x}^k = \mathbf{x}^0; \quad \mathbf{x}^k = \begin{pmatrix} x_1^k \\ \vdots \\ x_n^k \end{pmatrix} \quad \mathbf{x}^0 = \begin{pmatrix} x_1^0 \\ \vdots \\ x_n^0 \end{pmatrix}$

$$\begin{aligned} &\iff d(\mathbf{x}^k,\mathbf{x}^0) = \sqrt{\left(\sum_{i=1}^n (x_i^k - x_i^0)^2\right)} \overset{k\to\infty}{\to} 0 \\ &\iff (x_i^k - x_i^0)^2 \overset{k\to\infty}{\to} 0 \quad \forall i \\ &\iff x_i^k \overset{k\to\infty}{\to} x_i^0 \quad \forall i \end{aligned}$$

Convergence in the Euclidian Space is convergence by coordinates, is equivalent to convergence of all components.

Example 1.15 $\mathbb{X} = C[a,b], \quad d(\mathbf{x},\mathbf{y}) = \max_{a\le t\le b} |\mathbf{x}(t) - \mathbf{y}(t)|$:

$$\lim_{k\to\infty} \mathbf{x}_k(t) = \mathbf{x}_0(t) \iff d(\mathbf{x}_k,\mathbf{x}_0) = \max_{a\le t\le b} |\mathbf{x}_k(t) - \mathbf{x}_0(t)| \overset{k\to\infty}{\to} 0$$

$\curvearrowright$ $\forall\varepsilon > 0 \quad \exists n_0(\varepsilon) \in \mathbb{N} \mid |\mathbf{x}_k(t) - \mathbf{x}_0(t)| < \varepsilon \quad \forall k > n_0(\varepsilon), \quad \forall t \in [a,b]$

$\curvearrowright$ $n_0(\varepsilon)$ *is independent of* t.

Convergence of a sequence of functions in $C[a,b]$ *is uniform convergence with respect to* t.

Example 1.16 $\mathbb{X} = C[a,b], \quad d(\mathbf{x},\mathbf{y}) = \left(\int_a^b |\mathbf{x}(t) - \mathbf{y}(t)|^p dt\right)^{\frac{1}{p}}; \quad p \ge 1, \ p \in \mathbb{N}$, *fixed*:

$$\lim_{k\to\infty} \mathbf{x}_k(t) = \mathbf{x}_0(t) \iff d(\mathbf{x}_k,\mathbf{x}_0) = \left(\int_a^b |\mathbf{x}_k(t) - \mathbf{x}_0(t)|^p dt\right)^{\frac{1}{p}} | \overset{k\to\infty}{\to} 0$$

This is called convergence in the p-th mean. If $p = 2$ *the convergence is called convergence in quadratic mean.*

Notation 1.8 $A \subset \mathbb{X}; \quad \mathbf{x}_0 \in A^+ \iff \exists\{\mathbf{x}_n \mid x_n \neq x_0\}_{n=1}^\infty \subset A$ *with* $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}_0$

Definition 1.12 *Cauchy sequence*
A sequence $\{\mathbf{x}_n\}_{n=1}^\infty \subset \mathbb{X}$ *is said to be* ***Cauchy****, if given any* $\varepsilon > 0$ *there exists an integer* $n_0(\varepsilon)$ *such that* $d(x_n,x_m) < \varepsilon$ *whenever* $n,m > n_0(\varepsilon)$,
i.e. $\lim_{n,m\to\infty} d(x_n,x_m) = 0$.

Theorem 1.5 *Any convergent sequence in* $\mathbb{X}$ *is Cauchy.*

Proof. Assume that $\{\mathbf{x}_n\}_{n=1}^\infty$ converges and let be $\varepsilon > 0$:

$\curvearrowright$ $\exists\mathbf{x}_0 \in \mathbb{X}, \ n_0 \in \mathbb{N}, \ n_0 > 0 \mid d(\mathbf{x}_n,\mathbf{x}_0) < \frac{\varepsilon}{2} \quad \wedge \quad d(\mathbf{x}_m,\mathbf{x}_0) < \frac{\varepsilon}{2} \quad \forall n,m > n_0$

$\curvearrowright$ $d(\mathbf{x}_n,\mathbf{x}_m) \le d(\mathbf{x}_n,\mathbf{x}_0) + d(\mathbf{x}_0,\mathbf{x}_m) < \varepsilon \quad \forall n,m > n_0$ ■

Notation 1.9 *In general the converse of this theorem is not true! It is possible, that a Cauchy sequence tends to a limit which is not an element of the space* $\mathbb{X}$.

Example 1.17 $\mathbb{X} = (0;1); \quad d(\mathbf{x},\mathbf{y}) = |\mathbf{x}-\mathbf{y}|$
We choose $\mathbf{x}_n = \frac{1}{n} : \{\mathbf{x}_n\}_{n=1}^{\infty}$ *is Cauchy in* $\mathbb{X}$ *because we get* $n_0(\varepsilon) = \left[\frac{2}{\varepsilon}\right] + 1, \quad \forall\varepsilon > 0$ *(square brackets = whole number).*
Then

$$d(\mathbf{x}_n, \mathbf{x}_m) = \left|\frac{1}{n} - \frac{1}{m}\right| \leq \frac{2}{n_0} < \varepsilon \quad \forall n, m > n_0,$$

but the limit $\mathbf{x}_0 = 0 \notin \mathbb{X}$. $\implies$ $\{\mathbf{x}_n\}_{n=1}^{\infty}$ *is not convergent in* $\mathbb{X}$ *by definition.*

Definition 1.13 *A metric space* $\mathbb{X}$ *is* ***complete****, if every Cauchy sequence in* $\mathbb{X}$ *converges to a point of* $\mathbb{X}$.

Example 1.18 $\mathbb{X} = \mathbb{R}^n, \quad d(\mathbf{x},\mathbf{y}) = \sqrt{(\sum_{i=1}^{n}(x_i - y_i)^2)}$ *is complete:*
Let $\{\mathbf{x}^k\}_{k=1}^{\infty}$ *be Cauchy in* $\mathbb{R}^n \curvearrowright$

$$\forall\varepsilon > 0\ \exists n_0(\varepsilon) \mid d(\mathbf{x}^k, \mathbf{x}^l) = \sqrt{\sum_{i=1}^{n}(x_i^k - x_i^l)^2} < \varepsilon \quad \forall k, l > n_0 \quad \curvearrowright$$

$$\left|x_i^k - x_i^l\right| < \varepsilon \quad \forall k, l > n_0 \quad \wedge \quad i = 1, 2, ..., n$$

$\curvearrowright$ $\{x_i^k\}_{k=1}^{\infty}$ *is Cauchy in* $\mathbb{R}$ $\forall i$. $\mathbb{R}$ *is complete.*
$\curvearrowright$ $\exists$ *the limit* $x_i^0 \in \mathbb{R}$ *for* $i = 1, 2, .., n$

$\curvearrowright$ $\exists$ *the limit* $\mathbf{x}^0 = \begin{pmatrix} x_1^0 \\ \vdots \\ x_n^0 \end{pmatrix} \in \mathbb{R}^n$

Example 1.19 $\mathbb{X} = C[a,b], \quad d(\mathbf{x},\mathbf{y}) = \max\limits_{a\leq t\leq b}|\mathbf{x}(t) - \mathbf{y}(t)|$ *is complete:*
The proof consists of three parts :
1) We construct an element $\mathbf{x}_0(t)$ *which can be the limit of a Cauchy sequence.*
2) We show that $\mathbf{x}_0(t)$ *is the limit of the sequence.*
3) We prove that $\mathbf{x}_0(t) \in C[a,b]$.

1) Let $\{\mathbf{x}_k(t)\}_{k=1}^{\infty}$ *be Cauchy in* $C[a,b] \curvearrowright$

$$\forall\varepsilon > 0\ \exists n_0(\varepsilon) \mid d(\mathbf{x}_n(t), \mathbf{x}_m(t)) = \max_{a\leq t\leq b}|\mathbf{x}_n(t) - \mathbf{x}_m(t)| < \varepsilon \quad \forall n, m > n_0$$

$$\curvearrowright \quad |\mathbf{x}_n(t) - \mathbf{x}_m(t)| < \varepsilon \quad \forall n, m > n_0, \quad \forall t \in [a,b]$$

Let t *be fixed.* $\curvearrowright$ $\{\mathbf{x}_k(t)\}_{k=1}^{\infty} \subset \mathbb{R}$ *and*

$$\forall\varepsilon > 0\ \exists n_0(\varepsilon) \mid |\mathbf{x}_n(t) - \mathbf{x}_m(t)| < \varepsilon \quad \forall n, m > n_0$$

means $\{\mathbf{x}_k(t)\}_{k=1}^{\infty}$ *is Cauchy in* $\mathbb{R}$*.* $\mathbb{R}$ *is complete.*

$$\curvearrowright \quad \mathbf{x}_n(t) \quad \overset{n\to\infty}{\longrightarrow} \quad \mathbf{x}_0(t) \quad \forall t \in [a,b] \quad (*)$$

$\curvearrowright$ $\mathbf{x}_0(t)$ *is a function defined at* $[a,b]$*.*

2) $(*)$ *implies*

$$|\mathbf{x}_n(t) - \mathbf{x}_{n+k}(t)| < \varepsilon \quad \forall n > n_0, k \in \mathbb{N} \quad \forall t \in [a,b].$$

As k *tends to infinity we get*

$$|\mathbf{x}_n(t) - \mathbf{x}_0(t)| < \varepsilon \quad \forall n > n_0 \quad \forall t \in [a,b].$$

$\curvearrowright$ $\{\mathbf{x}_k(t)\}_{k=1}^{\infty}$ *is uniformly convergent at* $[a,b]$*. The limit is* $\mathbf{x}_0(t)$*.*

3) To show: $\mathbf{x}_0(t)$ *is a continuous function.*

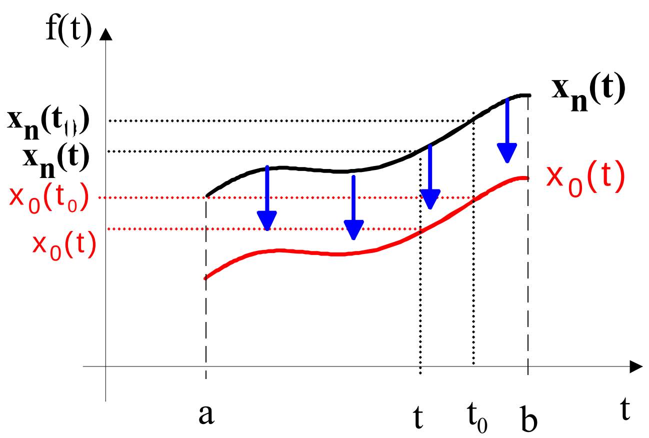
$\mathbf{x}_n(t)$ *is continuous. This means*

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \mid |\mathbf{x}_n(t) - \mathbf{x}_n(t_0)| < \varepsilon \quad \forall t \in [a,b]; \ |t - t_0| < \delta; \ \forall n$$

and we get

$$\begin{aligned} |\mathbf{x}_0(t) - \mathbf{x}_0(t_0)| &\leq |\mathbf{x}_0(t) - \mathbf{x}_n(t)| + |\mathbf{x}_n(t) - \mathbf{x}_n(t_0)| + |\mathbf{x}_n(t_0) - \mathbf{x}_0(t_0)| \\ &\leq 3\varepsilon \quad \forall |t - t_0| < \delta \end{aligned}$$

Thus $\mathbf{x}_0(t) \in C[a,b]$*.*



Example 1.20 $\mathbb{X} = \mathbb{Q}:$ *The set of all rational numbers with* $d(\mathbf{x},\mathbf{y}) = |\mathbf{x} - \mathbf{y}|$ *is not complete.*

Counterexample: $\left\{\mathbf{x}_n = \left(1 + \frac{1}{n}\right)^n\right\}_{n=1}^{\infty} \subset \mathbb{Q}$ *is Cauchy, but* $\lim \mathbf{x}_n = e \notin \mathbb{Q}$ *(Euler's constant)*

Example 1.21 $\mathbb{X} = C[0,1]$, $d(\mathbf{x},\mathbf{y}) = \left(\int_0^1 |\mathbf{x}(t)-\mathbf{y}(t)|^2 dt\right)^{\frac{1}{2}}$ *is not complete:*
Counterexample: Consider the sequence $\{\mathbf{x}_n(t)\}_{n=1}^{\infty}$:

$$\mathbf{x}_n(t) = \begin{cases} n^{1/3} & for \quad t \le \frac{1}{n} \\ t^{-1/3} & for \quad t > \frac{1}{n} \end{cases}$$

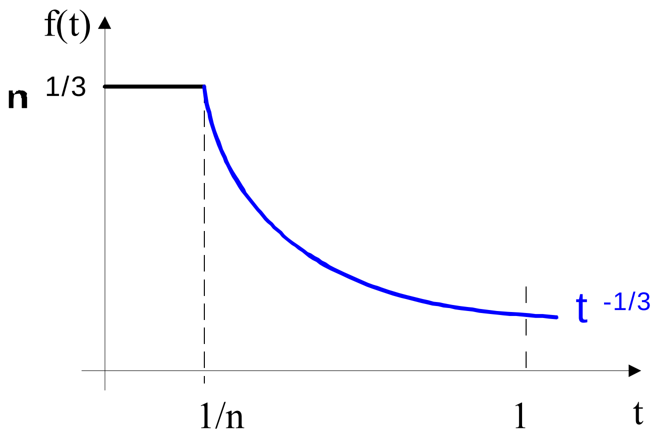
$\mathbf{x}_n(t)$ *is continuous at* $[0,1]$ *for all* n *because* $\lim\limits_{t\to 1/n} t^{-1/3} = \left(\frac{1}{n}\right)^{-\frac{1}{3}} = n^{\frac{1}{3}}$. *The limit of* $\{\mathbf{x}_n(t)\}_{n=1}^{\infty}$ *is*

$$\mathbf{x}_0(t) = t^{-\frac{1}{3}}$$

because of

$$\begin{aligned} (d(\mathbf{x}_n,\mathbf{x}_0))^2 &= \int_0^1 |\mathbf{x}_n(t)-\mathbf{x}_0(t)|^2 dt \\ &= \int_0^{1/n} |n^{1/3} - t^{-1/3}|^2 dt \\ &\overset{(*)}{\le} \int_0^{1/n} \left(n^{2/3} + t^{-2/3}\right) dt \\ &= n^{2/3} \cdot \frac{1}{n} + 3\left(\frac{1}{n}\right)^{1/3} \\ &= 4n^{-1/3} \quad \overset{n\to\infty}{\longrightarrow} \quad 0. \end{aligned}$$

using $(*)$: $(a-b)^2 = a^2 - 2ab + b^2 \le a^2 + b^2$ *if* $a,b > 0$.
But $\mathbf{x}_0(t)$ *is not continuous at* $[0,1]$.



1.4 Compact Sets

Definition 1.14 *Let $(\mathbb{X}, d)$ be a metric space. The subset $A \subset \mathbb{X}$ is called **sequentially compact**, if every sequence $\{\mathbf{x}_n\}_{n=1}^{\infty} \subset A$ contains a convergent subsequence $\{\widetilde{\mathbf{x}}_n\}_{n=1}^{\infty} \subset A$ with $lim_{n\to\infty}\widetilde{\mathbf{x}}_n = \mathbf{x} \in \mathbb{X}$.*

Definition 1.15 *The subset $A \subset \mathbb{X}$ is called **compact**, if every sequence $\{\mathbf{x}_n\}_{n=1}^{\infty} \subset A$ contains a convergent subsequence $\{\widetilde{\mathbf{x}}_n\}_{n=1}^{\infty} \subset A$ with $lim_{n\to\infty}\widetilde{\mathbf{x}}_n = \mathbf{x} \in A$.*

Notation 1.10 *$A \subset \mathbb{X}$ is compact $\iff$ $A \subset \mathbb{X}$ is sequentially compact and closed.*

Example 1.22 *$\mathbb{X} = \mathbb{R}$ with $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$: All bounded subsets $A \subset \mathbb{X}$ are sequentially compact.*
For any bounded subset $A \subset \mathbb{X}$ there exists a closed interval $I = [a, b]$ such that $A \subset [a, b] = I$.
Look at any sequence $\{\mathbf{x}_n\}_{n=1}^{\infty} \subset A$ and choose a subsequence $\{\widetilde{\mathbf{x}}_n\}$ by succesive bisection in the following way:
During the n-th step take $\widetilde{\mathbf{x}}_n$ from this half of the intervall in which the number of elements is infinite.

$$\begin{aligned}
&\Longrightarrow \quad d(\widetilde{\mathbf{x}}_n, \widetilde{\mathbf{x}}_m) \le \frac{b-a}{2^n} \quad for \quad m > n \\
&\Longrightarrow \quad \{\widetilde{\mathbf{x}}_n\} \text{ is Cauchy in } \mathbb{R} \\
&\Longrightarrow \quad \exists \mathbf{x} \in \mathbb{X} \mid \lim_{n\to\infty} \widetilde{\mathbf{x}}_n = \mathbf{x} \\
&\Longrightarrow \quad A \subset \mathbb{X} \text{ is sequentially compact}
\end{aligned}$$

Example 1.23 $\mathbb{X} = \mathbb{R}^n$, $\quad d(\mathbf{x}, \mathbf{y}) = \sqrt{(\sum_{i=1}^{n}(x_i - y_i)^2)}$:
All bounded subsets $A \subset \mathbb{X}$ are sequentially compact:

Let $A \subset \mathbb{X}$ be a bounded subset and let $\{\mathbf{x}^m\}_{m=1}^{\infty} \subset A$ be a sequence with $\mathbf{x}^m = \begin{pmatrix} x_1^m \\ \vdots \\ x_n^m \end{pmatrix}$.

Then there exists an element $\mathbf{x}^0 \in \mathbb{R}^n$ with $d\left(\mathbf{x}^0, \mathbf{x}^m\right) \le M < \infty \quad \forall m$.

$$\begin{aligned}
&\Longrightarrow \quad \left|x_i^0 - x_i^m\right| \le M \quad i = 1, 2, ..., n \\
&\Longrightarrow \quad \{\mathbf{x}_i^m\}_{m=1}^{\infty} \text{ is bounded in } \mathbb{R}, \quad i = 1, 2, ..., n
\end{aligned}$$

Then there exists a subsequence $\{\widetilde{\mathbf{x}}_i^m\} \subset \{\mathbf{x}_i^m\}$ such that $\widetilde{\mathbf{x}}_i^m \overset{m\to\infty}{\to} \mathbf{x}_i, \quad i = 1, 2, ..., n$ because of the example above.

$$\begin{aligned}
&\Longrightarrow \quad \widetilde{\mathbf{x}}^m \overset{m\to\infty}{\to} \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \\
&\Longrightarrow \quad A \subset \mathbb{R}^n \text{ is sequentially compact.}
\end{aligned}$$

Example 1.24 $\mathbb{X} = C[a,b], \quad d(\mathbf{x},\mathbf{y}) = \max_{a\le t\le b} |\mathbf{x}(t) - \mathbf{y}(t)|$:
Let $A \subset \mathbb{X}$ be bounded and closed:

$$\max_{a\le t\le b} |\mathbf{x}(t)| \le M < \infty \quad \forall \mathbf{x}(t) \in A$$

Let $A \subset \mathbb{X}$ be equicontinuous:

$$\forall \varepsilon > 0 \quad \exists \delta(\varepsilon) > 0 \mid \ |t-s| < \delta \text{ implies } |\mathbf{x}(s) - \mathbf{x}(t)| < \varepsilon \quad \forall \mathbf{x} \in A;$$

$\Longrightarrow$ *A is compact in $\mathbb{X}$. (Arzela-Ascoli Theorem)*
Proof: s. [3, p.68ff] or [4, p. 20-21]
For example: $A = \{\mathbf{x}(t) \mid |\mathbf{x}(t)| \le M_1; \ \left|\frac{d\mathbf{x}}{dt}\right| \le M_2; \quad a \le t \le b\}$
$\curvearrowright$ *A is bounded and closed.*
$\curvearrowright$ *The Mean Value Theorem implies $|\mathbf{x}(s) - \mathbf{x}(t)| \le M_2 |s-t|$.*
$\curvearrowright$ *A is equicontinuous.*
$\curvearrowright$ *A is compact.*

Conclusion 1.4 *In a metric space $\mathbb{X}$ every sequentially compact set A is bounded.*

Proof. We assume $A \subset \mathbb{X}$ is sequentially compact and not bounded.

$$\Longrightarrow \exists \{\mathbf{x}_n\}_{n=1}^{\infty} \subset A \mid d(\mathbf{x}_0, \mathbf{x}_n) \ge n; \quad \mathbf{x}_0 \in \mathbb{X}$$

Because of the sequentially compactness of A there must be a convergent subsequence. Contradiction ■

Conclusion 1.5 *The converse of the above conclusion is not true. A bounded set does not be necessarily sequentially compact.*

Example 1.25 $\mathbb{X} = L_2[0,1], \quad d(\mathbf{x},\mathbf{y}) = \sqrt{\int_0^1 |\mathbf{x}(t) - \mathbf{y}(t)|^2 dt}$
$A = \{\mathbf{g}_n(t) = \sin(n\pi t)\}_{n=1}^{\infty}$

$$\begin{aligned}
d(\mathbf{g}_n, \mathbf{g}_m) &= \sqrt{\int_0^1 |\mathbf{g}_n(t) - \mathbf{g}_m(t)|^2 dt} \\
&= \sqrt{\int_0^1 |\sin(n\pi t) - \sin(m\pi t)|^2 dt} \\
&= \sqrt{\int_0^1 \left(\sin^2(n\pi t) - 2\sin(n\pi t)\sin(m\pi t) + \sin^2(m\pi t)\right) dt} \\
&= \sqrt{\frac{1}{2} - 0 + \frac{1}{2}} = 1 \\
d(\mathbf{0}, \mathbf{g}_m) &= \sqrt{\int_0^1 \sin^2(m\pi t) dt} = \sqrt{\frac{1}{2}} \quad \forall m
\end{aligned}$$

A is bounded, but not sequentially compact.

Auxiliary calculation:
$\int \sin^2(n\pi t)dt = \int \sin^2(u)\frac{du}{n\pi} = \frac{1}{2n\pi}(u - \sin u \cdot \cos u) + C$

$= \frac{1}{2n\pi}(n\pi t - \sin(n\pi t) \cdot \cos((n\pi t))) + C$

$\Longrightarrow \int_0^1 \sin^2(n\pi t)dt = \frac{1}{2n\pi}(n\pi - 0) = \frac{1}{2}$

Example 1.26 *Let $\mathbb{X}$ be the set of all number sequences $\mathbf{x} = \{x_i\}_{i=1}^{\infty}$ with $\sum_{i=1}^{\infty}|x_i|^2 < \infty$ and $d(\mathbf{x},\mathbf{y}) = \left(\sum_{i=1}^{\infty}|x_i - y_i|^2\right)^{\frac{1}{2}}$.*

Let A be the unit ball: $A = \left\{\mathbf{x} \in \mathbb{X} \mid \sum_{i=1}^{\infty}|x_i|^2 \leq 1\right\}$.
A is bounded and closed. We have a look at the sequence $\boldsymbol{\xi} = \left\{\boldsymbol{\xi}^i\right\}_{i=1}^{\infty} \subset A$ with

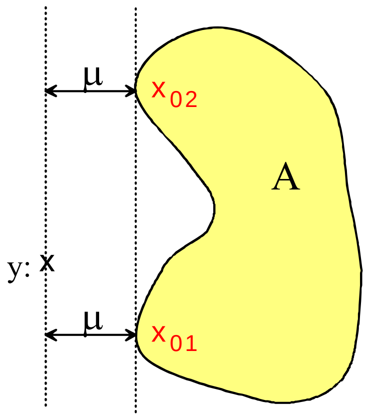
$$\begin{aligned} \boldsymbol{\xi}^i &= \left\{0, ..., 0, \underset{position\ i}{1}, 0, ...\right\} \\ &\Longrightarrow d(\boldsymbol{\xi}^i, \boldsymbol{\xi}^j) = \sqrt{2} \quad if \quad i \neq j \end{aligned}$$

That's why you cannot find a convergent subsequence in $\boldsymbol{\xi}$ and the unit ball is not sequentially compact!

Theorem 1.6 *WEIERSTRASS*
Let $f : A \to \mathbb{R}$ be a continuous function on the compact subset $A \subset \mathbb{X}$. Then f has a maximum and a minimum on A. That means: $\exists\ \underline{\mathbf{x}} \in A$ with $f(\underline{\mathbf{x}}) = \min_{\mathbf{x}\in A} f(\mathbf{x}) \quad \wedge$ $\exists\ \overline{\mathbf{x}} \in A$ with $f(\overline{\mathbf{x}}) = \max_{\mathbf{x}\in A} f(\mathbf{x})$

Proof. (for the minimum)
We set: $\alpha = \inf\limits_{x\in A} f(\mathbf{x})$ and have to prove: $\exists\ \underline{\mathbf{x}} \in A \mid f(\underline{\mathbf{x}}) = \alpha$.
Now we build the sequence $\{\mathbf{x}_n\}_{n=1}^{\infty} \subset A$ such that $f(\mathbf{x}_{n)} \leq \alpha + \frac{1}{n};\ n = 1, 2, ...$
A is compact $\Longrightarrow$ There exists a subsequence $\{\widetilde{\mathbf{x}}_n\}_{n=1}^{\infty} \subset A$ with $\widetilde{\mathbf{x}}_n \to \underline{\mathbf{x}} \in \mathbf{A}$ as n tends to infinity.
f is continuous at A and so we get $f(\widetilde{\mathbf{x}}_n) \to \alpha = f(\underline{\mathbf{x}})$ as n tends to infinity. ∎

Theorem 1.7 *If the subset $A \subset \mathbb{X}$ is compact then there exists a point $\mathbf{x}_0 \in A$ such that $d(\mathbf{x}_0, \mathbf{y}) = \min\limits_{x\in A} d(\mathbf{x}, \mathbf{y})$ for any given point $\mathbf{y} \in \mathbb{X}$. $\mathbf{x}_0$ is called the best approximation of $\mathbf{y}$ in A. (In general $\mathbf{x}_0$ is not unique!)*



Proof. $\mu = \inf_{x \in A} d(\mathbf{x}, \mathbf{y})$ implies that there exists a sequence $\{\mathbf{x}_n\}_{n=1}^{\infty} \subset A$ with

$$d(\mathbf{y}, \mathbf{x}_n) \leq \mu + \frac{1}{n}, \ n = 1, 2, ...$$

A is compact $\Longrightarrow$ There exists a subsequence $\{\widetilde{\mathbf{x}}_n\}_{n=1}^{\infty} \subset A$ with $\widetilde{\mathbf{x}}_n \to \mathbf{x}_0 \in \mathbf{A}$ as n tends to infinity.

$$d(\mathbf{y}, \mathbf{x}_0) \leq d(\mathbf{y}, \widetilde{\mathbf{x}}_n) + d(\widetilde{\mathbf{x}}_n, \mathbf{x}_0) \stackrel{n \to \infty}{\to} \mu$$

On the other hand we have

$$d(\mathbf{y}, \mathbf{x}_0) \geq \mu$$

because $x_0 \in A$. Thus $d(\mathbf{y}, \mathbf{x}_0) = \mu$ and $\mathbf{x}_0$ is the best approximation. ■

Summary:

Compactness is the generalisation of the terms „closed interval" or „bounded closed set". Compare this with the real one dimensional analysis: A continuous function has an extremum on a finite closed interval.

1.5 Operators

The generalisation of the function concept leads us to the definition of an operator. We consider two real (or complex) metric spaces $(\mathbb{X}, d_{\mathbb{X}}), (\mathbb{Y}, d_{\mathbb{Y}})$ and two subsets
$A \subseteq \mathbb{X}; \quad B \subseteq \mathbb{Y}$.

Definition 1.16 *A unique mapping from A to B, $T : A \to B$ is called an operator. That means, for every $\mathbf{x} \in A$ there exists a unique element $\mathbf{y} \in B$ such that $T\mathbf{x} = \mathbf{y}$. A is the domain of T.*

Definition 1.17 *The range of T is the set $T(A) = \{\mathbf{y} \in B \mid \exists \mathbf{x} \in A \text{ with } T\mathbf{x} = \mathbf{y}\}$.*

Example 1.27 *System of linear equations:*

$$\begin{pmatrix} 1 & 5 & 3 \\ 4 & 0 & 7 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} d \\ e \end{pmatrix} \Longleftrightarrow T\mathbf{x} = \mathbf{y} \quad \text{with } \mathbf{x} \in \mathbb{R}^3,\ \mathbf{y} \in \mathbb{R}^2,\ T : \mathbb{R}^3 \to \mathbb{R}^2$$

Example 1.28 $ln(\mathbf{x}) = \mathbf{y} \iff T\mathbf{x} = \mathbf{y} \quad \text{with } \mathbf{x} \in \mathbb{R}^+,\ \mathbf{y} \in \mathbb{R}^1,\ T : \mathbb{R}^+ \to \mathbb{R}^1$

Example 1.29 $\frac{d}{dt}\mathbf{x}(t) = \mathbf{y}(t) \iff T\mathbf{x} = \mathbf{y} \quad \text{with } \mathbf{x} \in C^1[a,b];\ \mathbf{y} \in C[a,b];$
$T : C^1[a,b] \to C[a,b]$: *differential operator*

Example 1.30 *The map*

$$R\mathbf{x} = \int_a^b \mathbf{x}(t)dt$$

defines an operator R from $C[a,b]$ to $\mathbb{R}$: an integral operator

Example 1.31 *The map*

$$S\mathbf{x}(t) = \int_a^b F(t, s, \mathbf{x}(s))ds; \quad t \in [a,b]$$

defines an operator S from $C[a,b]$ to $C[a,b]$; F is called the kernel of the operator.

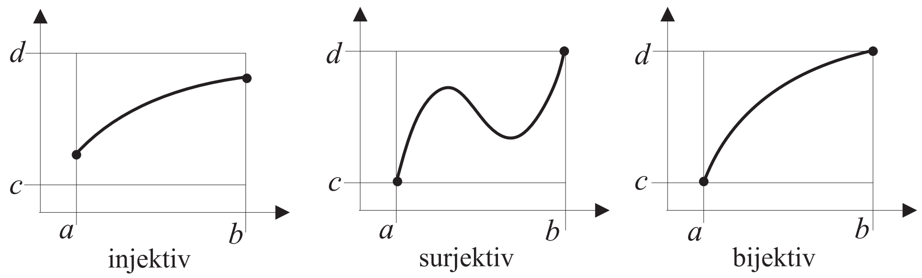
Integral and differential equations are rich areas of application for operator theory and they provided impetus for the early development of functional analysis.

Definition 1.18 *The operator $T : A \to B$ is called*

- **surjective (onto)** $\iff T(A) = B$
- **injective** (one-to-one) $\iff T\mathbf{x} = T\mathbf{y} \quad \curvearrowright \quad \mathbf{x} = \mathbf{y}$

- **bijective** $\iff$ T is surjective and injective.

Example 1.32 $\mathbf{X} = \mathbf{Y} = \mathbb{R}$, $\mathbf{A} = [a, b]$, $\mathbf{B} = [c, d]$, $\mathbf{T} : \mathbf{A} \subseteq \mathbf{X} \to \mathbf{B}$



Definition 1.19 *If the operator $T : A \to B$ is bijective, then there exists the **inverse operator** $T^{-1} : B \to A$, which is defined by $T^{-1}\mathbf{y} = \mathbf{x} \iff T\mathbf{x} = \mathbf{y}$.*

Definition 1.20 *The **operator** $T : A \subseteq \mathbb{X} \to B \subseteq \mathbb{Y}$ **is called continuous at** $\mathbf{x}_0 \in A$, if $\forall \varepsilon > 0 \quad \exists \delta = \delta(x_0, \varepsilon) \quad | \quad d_{\mathbb{Y}}(T\mathbf{x}, T\mathbf{x}_0) < \varepsilon \quad \forall \mathbf{x} \in A$ with $d_{\mathbb{X}}(\mathbf{x}, \mathbf{x}_0) < \delta$. If T is continuous at every point $\mathbf{x}_0 \in A$, then T is called **continuous on** A. In Addition, if $\delta(x_0, \varepsilon)$ is independent of $\mathbf{x}_0$ for all ε then T is called **uniformly continuous on** A.*

Conclusion 1.6 *T is uniformly continuous, that means:*
$\forall \mathbf{x} \in A,\ \forall \varepsilon > 0\ \exists \delta = \delta(\varepsilon) \mid d_{\mathbb{Y}}(T\mathbf{x}, T\mathbf{y}) < \varepsilon \quad \forall \mathbf{y} \in A$ *with* $d_{\mathbb{X}}(x, \mathbf{y}) < \delta$

Theorem 1.8 *Let $(\mathbb{X}, d_{\mathbb{X}}), (\mathbb{Y}, d_{\mathbb{Y}})$ be two metric spaces and let $T{:}A \subseteq \mathbb{X} \to \mathbb{Y}$ be a continuous operator from the compact set A to the space $\mathbb{Y}$, then T is uniformly continuous on A.*

Proof. Suppose T is not uniformly continuous.
$\curvearrowright \quad \exists \varepsilon_0 > 0$ and there exist sequences $\{\mathbf{x}_n\}_{n=1}^{\infty} \subset A,\ \{\mathbf{y}_n\}_{n=1}^{\infty} \subset A$ such that

$$d_{\mathbb{X}}(\mathbf{x}_n, \mathbf{y}_n) < \frac{1}{n} \text{ and } d_{\mathbb{Y}}(T\mathbf{x}_n, T\mathbf{y}_n) \geq \varepsilon_0 \quad (*).$$

A is compact. This implies that there exists a subsequence $\{\widetilde{\mathbf{x}}_n\}_{n=1}^{\infty}$ with $\widetilde{\mathbf{x}}_n \to \mathbf{x} \in A$ as n tends to infinity. Thus

$$d_{\mathbb{X}}(\mathbf{y}_n, \mathbf{x}) \leq d_{\mathbb{X}}(\mathbf{y}_n, \widetilde{\mathbf{x}}_n) + d_{\mathbb{X}}(\widetilde{\mathbf{x}}_n, \mathbf{x}) \overset{n\to\infty}{\to} 0.$$

$\curvearrowright \quad y_n \overset{n\to\infty}{\to} \mathbf{x}$. T is continuous. $\curvearrowright$

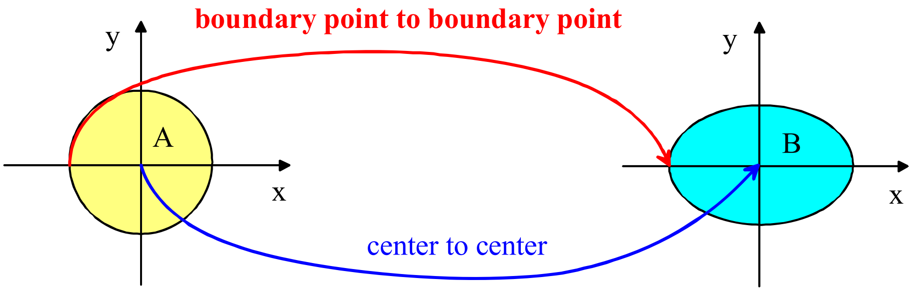
$$d_{\mathbb{Y}}(T\widetilde{\mathbf{x}}_n, T\mathbf{y}_n) \overset{n\to\infty}{\to} 0 : \quad \text{contradiction to } (*)$$

■

Definition 1.21 *A bijective continuous mapping $T : A \to B$, with a continuous inverse mapping is called a* ***homeomorphism****.*
Two set are called homeomorphic $\Leftrightarrow \exists$ homeomorphism $T : A \to B$.

Example 1.33 *The circle $A := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \le r^2\}$ and the ellipse $B := \{(x, y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1\}$ are homeomorphic.*
T is given by

$$T(x, y) = \left(\frac{a}{r}x; \frac{b}{r}y\right); \quad T : A \to B$$



1.6 The BANACH Fixed Point Theorem

The BANACH fixed point theorem has a very great importance, for example

- in the proofs of theorems of existance and unicity for several mathematical problems,
- for the solution of operator equations (see numerics).

Let $(\mathbb{X}, d)$ be a metric space and A be a subset of the space: $A \subseteq \mathbb{X}$.

Definition 1.22 *Let A be closed. The mapping $T : A \to A$ is called a* ***contraction mapping****, if there exists a number $0 < q < 1$ such that*

$$d(T\mathbf{x}, T\mathbf{y}) \le q \cdot d(\mathbf{x}, \mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in A.$$

Theorem 1.9 ***BANACH Fixed Point Theorem (FPT)***
Let A be a closed subset of the complete metric space $(\mathbb{X}, d)$ with a contraction mapping $T : A \to A$. Then T admits a unique fixed-point $\mathbf{x}^* \in A$, i.e. $\mathbf{x}^*$ is the solution of the fixed point equation $\mathbf{x} = T\mathbf{x}$.
Then the iterative sequence $\{\mathbf{x}_n\}_{n=0}^{\infty}$ which starts with an arbitrary element $\mathbf{x}_0 \in A$, defined by $\mathbf{x}_{n+1} = T\mathbf{x}_n$, tends to $\mathbf{x}^*$ as n tends to infinity.
The following inequalities are true and describe the speed of convergence:

- a priori : $d(\mathbf{x}_n, \mathbf{x}^*) \leq \frac{q^n}{1-q} d(\mathbf{x}_0, \mathbf{x}_1)$
- a posteriori : $d(\mathbf{x}_n, x^*) \leq \frac{q}{1-q} d(x_n, x_{n-1})$.

Notation 1.11 *The weakening of the conditions leads to generalisations of the BANACH FPT.*

Notation 1.12 *There are some analogous kinds of FPT which are of interest for the sake of applications, for example the SCHAUDER FPT (for BVP of pde) or the KAKUTANI FPT (for economics).*

Proof. Auxiliary calculation: $\mathbf{x}_0 \in A; \quad \mathbf{x}_{n+1} = T\mathbf{x}_n$ Thus

$$\begin{aligned} d(\mathbf{x}_n, \mathbf{x}_{n+1}) &= d(T\mathbf{x}_{n-1}, T\mathbf{x}_n) \leq q d(\mathbf{x}_{n-1}, \mathbf{x}_n) \\ &= q d(T\mathbf{x}_{n-2}, T\mathbf{x}_{n-1}) \leq q^2 d(\mathbf{x}_{n-2}, \mathbf{x}_{n-1}) = ... \\ &\leq q^{n-k} d(\mathbf{x}_k, \mathbf{x}_{k+1}) \quad for\ 0 \leq k \leq n \end{aligned}$$

Therefore we get

$$\begin{aligned} d(\mathbf{x}_n, \mathbf{x}_{n+m}) &\leq d(\mathbf{x}_n, \mathbf{x}_{n+1}) + d(\mathbf{x}_{n+1}, \mathbf{x}_{n+2}) + ... + d(\mathbf{x}_{n+m-1}, \mathbf{x}_{n+m}) \\ &\leq d(\mathbf{x}_n, \mathbf{x}_{n+1}) + d(T\mathbf{x}_n, T\mathbf{x}_{n+1}) + ... + d(T\mathbf{x}_{n+m-2}, T\mathbf{x}_{n+m-1}) \\ &\leq (1 + q + q^2 + ... + q^{m-1}) d(\mathbf{x}_n, \mathbf{x}_{n+1}) \\ &\leq (1 + q + q^2 + ... + q^{m-1}) q^{n-k} d(\mathbf{x}_k, \mathbf{x}_{k+1}) \\ &= \frac{1 - q^m}{1 - q} q^{n-k} d(\mathbf{x}_k, \mathbf{x}_{k+1}) \quad for\ 0 \leq k \leq n;\ m \geq 1. \quad (*) \end{aligned}$$

A) Existance: $(*)$ implies

$$\lim_{n\to\infty} d(\mathbf{x}_n, \mathbf{x}_{n+m}) = 0$$

$\curvearrowright$ $\{\mathbf{x}_n\}_{n=0}^{\infty}$ is Cauchy in A. A is closed, $A \subseteq \mathbb{X}$ and $\mathbb{X}$ is complete. That's why there exists an element $\mathbf{x}^*$, such that $\mathbf{x}^* = \lim_{n\to\infty} \mathbf{x}_n$.

$$\begin{aligned} d(\mathbf{x}^*, T\mathbf{x}^*) &\leq d(\mathbf{x}^*, \mathbf{x}_n) + d(\mathbf{x}_n, T\mathbf{x}^*) \\ &\leq d(\mathbf{x}^*, \mathbf{x}_n) + d(T\mathbf{x}_{n-1}, T\mathbf{x}^*) \\ &\leq d(\mathbf{x}^*, \mathbf{x}_n) + d(\mathbf{x}_{n-1}, \mathbf{x}^*) \stackrel{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

$$\curvearrowright \; T\mathbf{x}^* = \mathbf{x}^*$$

B) Unicity: We assume: $\exists \mathbf{x}^*,\ \mathbf{y}^* \mid \mathbf{x}^* \neq \mathbf{y}^*$ but $T\mathbf{x}^* = \mathbf{x}^*$ and $T\mathbf{y}^* = \mathbf{y}^*$. $\curvearrowright$

$$\begin{aligned} d(\mathbf{x}^*, \mathbf{y}^*) &= d(T\mathbf{x}^*, T\mathbf{y}^*) \leq q d(\mathbf{x}^*, \mathbf{y}^*) \\ &\Longrightarrow q \geq 1 : \textit{contradiction} \end{aligned}$$

C) Error estimations:
a posteriori: use $(*)$ with $k = n - 1$ and $m \to \infty$
a priori: use $(*)$ with $k = 0$ and $m \to \infty$ ■

Example 1.34 *Application of the BANACH FPT to the integral equation*

$$\mathbf{x}(t) = \lambda \int_a^b F(t, s, \mathbf{x}(s))ds + f(t); \quad a \leq t \leq b; \quad \lambda \in \mathbb{R} \qquad (+)$$

in compliance with the conditions:
1) $F : [a, b] \times [a, b] \times \mathbb{R} \to \mathbb{R}$ *is continuous,*
2) F_x *is continuous;* $|F_x(t, s, x)| \leq M_1, \quad \forall t, s \in [a, b], x \in \mathbb{R}$
3) $f : [a, b] \to \mathbb{R}$ *is continuous*
4) $\lambda \in \mathbb{R}$, *such that:* $(b - a)|\lambda|M_1 = q < 1$.

Theorem 1.10 *If the conditions 1) to 4) are satisfied, then the integral equation* $(+)$ *admits a unique solution* $x^*(t) \in C[a, b]$.

Proof. By using the integral operator

$$(T_\lambda \mathbf{x})(t) = \lambda \int_a^b F(t, s, \mathbf{x}(s))ds + f(t); \quad a \leq t \leq b$$

the integral equation $(+)$ can be written in the form of a fixpoint equation

$$\mathbf{x}(t) = (T_\lambda \mathbf{x})(t); \quad a \leq t \leq b.$$

A)

$$\begin{aligned} \mathbf{x} \in C[a, b] &\overset{1), 3)}{\Longrightarrow} (T_\lambda \mathbf{x})(t) \in C[a, b] \\ \curvearrowright \quad & T_\lambda : C[a, b] \to C[a, b] \end{aligned}$$

B) By the Mean Value Theorem of differential calculus we get
$\forall s, t \in [a, b] \quad \wedge \quad x_1, x_2 \in \mathbb{R} \quad \exists \xi \in \mathbb{R} \mid$

$$|F(t, s, x_1) - F(t, s, x_2)| = |F_x(t, s, \xi)|\, |x_1 - x_2| \overset{2)}{\leq} M_1 \, |x_1 - x_2| \qquad (*)$$

This implies

$$\begin{aligned}
d((T_\lambda \mathbf{x}_1)(t),(T_\lambda \mathbf{x}_2)(t)) &= \max_{a\le t\le b} |(T_\lambda \mathbf{x}_1)(t)-(T_\lambda \mathbf{x}_2)(t)| \\
&= \max_{a\le t\le b} |\lambda| \left| \int_a^b (F(t,s,\mathbf{x}_1(s))-F(t,s,\mathbf{x}_2(s)))\,ds \right| \\
&\overset{MVTI}{\le} |\lambda| \max_{a\le t\le b} |b-a|\cdot |F(t,\eta,\mathbf{x}_1(\eta))-F(t,\eta,\mathbf{x}_2(\eta))| \\
&\overset{MVTD(*)}{\le} |\lambda||b-a|M_1\,|\mathbf{x}_1(\eta)-\mathbf{x}_2(\eta)| \\
&\le |\lambda||b-a|M_1 \max_{a\le t\le b} |\mathbf{x}_1(t)-\mathbf{x}_2(t)| \\
&= q d(\mathbf{x}_1(t),\mathbf{x}_2(t))
\end{aligned}$$

Condition 4) delivers $q<1$ and therefore the operator T_λ is contractive. ■

Notation 1.13 *The corresponding iteration method is:*

$$\mathbf{x}_{n+1}(t)=\lambda \int_a^b F(t,s,\mathbf{x}_n(s))ds+f(t); \quad a\le t\le b; \quad n=0,1,2,...$$

It converges, for example with $x_0=1$, to the unique solution because of the BANACH FPT.

Notation 1.14 *A special case, in this example, is the case of linear integral equations: $F(t,s,x(s))=K(t,s)x(s)$; with a continuous function $K(t,s):[a,b]\times[a,b]\to\mathbb{R}$*

Example 1.35 *Given the linear integral equation with $K(t,s)=ts$,*

$$\mathbf{x}(t)=\lambda \int_{a=0}^{b=1} ts\mathbf{x}(s)\,ds+1 \qquad \text{for} \qquad 0\le t\le 1 \qquad (o)$$

That means the functions $f(t)\equiv 1$, $\mathbf{F}(t,s,\mathbf{x}(s))=ts\mathbf{x}(s)$ and $\mathbf{F}_x(t,s)=ts$ are continuous in their domains.

$$\max_{0\,\le\, t,s\,\le\, 1} |\,\mathbf{F}_x(t,s)\,|=1=M_1$$

Let λ be a real number such that λ: $(b-a)\,|\lambda|\,M_1=|\lambda|<1$. Then the conditions of the theorem above are fulfilled for the integral equation (o) with $q=|\lambda|$ .This means that in the case of $|\lambda|<1$ the equation (o) has a unique solution $\mathbf{x}^*(t)\in\mathbf{C}[0,1]$.

By successive approximation with $\mathbf{x}_0(t) \equiv 1$ *in* $[0,1]$ *we get the following sequence of approximate solutions* $\mathbf{x}_n(t)$*:*

$$\mathbf{x}_{n+1}(t) = \lambda \int_0^1 ts\mathbf{x}_n(s)\,ds + 1 \qquad \textit{for} \quad 0 \le t \le 1 \quad \textit{and} \quad n = 0,1,2,\ldots\ .$$

$$\begin{aligned}
\mathbf{x}_1(t) &= \lambda\, t \int_0^1 s\ ds + 1 = \frac{\lambda}{2}t + 1 \\
\mathbf{x}_2(t) &= \lambda\, t \int_0^1 \left(\frac{\lambda}{2}s^2 + s\right)\ ds + 1 = \frac{\lambda\, t}{2}\left(\frac{\lambda}{3} + 1\right) + 1 \\
\mathbf{x}_3(t) &= \lambda\, t \int_0^1 \left[\frac{\lambda}{2}\left(\frac{\lambda}{3} + 1\right)s^2 + s\right]\ ds + 1 = \frac{\lambda\, t}{2}\left(\frac{\lambda^2}{3^2} + \frac{\lambda}{3} + 1\right) + 1 \\
&\quad - \\
\mathbf{x}_{n+1}(t) &= \lambda\, t \int_0^1 s\ \mathbf{x}_n(s)\ ds + 1 = \frac{\lambda\, t}{2}\sum_{k=0}^{n}\left(\frac{\lambda}{3}\right)^k + 1 \\
&= \left(\frac{\lambda\, t}{2}\frac{1}{1 - \frac{\lambda}{3}}\right) + 1 \\
&\to \frac{3}{2}\left(\frac{\lambda}{3 - \lambda}\right)t + 1 = \mathbf{x}^*(t) \qquad \textit{for} \quad n \to \infty.
\end{aligned}$$

$\mathbf{x}^*(t)$ *is the solution of* (o) *for every* $|\lambda| < 1$*.*

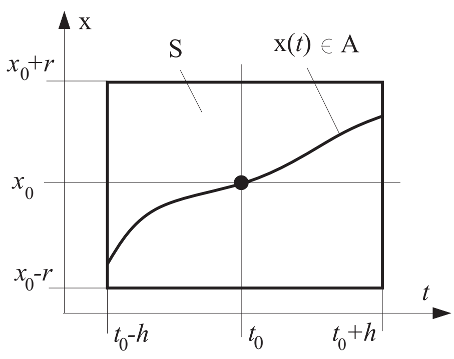
Application to an initial value problem (IVP)

Given the **IVP:**

$$\begin{aligned}
\mathbf{x}'(t) &= \mathbf{f}(t, \mathbf{x}(t)) \qquad |t - t_0| \le h \qquad (*) \\
\mathbf{x}(t_0) &= \mathbf{x}_0
\end{aligned}$$

We search $\mathbf{x}(t) \in \mathbf{A}$ with
$\mathbf{A} \subset \mathbf{X} = \mathbf{C}(t_0 - h, t_0 + h)$
$\mathbf{x}(t)$ differentiable

$$\mathbf{A} = \left\{ \mathbf{x}(t) \in \mathbf{X} \;\middle|\; \max_{|t - t_0| \le h} |\mathbf{x}(t) - \mathbf{x}_0| \le r \right\}$$



We transform the **IVP in to an equivalent integral equation:**

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}(s))\ ds, \qquad |t - t_0| \le h, \qquad \mathbf{x} \in \mathbf{A}$$

With the integral operator $\mathbf{T} : \mathbf{A} \to \mathbf{X}$ given by

$$\mathbf{Tx}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}(s))\ ds, \qquad |t - t_0| \le h$$

we get the fixpoint equation $\mathbf{x}(t) = \mathbf{Tx}(t)$.

Theorem 1.11 ***(PICARD-LINDELÖF)***

1. *Let* $\mathbf{f} : \mathbf{S} \to \mathbb{R}$ *and its partial derivative* $\mathbf{f}_x = \dfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(t, \mathbf{x}) : \mathbf{S} \to \mathbb{R}$ *be continuous on* $\mathbf{S}$*:*

$$\mathbf{S} = \left\{(t, \mathbf{x}) \in \mathbb{R}^2 \mid |t - t_0| \le h; \quad |\mathbf{x} - \mathbf{x}_0| \le r\right\}$$

2. *Let* $\mathbf{f}$ *and* $\mathbf{f}_x$ *be bounded on* $\mathbf{S}$*:*

$$\max_{(t,\mathbf{x}) \in \mathbf{S}} |\mathbf{f}(t, \mathbf{x})| = M \qquad\qquad \max_{(t,\mathbf{x}) \in \mathbf{S}} |\mathbf{f}_x(t, \mathbf{x})| = M_1$$

Suppose

$$hM \le r \qquad \text{and} \qquad hM_1 = q < 1.$$

Then the IVP $(*)$ *has a unique solution*

$$\mathbf{x}^*(t) \in \mathbf{A} \subset \mathbf{C}(t_0 - h, t_0 + h)$$

The sequence of approximate solutions $\mathbf{x}_n(t)$ *calculated by successive approximation*

$$\begin{aligned} \mathbf{x}_{n+1}(t) &= \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}_n(s))\ ds \qquad |t - t_0| \le h \\ \text{or} \qquad \mathbf{x}_{n+1}(t) &= \mathbf{Tx}_n(t) \qquad n = 0, 1, 2, \ldots \end{aligned}$$

tends to $\mathbf{x}^*(t)$ *as* n *tends to* ∞ *for every* $\mathbf{x}_0(t) \in \mathbf{A}$ *, i.e.* $\mathbf{x}_n(t) \overset{n\to\infty}{\to} \mathbf{x}^*(t)$.
For example take $\mathbf{x}_0(t) \equiv \mathbf{x}_0$ *for* $|t - t_0| \le h$ *as a starting point.*

Proof:
We have to show, that $\mathbf{T}$ is contractive on the complete metric space $\mathbf{X} = \mathbf{C}\left(h - t_0, h + t_0\right)$ under the assumptions 1. and 2. The BANACH Fixed Point theorem then implies the theorem wich we require.
a) We verify: $\mathbf{T} : \mathbf{A} \rightarrow \mathbf{A}$.
$\forall\ \mathbf{x}\left(t\right) \in \mathbf{A}$ we have

$$\begin{aligned} \left| \int_{t_0}^{t} \mathbf{f}\left(s, \mathbf{x}\left(s\right)\right)\ ds \right| &\leq \left|t - t_0\right| \max_{(t,\mathbf{x})\ \in\ \mathbf{S}} \left|\mathbf{f}\left(t, \mathbf{x}\right)\right| \leq hM \leq r \\ \text{with}\quad \left|t - t_0\right| &\leq h \end{aligned}$$

Because $\mathbf{Tx}$ is continuous for every $\mathbf{x} \in \mathbf{A}$ we get

$$d_{\mathbb{X}}\left(\mathbf{Tx}, \mathbf{x}_0\right) = \max_{|t-t_0|\leq\ h} \left| \int_{t_0}^{t} \mathbf{f}\left(s, \mathbf{x}\left(s\right)\right)\ ds \right| \leq r$$

and therefore $\mathbf{Tx}\left(t\right) \in \mathbf{A}$.
b) The mean value theorem of the differential calculus implies

$$\left|\mathbf{f}\left(t, \mathbf{x}_1\right) - \mathbf{f}\left(t, \mathbf{x}_2\right)\right| = \left|\mathbf{f}_x\left(t, \boldsymbol{\xi}\right)\right| \left|\mathbf{x}_1 - \mathbf{x}_2\right| \leq M_1 \left|\mathbf{x}_1 - \mathbf{x}_2\right| \quad \forall\left(t, \mathbf{x}_1\right) \in \mathbf{S} \wedge \forall\left(t, \mathbf{x}_2\right) \in \mathbf{S}$$

Thus

$$\begin{aligned} d_{\mathbb{X}}\left(\mathbf{Tx}_1, \mathbf{Tx}_2\right) &= \max_{|t-t_0|\leq\ h} \left| \int_{t_0}^{t} \left[\mathbf{f}\left(s, \mathbf{x}_1\left(s\right)\right) - \mathbf{f}\left(s, \mathbf{x}_2\left(s\right)\right)\right]\ ds \right| \\ &\leq hM_1 \max_{|t-t_0|\leq\ h} \left|\mathbf{x}_1\left(t\right) - \mathbf{x}_2\left(t\right)\right| \\ &= q \cdot d_{\mathbb{X}}\left(\mathbf{x}_1, \mathbf{x}_2\right) \end{aligned}$$

a) and b) together imply that $\mathbf{T}$ is contractive.

Example 1.36 *Given the IVP*

$$\mathbf{x}'\left(t\right) = t\ \mathbf{x}\left(t\right) \qquad \textit{with} \quad \mathbf{x}\left(0\right) = \mathbf{x}_0 = 1 \qquad\qquad (**)$$

The equivalent integral equation to $(**)$ *is the following:*

$$\mathbf{x}\left(t\right) = 1 + \int_{0}^{t} s\ \mathbf{x}\left(s\right)\ ds.$$

With $t_0 = 0$ *and* $\mathbf{x}_0 = 1$ *we get:* $\mathbf{S} = \{(t, \mathbf{x}) \in \mathbb{R}^2 \mid \ |t| \le h; \ \ |\mathbf{x}{-}1| \le r\}$,
and with $\mathbf{f}(t, \mathbf{x}) = t\mathbf{x}$, $\mathbf{f}_x(t) = t$ *we have:*

$$\max_{(t,\mathbf{x}) \in \mathbf{S}} |\mathbf{f}(t, \mathbf{x})| \le h(r+1) = M \qquad \max_{(t,\mathbf{x}) \in \mathbf{S}} |\mathbf{f}_x(t, \mathbf{x})| = h = M_1 .$$

The conditions of the theorem above are valid , if

$$hM = h^2(r+1) \le r \qquad \text{and} \qquad hM_1 = h^2 = q < 1.$$

(We used $|\mathbf{x}| \le r+1$*.) This means that in the case of* $h < 1$ *the IVP* $(**)$ *has a unique solution* $\mathbf{x}^*(t) \in \mathbf{C}(-h, h)$ *. Given* h *then* r *has to be chosen such that* $r \ge h^2/(1-h^2)$ *. By successive approximation with* $\mathbf{x}_0(t) \equiv 1$ *in* $[-h, h]$ *we get the following sequence of approximate solutions* $\mathbf{x}_n(t)$*:*

$$\begin{aligned}
\mathbf{x}_1(t) &= 1 + \int_0^t s \, ds = 1 + \frac{t^2}{2} \\
\mathbf{x}_2(t) &= 1 + \int_0^t s \left(1 + \frac{s^2}{2}\right) ds = 1 + \frac{t^2}{2} + \frac{1}{2}\left(\frac{t^2}{2}\right)^2 \\
\mathbf{x}_3(t) &= 1 + \int_0^t s \, \mathbf{x}_2(s) \, ds = 1 + \frac{t^2}{2} + \frac{1}{2!}\left(\frac{t^2}{2}\right)^2 + \frac{1}{3!}\left(\frac{t^2}{2}\right)^3 \\
& \quad - \\
\mathbf{x}_{n+1}(t) &= 1 + \int_0^t s \, \mathbf{x}_n(s) \, ds = \sum_{k=0}^{n+1} \frac{1}{k!}\left(\frac{t^2}{2}\right)^k \\
&\to \exp\left(\frac{t^2}{2}\right) = \mathbf{x}^*(t) \qquad \textit{for} \qquad (n \to \infty) .
\end{aligned}$$

$\mathbf{x}^*(t)$ *is the solution of* $(**)$ $\forall \ t \in \mathbb{R}$*.*

2 Linear Normed Spaces

The simplest one of the abstract topological spaces is the metric space, about which we spoke earlier. In a metric space you can do analysis because the metric is a distance function between the elements. But you can't compare the elements themselves. Thus you can't order the elements and so you don't have an algebraic structure. That's why you can't calculate in a metric space. But we need to be able to do all these things. First we take the vector space $\mathbb{V}^3$ as a model:

- The connection between $\mathbb{V}^3$ and $\mathbb{R}^3$ is the unique mapping ϕ :
$$\phi(\mathbf{x},\mathbf{y}) = \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$
which assigns every pair $\mathbf{x},\mathbf{y} \in \mathbb{R}^3$ to a vector $\mathbf{v} \in \mathbb{V}^3$: $(\mathbb{R}^3, \mathbb{V}^3, \phi)$ is an affine space.

- The starting point of our consideration is the metric space $\mathbb{R}^3$ with
$$\begin{aligned} d(\mathbf{x},\mathbf{y}) &= \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2} \\ &= \sqrt{v_1^2 + v_2^2 + v_3^2} \\ &= \|\mathbf{v}\|, \end{aligned}$$
such that $d(\mathbf{x},\mathbf{y})$ is the length of the vector $\mathbf{v}$: This is an Euclidian Space.

- The connection between $\mathbb{V}^3$ and $\mathbb{R}^3$ is the mapping ϕ :
$$\phi(\mathbf{x},\mathbf{y}) = \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$
which attaches every pair $\mathbf{x},\mathbf{y} \in \mathbb{R}^3$ to a vector $\mathbf{v} \in \mathbb{V}^3$. We say $(\mathbb{R}^3, \mathbb{V}^3, \phi)$ is an affine space.

- $\mathbb{V}^3$ is a linear space, arithmetic operations between vectors are defined.

The generalisation of the vector space $\mathbb{V}^3$ is a space with arbitrary mathematical objects which has a connection to a corresponding metric space. We obtain linear normed spaces where you can measure and calculate.

2.1 Linear Spaces

Now we recall some facts about linear spaces from the Linear Algebra:

Definition 2.1 *A vector space over a field* $\mathbb{K}$ *(* $\mathbb{K}=\mathbb{R}$ *or* $\mathbb{K}=\mathbb{C}$ *) is a nonempty set* $\mathbb{X}$ *together with two binary operations that satisfy the eight axioms listed below. (Elements of* $\mathbb{X}$ *are called vectors. Elements of* $\mathbb{K}$ *are called scalars.)*

(A) The first operation, **addition**, takes any two elements $\mathbf{x} \in \mathbb{X}$, $\mathbf{y} \in \mathbb{X}$ and assigns to them a third, unique element which is commonly written as $\mathbf{x}+\mathbf{y} \in \mathbb{X}$ and called the sum of these two elements.

(M) The second operation takes any scalar $\lambda \in \mathbb{K}$ and any vector $\mathbf{x} \in \mathbb{X}$ and gives another unique element $\lambda\mathbf{x} \in \mathbb{X}$. The multiplication is called the **scalar multiplication** of $\mathbf{x}$ by λ.

To qualify as a vector space, the set $\mathbb{X}$ and the operations of addition and scalar multiplication must adhere to a number of requirements called the
Axioms of the linear space.
Let $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ be arbitrary vectors in $\mathbb{X}$, and λ and μ be scalars in $\mathbb{K}$.

(A1) $\mathbf{x}+(\mathbf{y}+\mathbf{z}) = (\mathbf{x}+\mathbf{y})+\mathbf{z}$ (**Associativity of addition**)

(A2) $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$ (**Commutativity of addition**)

(A3) There exists a unique element $\mathbb{O} \in \mathbb{X}$, such that $\mathbf{x}+\mathbb{O} = \mathbf{x} \quad \forall \mathbf{x} \in \mathbb{X}$
(**zero/identity element of addition**)

(A4) For every $\mathbf{x} \in \mathbb{X}$, there exists a unique element $(-\mathbf{x}) \in \mathbb{X}$ such that $\mathbf{x}+(-\mathbf{x}) = \mathbb{O}$
(**Inverse element of addition**)

(M1) $(\lambda+\mu)\mathbf{x} = \lambda\mathbf{x}+\mu\mathbf{x}$ (**Distributivity of scalar multiplication with respect to field addition**)

(M2) $\lambda(\mathbf{x}+\mathbf{y}) = \lambda\mathbf{x}+\lambda\mathbf{y}$ (**Distributivity of scalar multiplication with respect to vector addition**)

(M3) $(\lambda\mu)\mathbf{x} = \lambda(\mu\mathbf{x})$ (**Compatibility of scalar multiplication with field multiplication**)

(M4) $1\mathbf{x} = \mathbf{x}$; ($1 \in \mathbb{K}$, **Identity element of scalar multiplication**)

Notation 2.1 *If* $\mathbb{K}=\mathbb{R}$ *(*$\mathbb{K}=\mathbb{C}$*) then we get a real (complex) linear space.*

Notation 2.2 *These axioms generalise properties of the vectors introduced in the beginning of this chapter.*

Example 2.1 $\mathbb{V}^n$: *vector space analogous to* $\mathbb{R}^n$

Example 2.2 *Space of all polynomials* $\mathbb{P}^n = \{\mathbf{P}_n(x)|n \in \mathbb{N}, x \in \mathbb{R}$ *(or.* $\mathbb{C}$)}:

$$\mathbf{P}_n(x) = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n; \quad a_i \in \mathbb{R} \text{ (or } a_i \in \mathbb{C}), \ i{=}0, 1, .., n$$

Addition corresponds to the standard addition of polynomials, scalar multiplcation corresponds to the multiplication of a polynomial by a real or complex number. There exists a connection between the polynomials $\mathbf{P}_n(x)$ *and the vectors of the coefficients of the polynomial:* $\begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{V}^{n+1}$

Example 2.3 $\mathbf{x}(t), \mathbf{y}(t) \in C[a, b]; \quad \lambda \in \mathbb{K}{=}\ \mathbb{C}$:
Addition: $(\mathbf{x} + \mathbf{y})(t) = \mathbf{x}(t) + \mathbf{y}(t)$
Multiplication by a scalar: $(\lambda\mathbf{x})(t) = \lambda\mathbf{x}(t)$
Zero Element: $\mathbf{0}(t) \equiv 0 \quad \forall t \in [a, b]$
Inverse element: $-\mathbf{x}(t)$

Example 2.4 *Space* l_p *of all real or complex number sequences* $\mathbf{x} = \{x_i\}_{i=1}^{\infty}$ *with* $\sum_{i=1}^{\infty} |x_i|^p < \infty; \ 1 \le p < \infty \quad \lambda \in \mathbb{K} = \mathbb{C}^+$:
Addition: $\mathbf{x} + \mathbf{y} = \{x_i + y_i\}_{i=1}^{\infty}$
Multiplication by a scalar: $\lambda\mathbf{x} = \{\lambda x_i\}_{i=1}^{\infty}$
Because of

$$\sum_{i=1}^{\infty} |\lambda x_i|^p = |\lambda|^p \sum_{i=1}^{\infty} |x_i|^p < \infty$$

and the Minkowski inequality

$$\left(\sum_{i=1}^{\infty} |x_i + y_i|^p\right)^{1/p} \le \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{\infty} |y_i|^p\right)^{1/p} < \infty$$

we get $\mathbf{x} + \mathbf{y} \in l_p$ *and* $\lambda\mathbf{x} \in l_p$
Zero Element: $\mathbf{0} = \{0_i\}_{i=1}^{\infty}$
Inverse element: $-\mathbf{x} = \{-x_i\}_{i=1}^{\infty}$

Example 2.5 $\mathbb{X} = \mathbb{R}^+ = (0, \infty)$ *with* $\mathbb{K} = \mathbb{R}$:
Addition: $\mathbf{x} + \mathbf{y} = x \cdot y$
Multiplication by a scalar: $\lambda\mathbf{x} = x^{\lambda}$
Zero element: $\mathbf{0} = 1 \in \mathbb{R}$
Inverse element: $\mathbf{x} + -\mathbf{x} = 1$ *corresponds to* $x \cdot \frac{1}{x} = 1 \curvearrowright -\mathbf{x} = \frac{1}{x}$
Homework: Test the axioms A1 to M4 of the linear space.

Now we adopt several definitions from the linear algebra:

Definition 2.2 *Let $\mathbb{U}$ be a subset of $\mathbb{X}$. Then $\mathbb{U}$ is a subspace if and ony if it satisfies the following conditions:*
a) If $\mathbf{x}$ and $\mathbf{y}$ are elements of $\mathbb{U}$, then the sum $\mathbf{x}+\mathbf{y}$ is an element of $\mathbb{U}$.
b) If $\mathbf{x}$ is an element of $\mathbb{U}$ and λ is a scalar from $\mathbb{K}$, then the scalar product $\lambda\mathbf{x}$ is an element of $\mathbb{U}$.

Conclusion 2.1 *$\mathbb{U}$ itself is a linear space over the field $\mathbb{K}$*

Definition 2.3 *Let $\mathbb{U}$ be a subspace of $\mathbb{X}$ and $x_0 \in \mathbb{X}$ then*

$$M = \{\mathbf{x}_0 + \mathbf{y} \mid \mathbf{y} \in \mathbb{U}\} \equiv \mathbf{x}_0 + \mathbb{U}$$

is called ***linear manifold*** *in $\mathbb{X}$.*

Definition 2.4 *Let A be a subset $A \subset \mathbb{X}$. The set of all finite linear combinations of elements of A*

$$spanA = \left\{ \sum_{k=1}^{m} \lambda_k \mathbf{x}_k \quad | \quad \mathbf{x}_k \in A, \quad \lambda_k \in \mathbb{K}, \quad m \in \mathbb{N} \right\}$$

is called the ***linear span (linear cover) of*** *A*

Definition 2.5 *Let $\mathbb{U}$ and $\mathbb{V}$ be subspaces of $\mathbb{X}$, then*

$$\mathbb{U} + \mathbb{V} = span(\mathbb{U} \cup \mathbb{V})$$

is called the ***sum of*** *$\mathbb{U}$* ***and*** *$\mathbb{V}$. Additionally, if $\mathbb{U} \cap \mathbb{V} = \{\mathbb{O}\}$,*
then $\mathbb{U}+\mathbb{V}$ is called the ***direct sum*** *$\mathbb{U}\oplus\mathbb{V}$. Every $\mathbf{z} \in \mathbb{U}\oplus\mathbb{V}$ has a unique representation in the form $\mathbf{z} = \mathbf{x} + \mathbf{y}$ with $\mathbf{x} \in \mathbb{U}$ and $\mathbf{y} \in \mathbb{V}$.*

Definition 2.6 *If $\mathbb{X} = \mathbb{U} \oplus \mathbb{V}$, then the subspaces $\mathbb{U} \subset \mathbb{X}$ and $\mathbb{V} \subset \mathbb{X}$ are called* ***complementary.***

Definition 2.7 *The set $\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m\} \subset \mathbb{X}$ is called linearly independent, if*

$$\sum_{k=1}^{m} \lambda_k \mathbf{x}_k = \mathbf{0} \Longleftrightarrow \lambda_1 = \lambda_2 = ... = \lambda_m = 0$$

Definition 2.8 *The set $B \subset \mathbb{X}$ is called* ***linearly independent****, if every finite subset of B is linearly independent.*

Definition 2.9 *A linearly independent subset* $B \subset \mathbb{X}$ *with* $\mathbb{X} = \overline{spanB}$ *is called a* ***basis*** *in* $\mathbb{X}$.

Definition 2.10 *If there exists a basis of* $\mathbb{X}$ *with* $|B| = n$*, then every basis of* $\mathbb{X}$ *consists of* n *elements:* $\dim \mathbb{X} = n$. *If there is no finite* n *then* $\mathbb{X}$ *is called* ***infinitely dimensional****.*

Definition 2.11 *Let* $\mathbb{X}$ *and* $\mathbb{Y}$ *be linear spaces over* $\mathbb{K}$. $\mathbb{X}$ *and* $\mathbb{Y}$ *are said to be* ***linear isomorphic****, if there exists a bijection* $f\colon \mathbb{X} \to \mathbb{Y}$ *with the property*

$$f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) \qquad \forall \mathbf{x}, \mathbf{y} \in \mathbb{X}; \quad \alpha, \beta \in \mathbb{K}$$

Example 2.6 *Dimension of* $\mathbb{P}^n$:

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + ... + \alpha_n x^n = 0 \;\Leftrightarrow\; \alpha_i = 0; \quad i = 0, 1, 2, ..., n$$

because a polynomial has exactly n *complex roots (Fundamental theorem of linear algebra)*
$\curvearrowright$ *Basis of* $\mathbb{P}^n = \{1, x, x^2, ..., x^n\}$ $\curvearrowright$ $\dim(\mathbb{P}^n) = n + 1$

Example 2.7 $C[a, b]$:

$$B = \left\{\mathbf{1}, \mathbf{x}, \mathbf{x}^2, ..., \mathbf{x}^n\right\} \subset C[a, b]; \quad n \in \mathbb{N}$$

Any finite linear combination of elements of B *is a polynomial.* $\curvearrowright$ B *is linear independent.*
Weierstrass approximation theorem: Any continuous function can be approximated by polynomials <u>*in arbitrary*</u> *accuracy.*
$\curvearrowright$ $C[a, b] = \overline{span(B)}$ $\curvearrowright$ B *is a basis.* $\curvearrowright$ $dim(C[a, b]) = \infty$

Example 2.8 l_p :
Basis: $= \left\{\mathbf{x}_k = (0, ..., 0, \underset{position\ k}{1}, 0, ...);\ k = 1, 2, ...\right\}$
l_p *is infinite dimensional.*

2.2 Normed Space - BANACH Space

Definition 2.12 *A (real)* ***normed linear space*** $(\mathbb{V}, \|.\|)$ *is a (real) linear space* $\mathbb{V}$ *over the field* $\mathbb{K}$ *together with a function* $\|.\|$, $\mathbb{V} \to \mathbb{R}$*, called the norm, satisfying the following 3 conditions for any* $\mathbf{x}, \mathbf{y} \in \mathbb{V}$*:*

(I) $\|\mathbf{x}\| \geq 0 \quad \wedge \quad \|\mathbf{x}\| = 0 \quad \Longleftrightarrow \quad \mathbf{x} = \mathbf{0}$ (nonnegativity and nondegeneracy)

(II) $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|\,,\quad \alpha\in\mathbb{K}$ (multiplicativity)

(III) $\|\mathbf{x}+\mathbf{y}\|\leq\|\mathbf{x}\|+\|\mathbf{y}\|$ (triangle inequality)

$\|\mathbf{x}\|$ *is called the norm of the element* $\mathbf{x}$.

Notation 2.3 *In any linear normed space* $\mathbb{V}$ *you can introduce the* ***canonical or induced metric by*** $d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|\quad\forall\mathbf{x},\mathbf{y}\in\mathbb{V}$. *Therefore every linear normed space is a metric space.*

Notation 2.4 *In linear normed spaces the metric properties are combined with algebraic structures. You can measure and calculate.*

Notation 2.5 *The revesal, that a linear metric space is a linear normed space, too, is true only if you can find in* $\mathbb{X}$ *a metric* $d(\mathbf{x},\mathbf{y})$ *which is homogeneous (uniform) and invariant by translations:*

(A) Invariance by Translations: $d(\mathbf{x}+\mathbf{z},\mathbf{y}+\mathbf{z}) = d(\mathbf{x},\mathbf{y})\quad\forall\mathbf{x},\mathbf{y},\mathbf{z}\in\mathbb{X}$

(B) Homogeneity: $d(\alpha\mathbf{x},\alpha\mathbf{y}) = |\alpha|\;d(\mathbf{x},\mathbf{y})\quad\forall\mathbf{x},\mathbf{y}\in\mathbb{X}\quad\alpha\in\mathbb{K}$
Then we set $\|\mathbf{x}\| = d(\mathbf{x},\mathbf{0})$ because

$$
\begin{aligned}
\|\mathbf{x}\| &= d(\mathbf{x},\mathbf{0})\geq 0\quad\wedge\quad\|\mathbf{x}\| = d(\mathbf{x},\mathbf{0}) = 0\Leftrightarrow\mathbf{x}=\mathbf{0}\\
\|\alpha\mathbf{x}\| &= d(\alpha\mathbf{x},\mathbf{0}) \overset{(B)}{=} |\alpha|d(\mathbf{x},\mathbf{0}) = |\alpha|\,\|\mathbf{x}\|\\
\|\mathbf{x}+\mathbf{y}\| &= d(\mathbf{x}+\mathbf{y},\mathbf{0}) = d(\mathbf{x}+\mathbf{y},-\mathbf{y}+\mathbf{y})\\
&\overset{(A)}{=} d(\mathbf{x},-\mathbf{y})\\
&\leq d(\mathbf{x},\mathbf{0})+d(\mathbf{0},-\mathbf{y})\\
&\overset{(B)}{=} d(\mathbf{x},\mathbf{0})+d(\mathbf{0},\mathbf{y})\\
&= d(x,\mathbf{0})+d(\mathbf{y},\mathbf{0})\\
&= \|\mathbf{x}\|+\|\mathbf{y}\|
\end{aligned}
$$

Notation 2.6 *The norm is a continuous function which satisfies the inequality:*

$$\big|\|\mathbf{x}\|-\|\mathbf{y}\|\big|\leq\|\mathbf{x}-\mathbf{y}\|\quad\forall\mathbf{x},\mathbf{y}\in\mathbb{V}$$

Proof. *I*)

$$
\begin{aligned}
\|\mathbf{x}\| &= \|\mathbf{x}-\mathbf{y}+\mathbf{y}\|\leq\|\mathbf{x}-\mathbf{y}\|+\|\mathbf{y}\|\\
\|\mathbf{x}\|-\|\mathbf{y}\| &\leq \|\mathbf{x}-\mathbf{y}\|
\end{aligned}
$$

II)

$$\begin{aligned} \|\mathbf{y}\| &= \|\mathbf{y}-\mathbf{x}+\mathbf{x}\| \leq \|\mathbf{y}-\mathbf{x}\| + \|\mathbf{x}\| \\ -\|\mathbf{x}\| + \|\mathbf{y}\| &\leq \|\mathbf{x}-\mathbf{y}\| \end{aligned}$$

Therefore we get

$$\big|\|\mathbf{x}\| - \|\mathbf{y}\|\big| \leq \|\mathbf{x}-\mathbf{y}\| \quad \forall \mathbf{x},\mathbf{y} \in \mathbb{V}$$

■

Definition 2.13 *A linear normed vector space* $\mathbb{V}$ *over the field* $\mathbb{K}$ *which is complete with respect to the metric* $d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|$, *induced by the norm, is called a* ***BANACH space***.

Example 2.9 *BANACH spaces are for example:*

1. $\mathbb{R}^n$ *with*

$$\|\mathbf{x}\|_p = \begin{cases} \left(\sum\limits_{k=1}^{n} |x_k|^p\right)^{1/p} & for \quad 1 \leq p < \infty \\ \max\limits_{1\leq k\leq n} |x_k| & for \quad p = \text{“}\infty\text{”} \end{cases}$$

2. l_p: *space of all bounded number sequences* $\{x_k\}_{k=1}^{\infty}$ *with*

$$\|\mathbf{x}\|_p = \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p} \qquad for \quad 1 \leq p < \infty$$

3. l_∞ : *space of all bounded number sequences* $\{x_k\}_{k=1}^{\infty}$ *with*

$$\|\mathbf{x}\|_\infty = \sup_k |x_k|$$

4. $C[a,b]$: *space of all continuous functions* $f : [a,b] \to \mathbb{C}$ *with*

$$\|\mathbf{f}\|_\infty = \max_{a\leq t\leq b} |\mathbf{f}(t)|$$

5. $C^m[a,b]$: *space of all m-times continuously differentiable functions* $f : [a,b] \to \mathbb{C}$ *with*

$$\|\mathbf{f}\|_\infty = \sum_{k=0}^{m} \max_{a\leq t\leq b} |\mathbf{f}^{(k)}(t)|$$

6. *$L_p[a,b]$: space of all measurable functions $f : [a,b] \to \mathbb{C}$ whose absolute value raised to the p^{th} power has a finite integral:*

$$\int_a^b |\mathbf{f}(t)|^p dt < \infty$$

with

$$\|\mathbf{f}\|_p = \left(\int_a^b |\mathbf{f}(t)|^p dt\right)^{1/p}$$

especially with $p = 2$.

Notation 2.7 *STEFAN BANACH. (1892-1945, Pole, died from lung cancer) His book "Theorie des Operations Linéaires"(1932) is the foundation of the fundamentials of the function analysis in normed spaces.*

2.3 Metric Properties of BANACH Spaces $\mathbb{B}$

Definition 2.14 *The set $K_r(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{B} \mid \|\mathbf{x} - \mathbf{x}_0\| < r\}$ is called an* ***open ball centered at*** $\mathbf{x}_0 \in B$ ***with the radius*** *r.*

Convergence in the BANACH space $\mathbb{B}$:

- Let $\{\mathbf{x}_n\}_{n=1}^{\infty}$ be a sequence in $\mathbb{B}$.
  $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}_0 \quad \Longleftrightarrow \quad \forall\varepsilon > 0\ \exists n_0(\varepsilon) \mid \|\mathbf{x}_n - \mathbf{x}_0\| < \varepsilon \quad \forall n \geq n_0$
- $\{\mathbf{x}_n\}_{n=1}^{\infty}$ is Cauchy in $\mathbb{B}$, if
  $\forall\varepsilon > 0\ \exists n_0(\varepsilon) \mid \|\mathbf{x}_n - \mathbf{x}_m\| < \varepsilon \quad \forall n, m \geq n_0$
- Because of the completeness of the BANACH space, every Cauchy sequence tends to a limit in the BANACH space.
- Convergence in normed spaces is called **norm convergence**.
- The set of all norm convergent sequences is linear:
  $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x} \wedge \lim_{n\to\infty} \mathbf{y}_n = \mathbf{y}$ implies $\lim_{n\to\infty}(\mathbf{x}_n + \mathbf{y}_n) = \mathbf{x} + \mathbf{y}$.
  $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x} \wedge \lim_{n\to\infty} \alpha_n = \alpha$ implies $\lim_{n\to\infty} \alpha_n\mathbf{x}_n = \alpha\mathbf{x}$.
  $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$ implies $\lim_{n\to\infty} \|\mathbf{x}_n\| = \|\mathbf{x}\|$.

Definition 2.15 *In a normed space $\mathbb{V}$ the norms $\|.\|_1$ and $\|.\|_2$ are said to be equivalent if $\exists m, M \in \mathbb{R}, \quad m > 0, M > 0 \mid m\|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2 \leq M\|\mathbf{x}\|_1 \quad \forall \mathbf{x} \in \mathbb{V}$.*

Example 2.10 *Equivalent norms are in the space*

a) $\mathbb{V} = \mathbb{R}^n (or\ \mathbb{V} = \mathbb{C}^n) : \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}; \quad x_k \in \mathbb{R}(\mathbb{C})$

$\|\mathbf{x}\|_\infty = \max_{1\le k\le n} |x_k|; \quad \|\mathbf{x}\|_1 = \sum_{k=1}^{n} |x_k|; \quad \|\mathbf{x}\|_2 = \sqrt{\sum_{k=1}^{n} |x_k|^2}$

A)

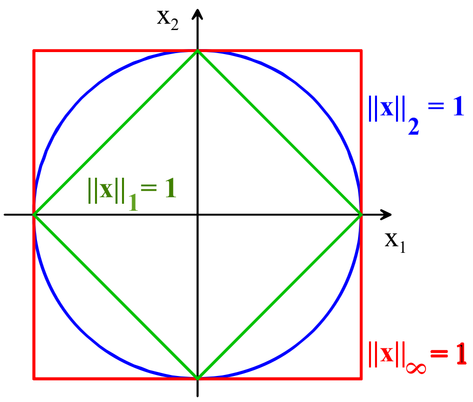
$$\|\mathbf{x}\|_\infty = \max_{1\le k\le n} |x_k| \le \sum_{k=1}^{n} |x_k| = \|\mathbf{x}\|_1 \le n \max_{1\le k\le n} |x_k| = n\,\|\mathbf{x}\|_\infty$$

B)

$$\|\mathbf{x}\|_\infty = \max_{1\le k\le n} |x_k| \le \sqrt{\sum_{k=1}^{n} |x_k|^2} = \|\mathbf{x}\|_2 \le \sqrt{n} \max_{1\le k\le n} |x_k| = \sqrt{n}\,\|\mathbf{x}\|_\infty$$

C)

$$\|\mathbf{x}\|_1 \le n\,\|\mathbf{x}\|_\infty \le n\,\|\mathbf{x}\|_2 \le n\sqrt{n}\,\|\mathbf{x}\|_1$$



b) $\mathbb{V} = C[a,b] : \quad \|\mathbf{f}\|_\infty = \max_{a\le t\le b} |f(t)|; \quad \|\mathbf{f}\|_2 = \max_{a\le t\le b} |e^{-rt} f(t)|;\ r > 0$

e^{-x} *is a strictly monotonic decreasing function. Therefore we get*

$$\begin{aligned} \min_{a\le t\le b} e^{-rt}\,\|\mathbf{f}\|_\infty &= e^{-rb} \max_{a\le t\le b} |\mathbf{f}(t)| \le e^{-rt} \cdot \max_{a\le t\le b} |\mathbf{f}(t)| \le \max_{a\le t\le b} |e^{-rt}\mathbf{f}(t)| \\ &= \|\mathbf{f}\|_2 \le \max_{a\le t\le b} e^{-rt} \cdot \max_{a\le t\le b} |\mathbf{f}(t)| = \max_{a\le t\le b} e^{-ra}\,\|\mathbf{f}\|_\infty \end{aligned}$$

Notation 2.8 *In a finite dimensional normed space* $\mathbb{V}$ *all norms are equivalent.*

Notation 2.9 *Every finite dimensional normed space $\mathbb{V}^n$ is a BANACH space.*

Proof. Let $\{\mathbf{e}_1, \mathbf{e}_2, ...\mathbf{e}_n\}$ be a basis in $V^n \Longrightarrow x = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i \quad \forall x \in \mathbb{V}^n$.
We introduce $\|\mathbf{x}\|_\infty = \max_{1 \le i \le n} |\alpha_i|$. Any other norm is equivalent to $\|.\|_\infty$.
Let $\{\mathbf{x}_k = \sum_{i=1}^{n} \alpha_{ki} \mathbf{e}_i\}_{k=1}^{\infty} \subset \mathbb{V}^n$ be Cauchy in $\mathbb{V}^n$. Thus

$$|\alpha_{mi} - \alpha_{ki}| \le \|\mathbf{x}_m - \mathbf{x}_k\|_\infty \le M \|\mathbf{x}_m - \mathbf{x}_k\|_. < \varepsilon \quad \forall m, k \ge n_0(\varepsilon),\ i = 1, 2, ..., n$$

implies $\{\alpha_{ki}\}_{k=1}^{\infty}$ is Cauchy in $\mathbb{R}$ and $\lim_{k\to\infty} \alpha_{ki} = \alpha_{0i} \quad \forall i$.
$\curvearrowright \quad \lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}_0 = \sum_{i=1}^{n} \alpha_{0i} \mathbf{e}_i \in \mathbb{V}^n$. ∎

Notation 2.10 *The change to an equivalent norm has no influence to convergent sequences. Maybe you have a computational advantage.*

Notation 2.11 *The equivalence of norms is an equivalence relation, that means it is reflexive, symmetric and transitive.*

Series in Normed Spaces:

Definition 2.16 *Let $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, ...$ be elements of a linear normed space $\mathbb{V}$,*
$\mathbf{s}_n = \sum_{k=1}^{n} \mathbf{x}_k$. *(partial sum)*
By definition the series $\sum_{k=1}^{\infty} \mathbf{x}_k$ converges to a limit $\mathbf{s} \in \mathbb{V}$ if and only if the associated sequence of partial sums $\{\mathbf{s}_n\}$converges to $\mathbf{s}$ i.e.

$$\mathbf{s} = \sum_{k=1}^{\infty} \mathbf{x}_k \Longleftrightarrow \lim_{n\to\infty} \mathbf{s}_n = \mathbf{s}$$

$\mathbf{s} = \sum_{k=1}^{\infty} \mathbf{x}_k$ is called the **sum of the series** in $\mathbb{V}$.

Definition 2.17 *The series $\sum_{k=1}^{\infty} \mathbf{x}_k$ is called* ***absolutely convergent****, if the number series $\sum_{k=1}^{\infty} \|\mathbf{x}_k\|$ is convergent.*

Notation 2.12 *In a BANACH space $\mathbb{B}$ every absolutely convergent series is convergent and the following inequality is satisfied:*

$$\left\| \sum_{k=1}^{\infty} \mathbf{x}_k \right\| \leq \sum_{k=1}^{\infty} \|\mathbf{x}_k\| .$$

Proof. *The sequence of the partial sums $\widetilde{s_n} = \sum_{k=1}^{\infty} \|\mathbf{x}_k\|$ of the absolutly convergent series is Cauchy. This implies*

$$\begin{aligned} \|s_n - s_m\| &= \left\| \sum_{k=n+1}^{m} \mathbf{x}_k \right\| \leq \sum_{k=n+1}^{m} \|\mathbf{x}_k\| < \varepsilon \quad \forall n, m \geq n_0(\varepsilon) \\ &\Longrightarrow \quad \exists\, s \in \mathbb{B} \mid \lim_{n\to\infty} s_n = s. \end{aligned}$$

Therefore we get

$$\begin{aligned} \left\| \sum_{k=1}^{n} \mathbf{x}_k \right\| &\leq \sum_{k=1}^{n} \|\mathbf{x}_k\| \leq \sum_{k=1}^{\infty} \|\mathbf{x}_k\| \quad \text{and} \\ \lim_{n\to\infty} \left\| \sum_{k=1}^{n} \mathbf{x}_k \right\| &= \|s\| = \left\| \sum_{k=1}^{\infty} \mathbf{x}_k \right\| \leq \sum_{k=1}^{\infty} \|\mathbf{x}_k\| \end{aligned}$$

■

2.4 Linear Operators

The consideration of linear problems in linear normed spaces makes sense because linear operators use the linear structure of these spaces.

Definition 2.18 *Let $\mathbb{X}, \mathbb{Y}$ be linear normed spaces over the (same) field $\mathbb{K}$. A mapping $\mathbf{A} : \mathbb{X} \to \mathbb{Y}$ is called a **linear operator** if:*

$$\begin{aligned} \mathbf{A}(\mathbf{u} + \mathbf{v}) &= \mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{v} & \forall \mathbf{u}, \mathbf{v} \in \mathbb{X} \\ \mathbf{A}(\alpha \mathbf{u}) &= \alpha A \mathbf{u} & \forall \alpha \in \mathbb{K} \wedge \forall \mathbf{u} \in \mathbb{X}. \end{aligned}$$

Image space of A: $R(\mathbf{A}) = \{\mathbf{y} \in \mathbb{Y} \mid \mathbf{y} = \mathbf{A}\mathbf{x};\ \mathbf{x} \in \mathbb{X}\}$
Null space (kernel) of A: $N(\mathbf{A}) = \{\mathbf{x} \in \mathbb{X} \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}$

Notation 2.13 $N(\mathbf{A}) = \{\mathbf{0}\} \iff \mathbf{A} : \mathbb{X} \to \mathbb{Y}$ *is injective.*

Definition 2.19 *The linear operator* $\mathbf{A} : \mathbb{X} \to \mathbb{Y}$ *is called* ***bounded*** *if there exists some finite positive constant* $C \in \mathbb{R}$ *such that* $\|\mathbf{Ax}\| \leq C \|\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{X}$.

Definition 2.20 *The number*

$$\begin{aligned} \|\mathbf{A}\| &= \inf \{C \in \mathbb{R} \,|\, \|\mathbf{Au}\| \leq C \|\mathbf{u}\|, \forall \mathbf{u} \in \mathbb{X}\} \\ &= \sup_{\|\mathbf{u}\|=1} \|\mathbf{Au}\| \end{aligned}$$

is called the ***norm of the operator.***

Notation 2.14 *:*

$$\begin{aligned} \|\mathbf{Ax}\| &\leq \|\mathbf{A}\| \|\mathbf{x}\| \quad \forall\, \mathbf{x} \in \mathbb{X} \quad \text{and} \\ \mathbf{u} &= \frac{\mathbf{x}}{\|\mathbf{x}\|} \quad \forall\, \mathbf{x} \neq \mathbf{0} \end{aligned}$$

implies

$$\begin{aligned} \|\mathbf{Au}\| &= \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}\| \frac{\|\mathbf{x}\|}{\|\mathbf{x}\|} = \|\mathbf{A}\| \|\mathbf{u}\| = \|\mathbf{A}\| \quad \text{and} \\ \|\mathbf{A}\| &= \sup_{\|\mathbf{u}\|=1} \|\mathbf{Au}\| \end{aligned}$$

Notation 2.15 *You can find linear operators* $\mathbf{A} : \mathbb{X} \to \mathbb{Y}$ *in linear equations of the form* $\mathbf{Ax} = \mathbf{y} \ \ (\mathbf{x} \in \mathbb{X},\ \mathbf{y} \in \mathbb{Y})$. *For example the operator* $\mathbf{A}$ *can be*

- *a system of algebraic equations*
- *an integral equation or*
- *an ordinary or partial differential equation with initial conditions or boundary values.*

Notation 2.16 *Therefore, it is possible to transfer the properties of solutions of systems of linear equations to solutions of linear operator equations.*

Example 2.11 *Let* $\mathbb{X} = \mathbb{R}^n$ *have the basis* $\mathbf{e}_1, ..., \mathbf{e}_n$. *Therefore, every element* $\mathbf{x} \in \mathbb{X}$ *can be described as* $\mathbf{x} = \sum_{k=1}^{n} x_k \mathbf{e}_k$.

Furthermore $\mathbb{Y} = \mathbb{R}^m$ *has the basis* $\mathbf{f}_1, ..., \mathbf{f}_m$ *and so* $\mathbf{y} = \sum_{i=1}^{m} y_i \mathbf{f}_i$, $\forall \mathbf{y} \in \mathbb{Y}$.

The operator $\mathbf{A} : \mathbb{X} \to \mathbb{Y}$ *is defined by:*

$$\begin{aligned} \mathbf{Ae}_k &= \sum_{i=1}^{m} a_{ik} \mathbf{f}_i \in \mathbb{Y} \quad k = 1, ..., n \quad (*) \ \text{and} \\ \mathbf{Ax} &= \sum_{k=1}^{n} x_k \mathbf{Ae}_k \quad (**) \quad \forall\, \mathbf{x} \in \mathbb{X} \\ &= \sum_{k=1}^{n} x_k \sum_{i=1}^{m} a_{ik} \mathbf{f}_i = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} x_k a_{ik} \right) \mathbf{f}_i = \sum_{i=1}^{m} y_i \mathbf{f}_i = \mathbf{y} \end{aligned}$$

Therefore we get

$$\sum_{k=1}^{n} x_k a_{ik} = y_i$$

$\curvearrowright$ *The definition* $(*), (**)$ *of the operator is equivalent to the matrix equation* $Ax = y$:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \Leftrightarrow A\mathbf{x} = \mathbf{y}$$

in the basis $\{\mathbf{e}_k\}$ *of* $\mathbb{X}$ *and in the basis* $\{\mathbf{f}_k\}$ *of* $\mathbb{Y}$. *If you have a look at* $(**)$ *or the matrix then you will see that the operator* A *is obviously linear. Now we define the operator norms in* $\mathbb{X}$ *and* $\mathbb{Y}$ *by*

$$\|\mathbf{x}\|_\infty = \max_{1\le k\le n} |x_k| \quad \textit{and} \quad \|\mathbf{y}\|_\infty = \max_{1\le k\le m} |y_k|$$

$\curvearrowright$

$$\begin{aligned} \|A\mathbf{x}\|_\infty &= \max_{1\le i\le m} |y_i| = \max_{1\le i\le m} \left| \sum_{k=1}^{n} a_{ik} x_k \right| \\ &\le \max_{1\le i\le m} \sum_{k=1}^{n} |a_{ik}| \cdot \max_{1\le k\le n} |x_k| = \left(\max_{1\le i\le m} \sum_{k=1}^{n} |a_{ik}| \right) \|\mathbf{x}\|_\infty \le C \|\mathbf{x}\|_\infty \end{aligned}$$

This means A *is bounded.*

Example 2.12 $\mathbb{X} = \left\{ \mathbf{f}(t) \in \mathbf{C}[a,b] \;\middle|\; \exists\ \dfrac{d\mathbf{f}}{dt} \in \mathbf{C}[a,b] \right\}; \quad \mathbb{Y} = \mathbf{C}[a,b].$
We introduce the differential operator $\mathbb{D} : \mathbb{X} \to \mathbb{Y}$. *It is linear because:*

$$\mathbf{D}(\alpha\mathbf{f} + \beta\mathbf{g}) = \alpha\mathbf{D}(\mathbf{f}) + \beta\mathbf{D}(\mathbf{g}) \qquad \forall \mathbf{f}, \mathbf{g} \in \mathbb{X} \quad \textit{and} \quad \alpha, \beta \in \mathbb{R}$$

For example, we obtain with $\mathbf{f} = f(t) = \sin t; \quad t \in [a,b] : \ \mathbf{D}(\mathbf{f}) = \dfrac{df}{dt} = \cos t.$
In $\mathbb{X} \subset \mathbf{C}[a,b]$ *and* $\mathbb{Y} = \mathbf{C}[a,b]$ *we introduce the norm* $\|\mathbf{f}\| = \max_{a\le\ t\ \le\ b} |\ f(t)|$. *Then* $\mathbf{D}$ *is an unbounded operator:*
Choose $\mathbf{f}_n(t) = \sin(nt) \quad \Longrightarrow \quad \mathbf{f}_n'(t) = n\cos(nt) \quad \wedge \quad \|\mathbf{f}_n\| = 1$

$$\begin{aligned} &\Longrightarrow \ D\mathbf{f}_n = n\cos(nt) \\ &\Longrightarrow \ \|D\mathbf{f}_n\| = n\|\mathbf{f}_n\|; \quad n = 1, 2, ... \end{aligned}$$

Example 2.13 $\mathbb{X} = \mathbb{Y} = \mathbf{C}[a,b]$; $\mathbf{x} = \mathbf{x}(t) \in \mathbf{C}[a,b]$ *and* $\|\mathbf{x}\|_\infty = \max\limits_{a \le t \le b} |\mathbf{x}(t)|$
Consider the integral operator:

$$(\mathbf{A}\mathbf{x})(t) = \int_a^b K(t,s)\,\mathbf{x}(s)\,ds; \qquad \mathbf{x} \in \mathbf{C}[a,b]$$

with the continuous integral kernel function $K(t,s) : [a,b] \times [a,b] \to \mathbb{R}$.
The operator $\mathbf{A} : \mathbf{C}[a,b] \to \mathbf{C}[a,b]$ *is linear and bounded.Linearity is obviuos. (homework)*

$$\begin{aligned}
\|\mathbf{A}\mathbf{x}\| &= \max_{a \le t \le b} \left| \int_a^b K(t,s)\mathbf{x}(s)ds \right| \\
&\le \max_{a \le t \le b} \int_a^b |K(t,s)| ds \cdot \max_{a \le s \le b} |\mathbf{x}(s)| \\
&\le (b-a) \max_{a \le t,s \le b} |K(t,s)| \cdot \|\mathbf{x}\| \le (b-a)M \|\mathbf{x}\| \\
\|\mathbf{A}\| &\le (b-a)M
\end{aligned}$$

Theorem 2.1 *The linear operator* $\mathbf{A} : \mathbb{X} \to \mathbb{Y}$ *is* ***continuous*** *if and only if it is bounded.*

Proof. I) Let $\mathbf{A}$ be continuous on $\mathbb{X} \Longrightarrow \mathbf{A}$ is continuous at $\mathbf{x}_0 = \mathbf{0}$.
Proof by contradiction: We assume that $\mathbf{A}$ is unbounded. $\Longrightarrow \exists \{\mathbf{x}_n\}_{n=1}^{\infty} \subset \mathbb{X}$ such that

$$\mathbf{x}_k \ne \mathbf{0} \ \wedge \ \|\mathbf{A}\mathbf{x}_k\| > k\,\|\mathbf{x}_k\|; \quad k \in \mathbb{N}$$

We set

$$\begin{aligned}
\mathbf{u}_k &= \frac{\mathbf{x}_k}{k\,\|\mathbf{x}_k\|} \in \mathbb{X} \\
\|\mathbf{A}\mathbf{u}_k\| &= \frac{\|A\mathbf{x}_k\|}{k\,\|\mathbf{x}_k\|} > 1 \quad k \in \mathbb{N}
\end{aligned}$$

On the other hand

$$\|\mathbf{u}_k\| = \frac{1}{k} \overset{k\to\infty}{\to} \mathbf{0} \quad \curvearrowright \quad \lim_{k\to\infty} \mathbf{u}_k = \mathbf{0}$$

But the continuity of $\mathbf{A}$ at $\mathbf{x}_0 = \mathbf{0}$ implies

$$\mathbf{A}\mathbf{u}_k \overset{k\to\infty}{\to} \mathbf{A}(\mathbf{0}) = \mathbf{A}(\mathbf{x}-\mathbf{x}) = \mathbf{0} \quad \text{contradiction}$$

$\curvearrowright$ $\mathbf{A}$ is bounded.
II) Precondition: $\mathbf{A}$ is bounded at an arbitrary $\mathbf{x}_0 \in \mathbb{X}$. If $\|\mathbf{x}-\mathbf{x}_0\| < \frac{\varepsilon}{C} = \delta$ then

$$\|\mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{x}_0\| = \|\mathbf{A}(\mathbf{x}-\mathbf{x}_0)\| \le C\,\|\mathbf{x}-\mathbf{x}_0\| < C\frac{\varepsilon}{C} = \varepsilon.$$

Therefore $\mathbf{A}$ is continuous at $\mathbf{x}_0$. ■

Definition 2.21 *A linear continuous operator* $\mathbf{A}:\mathbb{X}\to\mathbb{Y}$ *is called an* ***isomorphism*** *if it is bijective and if* $\mathbf{A}^{-1}$ *is continuous. That means: An isomorphism is a linear homeomorphism. Moreover if* $\|\mathbf{Ax}\| = \|\mathbf{x}\|\ \forall\mathbf{x}\in\mathbb{X}$ *is satisfied then* $\mathbf{A}$ *is called* ***isometric.*** *Normed spaces which are connected by an (isometric) isomorphism are called (isometricly) isomorph.*

Definition 2.22 *The* ***sum*** $\mathbf{T}+\mathbf{S}$ *of the linear operators* $\mathbf{T}$ *and* $\mathbf{S}$ *is defined by the equation* $(\mathbf{T}+\mathbf{S})\mathbf{x} = \mathbf{Tx}+\mathbf{Sx}\quad\forall\mathbf{x}\in\mathbb{X}$*, the* ***product of the operators*** $\mathbf{T}$ ***by*** $\lambda\in\mathbb{R}$ *or* $\lambda\in\mathbb{C}$ *is defined by* $(\lambda\mathbf{T})\mathbf{x} = \lambda(\mathbf{Tx})\quad\forall\mathbf{x}\in\mathbb{X}$*.*

Definition 2.23 *The collection* $\mathbb{L}(\mathbb{X},\mathbb{Y})$ *of all linear bounded operators* $\mathbf{A}:\mathbb{X}\to\mathbb{Y}$ *with the sum and the product defined above is called the* ***space of the linear bounded operators*** $\mathbb{L}(\mathbb{X},\mathbb{Y})$.

Notation 2.17 *The result of the sum and the product defined above is an element of* $\mathbb{L}(\mathbb{X},\mathbb{Y})$*.*

Notation 2.18 *If we introduce the operator norm* $\|\mathbf{A}\|$ *in* $\mathbb{L}(\mathbb{X},\mathbb{Y})$ *for every* $\mathbf{A}\in\mathbb{L}(\mathbb{X},\mathbb{Y})$ *then* $\mathbb{L}(\mathbb{X},\mathbb{Y})$ ***is a linear normed space.***

Example 2.14 *(for the space* $\mathbb{L}(\mathbb{X},\mathbb{Y})$ *)*

$$\mathbb{X}=\mathbb{R}^n \quad \text{with} \quad \|\mathbf{x}\| = \max_{1\le j\le n}|x_j| \qquad \mathbf{x} = (x_1,...,x_n)^T$$

$$\mathbb{Y}=\mathbb{R}^m \quad \text{with} \quad \|\mathbf{y}\| = \max_{1\le i\le m}|y_i| \qquad \mathbf{y} = (y_1,...,y_m)^T$$

Every linear operator $\mathbf{A}:\mathbb{R}^n\to\mathbb{R}^m$ *of* $\mathbb{L}(\mathbb{R}^n,\mathbb{R}^m)$ *can be represented by a matrix*

$$\mathbf{A} = \begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{m1} & ... & a_{mn} \end{pmatrix}.$$

The sum of two $m\times n$ *matrices and the product of a* $m\times n$ *matrix with an element of* $\mathbb{R}$ *or* $\mathbb{C}$ *are the same as the arithmetic operations between linear operators.*$\Longrightarrow$ *There exists a linear isomorphism between the set of all* $m\times n$ *matrices and* $\mathbb{L}(\mathbb{R}^n,\mathbb{R}^m)$.

$$\|\mathbf{Au}\| = \max_{1\le i\le m}\left|\sum_{j=1}^{n} a_{ij}u_j\right| \le \max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}|\max_{1\le j\le n}|u_j| \qquad (+)$$

implies

$$\|\mathbf{A}\| \le \max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}|.$$

With $\max\limits_{1\leq\ i\ \leq\ m}\sum\limits_{j=1}^{n}|a_{ij}| = \sum\limits_{j=1}^{n}|a_{kj}|$ *and* $\overline{\mathbf{u}} = (\overline{u}_1, \overline{u}_1, ..., \overline{u}_n)^T \in \mathbb{R}^n$, *such that* $\overline{u}_j = sign\,(a_{kj})$ *we get* $\|\overline{\mathbf{u}}\| = 1$ *and therefore*

$$\begin{aligned} a_{kj}\cdot sgn\,(a_{kj}) &= a_{kj}\cdot\overline{u}_j = |a_{kj}| \\ \|\mathbf{A}\overline{\mathbf{u}}\| &= \max_{1\leq\ i\ \leq\ m}\left|\sum_{j=1}^{n} a_{ij}\overline{u}_j\right| = \left(\max_{1\leq\ i\ \leq\ m}\sum_{j=1}^{n}|a_{ij}|\right)\|\overline{\mathbf{u}}\|\,. \end{aligned}$$

Thus there exists an element $\overline{\mathbf{u}} \in \mathbb{R}^n$, *such that the inequality* $(+)$ *is an equation for* $\mathbf{u} = \overline{\mathbf{u}}$*:*

$$\|\mathbf{A}\| = \max_{1\leq\ i\ \leq\ m}\sum_{j=1}^{n}|a_{ij}|\,.$$

Only if $n = m$ *the operators of* $\mathbb{L}\,(\mathbb{R}^n,\mathbb{R}^m)$ *can be invertible. In this case we get:*

$$\mathbb{L}_{inv}\,(\mathbb{R}^n,\mathbb{R}^n) = \{\mathbf{A} \in \mathbb{L}\,(\mathbb{R}^n,\mathbb{R}^n)\ \ |\ \det(\mathbf{A}) \neq 0\}\,.$$

Theorem 2.2 *If the image space* $\mathbb{Y}$ *of a linear operator is a BANACH space, then* $\mathbb{L}\,(\mathbb{X},\mathbb{Y})$ *is a BANACH space, too.*

Proof. *Let* $\{A_n\}_{n=1}^{\infty}$ *be a Cauchy sequence in* $\mathbb{L}(\mathbb{X},\mathbb{Y})$. $\curvearrowright$

$$\|\mathbf{A_n} - \mathbf{A}_m\| < \varepsilon \quad \forall n, m \geq n_0(\varepsilon)$$

This implies

$$\begin{aligned} \|\mathbf{A}_n\mathbf{x} - \mathbf{A}_m\mathbf{x}\| &\leq \|\mathbf{A}_n - \mathbf{A}_m\|\,\|\mathbf{x}\| < \varepsilon\,\|\mathbf{x}\| \quad (*) \quad \forall n, m \geq n_0(\varepsilon), \quad \forall \mathbf{x} \in \mathbb{X} \\ &\Longrightarrow \{\mathbf{A}_n\mathbf{x}\}_{n=1}^{\infty} \text{ is Cauchy in } \mathbb{Y} \end{aligned}$$

$\mathbb{Y}$ *is a Banach space* $\Longrightarrow$

$$\begin{aligned} \exists\ \mathbf{Ax} &= \lim_{n\to\infty}\mathbf{A}_n\mathbf{x} \in \mathbb{Y} \quad \forall \mathbf{x} \in \mathbb{X} \\ &\curvearrowright \left.\begin{array}{ccc} \mathbf{A}_n(\alpha\mathbf{r} + \beta\mathbf{s}) & = & \alpha\mathbf{A}_n\mathbf{r} + \beta\mathbf{A}_n\mathbf{s} \\ \downarrow & n\to\infty & \downarrow \\ \mathbf{A}(\alpha\mathbf{r} + \beta\mathbf{s}) & = & \alpha\mathbf{Ar}\ +\ \beta\mathbf{As} \end{array}\right\}\mathbf{A} \text{ is linear} \\ &\curvearrowright \|\mathbf{Ax}\| \leq \|\mathbf{Ax} - \mathbf{A}_n\mathbf{x}\| + \|\mathbf{A}_n\mathbf{x}\| \leq (\varepsilon + \|\mathbf{A}_n\|)\,\|\mathbf{x}\| \quad \forall n \geq n_0(\varepsilon), \quad \forall \mathbf{x} \in \mathbb{X} \\ &\curvearrowright \mathbf{A} \text{ is bounded} \end{aligned}$$

$(*)$ *implies as* $m \to \infty$

$$\begin{aligned} \|\mathbf{A}_n\mathbf{x} - \mathbf{Ax}\| &\leq \varepsilon\,\|\mathbf{x}\| \quad \forall n \geq n_0(\varepsilon), \quad \forall \mathbf{x} \in \mathbb{X} \\ \|\mathbf{A}_n - \mathbf{A}\| &\leq \varepsilon \quad \forall n \geq n_0(\varepsilon), \quad \forall \mathbf{x} \in \mathbb{X} \\ &\curvearrowright \lim_{n\to\infty}\mathbf{A}_n = \mathbf{A} \in \mathbb{L}(\mathbb{X},\mathbb{Y}) \end{aligned}$$

■

Definition 2.24 *The sequence* $\{\mathbf{A}_n\}_{n=1}^{\infty} \subset \mathbb{L}(\mathbb{X},\mathbb{Y})$ *is called* ***norm convergent*** *(strongly convergent) with the limit* $\mathbf{A} \in \mathbb{L}(\mathbb{X},\mathbb{Y})$ *if* $\lim_{n\to\infty} \|\mathbf{A}_n\mathbf{x} - \mathbf{A}\mathbf{x}\| = 0 \quad \forall \mathbf{x} \in \mathbb{X}$ *is satisfied. We write:*

$$\mathbf{A}_n \overset{n\to\infty}{\to} \mathbf{A} \quad \textit{or} \lim_{n\to\infty} \mathbf{A}_n = \mathbf{A}.$$

Notation 2.19 *Strong convergence implies the pointwise convergence.*

Definition 2.25 *Let* $\mathbb{X}$, $\mathbb{Y}$, $\mathbb{Z}$ *be linear spaces and let* $\mathbf{T} : \mathbb{X} \to \mathbb{Y}; \quad \mathbf{S} : \mathbb{Y} \to \mathbb{Z}$ *be linear operators. Then the product* $\mathbf{ST}$ *of the operators is defined by*
$(\mathbf{ST})\mathbf{x} = \mathbf{S}(\mathbf{T}\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{X}$.

Definition 2.26 *Let* $\mathbf{T} : \mathbb{X} \to \mathbb{Y}$ *be a linear operator. If there exists a linear operator* $\mathbf{S} : \mathbb{Y} \to \mathbb{X}$ *such that :*
$\mathbf{ST} = \mathbf{I}_{\mathbb{X}} \quad \wedge \quad \mathbf{TS} = \mathbf{I}_{\mathbb{Y}};$ *with* $\mathbf{I}_{\mathbb{X}}, \mathbf{I}_{\mathbb{Y}}$ *identity maps from* $\mathbb{X}$ *to* $\mathbb{X}$ *or* $\mathbb{Y}$ *to* $\mathbb{Y}$ *then* $\mathbf{S}$ *is the* ***inverse Operator*** *of* $\mathbf{T} : \mathbf{S} = \mathbf{T}^{-1}$.
The collection of all invertible operators $T \in \mathbb{L}(\mathbb{X},\mathbb{Y})$ *is called* $\mathbb{L}_{inv}(\mathbb{X},\mathbb{Y})$ *.*

Theorem 2.3 *Given an operator* $\mathbf{T}$*. If the inverse operator* $\mathbf{T}^{-1}$ *exists then it is unique.*

Proof. Assumption: $\exists\ \mathbf{T}_1^{-1} \ \wedge\ \mathbf{T}_2^{-1} \mid \mathbf{T}_1^{-1} \neq \mathbf{T}_2^{-1}$

$$\begin{aligned} \mathbf{T}\mathbf{T}_1^{-1} &= \mathbf{I}_{\mathbb{Y}} \ \wedge\ \mathbf{T}\mathbf{T}_2^{-1} = \mathbf{I}_{\mathbb{X}} \\ \mathbf{T}(\mathbf{T}_1^{-1} - \mathbf{T}_2^{-1}) &= \mathbf{0} \qquad \curvearrowright \\ \mathbf{T}_1^{-1} &= \mathbf{T}_2^{-1} \qquad \text{Contradiction} \end{aligned}$$

■

Notation 2.20 $\mathbb{L}_{inv}(\mathbb{X},\mathbb{Y})$ *is open in* $\mathbb{L}(\mathbb{X},\mathbb{Y})$ *with respect to the operator norm.*

3 HILBERT Spaces

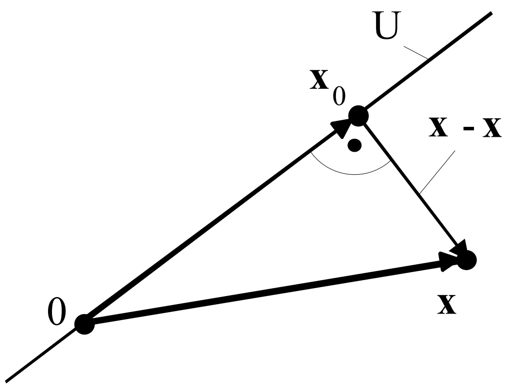
In the space $\mathbb{R}^3$ an inner product is defined:
$(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3 \quad \forall \mathbf{x} = (x_1, x_2, x_3)^T, \mathbf{y} = (y_1, y_2, y_3)^T \in \mathbb{R}^3$.
Furthermore we have $\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$.
The following condition for the inner product leads us to the orthgonality of two elements: $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ are orthogonal if and only if $(\mathbf{x}, \mathbf{y}) = 0$. Therefore we can define the position of any two elements of the space.
Orthogonality is the background for the solution of the following **approximation problem**:

Given:

$\mathbf{x} \in \mathbb{R}^3$ and a subspace $\mathbf{U} \subset \mathbb{R}^3$

We look for an element

$\mathbf{x}_0 \in \mathbf{U}$ with $\|\mathbf{x} - \mathbf{x}_0\| = \min_{\mathbf{y} \in \mathbf{U}} \|\mathbf{x} - \mathbf{y}\|$



$\mathbf{x}_0$ is the **best approximation** of $\mathbf{x}$ with respect to the subspace $\mathbf{U}$. $\mathbf{x}_0$ satisfies the condition

$$(\mathbf{x} - \mathbf{x}_0, \mathbf{y}) = 0 \qquad \forall \mathbf{y} \in \mathbf{U}.$$

Moreover using this we can specify the definition of a basis: If you take an orthonormal system (ONS) in $\mathbb{R}^3$, for example $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \mid (\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}\}$, then you can describe an element $\mathbf{x} \in \mathbb{R}^3$ by $\mathbf{x} = \sum_{k=1}^{3} (\mathbf{x}, \mathbf{e}_k)\mathbf{e}_k$.
We want to generalise these facts in order to abstract this concept to infinite dimensional spaces. If an abstract space with an inner product is complete then you can represent its elements by FOURIER series. This is the basis of numerics in such spaces. That's why we are looking first for an inner product space.

Definition 3.1 *An inner product space* $(\mathbb{H}, (.,.))$ *or pre-HILBERT space is a linear space* $\mathbb{H}$ *over the field* $\mathbb{K}$ *together with a function* $(.,.) : \mathbb{H} \times \mathbb{H} \to \mathbb{K}$*, called the* ***inner product*** *which satisfies the following conditions:*

1. $(\mathbf{x}, \mathbf{x}) \geq 0 \qquad \wedge \qquad (\mathbf{x}, \mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ *(nonnegativity and nondegeneracy)*

2. $(\mathbf{x},\mathbf{y}) = \overline{(\mathbf{y},\mathbf{x})}$ *(Hermitian symmetry)*

3. $(\alpha\mathbf{x}+\beta\mathbf{y},\mathbf{z}) = \alpha(\mathbf{x},\mathbf{z})+\beta(\mathbf{y},\mathbf{z}) \quad \forall \mathbf{x},\mathbf{y},\mathbf{z}\in\mathbb{H};\ \alpha,\beta\in\mathbb{K}$ *(linearity in the first argument)***.**

Theorem 3.1 *Every pre-HILBERT space is a normed space with the norm* $\|\mathbf{x}\| = \sqrt{(\mathbf{x},\mathbf{x})}\ \ \forall \mathbf{x}\in\mathbb{H}$ *. (Proof see below)*

Example 3.1 $\mathbb{H} = \mathbf{l}_2 : \mathbf{x} = \{x_1, x_2, ...\}\in\mathbb{H}, \quad \mathbf{y} = \{y_1, y_2, ...\}\in\mathbb{H}$

$$(\mathbf{x},\mathbf{y}) = \sum_{k=1}^{\infty} x_k\overline{y_k}; \quad \|\mathbf{x}\| = \left(\sum_{k=1}^{\infty}|x_k|^2\right)^{0.5}$$

Example 3.2 $\mathbb{H} = \mathbf{C}[a,b] :\ \mathbf{f},\mathbf{g}\in\mathbb{H}$

$$(\mathbf{f},\mathbf{g}) = \int_a^b \mathbf{f}(t)\overline{\mathbf{g}(t)}dt; \quad \|\mathbf{f}\| = \left(\int_a^b |\mathbf{f}|^2 dt\right)^{0.5}$$

Properties of the inner product:

1. $(\mathbf{u},\alpha\mathbf{v}) = \overline{\alpha}(\mathbf{u},\mathbf{v}) \qquad \forall \mathbf{u},\mathbf{v}\in\mathbb{H},\ \alpha\in\mathbb{K}$

2. $(\mathbf{u},\mathbf{v}+\mathbf{w}) = (\mathbf{u},\mathbf{v})+(\mathbf{u},\mathbf{w}) \qquad \forall \mathbf{u},\mathbf{v},\mathbf{w}\in\mathbb{H}$

3. $|(\mathbf{u},\mathbf{v})| \le \|\mathbf{u}\|\cdot\|\mathbf{v}\| \qquad \forall \mathbf{u},\mathbf{v}\in\mathbb{H}$ SCHWARZ's Inequality

Proof. 1.:

$$(\mathbf{u},\alpha\mathbf{v}) = \overline{\alpha}\overline{(\mathbf{v},\mathbf{u})} = \overline{\alpha}(\mathbf{u},\mathbf{v}) \quad \forall \mathbf{u},\mathbf{v}\in\mathbb{H},\ \alpha\in\mathbb{K}$$

2.:

$$\begin{aligned}(\mathbf{u},\mathbf{v}+\mathbf{w}) &= \overline{(\mathbf{v}+\mathbf{w},\mathbf{u})} = \overline{(\mathbf{v},\mathbf{u})}+\overline{(\mathbf{w},\mathbf{u})}\\ &= (\mathbf{u},\mathbf{v})+(\mathbf{u},\mathbf{w}) \qquad \forall \mathbf{u},\mathbf{v},\mathbf{w}\in\mathbb{H}\end{aligned}$$

3.:

$$\begin{aligned}0 &\le (\mathbf{u}-\lambda\mathbf{v},\mathbf{u}-\lambda\mathbf{v}) \qquad \forall \mathbf{u},\mathbf{v}\in\mathbb{H},\ \lambda\in\mathbb{K}\\ &= (\mathbf{u},\mathbf{u})-\lambda(\mathbf{v},\mathbf{u})-\overline{\lambda}(\mathbf{u},\mathbf{v})+\lambda\overline{\lambda}(\mathbf{v},\mathbf{v})\end{aligned}$$

We choose $\lambda = \frac{(\mathbf{u},\mathbf{v})}{(\mathbf{v},\mathbf{v})}$ and substitute:

$$\begin{aligned} 0 &\leq (\mathbf{u},\mathbf{v}) - \frac{(\mathbf{u},\mathbf{v})}{(\mathbf{v},\mathbf{v})}(\mathbf{v},\mathbf{u}) - \frac{\overline{(\mathbf{u},\mathbf{v})}}{(\mathbf{v},\mathbf{v})}(\mathbf{u},\mathbf{v}) + \frac{(\mathbf{u},\mathbf{v})\overline{(\mathbf{u},\mathbf{v})}}{(\mathbf{v},\mathbf{v})(\mathbf{v},\mathbf{v})}(\mathbf{v},\mathbf{v}) \\ &= \|\mathbf{u}\|^2 - \frac{|(\mathbf{u},\mathbf{v})|^2}{\|\mathbf{v}\|^2} \curvearrowright \\ |(\mathbf{u},\mathbf{v})| &\leq \|\mathbf{u}\| \cdot \|\mathbf{v}\| \end{aligned}$$

■

Proof. of theorem 3.1:

The norm properties 1 and 2 of $\|\mathbf{x}\| = \sqrt{(\mathbf{x},\mathbf{x})}$ are obvious. We prove only property 3:

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\|^2 &= (\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) \\ &= (\mathbf{x},\mathbf{x})+(\mathbf{x},\mathbf{y})+(\mathbf{y},\mathbf{x})+(\mathbf{y},\mathbf{y}) \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + (\mathbf{x},\mathbf{y}) + \overline{(\mathbf{x},\mathbf{y})} \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \mathbf{2}\,\mathrm{Re}(\mathbf{x},\mathbf{y}) \\ &\leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \mathbf{2}|(\mathbf{x},\mathbf{y})| \\ &\leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \mathbf{2}\,\|\mathbf{x}\|\,\|\mathbf{y}\| \quad (SCHWARZ) \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2 \end{aligned}$$

■

Theorem 3.2 *The inner product in a pre-HILBERT space is continuous, i.e.* $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$ *and* $\lim_{n\to\infty} \mathbf{y}_n = \mathbf{y}$ *imply* $\lim_{n\to\infty}(\mathbf{x}_n,\mathbf{y}_n) = (\mathbf{x},\mathbf{y})$.

Proof.

$$\begin{aligned} 0 &\leq |(\mathbf{x},\mathbf{y})-(\mathbf{x}_n,\mathbf{y}_n)| \\ &= |(\mathbf{x},\mathbf{y}) - (\mathbf{x_n},\mathbf{y})+(\mathbf{x_n},\mathbf{y})-(\mathbf{x}_n,\mathbf{y}_n)| \\ &\leq |(\mathbf{x},\mathbf{y}) - (\mathbf{x_n},\mathbf{y})| + |(\mathbf{x_n},\mathbf{y})-(\mathbf{x}_n,\mathbf{y}_n)| \\ &\leq \|\mathbf{x}-\mathbf{x_n}\|\,\|\mathbf{y}\| + \|\mathbf{x}_n\|\,\|\mathbf{y}-\mathbf{y}_n\| = \alpha_n \end{aligned}$$

α_n is a null sequence because of $\lim_{n\to\infty}(\mathbf{x}-\mathbf{x_n}) = \lim_{n\to\infty}(\mathbf{y}-\mathbf{y_n}) = \mathbf{0}$ and $\lim_{n\to\infty} \mathbf{x_n} = \mathbf{x}$. ■

Definition 3.2 *A* ***HILBERT space*** *is a complete pre-HILBERT space with respect to the metric* $d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|$, *induced by the norm* $\|\mathbf{x}\| = \sqrt{(\mathbf{x},\mathbf{x})}$.

Notation 3.1 *Every HILBERT space is a BANACH space. A BANACH space is a HILBERT space if and only if the norm of the BANACH space satisfies the parallelogramm idendity*

$$\|\mathbf{x}-\mathbf{y}\|^2 + \|\mathbf{x}+\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2).$$

Notation 3.2 *Therefore all definitions, theorems and facts about linear normed spaces are hold in HILBERT spaces.*

Example 3.3 $\mathbb{H} = \mathbb{C}$ (*or* $\mathbb{H} = \mathbb{R}$) :

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \overline{\mathbf{y}}; \quad \|\mathbf{x}\| = |\mathbf{x}| \quad \forall \mathbf{x} \in \mathbb{H}$$

Example 3.4 $\mathbb{H} = \mathbb{C}^n$ (*or* $\mathbb{H} = \mathbb{R}^n$) :

$$(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{n} x_k \overline{y_k}; \quad \|\mathbf{x}\| = \left(\sum_{k=1}^{n} |x_k|^2 \right)^{0.5} \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{H}$$

Example 3.5 $n \to \infty : \mathbb{H} = \mathbf{l}_2, \ \dim \mathbf{l}_2 = \infty$:

$$(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{\infty} x_k \overline{y_k}; \quad \|\mathbf{x}\| = \left(\sum_{k=1}^{\infty} |x_k|^2 \right)^{0.5} \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{H}$$

(space of square-summable sequences)

Example 3.6 $\mathbb{H} = \mathbb{L}_2[a, b] : \ -\infty < a < b < \infty$
is the set of all measurable functions $\mathbf{f} : [a, b] \to \mathbb{C}$ *such that* $(L) \int_a^b |\mathbf{f}(t)|^2 \, dt < \infty$

$$(\mathbf{f}, \mathbf{g}) = \int_a^b \mathbf{f}(t) \overline{\mathbf{g}(t)} dt \quad \forall \mathbf{f}, \mathbf{g} \in \mathbb{H}$$

But: $C[a, b]$ *with the inner product* $(\mathbf{f}, \mathbf{g}) = \int_a^b \mathbf{f}(t) \overline{\mathbf{g}(t)} dt$ *is not a HILBERT space!* $C[a, b]$ *with this inner product is a dense subspace in* $\mathbb{L}_2[a, b]$.

3.1 FOURIER Series in HILBERT Spaces

Definition 3.3 *Let* $\mathbb{H}$ *be a HILBERT space. The two elements* $\mathbf{u}, \mathbf{v} \in \mathbb{H}$ *are* ***orthogonal*** $(\mathbf{u} \perp \mathbf{v})$, *if* $(\mathbf{u}, \mathbf{v}) = 0$.

Definition 3.4 *Let* $\mathbb{H}$ *be a HILBERT space. A system* $\{\mathbf{e}_i\}_{i=1}^{\infty}$ *is called an* ***orthonormal system*** *(ONS) if* $(\mathbf{e}_i, \mathbf{e}_k) = \delta_{ik} = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases}$.
The ONS is called ***closed or complete***, *if* $\overline{\underset{i}{span}\{\mathbf{e}_i\}} = \mathbb{H}$.

Conclusion 3.1 $\forall \ \mathbf{u} \in \mathbb{H} \ , \forall \varepsilon > 0 \quad \exists u_i \in \mathbb{R} \quad \wedge \quad \exists n_0(\varepsilon) \quad |$

$$\left\| \mathbf{u} - \sum_{i=1}^{n} u_i \mathbf{e}_i \right\| < \varepsilon \quad \forall n > n_0$$

Definition 3.5 *A HILBERT space $\mathbb{H}$ is called* ***separable****, if $\mathbb{H}$ has a countable subset $M = \{\mathbf{u}_1, \mathbf{u}_2, ...\}$ which is dense in $\mathbb{H}$, i.e. $\overline{M} = \mathbb{H}$.*

Theorem 3.3 *In a separable HILBERT space $\mathbb{H}$, there exists at least one complete ONS (and the contrary is true, too).*

Proof. A) First we construct a basis of $\mathbb{H}$:
Let $M = \{u_i |\ i = 1, 2, ...\}$ be dense in $\mathbb{H}$ $\curvearrowright$ $\overline{M} = \mathbb{H}$
First step: We delete all linearly dependent elements of M and get
$M_u = \{u_k |\ k = 1, 2, ...\}$ such that

$$\overline{span(M)} = \overline{span(M_u)} = \mathbb{H}.$$

Second step: We orthogonalise and normalise the elements of M_u by the method of Erhard Schmidt and prove it by induction.
Beginning of induction:

$$\begin{aligned}
\mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}; \quad \|\mathbf{e}_1\| = 1 \\
\widetilde{\mathbf{e}}_2 &= \mathbf{u}_2 - (\mathbf{u}_2, \mathbf{e}_1)\mathbf{e}_1; \quad \mathbf{e}_2 = \frac{\widetilde{\mathbf{e}}_2}{\|\widetilde{\mathbf{e}}_2\|}; \quad \|\mathbf{e}_2\| = 1 \\
(\widetilde{\mathbf{e}}_2, \mathbf{e}_1) &= (\mathbf{u}_2 - (\mathbf{u}_2, \mathbf{e}_1)\mathbf{e}_1, \mathbf{e}_1) \\
&= (\mathbf{u}_2, \mathbf{e}_1) - (\mathbf{u}_2, \mathbf{e}_1)(\mathbf{e}_1, \mathbf{e}_1) = 0
\end{aligned}$$

We assume

$$\widetilde{\mathbf{e}}_k = \mathbf{u}_k - \sum_{i=1}^{k-1} (\mathbf{u}_k, \mathbf{e}_i)\mathbf{e}_i; \quad \mathbf{e}_k = \frac{\widetilde{\mathbf{e}}_k}{\|\widetilde{\mathbf{e}}_k\|} \ \textit{with} \ (\widetilde{\mathbf{e}}_k, \mathbf{e}_l) = (\mathbf{e}_k, \mathbf{e}_l) = 0; \quad l = 1, .., k-1 \quad (*)$$

Step of induction:

$$\begin{aligned}
\widetilde{\mathbf{e}}_{k+1} &= \mathbf{u}_{k+1} - \sum_{i=1}^{k} (\mathbf{u}_{k+1}, \mathbf{e}_i)\mathbf{e}_i; \quad \mathbf{e}_{k+1} = \frac{\widetilde{\mathbf{e}}_{k+1}}{\|\widetilde{\mathbf{e}}_{k+1}\|}; \quad \|\mathbf{e}_{k+1}\| = 1 \\
(\widetilde{\mathbf{e}}_{k+1}, \mathbf{e}_l) &= \left(\mathbf{u}_{k+1} - \sum_{i=1}^{k} (\mathbf{u}_{k+1}, \mathbf{e}_i)\mathbf{e}_i, \mathbf{e}_l \right); \quad l = 1, 2, .., k \\
&= (\mathbf{u}_{k+1}, \mathbf{e}_l) - \sum_{i=1}^{k} (\mathbf{u}_{k+1}, \mathbf{e}_i)(\mathbf{e}_i, \mathbf{e}_l) \\
&= (\mathbf{u}_{k+1}, \mathbf{e}_l) - (\mathbf{u}_{k+1}, \mathbf{e}_l) = 0 \quad \text{because of } (\mathbf{e}_i, \mathbf{e}_l) = \delta_{il} \text{ (see } (*))
\end{aligned}$$

$\curvearrowright$ $\mathbf{e}_{k+1}$ and $\mathbf{e}_l$ are orthogonal and normalised for all $l < k$.
B) $span\{\mathbf{e}_i\}$ is dense in $\mathbb{H}$ because we deleted the linearly dependent elements only. ■

Notation 3.3 *A HILBERT space with a complete ONS is separable. (= contrary to the theorem above)*

Examples for ONS and separable HILBERT spaces

Example 3.7 $\mathbb{H} = l_2$:

$$M = \{\mathbf{x} = \{r_1, r_2, ..., r_n, 0, 0...\} \mid n \in \mathbb{N};\;\; r_k \in \mathbb{Q};\;\; k = 1, 2, ..., n\}$$

Example 3.8 $\mathbb{H} = \mathbb{L}_2[a, b]$:

$$M = \{\mathbf{P}(t) = \sum_{k=0}^{n} r_k t^k \mid n \in \mathbb{N};\;\; r_k \in \mathbb{Q};\;\; k = 0, 1, ..., n\}$$

Example 3.9 $\mathbb{H} = \mathbb{L}_2[-\pi, \pi]$: *ONS:*

$$\mathbf{e}_0(t) = \frac{1}{\sqrt{2\pi}};\;\; \mathbf{e}_{2k-1}(t) = \frac{1}{\sqrt{\pi}}\cos(kt);\;\; \mathbf{e}_{2k}(t) = \frac{1}{\sqrt{\pi}}\sin(kt);\;\; k = 1, 2, ...$$

$$\begin{aligned} \mathbf{f} &\in \mathbb{L}_2[-\pi, \pi] \;\curvearrowright\; \mathbf{f}(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kt) + b_k \sin(kt)] \\ a_k &= \frac{1}{\sqrt{\pi}} \int\limits_{-\pi}^{\pi} \mathbf{f}(t)\cos(kt)dt; \quad b_k = \frac{1}{\sqrt{\pi}} \int\limits_{-\pi}^{\pi} \mathbf{f}(t)\sin(kt)dt \end{aligned}$$

Example 3.10 *Method of orthonormalisation by E. Schmidt*
$\mathbb{H} = L_2(-1, 1)$: $\quad M_n = \{\mathbf{f}_n(t) = t^n;\;\; n = 0, 1, 2, ...\}$; $\quad \overline{M_n} = \mathbb{L}_2[-1, 1]$
The functions $\mathbf{f}_n(t)$ are linearly independent.

$$\begin{aligned} \|\mathbf{f}_0(t)\| &= \sqrt{\int_{-1}^{1} 1dt} = \sqrt{2} \\ \mathbf{e}_1(t) &= \frac{\mathbf{f}_0(t)}{\|\mathbf{f}_0(t)\|} = \frac{1}{\sqrt{2}} \\ \widetilde{\mathbf{e}}_2(t) &= \mathbf{f}_1 - (\mathbf{f}_1, \mathbf{e}_1)\mathbf{e}_1 = t - \left(\int_{-1}^{1} t\frac{1}{\sqrt{2}}dt\right) \cdot \frac{1}{\sqrt{2}} \\ &= t - \frac{1}{2}\left[\frac{t^2}{2}\right]_{-1}^{1} = t - \frac{1}{2}\left[\frac{1}{2} - \frac{1}{2}\right] = t \\ \mathbf{e}_2(t) &= \frac{t}{\|t\|} = \frac{t}{\sqrt{(t, t)}} = \frac{t}{\sqrt{\int_{-1}^{1} t^2 dt}} = \frac{t}{\sqrt{\frac{1}{3}\left[t^3\right]_{-1}^{1}}} \\ &= \frac{t}{\sqrt{\frac{1}{3}[1 - (-1)]}} = \sqrt{\frac{3}{2}}t \end{aligned}$$

$$\begin{aligned}
\widetilde{\mathbf{e}}_3(t) &= \mathbf{f}_2 - (\mathbf{f}_2, \mathbf{e}_1)\mathbf{e}_1 - (\mathbf{f}_2, \mathbf{e}_2)\mathbf{e}_2 \\
&= t^2 - \left(t^2, \frac{1}{\sqrt{2}}\right) \cdot \frac{1}{\sqrt{2}} - \left(t^2, \sqrt{\frac{3}{2}}t\right) \cdot \sqrt{\frac{3}{2}}t \\
&= t^2 - \frac{1}{2}\int_{-1}^{1} t^2 dt - \sqrt{\frac{3}{2}}t \int_{-1}^{1} \sqrt{\frac{3}{2}} t^3 dt \\
&= t^2 - \frac{1}{2}\left[\frac{t^3}{3}\right]_{-1}^{1} - \sqrt{\frac{3}{2}}t\sqrt{\frac{3}{2}}\left[\frac{t^4}{4}\right]_{-1}^{1} = t^2 - \frac{1}{3} \\
\mathbf{e}_3(t) &= \frac{\widetilde{\mathbf{e}}_3(t)}{\|\widetilde{\mathbf{e}}_3(t)\|} = \frac{t^2 - \frac{1}{3}}{\left(\int_{-1}^{1}\left(t^2 - \frac{1}{3}\right) dt\right)^{0.5}} = \ldots \textit{(homework)} \\
&= \sqrt{\frac{5}{2}}\left(\frac{3}{2}t^2 - \frac{1}{2}\right)
\end{aligned}$$

The resulting ONS is called LEGENDRE polynomials. The functions $\mathbf{e}_n$ *can also be given by*

$$\begin{aligned}
\mathbf{e}_n &= \sqrt{\frac{2n+1}{2}}\frac{1}{(2n)!!}\frac{d^n}{dx^n}(x^2-1)^n; \quad n = 0, 1, 2, \ldots \\
(2n)!! &= 2n(2n-2)\ldots 2
\end{aligned}$$

Notation 3.4 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ *is a closed (= complete) ONS in the HILBERT space* $\mathbb{H}$
$\Longleftrightarrow (\mathbf{u}, \mathbf{e}_i) = 0 \quad \forall i \Longrightarrow \mathbf{u} = \mathbf{0}$

Notation 3.5 *Every* $\mathbf{u} \in \mathbb{H}$ *can be represented by* $\mathbf{u} = \sum_{i=1}^{\infty} u_i \mathbf{e}_i$
$\Longleftrightarrow$ *the ONS* $\{\mathbf{e}_i\}_{i=1}^{\infty}$ *is complete.*

Theorem 3.4 *Let* $\mathbb{H}$ *be a separable HILBERT space with the ONS* $\{\mathbf{e}_i\}_{i=1}^{\infty}$, $\mathbf{u} \in \mathbb{H}$, $\mathbf{s}_n = \sum_{i=1}^{n} \gamma_i \mathbf{e}_i$; $\gamma_i \in \mathbb{C}$. *Then we get:*

1. $\|\mathbf{u} - \mathbf{s}_n\|$ is minimal for $\gamma_i = u_i = (\mathbf{u}, \mathbf{e}_i) \quad \forall i$

2. $\lim_{n\to\infty} \sum_{i=1}^{n} u_i \mathbf{e}_i = \mathbf{s} \in \mathbb{H}$

3. The series $\sum_{i=1}^{\infty} |u_i|^2$ converges and BESSEL's inequality $\sum_{i=1}^{\infty} |u_i|^2 \leq \|\mathbf{u}\|^2$ is satisfied.

4. If the ONS $\{\mathbf{e}_i\}_{i=1}^{\infty}$ is complete, then $\mathbf{s} = \mathbf{u}$ and we get PARSEVAL's identity: $\sum_{i=1}^{\infty} |u_i|^2 = \|\mathbf{u}\|^2$

Proof. 1)

$$\begin{aligned}
\|\mathbf{u}-\mathbf{s}_n\|^2 &= \left(\mathbf{u}-\sum_{i=1}^{n}\gamma_i\mathbf{e}_i, \mathbf{u}-\sum_{j=1}^{n}\gamma_j\mathbf{e}_j\right)\\
&= \|\mathbf{u}\|^2-\sum_{i=1}^{n}\gamma_i(\mathbf{e}_i,\mathbf{u})-\sum_{i=1}^{n}\overline{\gamma}_i(\mathbf{u},\mathbf{e}_i)+\sum_{i=1}^{n}\gamma_i\overline{\gamma}_i\\
&= \|\mathbf{u}\|^2+\sum_{i=1}^{n}(u_i\overline{u}_i-\gamma_i\overline{u}_i-u_i\overline{\gamma}_i+\gamma_i\overline{\gamma}_i)-\sum_{i=1}^{n}u_i\overline{u}_i\\
&= \|\mathbf{u}\|^2+\sum_{i=1}^{n}|u_i-\gamma_i|^2-\sum_{i=1}^{n}|u_i|^2
\end{aligned}$$

We find the minimum in the case $u_i=\gamma_i$. Then $\mathbf{s}_n=\sum_{i=1}^{n}\gamma_i\mathbf{e}_i=\sum_{i=1}^{n}(\mathbf{u},\mathbf{e}_i)\mathbf{e}_i$ implies

$$0\le\|\mathbf{u}-\mathbf{s}_n\|^2=\|\mathbf{u}\|^2-\sum_{i=1}^{n}|(\mathbf{u},\mathbf{e}_i)|^2 \quad \text{Bessel's idendity}$$

3) $\curvearrowright$

$$\sum_{i=1}^{\infty}|(\mathbf{u},\mathbf{e}_i)|^2\le\|\mathbf{u}\|^2 \quad \text{Bessel's inequality } (*)$$

Therefore

$$\sum_{i=1}^{\infty}|(\mathbf{u},\mathbf{e}_i)|^2=\sum_{i=1}^{\infty}|u_i|^2 \text{ is convergent}$$

2) We show that $\{\mathbf{s}_n=\sum_{i=1}^{n}u_i\mathbf{e}_i\}_{n=1}^{\infty}$ is a Cauchy sequence.

$$\begin{aligned}
\|\mathbf{s}_n-\mathbf{s}_m\|^2 &= \left(\sum_{i=n+1}^{m}u_i\mathbf{e}_i, \sum_{j=n+1}^{m}u_j\mathbf{e}_j\right)\\
&= \sum_{i=n+1}^{m}|u_i|^2
\end{aligned}$$

The convergence of $\sum_{i=1}^{\infty}|u_i|^2$ implies $\|\mathbf{s}_n-\mathbf{s}_m\|^2<\varepsilon$ for $n,m\ge n_0(\varepsilon)$.
$\mathbb{H}$ is complete $\curvearrowright$ $\mathbf{lim}_{n\to\infty}\mathbf{s}_n=\mathbf{s}\in\mathbb{H}$
4) We verify that $\mathbf{lim}_{n\to\infty}\mathbf{s}_n=\mathbf{u}$.
$\overline{span\{\mathbf{e}_i\}}=\mathbb{H}$ because of the completness of the ONS (= precondition)

$$\curvearrowright \quad \forall\varepsilon>0 \quad \exists \text{ at least one } \mathbf{v}\in span\{\mathbf{e}_i\} \mid \|\mathbf{u}-\mathbf{v}\|<\varepsilon \ \wedge\ \mathbf{v}=\sum_{i=1}^{n}\beta_i\mathbf{e}_i\,;\ n=n(\varepsilon)$$

By 1) we get

$$\|\mathbf{u}-\mathbf{s}_n\|^2 = \|\mathbf{u}\|^2 - \sum_{i=1}^{n} |(\mathbf{u},\mathbf{e}_i)|^2 \overset{n\to\infty}{\longrightarrow} 0$$

$\curvearrowright$

$$\|\mathbf{u}\|^2 = \sum_{i=1}^{\infty} |u_i|^2$$

$\blacksquare$

Definition 3.6 *Let $\mathbb{H}$ be a HILBERT space with the ONS $\{\mathbf{e}_i\}_{i=1}^{\infty}$, $\mathbf{u} \in \mathbb{H}$.*
The series $\sum_{i=1}^{\infty}(\mathbf{u},\mathbf{e}_i)\mathbf{e}_i$ is called the ***FOURIER series*** *of $\mathbf{u}$ with respect to the ONS $\{\mathbf{e}_i\}_{i=1}^{\infty}$, the numbers $u_i = (\mathbf{u},\mathbf{e}_i)$ are called* ***FOURIER coefficients****.*

Example 3.11 *We know from doing analysis in the space $\mathbb{H} = \mathbb{L}_2[-\pi,\pi]$:*

$$\mathbf{f}(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty}[a_k\cos(kt) + b_k\sin(kt)]$$

Example 3.12 $\mathbb{H} = \mathbb{L}_2^c[0,T]$, *i.e.* $f : [0,T] \to \mathbb{C}$
A complete ONS is:

$$\phi_k(t) = \frac{1}{\sqrt{T}}e^{ik\omega t}; \quad k \in \mathbb{Z}; \quad \omega = \frac{2\pi}{T}$$

Then

$$\mathbf{f}(t) = \sum_{-\infty}^{\infty} c_k e^{ik\omega t}; \quad with \quad c_k = \frac{1}{\sqrt{T}}\int_0^T \mathbf{f}(t)e^{-ik\omega t}dt$$

Problem: The following is known:
$\mathbf{u} \in \mathbb{H}$ and $\{\mathbf{e}_i\}_{i=1}^{\infty}$ is an ONS in $\mathbb{H}$. This implies $(\mathbf{u}, e_i) = u_i;\quad \sum_{i=1}^{\infty}|u_i|^2 \le \|\mathbf{u}\|^2$. Is the opposite correct, too? Does any number sequence with $\sum_{i=1}^{\infty}|y_i|^2 < \infty$ define an element $\mathbf{y} \in \mathbb{H}$?

Theorem 3.5 ***RIESZ - FISCHER***
Let $\{\mathbf{e}_i\}_{i=1}^{\infty}$ be an ONS in the HILBERT space $\mathbb{H}$ and $\{\gamma_i\}_{i=1}^{\infty}$ be a number series with $\sum_{i=1}^{\infty}|\gamma_i|^2 < \infty \quad \Longrightarrow \quad \exists \mathbf{v} \in \mathbb{H} \mid (\mathbf{v},\mathbf{e}_i) = \gamma_i \quad \forall i \quad \wedge \quad \mathbf{v} = \sum_{i=1}^{\infty}\gamma_i\mathbf{e}_i$.
If $\{\mathbf{e}_i\}_{i=1}^{\infty}$ is closed in $\mathbb{H}$, then $\mathbf{v}$ is well defined.

Proof. 1) $s_n = \sum_{i=1}^{n}\gamma_i\mathbf{e}_i$ is a Cauchy sequence (see proof of theorem 3.5)
$\mathbb{H}$ is complete $\curvearrowright$ $\mathbf{lim}_{n\to\infty}\mathbf{s}_n = \mathbf{v} \in \mathbb{H}$ $\curvearrowright$ $\mathbf{v} = \sum_{i=1}^{\infty}\gamma_i\mathbf{e}_i$ $\curvearrowright$

$$(\mathbf{v},\mathbf{e}_k) = \mathbf{lim}_{n\to\infty}\left(\sum_{i=1}^{n}\gamma_i\mathbf{e}_i, e_k\right) = \gamma_k \quad \forall k$$

2) Let $\{\mathbf{e}_i\}_{i=1}^{\infty}$ be closed. We assume that

$$\exists \mathbf{u}, \mathbf{v} \in \mathbb{H} \mid (\mathbf{u}, \mathbf{e}_i) = (\mathbf{v}, \mathbf{e}_i) \quad \forall i \quad \wedge \quad \mathbf{u} \neq \mathbf{v}$$

$\mathbf{z} = \mathbf{u} - \mathbf{v}$ implies $(\mathbf{z}, \mathbf{e}_i) = 0 \quad \forall i \quad \wedge \quad \mathbf{z} \neq \mathbf{0}$.
Parseval's identity implies

$$\sum_{i=1}^{\infty} |(\mathbf{z}, \mathbf{e}_i)|^2 = \|\mathbf{z}\|^2 = 0 \curvearrowright \mathbf{z} = \mathbf{0} \quad \text{contradiction}$$

■

Summary:

1. Every element of a separable HILBERT space can be represented by a FOURIER series:
$$\mathbf{u} = \sum_{i=1}^{\infty} u_i \mathbf{e}_i; \quad u_i = (\mathbf{u}, \mathbf{e}_i).$$

2. Every number sequence $\{\gamma_i\}_{i=1}^{\infty}$ with $\sum_{i=1}^{\infty} |\gamma_i|^2 < \infty$ defines an element $\mathbf{v}$ of a HILBERT space uniquely:
$$\mathbf{v} = \sum_{i=1}^{\infty} \gamma_i \mathbf{e}_i$$

3. $$\mathbf{s}_n = \sum_{i=1}^{n} u_i \mathbf{e}_i; \quad \text{with } u_i = (\mathbf{u}, \mathbf{e}_i)$$
is the best approximation of the element $\mathbf{u}$ with respect to the finite basis $\{\mathbf{e}_i\}_{i=1}^{n}$ (see theorem 3.5)

3.2 Special HILBERT Spaces

3.2.1 The Space $\mathbb{C}^n(\mathbb{R}^n)$

Elements: Vectors $\mathbf{x} = (x_1 \ x_2 \ ... x_n)^T \quad x_i \in \mathbb{C}(\mathbb{R})$

Arithmetic operations: addition and multiplication by numbers are known from linear algebra.

Inner product/norm: $(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} x_i \overline{y}_i, \quad \|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}$

SCHWARTZ's inequality: $|(\mathbf{x},\mathbf{y})| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$, i.e.

$$\left|\sum_{i=1}^{n} x_i \overline{y_i}\right| \leq \sqrt{\sum_{i=1}^{n} |x_i|^2} \cdot \sqrt{\sum_{i=1}^{n} |y_i|^2}$$

Basis system: $\{\mathbf{e}_i\}_{i=1}^{n} = \{(0 \, ...0 \; \mathbf{e}_i^i = 1 \; 0...0)^T, \quad i = 1, 2, ..., n\}$
$\{\mathbf{e}_i\}_{i=1}^{n}$ is an ONS, $span\{\mathbf{e}_i\}_{i=1}^{n} = \mathbb{C}^n, \quad \dim \mathbb{C}^n = n$

Properties:
$\mathbb{C}^n$ is separable: If you calculate $span\{\mathbf{e}_i\}_{i=1}^{n}$ with rational coefficients then you get a countable subset which is dense in $\mathbb{C}^n$.
FOURIER series:

- $\mathbf{x} = \sum_{i=1}^{n} (\mathbf{x}, \mathbf{e}_i)\mathbf{e}_i; \quad x_i = (\mathbf{x}, \mathbf{e}_i) \quad \forall \mathbf{x} \in \mathbb{C}^n$
- $\|\mathbf{x}\|^2 = \sum_{i=1}^{n} |x_i|^2$: PARSEVAL's Equation
- Every n-tuple $\{x_i\}_{i=1}^{n}$ can be attached to an element $\mathbf{x} \in \mathbb{C}^n$ (RIESZ-FISCHER theorem).

Remarks:

1. other inner products can be defined:
   for example, if $A = A^*$: $\quad (A\mathbf{x}, \mathbf{x}) \geq 0 \quad \forall \mathbf{x} \neq \mathbf{0}$
   $\curvearrowright \quad (\mathbf{x}, \mathbf{y})_A = (A\mathbf{x}, \mathbf{y})$ is an inner product
   $\curvearrowright \quad |(\mathbf{x}, \mathbf{y})_A| \leq \|\mathbf{x}\|_A \, \|\mathbf{y}\|_A$, i.e.

$$\left|\sum_{j,k=1}^{n} a_{kj} x_j \overline{y_k}\right|^2 \leq \sum_{j,k=1}^{n} a_{kj} x_j \overline{x_k} \cdot \sum_{i=1}^{n} a_{kj} y_j \overline{y_k}$$

2. An example of another finite dimensional HILBERT space is the following:
   $P^n = \{\mathbf{X}(t) \mid \mathbf{X}(t) = a_0 + a_1 t + a_2 t^2 + ... + a_n t^n; \quad t \in [a, b], \quad a_i \in \mathbb{C}\}$
   $(\mathbf{X}(t), \mathbf{Y}(t)) = \int_a^b \mathbf{X}(t)\overline{\mathbf{Y}(t)}dt$

3.2.2 The space l_2

Elements: number series $\mathbf{x} = \{x_k\}_{k=1}^{\infty} \;\; \mathbf{y} = \{y_k\}_{k=1}^{\infty}$
with $\sum_{k=1}^{\infty} |x_k|^2 < \infty; \quad \sum_{k=1}^{\infty} |y_k|^2 < \infty; \quad x_i, y_i \in \mathbb{C}(\mathbb{R})$
The elements are sequences: $\mathbf{x} = \{x_1, x_2, ...\}$ with the components x_i

Arithmetic operations: $\mathbf{x} + \mathbf{y} = \{x_1 + y_1, x_2 + y_2, ...\}$
$\lambda\mathbf{x} = \{\lambda x_1, \lambda x_2, ...\}; \quad \lambda \in \mathbb{C}(\mathbb{R})$

Inner product/Norm:

$$(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{\infty} x_k \overline{y_k} \qquad \|\mathbf{x}\| = \sqrt{\sum_{k=1}^{\infty} |x_k|^2}$$

SCHWARTZ's inequality: $|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$, i.e.

$$\left|\sum_{k=1}^{\infty} x_k \overline{y_k}\right| \leq \sqrt{\sum_{k=1}^{\infty} |x_k|^2} \cdot \sqrt{\sum_{k=1}^{\infty} |y_k|^2}$$

(The convergence of the series $\sum\limits_{k=1}^{\infty} x_k \overline{y_k}$ must be shown using a different method:
$(|a| - |b|)^2 \geq 0$ implies $|a|^2 - 2|a||b| + |b|^2 \geq 0$ and therefore $2|a||b| \leq |a|^2 + |b|^2$.
$\curvearrowright \quad \left|\sum\limits_{k=1}^{\infty} x_k \overline{y_k}\right| \leq \sum\limits_{k=1}^{\infty} |x_k \overline{y_k}| \leq \sum\limits_{k=1}^{\infty} \frac{1}{2}(|x_k|^2 + |y_k|^2) < \infty$)

Basis systems:

- $\left\{\mathbf{e}_k = \{0, ..., 0, e_k^k = 1, 0, ...\}\right\}_{k=1}^{\infty}$ is an ONS.
- $\{\mathbf{e}_k\}_{k=1}^{\infty}$ is closed.
- l_2 is infinite-dimensional.

Properties: l_2 is separable. (without proof)
FOURIER series:

- $\mathbf{x} = \sum_{i=1}^{\infty} (\mathbf{x}, \mathbf{e}_i)\mathbf{e}_i; \quad x_i = (\mathbf{x}, \mathbf{e}_i) \quad \forall \mathbf{x} \in l_2$
- $\|\mathbf{x}\|^2 = \sum_{i=1}^{\infty} |x_i|^2$: PARSEVAL's Equation

Remarks about the space $\mathbf{l}_p$; $p \geq 1$; $p \neq 2$:

- Set of all number sequences $\mathbf{x} = \{x_k\}_{k=1}^{\infty}$ with $\sum_{i=1}^{\infty} |x_i|^p < \infty$
- normed by: $\|\mathbf{x}\| = \sqrt[\frac{1}{k}]{\sum_{i=1}^{\infty} |x_i|^p}$
- complete
- separable
- But $\mathbf{l}_p$ is not a HILBERT space because there isn't any inner product with: $\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$.

3.2.3 The space $\mathbb{L}_2(a,b)$

Elements: measurable functions $\mathbf{f}:\ (a,b)\to\mathbb{C}$, with $(L)\int\limits_a^b |\mathbf{f}(t)|^2\,dt<\infty$

These functions are called square-integrable.
Arithmetic operations: $\mathbf{f}(t)+\mathbf{g}(t)$ and. $\lambda\mathbf{f}(t),\ \lambda\in\mathbb{C}$ are calculated pointwise

Inner product/Norm: $(\mathbf{f},\mathbf{g})=(L)\int\limits_a^b \mathbf{f}(t)\overline{\mathbf{g}(t)}dt,\quad \|\mathbf{f}\|^2=(L)\int\limits_a^b |\mathbf{f}(t)|^2\,dt$

SCHWARTZ's inequality: $|(\mathbf{f},\mathbf{g})|\le\|\mathbf{f}\|\,\|\mathbf{g}\|$, i.e.

$$\left|(L)\int\limits_a^b \mathbf{f}(t)\overline{\mathbf{g}(t)}dt\right|^2\le\left((L)\int\limits_a^b |\mathbf{f}(t)|^2\,dt\right)\cdot\left((L)\int\limits_a^b |\mathbf{g}(t)|^2\,dt\right)$$

Basis systems:

A) $(a,b)=(-1,1)$: $\qquad \widetilde{\mathbf{e}_n}(t)=t^n;\quad n=0,1,2,...$
$\{\widetilde{\mathbf{e}_n}\}_{n=0}^{\infty}$ is linearly independent and complete in $\mathbb{L}_2(-1,1)$.
Orthonormalising by SCHMIDT implies:

$$\mathbf{e}_n(t)=\left(\frac{2n+1}{2}\right)^{1/2}\mathbf{L}_n(t)\qquad n=0,1,2,...$$

with the LEGENDRE polynomials

$$\mathbf{L}_n(t)=\frac{1}{n!2^n}\frac{d^n}{dt^n}\left(t^2-1\right)^n$$

for example:

$$\begin{aligned}\mathbf{L}_0(t)&=1,\qquad \mathbf{L}_1(t)=t,\qquad \mathbf{L}_2(t)=\frac{1}{2}\left(3t^2-1\right)\\ \mathbf{L}_3(t)&=\frac{1}{2}\left(5t^3-3t\right),\qquad \mathbf{L}_4(t)=\frac{1}{8}\left(35t^4-30t^2+3\right)\end{aligned}$$

B) $(a,b)=(-\pi,\pi)$:
complete ONS:

$$\boldsymbol{\varphi}_0(t)=\frac{1}{\sqrt{2\pi}}\ ,\ \boldsymbol{\varphi}_{2k-1}(t)=\frac{1}{\sqrt{\pi}}\cos(kt)\ ,\ \boldsymbol{\varphi}_{2k}(t)=\frac{1}{\sqrt{\pi}}\sin(kt)\qquad k=1,2,...$$

Thus:

$$\mathbf{f}(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kt) + b_k \sin(kt)] \quad ,$$

$$\text{with} \quad a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} \mathbf{f}(t)\cos(kt)\,dt \qquad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} \mathbf{f}(t)\sin(kt)\,dt$$

C) $(a,b) = (0,\pi)$:
complete ONS I :

$$\boldsymbol{\varphi}_0(t) = \frac{1}{\sqrt{\pi}} \ , \ \boldsymbol{\varphi}_k(t) = \sqrt{\frac{2}{\pi}}\cos(kt)\,, \quad k = 1,2,...$$

complete ONS II :

$$\boldsymbol{\varphi}_k(t) = \sqrt{\frac{2}{\pi}}\sin(kt)\,, \quad k = 1,2,...$$

D) $(a,b) = (0,T)$:
complete ONS:

$$\boldsymbol{\varphi}_k(t) = \frac{1}{\sqrt{T}}\exp(ik\omega t) \qquad k = 0,\pm1,\pm2,... \qquad \omega = \frac{2\pi}{T}$$

Thus

$$\mathbf{f}(t) = \sum_{k=-\infty}^{\infty} c_k \exp(ik\omega t) \ , \quad \text{with} \quad c_k = \frac{1}{T}\int_0^T \mathbf{f}(t)\exp(-ik\omega t)\,dt$$

Properties:

- The space $\mathbb{L}_2(a,b)$ is infinite-dimensional.
  Proof: Assume $[c,d] \subset (a,b)$ with $-\infty < c < d < \infty$ and

  $$\mathbf{f}_m(t) = \begin{cases} t^m & \text{for} \quad t \in [c,d] \\ 0 & \quad\text{otherwise} \end{cases} .$$

  then $\mathbf{f}_m \in \mathbb{L}_2(a,b)$ and the set $\{\mathbf{f}_0(t),...,\mathbf{f}_m(t)\}$ is linearly independent in $\mathbb{L}_2(a,b)$ for any $m \in \mathbb{N}$. Therefore $\dim(\mathbb{L}_2(a,b)) = \infty$.

- $\mathbb{L}_2(a,b)$ is complete and separable. If (a,b) is a finite interval then the set of polynomials with rational coefficients is dense in $\mathbb{L}_2(a,b)$. Every $\mathbf{f}(t) \in \mathbb{L}_2(a,b)$ can be approximated by such a polynomial with any accuracy.

- The elements of $\mathbb{L}_2(a,b)$ are classes of functions. $\mathbf{f}_1(t)$ and $\mathbf{f}_2(t)$ belong to the same class if $\mathbf{f}_1(t) = \mathbf{f}_2(t)$ almost everywhere over (a,b), i.e. if $\mathbf{f}_1(t) \neq \mathbf{f}_2(t)$ on a set of measure zero or $(L)\int\limits_a^b |\mathbf{f}_1(t) - \mathbf{f}_2(t)|\, dt = 0$

FOURIER series:

Using the basis systems B), C) and D) we get the common FOURIER series.
Remarks:

- about the space $\mathbb{L}_p$:
  The space $\mathbb{L}_p$ of all measurable functions $\mathbf{f}(t),\ p \geq 1;\ p \neq 2$, whose absolute value raised to the p-th power has a finite integral is not a HILBERT space. (For $p \neq 2$ there doesn' t exist an inner product such that $\|\mathbf{f}\| = \sqrt{(\mathbf{f},\mathbf{f})}$.) That means
  $(L)\int_a^b |\mathbf{f}(t)|^p\, dt < \infty; \quad \|\mathbf{f}\| = \sqrt[\frac{1}{p}]{(L)\int_a^b |\mathbf{f}(t)|^p\, dt}.$
  These functions are called: "to the pth power integrable functions".

- The set $\mathbf{C}_*\,[a,b]$ of all continuous functions over $[a,b]$ with the inner product

$$(\mathbf{f},\mathbf{g}) = \int\limits_a^b \mathbf{f}\,(t) \cdot \overline{\mathbf{g}\,(t)} dt \qquad \text{for any } \ \mathbf{f},\mathbf{g} \in \mathbf{C}\,[a,b]$$

  is a pre-HILBERT space, but not a HILBERT space. But: $\mathbf{C}_*\,[a,b]$ with this inner product is a dense subspace in $\mathbb{L}_2\,(a,b)$.

- The reason for the definition of the inner product in the space $\mathbb{L}_2\,(a,b)$ by LEBESGUE integrals is, that the definition by RIEMANN integrals leads only to a pre-HILBERT space. The LEBESGUE integral is more general than the RIEMANN integral. For example, you need more stringent requirements for changing integral and limit

$$\lim_{n\to\infty} \int\limits_a^b \mathbf{f}_n\,(t)\, dt = \int\limits_a^b \lim_{n\to\infty} \mathbf{f}_n\,(t)\, dt$$

  in the RIEMANN integral than in the LEBESGUE integral.

3.2.4 The space $\mathbb{L}_2(G)$; $G \subseteq \mathbb{R}^n$; $G \neq \varnothing$; G measurable

Elements: all on G defined functions $\mathbf{f}: G \to \mathbb{R}$ (or $\mathbb{C}$) with real or complex values, which are measurable, with
$(L)\int\limits_G |\mathbf{f}(\mathbf{x})|^2 \, dG < \infty$. These functions are called square-integrable on G.
Arithmetic operations:

$$\begin{aligned} \alpha\mathbf{f}+\beta\mathbf{g} &\in \mathbb{L}_2(G) \quad \text{because} \quad (\alpha\mathbf{f}+\beta\mathbf{g})(\mathbf{x}) = \alpha\mathbf{f}(\mathbf{x})+\beta\mathbf{g}(\mathbf{x}) \quad \mathbf{x}\in G \\ \forall \mathbf{f},\mathbf{g} &\in \mathbb{L}_2(G) \ \text{and} \ \forall\alpha,\beta\in\mathbb{R} \,(\text{or } \forall\alpha,\beta\in\mathbb{C}) \end{aligned}$$

Inner product/norm:

$$(\mathbf{f},\mathbf{g}) = \int\limits_G \mathbf{f}(\mathbf{x})\,\overline{\mathbf{g}(\mathbf{x})}\,dG \qquad \|\mathbf{f}\|_{L_2} = \sqrt{\int\limits_G |\mathbf{f}(\mathbf{x})|^2\,dG}$$

SCHWARTZ's inequality: $|(\mathbf{f},\mathbf{g})| \leq \|\mathbf{f}\|\,\|\mathbf{g}\|$, i.e.

$$\left|\int\limits_G \mathbf{f}(\mathbf{x})\,\overline{\mathbf{g}(\mathbf{x})}\,dG\right| \leq \sqrt{\int\limits_G |\mathbf{f}(\mathbf{x})|^2\,dG}\cdot\sqrt{\int\limits_G |\mathbf{g}(\mathbf{x})|^2\,dG}$$

Basis systems: $\mathbb{H}=\mathbb{L}_2(-\infty,\infty)$: $\mathbf{f}_n(t) = t^n \exp\left(-\frac{1}{2}t^2\right) \quad n = 0,1,2,...$
$\{\mathbf{f}_n\}$ is linear independent and complete in $\mathbb{L}_2(-\infty,\infty)$.
Orthonormalising by SCHMIDT implies:

$$\begin{aligned} \boldsymbol{\varphi}_n(t) &= \frac{1}{\alpha_n}\exp\left(-\frac{1}{2}t^2\right)\mathbf{H}_n(t) \qquad n = 0,1,2,... \\ \text{with} \quad \alpha_n &= \sqrt{2^n n!\sqrt{\pi}} \end{aligned}$$

and the HERMITE polynomials

$$\mathbf{H}_n(t) = (-1)^n \exp\left(t^2\right)\frac{d^n}{dt^n}\left(\exp\left(-t^2\right)\right)$$

for example:

$$\mathbf{H}_0(t) = 1, \qquad \mathbf{H}_1(t) = 2t, \qquad \mathbf{H}_3(t) = 4t^2 - 2$$

Properties:

- $\mathbb{L}_2(G)$ is a separable HILBERT space.

- The elements of $\mathbb{L}_2(G)$ are classes of functions like in the space $\mathbb{L}_2(a,b)$.
- If $G \subset \mathbb{R}^n$ is open then $\dim(\mathbb{L}_2(G)) = \infty$.
  If G is open, then there exists a cube

  $$\mathbf{C} = \{\mathbf{x} = (\xi_1, ..., \xi_n) \in \mathbb{R}^n \mid -\infty < a_k < \xi_k < b_k < \infty\}$$

  with $C \subset G$. We define

  $$\mathbf{f}_m(\mathbf{x}) = \begin{cases} \xi_1^m & \text{for} \quad \mathbf{x} = (\xi_1, \xi_2, ..., \xi_n) \in \mathbf{C} \\ 0 & \text{for} \quad \mathbf{x} \in \mathbb{R}^n \setminus \mathbf{G}. \end{cases} .$$

  This implies that $\mathbf{f}_m(\mathbf{x}) \in \mathbb{L}_2(G)$ and so the set $\{\mathbf{f}_0(\mathbf{x}), ..., \mathbf{f}_m(\mathbf{x})\}$ is independent in $\mathbb{L}_2(G)$ for any $m \in \mathbb{N}$ i.e. $\dim(\mathbb{L}_2(G)) = \infty$.

Remarks:

1. $\mathbf{C}(G)-$ space of all continuous functions which are defined on G:
   $\mathbf{C}(G) \subset \mathbb{L}_2(G)$.
2. $\mathbf{C}^m(G)-$ space of all functions defined on G which have a continuous partial derivative of order $k = 0, 1, ..., m$.
3. $\mathbf{C}_0(G)-$ space of all continuous **functions** φ defined on $\mathbb{R}^n$ **wich have a compact support in** G:
   i.e.
   $$\mathbf{C}_0(G) = \{\boldsymbol{\varphi} \in \mathbf{C}(G) \mid \boldsymbol{\varphi}(\mathbf{x}) = 0 \quad \text{for} \quad \forall\, \mathbf{x} \notin G\}.$$
4. $\mathbf{C}_0^m(G)-$ space of all functions $\varphi \in \mathbf{C}_0(G)$, which have continuous partial derivatives of order $k = 0, 1, ..., m$.
   $$\mathbf{C}_0^m(G) = \left\{ \boldsymbol{\varphi} \in \mathbf{C}_0(G) \;\middle|\; \frac{\partial^{\alpha_1 + ... + \alpha_n} \varphi(\mathbf{x})}{\partial x_1^{\alpha_1} ... \partial x_n^{\alpha_n}} \in \mathbf{C}_0(G) \quad \begin{matrix} \sum_{k=1}^{n} \alpha_k \le m \\ \mathbf{x} = (x_1, ..., x_n) \end{matrix} \right\}$$
5. $\mathbf{C}_0^\infty(G)-$ space of all functions $\varphi \in \mathbf{C}_0(G)$, which have continuous partial derivatives of any order, i.e. $\boldsymbol{\varphi} \in \mathbf{C}_0^\infty(G)$, if $\boldsymbol{\varphi} \in \mathbf{C}_0^m(G)$ for $m = 0, 1, ...$

3.3 Isometry of HILBERT spaces

Definition 3.7 *The HILBERT spaces $\mathbb{H}_1$, $\mathbb{H}_2$ are called isometric, if there exists a unique mapping $f : \mathbb{H}_1 \to \mathbb{H}_2$ with:*

- $f(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha f(\mathbf{u}) + \beta f(\mathbf{v})$
- $(\mathbf{u}, \mathbf{v}) = (f(\mathbf{u}), f(\mathbf{v})) \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{H}_1, \forall \alpha, \beta \in \mathbb{C}$

Conclusion 3.2 *This mapping is even a one-to-one mapping:*

Proof. Assumption:

$$f(\mathbf{u}_1) = \mathbf{x};\ f(\mathbf{u}_2) = \mathbf{x} \quad \wedge \quad \mathbf{u}_1 \neq \mathbf{u}_2$$

This implies

$$\begin{aligned} f(\mathbf{u}_1 - \mathbf{u}_2) &= f(\mathbf{u}_1) - f(\mathbf{u}_2) = \mathbf{0} \\ \|\mathbf{u}_1 - \mathbf{u}_2\|^2 &= \|f(\mathbf{u}_1) - f(\mathbf{u}_2)\|^2 = \|\mathbf{0}\| = 0 \\ \mathbf{u}_1 &= \mathbf{u}_2 \quad \text{contradiction} \end{aligned}$$

■

Notation 3.6 *In the application of this isometry, all charactristic properties of a normed space are kept. Thus such spaces are considered as the same, they are identical. (see for example the real numbers and the points on the number line or the complex numbers $\mathbb{C}$ and the vectors of $\mathbb{R}^2$).*

Theorem 3.6 *Any two separable infinite dimensional HILBERT spaces $\mathbb{H}_1$ and $\mathbb{H}_2$ are isometric.*

Proof. $\mathbf{u}, \mathbf{v} \in \mathbb{H}_1$, $\mathbb{H}_1$ is separable
$\curvearrowright \mathbf{u} = \sum_{i=1}^{\infty} u_i\mathbf{e}_i; \quad u_i = (\mathbf{u}, \mathbf{e}_i); \quad \mathbf{v} = \sum_{i=1}^{\infty} v_i\mathbf{e}_i; \quad v_i = (\mathbf{u}, \mathbf{e}_i);$
$\{\mathbf{e}_i\}_{i=1}^{\infty}$ is a complete ONS

$$\sum_{i=1}^{\infty} |u_i|^2 \leq \|\mathbf{u}\|^2 < \infty \quad \text{(Bessel's inequality)}$$

The Riesz-Fischer theorem says that then exists an element $\mathbf{x} \in \mathbb{H}_2$ with the Fourier coefficients u_i :

$$\mathbf{x} = \sum_{i=1}^{\infty} u_i\mathbf{g}_i; \quad \{\mathbf{g}_i\}_{i=1}^{\infty} \text{ is a complete ONS of } \mathbb{H}_2$$

Now we define a mapping

$$f(\mathbf{u}) = \mathbf{x} = \sum_{i=1}^{\infty} u_i \mathbf{g}_i \quad (*)$$

Further let be $f(\mathbf{v}) = \mathbf{y} = \sum_{i=1}^{\infty} v_i \mathbf{g}_i$
Mapping $(*)$ is an isometry: Property 1:

$$\begin{aligned}
f(\alpha\mathbf{u} + \beta\mathbf{v}) &= f\left(\alpha \sum_{i=1}^{\infty} u_i \mathbf{e}_i + \beta \sum_{i=1}^{\infty} v_i \mathbf{e}_i\right) \\
&= f\left(\alpha \lim_{n\to\infty} \sum_{i=1}^{n} u_i \mathbf{e}_i + \beta \lim_{n\to\infty} \sum_{i=1}^{n} v_i \mathbf{e}_i\right) \\
&= f\left(\lim_{n\to\infty} \sum_{i=1}^{n} (\alpha u_i + \beta v_i)\mathbf{e}_i\right) \\
&\overset{(*)}{=} \sum_{i=1}^{\infty} (\alpha u_i + \beta v_i)\mathbf{g}_i \\
&= \alpha \sum_{i=1}^{\infty} u_i \mathbf{g}_i + \beta \sum_{i=1}^{\infty} v_i \mathbf{g}_i \\
&= \alpha f(\mathbf{u}) + \beta f(\mathbf{v})
\end{aligned}$$

Property 2:

$$\begin{aligned}
(f(\mathbf{u}), f(\mathbf{v})) &= \left(\sum_{i=1}^{\infty} u_i \mathbf{g}_i, \sum_{j=1}^{\infty} v_j \mathbf{g}_j\right) = \sum_{i=1}^{\infty} u_i \overline{v}_i \\
&= \left(\sum_{i=1}^{\infty} u_i \mathbf{e}_i, \sum_{j=1}^{\infty} v_j \mathbf{e}_j\right) \\
&= (\mathbf{u}, \mathbf{v})
\end{aligned}$$

■

Notation 3.7 *The classification depends on the choice of the ONSs. For example,* $\mathbf{l}_2$ *and* $\mathbb{L}_2$ *are only two different implementations of the infinite dimensional HILBERT space.*

3.4 Orthogonality and Subspaces

Definition 3.8 *The proper subspaces* $M_1 \subset \mathbb{H}$ *and* $M_2 \subset \mathbb{H}$ *of a HILBERT space* $\mathbb{H}$ *are called* ***orthogonal*** *if and only if the inner product* $(\mathbf{u}, \mathbf{v}) = 0 \quad \forall \mathbf{u} \in M_1, \forall \mathbf{v} \in M_2$.

This means that all elements of M_1 are orthogonal to every element of M_2.

Definition 3.9 *For any subset M of $\mathbb{H}$ the set $M^{\perp} = \{\mathbf{v} \in \mathbb{H} \mid \quad (\mathbf{u}, \mathbf{v}) = 0 \quad \forall \mathbf{u} \in M\}$ is called the* ***orthogonal space*** *with respect to M.*

Theorem 3.7 *Let $M \subset \mathbb{H}$ be a (proper) subset of the HILBERT space $\mathbb{H}$. Then $M^{\perp}$ is a closed subspace of $\mathbb{H}$.*

Proof. 1) We show: $M^{\perp}$ is a linear space:
$\mathbf{u}_1, \mathbf{u}_2 \in M^{\perp}, \ \mathbf{v} \in M \quad \curvearrowright$

$$\begin{aligned} (\mathbf{u}_1, \mathbf{v}) &= 0 = (\mathbf{u}_2, \mathbf{v}) \\ (\alpha \mathbf{u}_1 + \beta \mathbf{u}_2, \mathbf{v}) &= \alpha(\mathbf{u}_1, \mathbf{v}) + \beta(\mathbf{u}_2, \mathbf{v}) = 0 \end{aligned}$$

$$\curvearrowright \quad \alpha \mathbf{u}_1 + \beta \mathbf{u}_2 \in M^{\perp}$$

2) We show: $M^{\perp}$ is closed.:
$\mathbf{u} \in (M^{\perp})^{+}$ implies

$$\begin{aligned} \exists \{\mathbf{u}_n\}_{n=1}^{\infty} \mid \mathbf{u}_n &\in M^{\perp} \wedge \lim_{n \to \infty} \mathbf{u}_n = \mathbf{u} \\ (\mathbf{u}_n, \mathbf{v}) &= 0 \quad \forall n, \quad \forall \mathbf{v} \in M \end{aligned}$$

Because of the continuity of the inner product we get

$$(\mathbf{u}, \mathbf{v}) = \lim_{n \to \infty} (\mathbf{u}_n, \mathbf{v}) = 0 \quad \curvearrowright \quad \mathbf{u} \in M^{\perp}$$

■

Conclusion 3.3 *If M is a subspace of $\mathbb{H}$ then the intersection $M \cap M^{\perp} = \{\mathbf{0}\}$.*

Conclusion 3.4 *If $M = \mathbb{H}$ then $M^{\perp} = \{\mathbf{0}\}$.*

Conclusion 3.5 *Let $M \subset \mathbb{H}$ be a (proper) subset of the HILBERT space $\mathbb{H}$, $\mathbf{u} \in \mathbb{H}$ and $\{\mathbf{u}\} \perp M$, then $\{\mathbf{u}\} \perp \overline{spanM}$.*

Theorem 3.8 ***PYTHAGOREAN theorem***
If $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n \in \mathbb{H}$ are pairwise orthogonal elements of the HILBERT space $\mathbb{H}$, i.e. $(\mathbf{u}_i, \mathbf{u}_j) = 0$ for $i \neq j$, then

$$\|\mathbf{u}_1 + \mathbf{u}_2 + ... + \mathbf{u}_n\|^2 = \|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + ... + \|\mathbf{u}_n\|^2$$

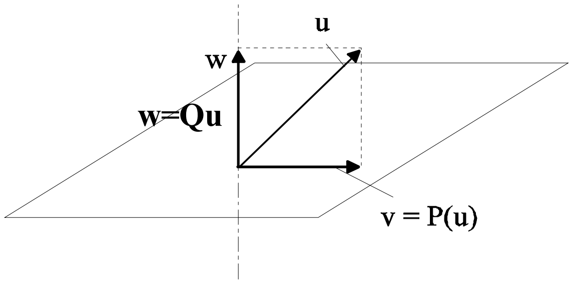
Definition 3.10 *Let $\mathbb{V}, \mathbb{W} \subset \mathbb{H}$ be closed subspaces of the HILBERT space $\mathbb{H}$. If every $\mathbf{u} \in \mathbb{H}$ is uniquely described as the sum $\mathbf{u} = \mathbf{v} + \mathbf{w}$ with $\mathbf{v} \in \mathbb{V}$ and $\mathbf{w} \in \mathbb{W}$ then $\mathbb{H}$ is called the* ***direct sum*** *of the subspaces $\mathbb{V}$ and $\mathbb{W}$: $\mathbb{H} = \mathbb{V} \oplus \mathbb{W}$.*

Definition 3.11 *$\mathbb{W} \subset \mathbb{H}$ is called an **orthogonal complement** of the closed subspace $\mathbb{V} \subset \mathbb{H}$ if and only if $\mathbb{W} \perp \mathbb{V} \quad \wedge \quad \mathbb{H} = \mathbb{V} \oplus \mathbb{W}$.*

Theorem 3.9 *The orthogonal complement $\mathbb{W}$ of a closed subspace $\mathbb{V} \subset \mathbb{H}$ of the HILBERT space $\mathbb{H}$ is unique.*

Theorem 3.10 *Let $\mathbb{W}$ be the orthogonal complement of $\mathbb{V}$, that means $\mathbb{W} \perp \mathbb{V} \quad \wedge \quad \mathbb{H} = \mathbb{V} \oplus \mathbb{W}$, $\mathbb{H}$: HILBERT space. Then every $\mathbf{u} \in \mathbb{H}$ may be decomposed uniquely into the sum $\mathbf{u} = \mathbf{v} + \mathbf{w}$ of an element $\mathbf{v} \in \mathbb{V}$ and an element $\mathbf{w} \in \mathbb{W}$ such that $(\mathbf{v}, \mathbf{w}) = 0$.*

*$\mathbf{v}$ is called the **orthogonal projection** of $\mathbf{u}$ onto $\mathbb{V}$ and $\mathbf{w}$ is called the **orthogonal projection** of $\mathbf{u}$ onto $\mathbb{W}$.*
*The mappings $\mathbf{P} : \mathbb{H} \to \mathbb{V}$ with $\mathbf{Pu} = \mathbf{v}$ and $\mathbf{Q} : \mathbb{H} \to \mathbb{W}$ with $\mathbf{Qu} = \mathbf{w}$ are called **orthogonal projectors** (orthoprojector) onto $\mathbb{V}$ or onto $\mathbb{W}$.*



$\mathbb{H} = \mathbb{R}^3, \quad \mathbb{V} = \mathbb{R}^2, \quad \mathbb{W} = \mathbb{R}$

Remark 3.11 *The orthogonal projector $\mathbf{P}$ is a **linear bounded operator** with $\|\mathbf{P}\| = 1$.*

Proof. $\mathbf{u}_i = \mathbf{v}_i + \mathbf{w}_i, \quad i = 1, 2$

$$\begin{aligned} \mathbf{P}(\alpha \mathbf{u}_1 + \beta \mathbf{u}_2) &= \mathbf{P}(\alpha(\mathbf{v}_1 + \mathbf{w}_1) + \beta(\mathbf{v}_2 + \mathbf{w}_2)) \\ &= \mathbf{P}((\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) + (\alpha \mathbf{w}_1 + \beta \mathbf{w}_2)) \\ &= \alpha \mathbf{v}_1 + \beta \mathbf{v}_2 = \alpha \mathbf{P}\mathbf{u}_1 + \beta \mathbf{P}\mathbf{u}_2; \quad \forall \mathbf{u}_1, \mathbf{u}_2 \in \mathbb{H} \end{aligned}$$

$$\|\mathbf{Pu}\|^2 = \|\mathbf{v}\|^2 \le \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 = \|\mathbf{u}\|^2 \quad \forall \mathbf{u} = \mathbf{v} + \mathbf{w} \in \mathbb{H}$$

On the other hand $\|\mathbf{Pu}\| = \|\mathbf{u}\|$ is satisfied for every $\mathbf{u} = \mathbf{v} \in \mathbb{V}$, and hence $\|\mathbf{P}\| = 1$.
■

Remark 3.12 $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ *($\mathbf{I}-$ identity operator)*

BEST APPROXIMATION IN HILBERT SPACES

Theorem 3.13 *Let $\mathbb{U}$ be a closed subspace of the HILBERT space $\mathbb{H}$ and $\mathbf{u}$ be an arbitrary element of $\mathbb{H}$.*
There exists a unique element $\mathbf{u}_0 \in \mathbb{U}$ with

$$\begin{aligned} a)\ \|\mathbf{u}-\mathbf{u}_0\| &= \min_{\mathbf{v}\,\in\,\mathbf{U}} \|\mathbf{u}-\mathbf{v}\| \ \text{ and} \\ b)\ (\mathbf{u}-\mathbf{u}_0,\mathbf{v}) &= 0 \qquad \text{for} \quad \forall \mathbf{v} \in \mathbb{U}, \quad \text{i.e. } \mathbf{u}-\mathbf{u}_0 \in \mathbb{U}^{\perp}. \end{aligned}$$

$\mathbf{u}_0$ *is called the* ***best approximation of*** $\mathbf{u} \in \mathbb{H}$ *with respect to the subspace* $\mathbb{U}$*.*
Proof: see literature (Kantorowitsch/Akilow)

Meaning: $\mathbf{u}_0$ minimises the distance between $\mathbf{u}$ and the elements in $\mathbb{U}$. The proof consists in showing that every minimising sequence $\{\mathbf{u}_n\} \subset \mathbb{U} \setminus \{\mathbf{u}_0\}$ is Cauchy, and hence converges to a point in $\mathbb{U} \setminus \{\mathbf{u}_0\}$ that has minimal norm.

Remark 3.14 *The minimisation problems* $\min_{\mathbf{v}\,\in\,\mathbf{U}} \|\mathbf{u}-\mathbf{v}\|$ *and* $\min_{\mathbf{v}\,\in\,\mathbf{U}} \mathbf{F}(\mathbf{v})$ *with the quadratic functional*

$$\mathbf{F}(\mathbf{v}) = \frac{1}{2}(\mathbf{v},\mathbf{v}) - (\mathbf{u},\mathbf{v})$$

are equivalent in HILBERT spaces over real fields because of:

$$\|\mathbf{u}-\mathbf{v}\|^2 = (\mathbf{u},\mathbf{u}) - 2(\mathbf{u},\mathbf{v}) + (\mathbf{v},\mathbf{v}) = 2\left[\frac{1}{2}(\mathbf{v},\mathbf{v}) - (\mathbf{u},\mathbf{v})\right] + \|\mathbf{u}\|^2 = 2\mathbf{F}(\mathbf{v}) + \|\mathbf{u}\|^2 .$$

Theorem 3.15 *Let $\mathbb{U}_n$ be an $n-$dimensional subspace of the HILBERT-space $\mathbb{H}$ and $\mathbf{e}_1, ..., \mathbf{e}_n$ be a basis in $\mathbb{U}_n$. The best approximation $\mathbf{u}_0 \in \mathbb{U}_n$ of an arbitrary element $\mathbf{u} \in \mathbb{H}$ can be written as a linear combination of the basis elements such that*

$$\mathbf{u}_0 = \sum_{k=1}^{n} \alpha_k \mathbf{e}_k .$$

The coefficients α_k are the unique solution of the system of linear equations

$$(\mathbf{u}-\mathbf{u}_0,\mathbf{e}_i) = (\mathbf{u},\mathbf{e}_i) - \sum_{k=1}^{n} \alpha_k (\mathbf{e}_k,\mathbf{e}_i) = 0, \qquad i = 1, ..., n.$$

Proof. Using the theorem above it is sufficient to remark that the condition

$(\mathbf{u}-\mathbf{u}_0,\mathbf{v}) = 0$ for every $\mathbf{v} \in \mathbb{U}_n$ is satisfied if and only if this condition is satisfied for all elements of the basis, i.e. if and only if $(\mathbf{u}-\mathbf{u}_0,\mathbf{e}_i) = 0$ $(i = 1, ..., n)$. The system of linear equations arises by substitution of the ansatz $\mathbf{u}_0 = \sum_{k=1}^{n} \alpha_k \mathbf{e}_k$ into the condition $(\mathbf{u}-\mathbf{u}_0,\mathbf{v}) = 0 \ \forall\ \mathbf{v} \in \mathbb{U}_n$. ■

Notation 3.8 *This theorem is the foundation of the numerical solution of maximum and minimum problems (quadratic variational problems).*

Example 3.13 $\mathbb{H}=\mathbb{L}_2[0,1]$; *Let* $\mathbb{U}_n \subset \mathbb{H}$ *be the subspace with the basis functions* $\boldsymbol{\varphi}_k(t) = t^k \quad (k = 0, 1, ..., n)$. *Hence* $dim\mathbb{U}_n = n + 1$. *Let* $\mathbf{f}(t)$ *be an element of* $\mathbb{L}_2[0,1]$. *The ansatz of the best approximation* $\mathbf{f}_n(t)$ *is then*

$$\mathbf{f}_n(t) = \sum_{k=0}^{n} \alpha_k \boldsymbol{\varphi}_k(t) .$$

↷

$$\begin{aligned}
(\mathbf{u} - \mathbf{u}_0, \mathbf{e}_i) &= (\mathbf{u}, \mathbf{e}_i) - \sum_{k=0}^{n} \alpha_k (\mathbf{e}_k, \mathbf{e}_i) = 0 \qquad i = 0, 1, ..., n \\
&\iff \\
\sum_{k=0}^{n} \alpha_k (\boldsymbol{\varphi}_k, \boldsymbol{\varphi}_i) &= (\mathbf{f}, \boldsymbol{\varphi}_i) \qquad i = 0, 1, ..., n
\end{aligned}$$

$$\begin{pmatrix} (\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_0) & \cdots & (\boldsymbol{\varphi}_n, \boldsymbol{\varphi}_0) \\ \vdots & & \vdots \\ (\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_n) & \cdots & \boldsymbol{\varphi}_n, \boldsymbol{\varphi}_n \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} (\mathbf{f}, \boldsymbol{\varphi}_0) \\ \vdots \\ (\mathbf{f}, \boldsymbol{\varphi}_n) \end{pmatrix} \qquad (*)$$

with

$$\begin{aligned}
(\boldsymbol{\varphi}_k, \boldsymbol{\varphi}_i) &= \int_0^1 \boldsymbol{\varphi}_k(t)\overline{\boldsymbol{\varphi}_i(t)}dt = \int_0^1 \mathbf{t}^{k+i} dt = \frac{1}{k+i+1} \\
(\mathbf{f}, \boldsymbol{\varphi}_i) &= \int_0^1 \mathbf{f}(t)\overline{\boldsymbol{\varphi}_i(t)}dt = \int_0^1 f(t)\mathbf{t}^i dt.
\end{aligned}$$

$(*)$ *is the system of linear equations for the continuous approximation of* $f(t)$ *in the quadratic mean (method of least squares, see numerics). We get*

$$\lim_{n\to\infty} \int_0^1 (\mathbf{f}_n(t) - \mathbf{f}(t))^2 dt = 0\text{: convergence in } \mathbb{L}_2[0,1]$$

Notation 3.9 *If* $\mathbf{e}_1, ..., \mathbf{e}_n$ *is an orthogonal basis system in* $\mathbb{U}_n$, *i.e.* $(\mathbf{e}_k, \mathbf{e}_i) = \delta_{ki} E_i$, *then* $\alpha_i = (\mathbf{u}, \mathbf{e}_i)/E_i$ *and therefore*

$$\mathbf{u}_0 = \sum_{k=1}^{n} \frac{(\mathbf{u}, \mathbf{e}_k)}{E_k} \mathbf{e}_k.$$

This is the beginning of a FOURIER series, i.e. the FOURIER series leads us to the best approximation.

Example 3.14 $\mathbb{H} = \mathbb{L}_2[0, 2\pi]$ *and* $\mathbf{f}(t) \in \mathbb{H}$ *is given.*

$\mathbb{U}_n = span\left\{\frac{1}{\sqrt{2\pi}}\ ,\ \frac{1}{\sqrt{\pi}}\cos(kt)\ ,\ \frac{1}{\sqrt{\pi}}\sin(kt) \qquad k = 1, 2, ...n\right\}$

$$\begin{aligned}
\mathbf{f}_n(t) &= \frac{\alpha_0}{\sqrt{2\pi}} + \sum_{k=1}^{n}\left[\alpha_k \frac{\cos(kt)}{\sqrt{\pi}} + \beta_k \frac{\sin(kt)}{\sqrt{\pi}}\right] \\
\alpha_0 &= \left(\mathbf{f}, \frac{1}{\sqrt{2\pi}}\right) = \int_0^{2\pi} \mathbf{f}(t)\frac{1}{\sqrt{2\pi}}dt \\
\alpha_k &= \left(\mathbf{f}, \frac{\cos(kt)}{\sqrt{\pi}}\right) = \frac{1}{\sqrt{\pi}}\int_0^{2\pi} \mathbf{f}(t)\cos(kt)dt \\
\beta_k &= \left(\mathbf{f}, \frac{\sin(kt)}{\sqrt{\pi}}\right) = \frac{1}{\sqrt{\pi}}\int_0^{2\pi} \mathbf{f}(t)\sin(kt)dt;\ \ k = 1, 2, ..., n
\end{aligned}$$

In general the factors are hidden in the coefficients $\alpha_k,\ \beta_k$ and we get the usual description of a Fourier series such that

$$\lim_{n\to\infty}\int_0^{2\pi}(\mathbf{f}_n(t) - \mathbf{f}(t))^2 dt = 0\text{: convergence in } \mathbb{L}_2[0, 2\pi]$$

3.5 Linear Operators in HILBERT Spaces

3.5.1 Adjoint, symmetric and monotonic Operators

If we are interested in the solutions of linear time independent operator equations $Au = f$ then we have to study the properties of the operators. Examples of such problems are

- the calculation of static electrical fields,
- research involving the balance of forces in mechanical systems,
- all problems which lead us to systems of linear equations.

We want to compose classes of problems whose solutions have similar properties. If the operator is a matrix then the properties of the matrix determine the properties of the solution. For example you know what it means when a matrix is symmetric or positve definite. Therefore we must generalise such matrix properties to abstract operators in the HILBERT space.

Definition 3.12 *Given a linear operator in the HILBERT space* $\mathbb{H}$ *with the domain* $D(\mathbf{A}) \subseteq \mathbb{H}$ *and the range* $\mathbb{H}$*;* $\overline{D(\mathbf{A})} = \mathbb{H}$*. The set*

$$D(\mathbf{A}^*) = \{\mathbf{x} \in \mathbb{H} \mid \quad \exists \mathbf{y} \in \mathbb{H} \text{ with } (\mathbf{Au}, \mathbf{x}) = (\mathbf{u}, \mathbf{y}) \quad \forall \mathbf{u} \in D(\mathbf{A})\}$$

is a subset of $\mathbb{H}$*. Then the operator* $\mathbf{A}^* : D(\mathbf{A}^*) \subseteq \mathbb{H} \to \mathbb{H}$ *with* $\mathbf{A}^*\mathbf{x} = \mathbf{y}$ *is called the* ***adjoint operator*** *(or Hermitian conjugate) of* $\mathbf{A}$*. Thus*

$$(\mathbf{Au}, \mathbf{x}) = (\mathbf{u}, \mathbf{A}^*\mathbf{x}) \ \mathit{for} \ \forall \mathbf{u} \in D(\mathbf{A}), \forall \mathbf{x} \in D(\mathbf{A}^*).$$

Thus the adjoint operator is the generalisation of conjugate transposes of square matrices.

Definition 3.13 *Let* $\mathbb{H}$ *be a HILBERT space. Given the linear operator* $\mathbf{A} : D(\mathbf{A}) \subseteq \mathbb{H} \to \mathbb{H}$ *with* $\overline{D(\mathbf{A})} = \mathbb{H}$*.* $\mathbf{A}$ *is called*

- **symmetric** $\iff (\mathbf{Au}, \mathbf{x}) = (\mathbf{u}, \mathbf{Ax}) \quad \forall \mathbf{u}, \mathbf{x} \in D(\mathbf{A})$,
  i.e. $\mathbf{A}^*\mathbf{x} = \mathbf{Ax} \quad \forall \mathbf{x} \in D(\mathbf{A}) \quad \wedge \quad D(\mathbf{A}) \subseteq D(\mathbf{A}^*)$
- **selfadjoint** $\iff \mathbf{A} = \mathbf{A}^*$, i.e. $\mathbf{A}^*\mathbf{x} = \mathbf{Ax} \quad \forall \mathbf{x} \in D(\mathbf{A}) = D(\mathbf{A}^*)$
- **skew symmetric** $\iff (\mathbf{Au}, \mathbf{x}) = -(\mathbf{u}, \mathbf{Ax}) \quad \forall \mathbf{u}, \mathbf{x} \in D(\mathbf{A})$
- **skew adjoint** $\iff \mathbf{A} = -\mathbf{A}^*$
- **compact** $\iff \forall \{\mathbf{x}_n\} \subset D(\mathbf{A}) \mid \|\mathbf{x}_n\| < M < \infty \ \forall n \quad \exists \{Ax_{\tilde{n}}\} \subset \{Ax_n\}_{n=1}^{\infty} \mid \lim_{\tilde{n}\to\infty} Ax_{\tilde{n}} = \tilde{\mathbf{x}} \in image(A)$

Example 3.15 $\mathbb{H} = \mathbb{R}^n; \quad \mathbf{A} = (a_{ij})_{i,j=1}^n; \quad \mathbf{Au} = \left(\sum_{j=1}^n a_{ij}u_j\right)_{i=1}^n; \quad \overline{D(\mathbf{A})} = \mathbb{H}$

$$\begin{aligned} (\mathbf{Au}, \mathbf{x}) &= \left(\sum_{j=1}^n a_{ij}u_j, \mathbf{x}\right) = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij}u_j\right) \cdot x_i \\ &= \sum_{j=1}^n \left(\sum_{i=1}^n a_{ji}u_i\right) \cdot x_j \qquad (i \longleftrightarrow j) \\ &= \sum_{i=1}^n u_i \left(\sum_{j=1}^n a_{ji}x_j\right) = (\mathbf{u}, \mathbf{A}^*\mathbf{x}) \\ \mathbf{A}^* &= (a_{ij}^*)_{i,j=1}^n = (a_{ji})_{i,j=1}^n \end{aligned}$$

In case of real matrices the operator $\mathbf{A}^*$ *is the matrix* $\mathbf{A}^T$*. In case of complex matrices the operator* $\mathbf{A}^*$ *is the conjugate transposed matrix* $\mathbf{A}^*$*.*

Example 3.16 $\mathbb{H} = \mathbb{L}_2[a,b]$; $(\mathbf{Au})(t) = \int_a^b K(t,s)\mathbf{u}(s)ds$; $D(\mathbf{A}) = \mathbb{H}$; $\mathbb{K} = \mathbb{R}(= \mathbb{C})$

$$
\begin{aligned}
(\mathbf{Au},\mathbf{x}) &= \int_a^b \left[\int_a^b K(t,s)\mathbf{u}(s)ds\right]\cdot \mathbf{x}(t)dt \stackrel{!}{=} \int_a^b \mathbf{u}(s)\mathbf{y}(s)ds \\
&\Longleftrightarrow \\
\int_a^b \left[\int_a^b K(t,s)\mathbf{x}(t)dt\right]\cdot \mathbf{u}(s)ds &= \int_a^b \mathbf{u}(s)\mathbf{y}(s)ds \\
&\Longleftrightarrow \\
\mathbf{y}(s) &= \int_a^b K(t,s)\mathbf{x}(t)dt = \mathbf{A}^* x; \\
\mathbf{A}(\mathbf{u}(\mathbf{s})) &= \int_a^b K(t,s)\mathbf{u}(s)ds
\end{aligned}
$$

$\curvearrowright$

$$
\begin{aligned}
\mathbf{A}^*(\mathbf{u}(\mathbf{s})) &= \int_a^b K(s,t)\mathbf{u}(s)ds \\
\mathbf{A} &= \mathbf{A}^* \Longleftrightarrow K(s,t) = K(t,s)
\end{aligned}
$$

In the case of complex functions and a complex field we get $K(s,t) = \overline{K(t,s)}$
Properties of the adjoint operator:

1. $\mathbf{A}^*$ is linear.
2. $(\alpha\mathbf{A})^* = \overline{\alpha}\mathbf{A}^*$
3. Let $\mathbf{T} : \mathbb{H} \to \mathbb{H}$ be a linear bounded operator. $\Longrightarrow$ $\exists$ a unique operator $\mathbf{T}^*$, which is linear and bounded, too, with $\|\mathbf{T}\| = \|\mathbf{T}^*\|$.
4. $(\mathbf{A}^*)^* = \mathbf{A}$, because of:

$$
\begin{aligned}
(\mathbf{A}^*\mathbf{u},\mathbf{x}) &= (\mathbf{u},(\mathbf{A}^*)^*\mathbf{x}) \\
&= \overline{(\mathbf{x},\mathbf{A}^*\mathbf{u})} = \overline{(\mathbf{Ax},\mathbf{u})} = (\mathbf{u},\mathbf{Ax})
\end{aligned}
$$

Definition 3.14 *The linear operator* $\mathbf{A} : D(\mathbf{A}) \subseteq \mathbb{H} \to \mathbb{H}$ *is called* ***positive definite****, if and only if:* $(\mathbf{Au},\mathbf{u}) \geq C\,\|\mathbf{u}\|^2$ *for* $\forall \mathbf{u} \in D(\mathbf{A})$, $C \in \mathbb{R}$, $C > 0$

3.5.2 Eigenvalues of Operators

Definition 3.15 *The complex number* $\lambda \in \mathbb{C}$ *is called the* ***eigenvalue of the operator*** $\mathbf{A} : \mathbb{H} \to \mathbb{H}$*, if there exists an element* $\mathbf{x} \in \mathbb{H}$, $\mathbf{x} \neq \mathbf{0}$*, such that* $\mathbf{Ax} = \lambda\mathbf{x}$*. Every* $\mathbf{x} \neq \mathbf{0}$ *with* $\mathbf{Ax} = \lambda\mathbf{x}$ *is called an* ***eigenelement*** *of the eigenvalue* λ*.*

Properties :

1. $\mathbf{A} : \mathbb{H} \to \mathbb{H}; \quad \mathbf{A} = \mathbf{A}^*; \quad \curvearrowright \quad (\mathbf{Ax}, \mathbf{x}) \in \mathbb{R} \quad \forall \mathbf{x}$
   Proof:
   $$(\mathbf{Ax}, \mathbf{x}) = (\mathbf{x}, \mathbf{Ax}) = \overline{(\mathbf{Ax}, \mathbf{x})}$$

2. $\mathbf{A} : \mathbb{H} \to \mathbb{H}; \quad \mathbf{A} = \mathbf{A}^*; \quad \curvearrowright$ All eigenvalues of $\mathbf{A}$ are real numbers.
   Proof: $\mathbf{Ax} = \lambda \mathbf{x} \quad \curvearrowright$
   $$\begin{aligned} (\mathbf{Ax}, \mathbf{x}) &= (\lambda \mathbf{x}, \mathbf{x}) \\ \lambda &= \frac{(\mathbf{Ax}, \mathbf{x})}{\|\mathbf{x}\|^2} \overset{1.}{\in} \mathbb{R} \end{aligned}$$

3. $\mathbf{A} : \mathbb{H} \to \mathbb{H}; \quad \mathbf{A} = \mathbf{A}^*; \quad \curvearrowright$ Eigenelements which belong to different eigenvalues are orthogonal.
   Proof: $\mathbf{Ax}_1 = \lambda_1 \mathbf{x}_1;\ \mathbf{Ax}_2 = \lambda_2 \mathbf{x}_2;\ \lambda_1, \lambda_2 \in \mathbb{R};\ \lambda_1 \neq \lambda_2$
   $\curvearrowright$
   $$\begin{aligned} (\mathbf{Ax}_1, \mathbf{x}_2) &= \lambda_1 (\mathbf{x}_1, \mathbf{x}_2) \\ (\mathbf{Ax}_1, \mathbf{x}_2) &= (\mathbf{x}_1, \mathbf{Ax}_2) = \lambda_2 (\mathbf{x}_1, \mathbf{x}_2) \quad \text{because of } A = A^* \end{aligned}$$
   Therefore we get
   $$\begin{aligned} \lambda_1 (\mathbf{x}_1, \mathbf{x}_2) &= \lambda_2 (\mathbf{x}_1, \mathbf{x}_2) \\ (\lambda_1 - \lambda_2)(\mathbf{x}_1, \mathbf{x}_2) &= 0 \end{aligned}$$
   Because of $\lambda_1 \neq \lambda_2$ we get $(\mathbf{x}_1, \mathbf{x}_2) = 0$.

4. $\mathbf{A} : \mathbb{H} \to \mathbb{H}; \quad \mathbf{A} = \mathbf{A}^*; \quad \curvearrowright \quad \|\mathbf{A}\| = \sup_{\|\mathbf{x}\|=1} |(\mathbf{Ax}, \mathbf{x})|$ (without proof)

Example 3.17 $\mathbb{H} = \mathbb{C}^n; \quad T : \mathbb{H} \to \mathbb{H}; \quad \mathbf{T} = (t_{ij})_{i,j=1}^n;$

$$\begin{aligned} \mathbf{Tx} &= \mathbf{y} \textit{ with } y_k = \sum_{j=1}^n t_{kj} x_j; \quad k = 1, 2, ..., n \\ \mathbf{Tx} &= \lambda \mathbf{x} \iff (\mathbf{T} - \boldsymbol{\lambda} \mathbf{I})\mathbf{x} = \mathbf{0} \\ \mathbf{x} &\neq \mathbf{0} \iff \det(\mathbf{T} - \boldsymbol{\lambda} \mathbf{I}) = 0, \end{aligned}$$

$\mathbf{I} = (\delta_{ij})_{i,j=1}^n$ *: identity operator; The eigenvalues of the matrix* $(t_{ij})_{i,j=1}^n$ *are the eigenvalues of the operator* $\mathbf{T}$.

Example 3.18 *Consider the differential operator* $\mathbf{D} = \frac{d^2}{dt^2}$; $\mathbf{x}(t)$ *is at least twice continuously differentiable and*

$\mathbf{x}(0) = \mathbf{x}(\pi) = 0; \quad \Longrightarrow$ *That means* $\mathbf{x}(t) \in \overset{o}{\mathbf{C}}{}^2[0,\pi]$. *But this space is only a BANACH space, not a HILBERT space!!*

$$\mathbf{Dx} = \frac{d^2}{dt^2}\mathbf{x}(\mathbf{t}) \overset{!}{=} \lambda\mathbf{x}(t) \quad \wedge \quad \mathbf{x}(0) = \mathbf{x}(\pi) = 0$$

The eigenvalue problem corresponds to a boundary value problem of an ordinary differential equation of second order with constant coefficients. With the ansatz $\mathbf{x}(t) = e^{\alpha t}$ *we get the characteristic equation*

$$\alpha^2 - \lambda = 0 \quad \curvearrowright \quad \alpha_{1/2} = \pm\sqrt{\lambda}$$

therefore we get the general solutions:
Case 1: $\lambda > 0$:

$$\begin{aligned}
\mathbf{x}(t) &= C_1 e^{\sqrt{\lambda}t} + C_2 e^{-\sqrt{\lambda}t} \quad \wedge \quad \mathbf{x}(0) = \mathbf{x}(\pi) = 0 \\
\mathbf{x}(0) &= C_1 + C_2 \overset{!}{=} 0 \quad \curvearrowright \quad C_2 = -C_1 \\
\mathbf{x}(\pi) &= C_1 e^{\sqrt{\lambda}l} + C_2 e^{-\sqrt{\lambda}l} \\
&= C_1 e^{\sqrt{\lambda}l} - C_1 e^{-\sqrt{\lambda}l} \overset{!}{=} 0 \Longleftrightarrow C_1 = 0 \Longleftrightarrow \mathbf{x}(t) = \mathbf{0}
\end{aligned}$$

Case 2: $\lambda = 0$:

$$\begin{aligned}
\mathbf{x}(t) &= C_1 + C_2 t \quad \wedge \quad \mathbf{x}(0) = \mathbf{x}(\pi) = 0 \\
\mathbf{x}(0) &= C_1 \overset{!}{=} 0 \\
\mathbf{x}(\pi) &= C_2 \cdot \pi \overset{!}{=} 0 \Longleftrightarrow C_2 = 0 \Longleftrightarrow \mathbf{x}(t) = \mathbf{0}
\end{aligned}$$

Case 3: $\lambda < 0$:

$$\begin{aligned}
\mathbf{x}(t) &= C_1 \cos(\sqrt{-\lambda}t) + C_2 \sin(\sqrt{-\lambda}t) \quad \wedge \quad \mathbf{x}(0) = \mathbf{x}(\pi) = 0 \\
\mathbf{x}(0) &= C_1 \cos 0 = C_1 \overset{!}{=} 0 \\
\mathbf{x}(\pi) &= C_2 \sin(\sqrt{-\lambda}\pi) \overset{!}{=} 0 \Longleftrightarrow C_2 = 0 \quad \vee \quad \sin(\sqrt{-\lambda}\pi) = 0
\end{aligned}$$

$C_2 = 0$ *is not constructive. Therefore*

$$\begin{aligned}
\sin(\sqrt{-\lambda}\pi) &= 0 \Longleftrightarrow \sqrt{-\lambda}\pi = k\pi \\
&\Longleftrightarrow \lambda_k = -k^2 \\
&\Longleftrightarrow \mathbf{x}_k(t) = C_k \sin(kt) \quad k \in \mathbb{Z}.
\end{aligned}$$

Only for the eigenvalues $\lambda_k = -k^2$ *we get nontrivial solutions.*

That means $\mathbf{x}(t) \in$. *But* $\overset{o}{\mathbf{C}}{}^2[0,\pi]$ *is only a BANACH space, not a HILBERT space and we do not know an inner product in this space. That's why we can not use the Fourier series.*

Theorem 3.16 *Let* $\mathbf{A} : D(\mathbf{A}) \subseteq \mathbb{H} \to \mathbb{H}$ *be a self adjoint, positive definite operator with a countable system of eigensolutions* $\{\lambda_k, \mathbf{u}_k\}_{k=1}^{\infty}$, *such that the eigenelements* $\{\mathbf{u}_k\}_{k=1}^{\infty}$ *are a complete ONS in* $\mathbb{H}$. *Then*

1.
$$\begin{aligned} \mathbf{Au} &= \sum_{k=1}^{\infty} \lambda_k (\mathbf{u}, \mathbf{u}_k) \mathbf{u}_k \quad for \\ \forall \mathbf{u} &\in D(A) = \{\mathbf{u} \in \mathbb{H} \mid \sum_{k=1}^{\infty} |\lambda_k (\mathbf{u}, \mathbf{u}_k)|^2 < \infty\} \end{aligned}$$

2. $\exists\ \mathbf{A}^{-1} : \mathbb{H} \to \mathbb{H}; \quad \mathbf{A}^{-1}$ *is linear and bounded.*
$$\begin{aligned} \mathbf{A}^{-1}\mathbf{f} &= \sum_{k=1}^{\infty} \frac{1}{\lambda_k} (\mathbf{f}, \mathbf{u}_k) \mathbf{u}_k \quad \forall \mathbf{f} \in H \\ \left\|\mathbf{A}^{-1}\mathbf{f}\right\| &\leq \frac{1}{C} \|\mathbf{f}\| \quad with \quad 0 < C \leq \lambda_1 \leq \lambda_2 \leq ... \end{aligned}$$

3. *The solution of the operator equation* $\mathbf{Au} = \mathbf{f},\ \mathbf{f} \in \mathbb{H}$ *can be written as*
$$\mathbf{u} = \mathbf{A}^{-1}\mathbf{f} = \sum_{k=1}^{\infty} \frac{1}{\lambda_k} (\mathbf{f}, \mathbf{u}_k) \mathbf{u}_k.$$

(without proof)

Notation 3.10 *To prove and use this theorem a HILBERT space is absolutely necessary! If you don't have a HILBERT space then look for an embedding of your problem in a HILBERT space.*

Notation 3.11 *The following conditions are equivalent:*

1. $\mathbf{u} \in D(\mathbf{A})$
2. $\sum_{k=1}^{\infty} \lambda_k (\mathbf{u}, \mathbf{u}_k) \mathbf{u}_k$ is a convergent series in $\mathbb{H}$.
3. $\sum_{k=1}^{\infty} |\lambda_k (\mathbf{u}, \mathbf{u}_k)|^2$ is a convergent number series.

3.5.3 Linear Functionals

Definition 3.16 *Operators in a HILBERT space* $\mathbb{H}$ *over the field* $\mathbb{K}$ *(*$\mathbb{R}$ *or* $\mathbb{C}$*) which map from* $\mathbb{H}$ *to* $\mathbb{K}$ *are called functionals or linear forms:* $\mathbf{f}:\mathbb{H}\to\mathbb{K}$.

Definition 3.17 *The functional* $\mathbf{f}:\mathbb{H}\to\mathbb{K}$ *is called:*

- **linear**, *if*: $\mathbf{f}(\alpha\mathbf{u}+\beta\mathbf{v})=\alpha f(\mathbf{u})+\beta f(\mathbf{v}) \quad \forall\mathbf{u},\mathbf{v}\in\mathbb{H};\ \forall\alpha,\beta\in\mathbb{K}$
- **bounded**, *if:* $\exists M\in\mathbb{R},\ M>0 \quad | \quad |f(\mathbf{u})|\leq M\,\|\mathbf{u}\| \quad \forall\mathbf{u}\in\mathbb{H}$
- **continuous**, *if:* $\lim_{n\to\infty}\mathbf{u}_n=\mathbf{u}$ implies $\lim_{n\to\infty}\mathbf{f}(\mathbf{u}_n)=\mathbf{f}(\mathbf{u})$; $\mathbf{u}_n,\mathbf{u}\in\mathbb{H}$.

Definition 3.18 *Given a linear bounded functional* $\mathbf{f}:\mathbb{H}\to\mathbb{K}$. *The number*

$$\begin{aligned}\|\mathbf{f}\| &= \sup_{\|\mathbf{u}\|=1}|\mathbf{f}(\mathbf{u})|, \quad \mathbf{u}\in\mathbb{H}\\ &= \inf\{M\in\mathbb{R}\mid\ |\mathbf{f}(\mathbf{u})|\leq M\,\|\mathbf{u}\| \quad \forall\mathbf{u}\in\mathbb{H}\}\end{aligned}$$

is called the ***norm of the functional****.*

Analogous to the operator theory we can prove the following theorem (homework):

Theorem 3.17 *A linear functional in the HILBERT space* $\mathbb{H}$ *is continuous if and only if it is bounded.*

Theorem 3.18 *Let* $\mathbb{H}$ *be a HILBERT space* $\mathbb{H}$ *over the field* $\mathbb{K}$ *and* $\mathbf{u}_0$ *be any fixed element of* $\mathbb{H}$. *Then* $\mathbf{f}(\mathbf{x})=(\mathbf{x},\mathbf{u}_0)$ *defines a linear functional in* $\mathbb{H}$.

Proof.

$$\begin{aligned}\mathbf{f}(\alpha\mathbf{x}+\beta\mathbf{y}) &= (\alpha\mathbf{x}+\beta\mathbf{y},\mathbf{u}_0)\\ &= \alpha(\mathbf{x},\mathbf{u}_0)+\beta(\mathbf{y},\mathbf{u}_0)\\ &= \alpha\mathbf{f}(\mathbf{x})+\boldsymbol{\beta}\mathbf{f}(\mathbf{y})\end{aligned}$$

■

Notation 3.12 *A in that wayl defined linear functional is continuous:*

1) $|\mathbf{f}(\mathbf{x})| = |(\mathbf{x},\mathbf{u}_0)|\leq\|\mathbf{x}\|\,\|\mathbf{u}_0\| \quad \curvearrowright \quad \frac{|\mathbf{f}(\mathbf{x})|}{\|\mathbf{x}\|}\leq\|\mathbf{u}_0\|$

2) $\dfrac{|\mathbf{f}(\mathbf{x})|}{\|\mathbf{x}\|} \leq M=\inf\{C\in\mathbb{R}| \quad |f(x)|\leq C\,\|\mathbf{x}\|\}=\|\mathbf{f}\| \quad \curvearrowright \quad \|\mathbf{f}\|\leq\|\mathbf{u}_0\|$

3) $|\mathbf{f}(\mathbf{u}_0)| = |(\mathbf{u}_0,\mathbf{u}_0)|=\|\mathbf{u}_0\|^2$ *implies* $\frac{|\mathbf{f}(\mathbf{u}_0)|}{\|\mathbf{u}_0\|}=\|\mathbf{u}_0\|$

Therefore we get $\|\mathbf{f}\|=\|\mathbf{u}_0\|$. *This means* $\mathbf{f}$ *is bounded.* $\curvearrowright$ $\mathbf{f}$ *is continuous.*

Theorem 3.19 ***The HILBERT space representation theorem (RIESZ)***
Let $\mathbb{H}$ be a HILBERT space $\mathbb{H}$ over the field $\mathbb{K}$ with the inner product $(.,.)$ and let $\mathbf{f}(\mathbf{x})$ be a continuous linear functional in $\mathbb{H}$. Then there exists a fixed unique element $\mathbf{u}_0 \in \mathbb{H}$ such that $\mathbf{f}(\mathbf{x}) = (\mathbf{x}, \mathbf{u}_0) \quad \forall \mathbf{x} \in \mathbb{H}$ and $\|\mathbf{f}\| = \|\mathbf{u}_0\|$.(without proof)

Example 3.19 *a)* $\mathbb{H} = \mathbb{R}^n$ *with* $(\mathbf{x}, \mathbf{y}) = \sum\limits_{k=1}^{n} x_k y_k; \quad \mathbf{f} : \mathbb{R}^\mathbf{n} \to \mathbb{R}$

$$\mathbf{f}(\mathbf{x}) = (\mathbf{x}, \mathbf{a}) = \sum_{k=1}^{n} x_k a_k; \quad \mathbf{a} \in \mathbb{R}^n; \ \mathbf{a} \textit{ is fixed}$$

b) $\mathbb{H} = \mathbf{l}_2$ *with* $(\mathbf{x}, \mathbf{y}) = \sum\limits_{k=1}^{\infty} x_k \overline{y}_k; \quad \mathbf{f} : \mathbf{l_2} \to \mathbb{C}$

$$\mathbf{f}(\mathbf{x}) = (\mathbf{x}, \mathbf{b}) = \sum_{k=1}^{\infty} x_k \bar{b}_k; \quad \mathbf{b} = \{b_1, b_2, ...\}; \ \mathbf{b} \textit{ is fixed}$$

c) $\mathbb{H} = \mathbb{L}_2[a,b]$ *with* $(\mathbf{x}, \mathbf{y}) = \int_a^b \mathbf{x}(t)\overline{\mathbf{y}(t)}dt; \quad \mathbf{f} : \mathbb{L}_\mathbf{2}[\mathbf{a}, \mathbf{b}] \to \mathbb{C}$

$$\mathbf{f}(\mathbf{x}) = \int_a^b \mathbf{x}(t)\overline{\mathbf{g}_0(t)}dt; \quad \mathbf{g}_0(t) \in \mathbb{L}_2[a,b]; \ \mathbf{g}_0(t) \textit{ is fixed}$$

Notation 3.13 *In the spaces $\mathbb{L}_p$ or. $\mathbf{l}_p$ with $p \neq 2$ the common shape of continuous linear functionals is given by $f(\mathbf{x}) = (\mathbf{x}, \mathbf{u})$, too, with $\mathbf{u} \in \mathbb{L}_q$ (or $\mathbf{u} \in \mathbf{l}_q$) and $\frac{1}{p} + \frac{1}{q} = 1; \quad 1 < p < \infty$. To prove this we need the Hölder inequality:*

$$\mathbf{l}_p \quad : \quad \left|\sum\nolimits_{k=1}^{\infty} x_k y_k\right| \leq \left(\sum\nolimits_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \left(\sum\nolimits_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}} < \infty$$

$$\mathbb{L}_p \quad : \quad \left|\int_a^b \mathbf{fg}dx\right| \leq \left(\int_a^b |\mathbf{f}|^p \, dx\right)^{\frac{1}{p}} \left(\int_a^b |\mathbf{g}|^q \, dx\right)^{\frac{1}{q}} < \infty.$$

Definition 3.19 *Given a sequence $\{\mathbf{x}_n\}_{n=1}^{\infty}$ in the HILBERT space $\mathbb{H}$.*
If $\lim_{n\to\infty} \mathbf{f}(\mathbf{x}_n) = \mathbf{f}(\mathbf{x})$ for every linear continuous functional $\mathbf{f}$ on $\mathbb{H}$ then $\{\mathbf{x}_n\}_{n=1}^{\infty}$ ***tends weakly*** *to $\mathbf{x} \in \mathbb{H}$ as n tends to infinity. We write: $\mathbf{x}_n \rightharpoonup \mathbf{x}$*

Notation 3.14 *Norm convergent sequences $\{\mathbf{x}_n\}_{n=1}^{\infty} \subset \mathbb{H}$ are called strongly convergent. That means: $\|\mathbf{x}_n - \mathbf{x}\| \overset{n\to\infty}{\to} 0$.*

Theorem 3.20 *The strong convergence implies the weak convergence.*

Proof.

$$|\mathbf{f}(\mathbf{x}_n) - \mathbf{f}(\mathbf{x})| = |\mathbf{f}(\mathbf{x}_n - \mathbf{x})| = |(\mathbf{x}_n - \mathbf{x}, \mathbf{u}_0)| \leq \|\mathbf{u}_0\| \, \|\mathbf{x}_n - \mathbf{x}\| = \|\mathbf{f}\| \, \|\mathbf{x}_n - \mathbf{x}\| \overset{n\to\infty}{\longrightarrow} 0$$

■

Notation 3.15 *The converse is not true.*

Example 3.20 $\mathbb{H} = \mathbb{L}_2[0,\pi]$ *with* $\mathbf{f}(\mathbf{x}) = \int_0^\pi \mathbf{x}(t)\overline{\mathbf{y}(t)}dt; \quad \mathbf{y} \in \mathbb{L}_\mathbf{2}[\mathbf{a},\mathbf{b}];$ $\mathbf{y}$ *is fixed and real.*
We consider the sequence $\mathbf{x}_n = \mathbf{x}_n(t) = \sin(nt); \quad n \in \mathbb{N}$

$$\mathbf{f}(\mathbf{x}_n) = \int_0^\pi \sin(nt)\overline{\mathbf{y}(t)}dt$$

$\mathbf{f}(\mathbf{x}_n)$ *are the Fourier coefficients of* $\overline{y(t)}$ *without the constant factor. Thus:* $\mathbf{f}(\mathbf{x}_n) \overset{n\to\infty}{\longrightarrow} 0$ *and*

$$\mathbf{f}(\mathbf{x}_n) \overset{n\to\infty}{\longrightarrow} 0 = \int_0^\pi 0 \cdot \overline{\mathbf{y}(t)}dt = \mathbf{f}(\mathbf{0}) \quad \forall \mathbf{y} \in \mathbb{L}_2[0,\pi],$$

i.e. for all functionals f*. On the other hand we get*

$$\|\mathbf{x}_n - \mathbf{0}\|^2 = \int_0^\pi (\sin(nt) - \mathbf{0})^2 dt = \frac{\pi}{2} \quad \textit{contradiction}$$

$\Longrightarrow \{\mathbf{x}_n\}$ *tends weakly to* $\mathbf{0}$*, but not strongly.*

3.5.4 Bilinear Forms

Definition 3.20 *A bilinear form on a HILBERT space* $\mathbb{H}$ *over the field* $\mathbb{K}$ *is a bilinear mapping* $\mathbf{a} : \mathbb{H} \times \mathbb{H} \to \mathbb{K}$*. That means:*

$$\begin{aligned} \mathbf{a}(\alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w}) &= \alpha\mathbf{a}(\mathbf{u}, \mathbf{w}) + \beta\mathbf{a}(\mathbf{v}, \mathbf{w}) \\ \mathbf{a}(\mathbf{w}, \alpha\mathbf{u} + \beta\mathbf{v}) &= \alpha\mathbf{a}(\mathbf{w}, \mathbf{u}) + \beta\mathbf{a}(\mathbf{w}, \mathbf{v}) \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{H}, \ \forall \alpha, \beta \in \mathbb{K}. \end{aligned}$$

Definition 3.21 *The bilinear form* $\mathbf{a} : \mathbb{H} \times \mathbb{H} \to \mathbb{K}$ *is called*

- **bounded**, if there $\exists\, C > 0; \ C \in \mathbb{R} \quad | \quad |\mathbf{a}(\mathbf{u}, \mathbf{v})| \le C\, \|\mathbf{u}\|\, \|\mathbf{v}\| \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{H}$
- **symmetric**, if $\mathbf{a}(\mathbf{u}, \mathbf{v}) = \mathbf{a}(\mathbf{v}, \mathbf{u}) \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{H}$
- **positive semidefinite**, if $\mathbf{a}(\mathbf{u}, \mathbf{u}) \ge 0 \quad \forall \mathbf{u} \in \mathbb{H}$
- **positive definite**, if $\mathbf{a}(\mathbf{u}, \mathbf{u}) \ge C\, \|\mathbf{u}\|^2 \quad \forall \mathbf{u} \in \mathbb{H} \ \wedge \ C > 0, \ C = const.$

Example 3.21 $\mathbb{H} = \mathbb{R}^n$ *with* $(\mathbf{u}, \mathbf{v}) = \sum\limits_{k=1}^n u_k v_k; \quad \mathbf{a}(.,.) : \mathbb{R}^\mathbf{n} \times \mathbb{R}^\mathbf{n} \to \mathbb{R}$

$$\mathbf{a}(\mathbf{u}, \mathbf{v}) = (A\mathbf{u}, \mathbf{v}) = \sum_{i,j=1}^n a_{ij} u_j v_i; \quad A = (a_{ij})_{i,j=1}^n$$

$\mathbf{a}$ *is symmetric (positive definit) if and only if the matrix* A *is symmetric (positive definit).*

4 Variational Calculus

From the classical analysis the search of local extrema of a real function is known: We are looking for real numbers x_0, for which a given a real function $f(x)$ has a local (or relative) maximum or minimum point. That means:

Definition 4.1 *Let the domain* $\mathbb{X}$ *of* $f(x)$ *be a metric space. Then* f *is said to have a local (or relative) maximum point at the point* x_0 *if* $\exists\ \varepsilon > 0 \mid f(x_0) \geq f(x)\ \forall x \in \mathbb{X}$ *with* $d(x_0, x) < \varepsilon$. *Similarly, the function has a local minimum point at* x_0 *if* $\exists\ \varepsilon > 0 \mid f(x_0) \leq f(x)\ \forall x \in \mathbb{X}$ *within distance* ε *of* x_0.

If $f(x)$ is a differentiable function the critical points can be found by the **necessary requirement** $f'(x_0) = 0$. To decide wether x_0 is a maximum or a minimum we take the second derivative test or the higher order derivative test which use the values of higher order derivatives at x_0.. **(sufficient requirement)**.
Calculus of variations deals with maximising or minimising functionals in common function spaces $\mathbb{X}$ and it is possible to generalise the way above.
Given a functional $\mathbf{f}(\mathbf{u}) : \mathbb{X} \to \mathbb{R}$. Wanted functions $\mathbf{u}_0 \in \mathbb{X}$, such that $\mathbf{f}(\mathbf{u}_0)$ is a minimum or a maximum with respect to the elements of an environment of $\mathbf{u}_0$. Solving this problem we have to generalise the derivatives f', f'', ... of a real function $f(x)$ and the necessary and sufficient conditions for an extremum.

4.1 Variation and Derivatives of Functionals

Let $\mathbb{X}$ be a normed space over the field $\mathbb{K} = \mathbb{R}$ and $\mathbf{f} : U(\mathbf{u}_0) \subseteq \mathbb{X} \to \mathbb{R}$ a functional, which is defined in the environment $U(\mathbf{u}_0)$ of $\mathbf{u}_0$.
Then $\varphi(t) = \mathbf{f}(\mathbf{u}_0 + t\mathbf{h})$ is for every $\mathbf{h} \in \mathbb{X}$ a real function of the parameter $t \in \mathbb{R}$.

Definition 4.2 $\delta\mathbf{f}(\mathbf{u}_0, \mathbf{h}) \equiv \varphi'(0)$ *is called **first variation** of* $\mathbf{f}$ *at* $\mathbf{u}_0$ *in the direction* $\mathbf{h}$. *The* $\boldsymbol{n}^{\mathbf{th}}$ ***variation*** $\delta^n\mathbf{f}(\mathbf{u}_0, \mathbf{h})$ *of the functional* $\mathbf{f}$ *at* $\mathbf{u}_0$ *in the direction* $\mathbf{h}$ *is the* n^{th} *derivative of* $\varphi(t)$ *at* $t = 0$:

$$\delta^n\mathbf{f}(\mathbf{u}_0, \mathbf{h}) \equiv \varphi^{(n)}(0) \qquad n = 1, 2, ...$$

Definition 4.3 *If the first variation* $\delta\mathbf{f}(\mathbf{u}_0, \mathbf{h})$ *exists for every* $\mathbf{h} \in \mathbb{X}$ *and if it is linear and continuous then the **GÂTEAUX derivative*** $\mathbf{f}'(\mathbf{u}_0)(\cdot) : \mathbb{X} \to \mathbb{R}$ *of the functional* $\mathbf{f}$ *at* $\mathbf{u}_0$ *is defined by*

$$\mathbf{f}'(\mathbf{u}_0)(\mathbf{h}) = \delta\mathbf{f}(\mathbf{u}_0, \mathbf{h}) \quad \textit{für} \ \ \forall\ \mathbf{h} \in \mathbb{X}.$$

Definition 4.4 *The GÂTEAUX derivative* $\mathbf{f}'(\mathbf{u}_0)$ *is called* ***FRÉCHET derivative,*** *if*

$$\begin{aligned} \mathbf{f}(\mathbf{u}_0+\mathbf{h})-\mathbf{f}(\mathbf{u}_0) &= \mathbf{f}'(\mathbf{u}_0)(\mathbf{h})+\|\mathbf{h}\|\,e(\mathbf{h}) \\ \textit{with} \quad \lim_{\mathbf{h}\to\mathbf{0}} e(\mathbf{h}) &= 0 \quad \textit{für} \quad \forall\ \mathbf{h}\in\mathbb{X}. \end{aligned}$$

Example 4.1 $\mathbb{X}=\mathbb{R}$, $\mathbf{f}(\mathbf{u})$ *is a continuous and differentiable function:* $\mathbf{u}=\mathbf{x}\in(a,b)$

$$\mathbf{f}:(a,b)\subseteq\mathbb{R}\to\mathbb{R}.$$

***1**st **variation** of* $\mathbf{f}$*:*

$$\delta\mathbf{f}(\mathbf{u},\mathbf{h})=\frac{d}{dt}\,\mathbf{f}(\mathbf{x}+t\mathbf{h})\big|_{t=0}=\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\mathbf{h}\equiv d\mathbf{f}(\mathbf{x})$$

***2**nd **variation** of* $\mathbf{f}$*:*

$$\delta^2\mathbf{f}(\mathbf{u},\mathbf{h})=\frac{d^2}{dt^2}\,\mathbf{f}(\mathbf{x}+t\mathbf{h})\big|_{t=0}=\frac{d^2\mathbf{f}(\mathbf{x})}{d\mathbf{x}^2}\mathbf{h}^2$$

That means the ***GATEAUX derivative*** *is the ordinary derivative* $\frac{d}{d\mathbf{x}}\mathbf{f}(\mathbf{x})$.
If $\mathbf{f}(\mathbf{u})$ *is twice continuously differentiable the* ***FRÉCHET derivative*** *exists. From the TAYLOR series we get then for* $\mathbf{f}(\mathbf{u})$ *at* $\mathbf{u}=\mathbf{x}$

$$\begin{aligned} \mathbf{f}(\mathbf{u}+\mathbf{h})-\mathbf{f}(\mathbf{u}) &= \frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\mathbf{h}+\frac{d^2\mathbf{f}(\boldsymbol{\xi})}{d\mathbf{x}^2}\frac{\mathbf{h}^2}{2}=\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\mathbf{h}+e(\mathbf{h})\,\mathbf{h} \\ \textit{with} \quad |e(\mathbf{h})| &\to 0 \quad \textit{as} \quad \mathbf{h}\to 0. \end{aligned}$$

Example 4.2 $\mathbb{X}=\mathbb{R}^n$, $\mathbf{u}=\overrightarrow{\mathbf{x}}=(x_1,x_2,...,x_n)^T$, $\mathbf{f}(\mathbf{u})=\mathbf{f}(x_1,...,x_n)$ *is a continuously differentiable function*

$$\mathbf{f}:\mathbf{M}\subseteq\mathbb{R}^n\to\mathbb{R}.$$

***1**st **variation** of* $\mathbf{f}$*: is the directional derivative of* $\mathbf{f}(\mathbf{u})$ *at* $\mathbf{u}=\overrightarrow{\mathbf{x}}\in\mathbf{M}$ *in the direction* $\mathbf{h}=\overrightarrow{\mathbf{h}}=(h_1,...,h_n)^T\,T$. *If* $\mathbf{h}$ *is normelised we get:*

$$\begin{aligned} \delta\mathbf{f}(\mathbf{u},\mathbf{h}) &= \frac{d}{dt}\,\mathbf{f}(x_1+th_1,...,u_n+th_n)\big|_{t=0} \\ &= \sum_{i=1}^{n}\frac{\partial\mathbf{f}(\overrightarrow{\mathbf{x}})}{\partial x_i}h_i\equiv grad\left(\mathbf{f}(\overrightarrow{\mathbf{x}})\right)\cdot\overrightarrow{\mathbf{h}}. \end{aligned}$$

2nd variation *of* $\mathbf{f}$:

$$\begin{aligned}\delta^2\mathbf{f}(\mathbf{u},\mathbf{h}) &= \left.\frac{d}{dt}\sum_{i=1}^{n}\frac{\partial\mathbf{f}\left(\overrightarrow{\mathbf{x}}+t\,\overrightarrow{\mathbf{h}}\right)}{\partial x_i}h_i\right|_{t=0}\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2\mathbf{f}\left(\overrightarrow{\mathbf{x}}\right)}{\partial x_i\partial x_j}h_ih_j = \left(\mathbf{G}\left(\overrightarrow{\mathbf{x}}\right)\overrightarrow{\mathbf{h}},\overrightarrow{\mathbf{h}}\right)\end{aligned}$$

with the functional matrix $\mathbf{G}\left(\overrightarrow{\mathbf{x}}\right) = \left(\frac{\partial^2\mathbf{f}\left(\overrightarrow{\mathbf{u}}\right)}{\partial x_i\partial x_j}\right)_{i,j=1}^{n}$.

GATEAUX derivative: $\mathbf{f}'(\mathbf{u}) = \mathbf{f}'\left(\overrightarrow{\mathbf{x}}\right) = grad\left(\mathbf{f}\left(\overrightarrow{\mathbf{x}}\right)\right)$

$$\mathbf{f}'(\mathbf{u})(\mathbf{h}) = \left(\mathbf{f}'\left(\overrightarrow{\mathbf{x}}\right),\overrightarrow{\mathbf{h}}\right) = grad\left(\mathbf{f}\left(\overrightarrow{\mathbf{x}}\right)\right)\cdot\overrightarrow{\mathbf{h}}$$

with the estimation

$$\left|\mathbf{f}'(\mathbf{u})(\mathbf{h})\right| \le \left\|grad\left(\mathbf{f}\left(\overrightarrow{\mathbf{x}}\right)\right)\right\|\left\|\overrightarrow{\mathbf{h}}\right\|.$$

If $\mathbf{f}$ *is continuously differentiable in an environment of* $\overrightarrow{\mathbf{x}}$ *then the GATEAUX derivative exists at* $\overrightarrow{\mathbf{x}}$*.*

FRÉCHET derivative: *From the TAYLOR series we get then:*

$$\begin{aligned}\mathbf{f}(\mathbf{u}+\mathbf{h})-\mathbf{f}(\mathbf{u}) &= grad\left(\mathbf{f}\left(\overrightarrow{\mathbf{x}}\right)\right)\cdot\overrightarrow{\mathbf{h}}+\frac{1}{2}\left(\mathbf{G}\left(\overrightarrow{\boldsymbol{\xi}}\right)\overrightarrow{\mathbf{h}},\overrightarrow{\mathbf{h}}\right)\\ &= \mathbf{f}'(\mathbf{u})\left(\overrightarrow{\mathbf{h}}\right)+\left\|\overrightarrow{\mathbf{h}}\right\|e\left(\overrightarrow{\mathbf{h}}\right)\end{aligned}$$

with

$$\left|e\left(\overrightarrow{\mathbf{h}}\right)\right| = \frac{1}{2\left\|\overrightarrow{\mathbf{h}}\right\|}\left|\left(\mathbf{G}\left(\overrightarrow{\boldsymbol{\xi}}\right)\overrightarrow{\mathbf{h}},\overrightarrow{\mathbf{h}}\right)\right| \le \frac{1}{2}\left\|\mathbf{G}\left(\overrightarrow{\boldsymbol{\xi}}\right)\right\|\left\|\overrightarrow{\mathbf{h}}\right\| \to 0 \quad for \quad \left\|\overrightarrow{\mathbf{h}}\right\| \to 0.$$

If $\mathbf{f}$ *is twice continuously differentiable in an environment of* $\overrightarrow{\mathbf{x}}$ *the FRÉCHET derivative at* $\overrightarrow{\mathbf{x}}$ *exists.*

Example 4.3 *Let* $\mathbb{X} = \mathbb{H}$ *be a HILBERT space and*

$$\mathbf{f}(\mathbf{u}) = \frac{1}{2}\mathbf{a}(\mathbf{u},\mathbf{u}) - (\mathbf{b},\mathbf{u})$$

be a ***quadratic functional*** $\mathbf{f}:\mathbb{H}\to\mathbb{R}$ *with the* ***symmetric bounded bilinear form*** $\mathbf{a}(\cdot,\cdot)$

$$|\mathbf{a}(\mathbf{u},\mathbf{v})| \le C\,\|\mathbf{u}\|\,\|\mathbf{v}\| \qquad \forall\ \mathbf{u},\mathbf{v}\in\mathbf{H}$$

*and the **linear functional*** $(\mathbf{b}, \mathbf{u})$ *with* $\mathbf{b} \in \mathbb{H}$ *fixed.*
We define

$$\begin{aligned}
\varphi(t) &= \mathbf{f}(\mathbf{u}+t\mathbf{h}) = \frac{1}{2}\mathbf{a}(\mathbf{u}+t\mathbf{h}, \mathbf{u}+t\mathbf{h}) - (\mathbf{b}, \mathbf{u}+t\mathbf{h}) \\
&= \frac{1}{2}\mathbf{a}(\mathbf{u}, \mathbf{u}) + \frac{1}{2}\mathbf{a}(\mathbf{u}, t\mathbf{h}) + \frac{1}{2}\mathbf{a}(t\mathbf{h}, \mathbf{u}) + \frac{1}{2}\mathbf{a}(t\mathbf{h}, t\mathbf{h}) - (\mathbf{b}, \mathbf{u}) - (\mathbf{b}, t\mathbf{h}) \\
&= \frac{t^2}{2}\mathbf{a}(\mathbf{h}, \mathbf{h}) + t\left[\mathbf{a}(\mathbf{u}, \mathbf{h}) - (\mathbf{b}, \mathbf{h})\right] + \mathbf{f}(\mathbf{u})
\end{aligned}$$

***1**st **variation** of* $\mathbf{f}$*:*

$$\delta\mathbf{f}(\mathbf{u}, \mathbf{h}) = \varphi'(0) = \mathbf{a}(\mathbf{u}, \mathbf{h}) - (\mathbf{b}, \mathbf{h})$$

***2**nd **variation** of* $\mathbf{f}$*:*

$$\delta^2\mathbf{f}(\mathbf{u}, \mathbf{h}) = \varphi''(0) = \mathbf{a}(\mathbf{h}, \mathbf{h})$$

GATEAUX derivative:

$$\begin{aligned}
\mathbf{f}'(\mathbf{u})(\mathbf{h}) &= \mathbf{a}(\mathbf{u}, \mathbf{h}) - (\mathbf{b}, \mathbf{h}) \qquad \textit{and} \\
|\mathbf{f}'(\mathbf{u})(\mathbf{h})| &\leq (C\|\mathbf{u}\| + \|\mathbf{b}\|)\|\mathbf{h}\|
\end{aligned}$$

If there exists a linear operator $\mathbf{A} : \mathbb{H} \rightarrow \mathbb{H}$ *with*

$$(\mathbf{A}\mathbf{u}, \mathbf{h}) = \mathbf{a}(\mathbf{u}, \mathbf{h}) \quad \forall\, \mathbf{h} \in \mathbb{H},$$

*then **GATEAUX derivative** is* $\mathbf{f}'(\mathbf{u}) = \mathbf{A}\mathbf{u} - \mathbf{b}$.
*The **FRÉCHET derivative** exists because*

$$\begin{aligned}
\mathbf{f}(\mathbf{u}+\mathbf{h}) - \mathbf{f}(\mathbf{u}) &= \mathbf{a}(\mathbf{u}, \mathbf{h}) - (\mathbf{b}, \mathbf{h}) + \frac{1}{2}\mathbf{a}(\mathbf{h}, \mathbf{h}) \\
&= \mathbf{f}'(\mathbf{u})(\mathbf{h}) + \|\mathbf{h}\|\, e(\mathbf{h})
\end{aligned}$$

with

$$|e(\mathbf{h})| = \frac{1}{2\|\mathbf{h}\|}|\mathbf{a}(\mathbf{h}, \mathbf{h})| \leq \frac{1}{2}C\|\mathbf{h}\| \rightarrow 0 \quad \textit{für} \quad \|\mathbf{h}\| \rightarrow 0.$$

Example 4.4 *In typical onedimensional variational problems the functionals are definit integrals:*

$$\mathbf{f} : \mathbb{X} \rightarrow \mathbb{R} \quad \textit{mit} \quad \mathbf{f} = \mathbf{f}(\mathbf{u}) = \int_a^b F(x, \mathbf{u}(x), \mathbf{u}'(x))\; dx$$

*with the so called **LAGRANGE function***

$$F = F(x, \mathbf{u}(x), \mathbf{u}'(x)); \qquad F : [a, b] \times \mathbb{R}^2 \rightarrow \mathbb{R} \quad \textit{für} \quad \forall\, \mathbf{u}(x) \in \mathbb{X} = \mathbf{C}^1(a, b),$$

Let F *be twice continuously differentiable with respect to all variables and* $\mathbf{u}(a) = u_a$ *plus* $\mathbf{u}(b) = u_b$

Now we calculate the first variation of $\mathbf{f}(\mathbf{u})$ *with* $\mathbf{h} \in \overset{o}{\mathbf{C}}{}^1(a,b)$ *(uniform convergence!)*

$$\varphi(t) = \mathbf{f}(\mathbf{u}_0 + t\mathbf{h}) = \int_a^b F(x, \mathbf{u}_0(x) + t\mathbf{h}(x), \mathbf{u}_0'(x) + t\mathbf{h}'(x))\ dx,$$

$\curvearrowright$

$$\left.\varphi'(t)\right|_{t=0} = \int_a^b \left[F_{u'}(x, \mathbf{u}_0, \mathbf{u}_0')\,\mathbf{h}'(x) + F_u(x, \mathbf{u}_0, \mathbf{u}_0')\,\mathbf{h}(x)\right]\ dx.$$

Because of $\mathbf{h}(a) = \mathbf{h}(b) = 0$ *we get by integration by parts*

$$\begin{aligned} \delta\mathbf{f}(\mathbf{u}_0, \mathbf{h}) &= \left.\varphi'(t)\right|_{t=0} \\ &= \int_a^b \left[-\frac{d}{dx}F_{u'}(x, \mathbf{u}_0, \mathbf{u}_0') + F_u(x, \mathbf{u}_0, \mathbf{u}_0')\right]\mathbf{h}(x)\ dx. \end{aligned}$$

Because F *is twice continuously differentiable, we get*

$$\begin{aligned} \mathbf{f}(\mathbf{u}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{u}_0) &= \int_a^b \left[F(x, \mathbf{u}_0(x) + \mathbf{h}(x), \mathbf{u}_0'(x) + \mathbf{h}'(x)) - F(x, \mathbf{u}_0(x), \mathbf{u}_0'(x))\right]\ dx \\ &= \int_a^b \left[F_u(x, \mathbf{u}_0, \mathbf{u}_0')\,\mathbf{h}(x) + F_{u'}(x, \mathbf{u}_0, \mathbf{u}_0')\,\mathbf{h}'(x) + \mathbf{h}e(\mathbf{h})\right]\ dx \\ \textit{with}\quad \|e(\mathbf{h})\| &\rightarrow 0 \quad \textit{for}\quad \mathbf{h} \rightarrow 0.\ \textit{By integration by parts we get} \\ &= \int_a^b \left[F_u(x, \mathbf{u}_0, \mathbf{u}_0') - \frac{d}{dx}F_{u'}(x, \mathbf{u}_0, \mathbf{u}_0')\right]\mathbf{h}\ dx + \int_a^b \mathbf{h}e(\mathbf{h})\ dx, \end{aligned}$$

If $\mathbf{h} \rightarrow 0$ *then the* 2^{nd} *integral tends to zero, too.* $\curvearrowright$ *The FRÉCHET derivative exists.*

$$\mathbf{f}'(\mathbf{u}_0) = F_u(x, \mathbf{u}_0, \mathbf{u}_0') - \frac{d}{dx}F_{u'}(x, \mathbf{u}_0, \mathbf{u}_0')$$

4.2 Extrema and Variational problems

We concern functionals $\mathbf{f}$ on the Banach space $\mathbb{X} = \mathbb{B}$ (Hilbert space $\mathbb{X} = \mathbb{H}$) and look for their extrema.

Theorem 4.1 ***Necessary requirement for an extremum:***
If there is at $\mathbf{u}_0 \in \overset{o}{\mathbf{D}}$ *a local extremum of* $\mathbf{f} : \mathbf{D} \subset \mathbb{X} \to \mathbb{R}$, $\mathbf{D}$ *open, then*

$$\delta \mathbf{f}\left(\mathbf{u}_0; \mathbf{h}\right) = 0 \quad \forall\ \mathbf{h} \in \mathbb{X}. \qquad (*)$$

Definition 4.5 *Points* $\mathbf{u}_0 \in \overset{o}{\mathbf{D}}$ *where the condition* $(*)$ *is satisfied are called* ***critical or stationary points of*** $\mathbf{f}$. ***A saddle point*** *is a stationary point which is not an extremum.*

Definition 4.6 *The variational problem on* $\mathbb{X}$ *is: Wanted the stationary points of* $\mathbf{f}$.

If there exists the FRECHET derivative $\mathbf{f}'\left(\mathbf{u}_0\right)$ then you can prove the following theorem:

Theorem 4.2 *If* $\mathbf{f} : \mathbf{D} \subset \mathbb{X} \to \mathbb{R}$ *is a FRECHET-differentiable functional and is* $\mathbf{u}_0 \in \overset{o}{\mathbf{D}}$ *the place of a local extremum of* $\mathbf{f}$ *then its FRECHET derivative vanishes:*

$$\mathbf{f}'\left(\mathbf{u}_0\right) = 0.$$

This equation ist called EULER-LAGRANGE equation or EULER equation.

Proof. Let $\mathbf{u}_0 \in \overset{o}{\mathbf{D}}$ be a local minimum point of $\mathbf{f}$.

$$\curvearrowright \mathbf{f}\left(\mathbf{u}\right) \geq \mathbf{f}\left(\mathbf{u}_0\right) \quad \forall \mathbf{u} \in U(\mathbf{u}_0)$$

We define

$$\mathbf{u} = \mathbf{u}_0 + t\mathbf{h}; \quad t > 0; \quad \mathbf{h} \in \mathbb{X}; \quad \|\mathbf{h}\| = 1$$

Because of the existance of the FRECHET derivative we get

$$\mathbf{f}\left(\mathbf{u}_0 + t\mathbf{h}\right) = \mathbf{f}\left(\mathbf{u}_0\right) + \mathbf{f}'\left(\mathbf{u}_0\right)\left(t\mathbf{h}\right) + \|t\mathbf{h}\|\, e\left(t\mathbf{h}\right) \qquad \text{with} \qquad \lim_{\mathbf{t}\to\mathbf{0}} e\left(t\mathbf{h}\right) = 0$$

Thus

$$\begin{aligned} 0 &\leq \frac{\mathbf{f}\left(\mathbf{u}_0 + t\mathbf{h}\right) - \mathbf{f}\left(\mathbf{u}_0\right)}{t} = \mathbf{f}'\left(\mathbf{u}_0\right)\left(\mathbf{h}\right) + \|\mathbf{h}\|\, e\left(t\mathbf{h}\right) = \mathbf{f}'\left(\mathbf{u}_0\right)\left(\mathbf{h}\right) + e\left(t\mathbf{h}\right) \\ 0 &\leq \mathbf{f}'\left(\mathbf{u}_0\right)\left(\mathbf{h}\right) \quad \forall \mathbf{h} \in \mathbb{X}\ \textit{mit}\ \|\mathbf{h}\| = 1,\ \textit{da}\ \lim_{\mathbf{t}\to\mathbf{0}} e\left(t\mathbf{h}\right) = 0 \end{aligned}$$

Using $-\mathbf{h}$ instaed of $\mathbf{h}$ we will get:

$$0 \leq \mathbf{f}'(\mathbf{u}_0)(-\mathbf{h}) = -\mathbf{f}'(\mathbf{u}_0)(\mathbf{h}).$$

Therefore

$$\mathbf{f}'(\mathbf{u}_0)(\mathbf{h}) = 0 \quad \forall \mathbf{h} \in \mathbb{X}; \quad \|\mathbf{h}\| = 1.$$

Arbitrary elements of $\mathbb{X}$ can be represented by $\mathbf{v} = \lambda\mathbf{h}$ with $\lambda \in \mathbb{R}$.$\curvearrowright$

$$\begin{aligned} \lambda\mathbf{f}'(\mathbf{u}_0)(\mathbf{h}) &= 0 \quad \forall\, \mathbf{h} \in \mathbb{X} \\ \mathbf{f}'(\mathbf{u}_0)(\mathbf{v}) &= 0 \quad \forall\, \mathbf{v} \in \mathbb{X} \end{aligned}$$

■

Example 4.5 *Formulation of the variational problems for the previous examples:*
Example 4.1: The EULER-LAGRANGE equation $\frac{d}{d\mathbf{x}}\mathbf{f}(\mathbf{x}) = 0$ *corresponds to the necessary requirement for an extremum of* $\mathbf{f}(\mathbf{x})$.
Example 4.2: If the EULER-LAGRANGE equation is satisfied:

$$grad\,(\mathbf{f}(\mathbf{u}_0)) = \mathbf{0}$$

then $\mathbf{u}_0$ *is a stationary point of* $\mathbf{f}$ *(necessary requirement for an extremum of* $\mathbf{f}(\mathbf{x})$*).*
Example 4.3: If the EULER-LAGRANGE equation is satisfied:

$$\mathbf{A}\mathbf{u}_0 - \mathbf{b} = \mathbf{0}$$

then $\mathbf{u}_0 \in \mathbb{H}$ *is a stationary point of* $\mathbf{f}$.

In the example 4.4 the domain $\mathbb{X}$ of $\mathbf{f}$ is not a space or a subspace, but a manifold of a BANACH space.

4.3 Constrained Extrema

Let $\mathbb{X} = \mathbb{B}$ be a BANACH space (or a HILBERT space $\mathbb{X} = \mathbb{H}$). $\mathbb{V}$ is then a subspace of $\mathbb{X}$, $\mathbf{f}$: $\mathbb{X} \to \mathbb{R}$ a functional, $\mathbf{u}^* \in \mathbb{X}$. Then

$$\mathbb{M} = \mathbf{u}^* + \mathbb{V} := \{\mathbf{u}^* + \mathbf{v} \mid \mathbf{v} \in \mathbb{V}\}$$

is a linear manifold. We concern the constraint $\mathbf{f}|_{\mathbf{M}}$ and look for their extrema, i.e. $\mathbf{u_0}$ is an extremum of $\mathbf{f}$ with the additional condition $\mathbf{u_0} \in \mathbb{M}$. Let the functional $\mathbf{f}$ be FRÉCHET differentiable.

Definition 4.7 *Points* $\mathbf{u}_0 \in \mathbb{M}$ *with* $\mathbf{f}'[\mathbf{u}_0] = 0 \quad \forall \mathbf{h} \in \mathbb{V}$ *are called constrained extrema of* $\mathbf{f}$ *with the additional condition* $\mathbf{u}_0 \in \mathbb{M}$.

Definition 4.8 *Variational problems on a manifold:*
Wanted the stationary point $\mathbf{u}_0$ *of* $\mathbf{f}$ *with the additional condition* $\mathbf{u}_0 \in \mathbb{M}$.

Theorem 4.3 $\mathbf{f} : \mathbf{D} \subset \mathbb{X} \to \mathbb{R}$ *is a FRÉCHET differentiable functional,* $\mathbb{V}$ *a subspace of* $\mathbb{X}$, $\mathbf{u}^*$ *an arbitrary fixed point of* $\mathbb{X}$ *and* $\mathbb{M} = \mathbf{u}^* + \mathbb{V}$ *a linear manifold of* $\mathbb{X}$. *If the constraint* $\mathbf{f}|_{\mathbb{M}}$ *has an extremum at* $\mathbf{u}_0$ *then the EULER-LAGRANGE equation is satisfied there:*

$$\mathbf{f}'(\mathbf{u}_0) = 0.$$

Proof. We define the functional

$$\widehat{\mathbf{f}}(\mathbf{v}) = \mathbf{f}(\mathbf{u}^* + \mathbf{v}) \quad \forall \mathbf{v} \in \mathbb{V}.$$

which maps $\mathbb{V}$ in $\mathbb{R}$. Because $\mathbf{f}$ is Fréchet differentiable we get for every $\mathbf{v}, \mathbf{h} \in \mathbb{V}$:

$$\begin{aligned}
\widehat{\mathbf{f}}(\mathbf{v}+\mathbf{h}) &= \mathbf{f}(\mathbf{u}^* + \mathbf{v} + \mathbf{h}) \\
&= \mathbf{f}(\mathbf{u}^* + \mathbf{v}) + \mathbf{f}'(\mathbf{u}^* + \mathbf{v})(\mathbf{h}) + \|\mathbf{h}\|\, e(\mathbf{u}^* + \mathbf{v}, \mathbf{h}) \text{ with } \lim_{\mathbf{h}\to\mathbf{0}} e(\mathbf{u}^* + \mathbf{v}, \mathbf{h}) = 0
\end{aligned}$$

With $\widehat{e}(\mathbf{v}, \mathbf{h}) = e(\mathbf{u}^* + \mathbf{v}, \mathbf{h})$ we obtain

$$\widehat{\mathbf{f}}(\mathbf{v}+\mathbf{h}) = \widehat{\mathbf{f}}(\mathbf{v}) + \mathbf{f}'(\mathbf{u}^* + \mathbf{v})(\mathbf{h}) + \|\mathbf{h}\|\, \widehat{e}(\mathbf{v}, \mathbf{h}) \quad \forall \mathbf{v} \in \mathbb{V}$$

and therefore

$$\mathbf{f}'(\mathbf{u}^* + \mathbf{v})(\mathbf{h}) = \widehat{\mathbf{f}}'(\mathbf{v})(\mathbf{h})$$

The theorem above implies $\widehat{\mathbf{f}}'(\mathbf{v}_0)(\mathbf{h}) = 0$, i.e. $\mathbf{f}'(\mathbf{u}^* + \mathbf{v}_0)(\mathbf{h}) = 0 \quad \forall \mathbf{h} \in \mathbb{V}$
for every extemum $\mathbf{v}_0 \in \mathbb{V}$ of $\widehat{\mathbf{f}}$. With

$$\mathbf{u}_0 = \mathbf{u}^* + \mathbf{v}_0$$

we get the statement $\mathbf{f}'(\mathbf{u}_0) = 0$. ■

Example 4.6 *Consider example 4.4. The condition there implies*

$$\begin{aligned}
\delta F(\mathbf{u}_0, \mathbf{h}) &= 0 \\
&= \int_a^b \left[-\frac{d}{dx} F_{u'}(x, \mathbf{u}_0, \mathbf{u}_0') + F_u(x, \mathbf{u}_0, \mathbf{u}_0') \right] \mathbf{h}(x)\ dx.
\end{aligned}$$

for every $\mathbf{h} \in \overset{o}{\mathbf{C}}{}^1(a,b)$. *Thus we get for a critical point of* $\mathbf{f}$ *the EULER-LAGRANGE equation*

$$\frac{d}{dx}\left(F_{\mathbf{u}'}(x,\mathbf{u},\mathbf{u}')\right) = F_{\mathbf{u}}(x,\mathbf{u},\mathbf{u}') \qquad (EG)_1$$

or (if we differentiate):

$$F_{\mathbf{u}} - F_{\mathbf{u}'x} - F_{\mathbf{u}\mathbf{u}'}\mathbf{u}' - F_{\mathbf{u}'\mathbf{u}'}\mathbf{u}'' = 0 \qquad (EG)_2 .$$

In this case the EULER-LAGRANGE equation is in general a nonlinear ***ordinary differential equation of 2***$^{\text{nd}}$ ***order for the function*** $\mathbf{u}(x)$ *(with boundary values).*

Example 4.7 *Wanted the beeline between the points* $P_0(x_0,y_0)$ *and* $P_1(x_1,y_1)$ *(with* $x_0 \neq x_1$*) in a plane. Let* $\mathbf{u} = \mathbf{u}(x)$ *be the function of the curve which connect the points* P_0 *and* $P(x,y)$. *The length* $\mathbf{f}(\mathbf{u})$ *of this curve is then:*

$$\mathbf{f}(\mathbf{u}) = \int_{x_0}^{x_1} \sqrt{1 + (\mathbf{u}'(x))^2}\,dx$$

with the LAGRANGE function

$$F = F(\mathbf{u}') = \sqrt{1 + (\mathbf{u}'(x))^2}.$$

Using

$$F_{\mathbf{u}} = 0 \qquad \textit{and} \qquad F_{\mathbf{u}'} = \frac{\mathbf{u}'}{\sqrt{1 + (\mathbf{u}'(x))^2}}$$

we get the EULER-LAGRANGE equation

$$\frac{d}{dx}\left(\frac{\mathbf{u}'}{\sqrt{1 + (\mathbf{u}'(x))^2}}\right) = 0$$

$\curvearrowright$

$$\begin{aligned}
\frac{\mathbf{u}'}{\sqrt{1 + (\mathbf{u}'(x))^2}} &= C \\
(\mathbf{u}')^2 &= C^2\left(1 + (\mathbf{u}')^2\right) \\
(\mathbf{u}')^2\left(1 - C^2\right) &= C^2 \\
\mathbf{u}'(x) &= \pm\sqrt{\frac{C^2}{1 - C^2}} \equiv a = const.
\end{aligned}$$

Therefore the beeline is the straight line:

$$\mathbf{u}(x) = ax + b$$

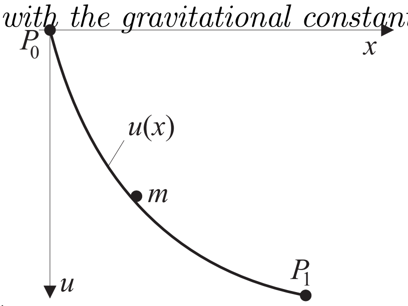
as it was to be expected. The line must go through the two points $P_0(x_0, y_0)$ *and* $P_1(x_1, y_1)$ *with* $x_0 \neq x_1$.
Using this boundary conditions we get:

$$\begin{aligned} \mathbf{u}(x_0) = ax_0 + b = y_0 \\ \mathbf{u}(x_1) = ax_1 + b = y_1 \end{aligned} \quad \Rightarrow \quad \begin{pmatrix} x_0 & 1 \\ x_1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}$$

$$\Rightarrow \quad a = \frac{y_0 - y_1}{x_0 - x_1} \qquad b = \frac{x_0 y_1 - x_1 y_0}{x_0 - x_1}.$$

Example 4.8 ***Brachistochrone Curve*** *(curve of fastest descent)*
Wanted a curve lying on a plane, going from $P_0(0,0)$ *to* $P_1(a,b)$ *(with* $a \neq 0,\ b < 0$*) which a particle with the mass* m *slides frictionlessly in the shortest time under the influence of a uniform gravitational field with the gravitational constant* g.

$\mathbf{u} = \mathbf{u}(x)$	*function of the curve from* P_0 *to* P_1
$s = s(x)$	*parameter of the length of the curve from* P_0 *to* P_1 *with* $s(0) = 0$ *and* $ds = \sqrt{1 + (u')^2}dx$



Using the law of conservation of energy $\frac{1}{2}mv^2 - mgu = 0$ *we get for the velocity* v *of the particle*

$$v = \frac{ds}{dt} = \sqrt{2g\mathbf{u}}$$

Thus

$$\mathbf{f}(\mathbf{u}) = \int_0^T dt = \int_0^L \frac{1}{v} ds = \int_0^a \sqrt{\frac{1 + (\mathbf{u}')^2}{2g\mathbf{u}}} dx$$

with the LAGRANGE function

$$F = F(\mathbf{u}, \mathbf{u}') = \sqrt{\frac{1 + (\mathbf{u}')^2}{2g\mathbf{u}}} = \frac{1}{\sqrt{2g}} \sqrt{\frac{1 + (\mathbf{u}')^2}{\mathbf{u}}}$$

Variational problem:
We look for a function $\mathbf{u}(x) \in \mathbf{C}^1(0,a)$ *with* $\mathbf{f}(\mathbf{u}) \to \min,\qquad \mathbf{u}(0) = 0$ *and* $\mathbf{u}(a) = b$.
Because the LAGRANGE function only depends on $\mathbf{u}$ *and* $\mathbf{u}'$ *the EULER equation*

(EG2) is :

$$\begin{aligned} 0 &= F_u - F_{\mathbf{u}\mathbf{u}'}\mathbf{u}' - F_{\mathbf{u}'\mathbf{u}'}\mathbf{u}'' \\ &= \mathbf{u}'\left(F_u - F_{\mathbf{u}\mathbf{u}'}\mathbf{u}' - F_{\mathbf{u}'\mathbf{u}'}\mathbf{u}''\right) \\ &= \frac{d}{dx}\left(F - \mathbf{u}'F_{\mathbf{u}'}\right) \end{aligned}$$

Thus the EULER equation has the workaround

$$F - u'F_{u'} = \frac{1}{\sqrt{2g}}\left[\sqrt{\frac{1+(\mathbf{u}')^2}{\mathbf{u}}} - \frac{(\mathbf{u}')^2}{\sqrt{\mathbf{u}\left(1+(\mathbf{u}')^2\right)}}\right] = \widetilde{C}$$

Multiplication by $\sqrt{2g}$ *and expansion by* $\sqrt{\left(1+(\mathbf{u}')^2\right)}$ *give*

$$\frac{1+(\mathbf{u}')^2-(\mathbf{u}')^2}{\sqrt{\mathbf{u}\left(1+(\mathbf{u}')^2\right)}} = C, \qquad \textit{with } C = \sqrt{2g}\widetilde{C}.$$

Therefore we get

$$\mathbf{u}\left(1+(\mathbf{u}')^2\right) = D \qquad \textit{with } D = \frac{1}{C^2}.\ .$$

Now let be $\mathbf{u}' = \cot t$. *Then*

$$\mathbf{u} = \frac{D}{1+\cot^2 t} = D\sin^2 t = \frac{D}{2}\left(1-\cos 2t\right).$$

Further

$$\begin{aligned} dx &= \frac{d\mathbf{u}}{\mathbf{u}'} = \frac{2D\sin t\cos t dt}{\cot t} = 2D\sin^2 t dt \\ &= D\left(1-\cos 2t\right)dt \\ x &= D\left(t - \frac{1}{2}\sin 2t\right) + E \\ &= \frac{D}{2}\left(2t - \sin 2t\right) + E. \end{aligned}$$

Using the boundary conditions we get

$$\begin{aligned} x(0) &= 0 \qquad \curvearrowright \qquad E = 0 \\ \mathbf{u}(0) &= 0 = D\cdot 0 \end{aligned}$$

Therefore the constant D must be calculated by using the end point of the curve. Setting $\phi = 2t$ we get a set of cycloids with $\frac{D}{2} = r$. r is the radius of the unrolling circle. ($\mathbf{u}$-axis points downward!):

$$\begin{aligned} x &= \frac{D}{2}(\phi - \sin\phi) \\ \mathbf{u} &= \frac{D}{2}(1 - \cos\phi) \end{aligned}$$

4.4 Generalisations

Given the LAGRANGE function: $F : [a,b] \times \mathbb{R}^{n+1} \to \mathbb{R}$ with

$$\begin{aligned} F &= F\left(x, \mathbf{u}(x), \mathbf{u}'(x), ..., \mathbf{u}^{(n)}(x)\right) \\ \text{für} \quad \forall\, \mathbf{u}(x) &\in \mathbf{X} = \mathbf{C}^n(a,b). \end{aligned}$$

We assume that F is often enough continuously differentiable with respect to all variables.
Using the LAGRANGE function we construct the functional:

$$\mathbf{f} = \mathbf{f}(\mathbf{u}) = \int_a^b F\left(x, \mathbf{u}(x), \mathbf{u}'(x), ..., \mathbf{u}^{(n)}(x)\right) dx$$

Definition 4.9 *Variational problem on a manifold*
Wanted the critical points $\mathbf{u}(x) \in \mathbf{C}^n(a,b)$ of $\mathbf{f}(\mathbf{u})$ with the additional conditions $\mathbf{u}^{(k)}(a) = \alpha_k$ and $\mathbf{u}^{(k)}(b) = \beta_k$; $\alpha_k \in \mathbb{R}$, $\beta_k \in \mathbb{R}$ $\quad k = 0, 1, ..., n-1$

Theorem 4.4 *Every solution $\mathbf{u}_0(x)$ of this variations problem is a solution of the EULER-LAGRANGE equation, too.*

$$F_u - \frac{d}{dx}F_{u'} + \frac{d^2}{dx^2}F_{u''} \mp ... + (-1)^n \frac{d^n}{dx^n}F_{u^{(n)}} = 0 \qquad (EG)_3$$

We calculate the first variation of $\mathbf{f}(\mathbf{u})$ with $\mathbf{h} \in \overset{o}{\mathbf{C}^n}(a,b)$,
$\mathbf{h}^{(k)}(a) = \mathbf{h}^{(k)}(b) = 0 \quad k = 0, 1, ..., n-1$ and $t \in \mathbb{R}$:

$$\begin{aligned} \varphi(t) &= \mathbf{f}(\mathbf{u}_0 + t\mathbf{h}) = \int_a^b F\left(x, \mathbf{u}_0 + t\mathbf{h}, \mathbf{u}_0' + t\mathbf{h}', ..., \mathbf{u}_0^{(n)} + t\mathbf{h}^{(n)}\right) dx \\ \varphi'(t)\big|_{t=0} &= \int_a^b \left[F_u\mathbf{h} + F_{u'}\mathbf{h}' + ... + F_{u^{(n)}}\mathbf{h}^{(n)}\right] dx \end{aligned}$$

For $k = 1, ..., n$ we integrate by parts $k-$times and get

$$\begin{aligned}
&\int_a^b \left(F_{u^{(k)}}\right) \mathbf{h}^{(k)} dx \\
&= \left[\left(F_{u^{(k)}}\right) \mathbf{h}^{(k-1)}\right]_a^b - \int_a^b \left(\frac{d}{dx} F_{u^{(k)}}\right) \mathbf{h}^{(k-1)} dx \\
&= \left[\left(F_{u^{(k)}}\right) \mathbf{h}^{(k-1)}\right]_a^b - \left[\left(\frac{d}{dx} F_{u^{(k)}}\right) \mathbf{h}^{(k-2)}\right]_a^b + \int_a^b \left(\frac{d^2}{dx^2} F_{u^{(k)}}\right) \mathbf{h}^{(k-2)} dx \\
&= \ldots \\
&= \left[\sum_{i=0}^{k-1} (-1)^i \left(\frac{d^i}{dx^i} F_{u^{(k)}}\right) \mathbf{h}^{(k-1-i)}\right]_a^b + (-1)^k \int_a^b \left(\frac{d^k}{dx^k} F_{u^{(k)}}\right) \mathbf{h}\ dx
\end{aligned}$$

Using

$$\left[\sum_{i=0}^{k-1} (-1)^i \left(\frac{d^i}{dx^i} F_{u^{(k)}}\right) \mathbf{h}^{(k-1-i)}\right]_a^b = 0$$

we obtain the EULER-LAGRANGE equation

$$\begin{aligned}
\varphi'(t)\big|_{t=0} &= \int_a^b \sum_{k=0}^{n} (-1)^k \left(\frac{d^k}{dx^k} F_{u^{(k)}}\right) \mathbf{h}\ dx \stackrel{!}{=} 0 \quad \forall \mathbf{h} \in \overset{o}{\mathbf{C}}{}^n (a, b)\,, \ \mathbf{i.e.} \\
\sum_{k=0}^{n} (-1)^k \left(\frac{d^k}{dx^k} F_{u^{(k)}}\right) &= 0
\end{aligned}$$