Chapter 18

Topics of Functional Analysis

In Chapter 14 there have been introduced the important concepts such as

1) *Lineality* of a space of elements,
2) *Metric* (or norm) in a space,
3) *Compactness*, *convergence* of a sequence of elements and *Cauchy sequences*,
4) *Contraction principle.*

As the examples we have considered in details the finite dimensional spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ of real and complex vectors (numbers). But the same definitions of lineality and norms remain true if we consider as another example a functional space (where an element is a function) or a space of sequences (where an element is a sequence of real or complex vectors). The specific feature of such spaces is that all of them are *infinite dimensional.* This chapter deals with the analysis of such spaces which is called "*Functional Analysis*".

Let us introduce two important additional concept which will be use below.

Definition 18.1 *The subset $\mathcal{V}$ of a linear normed space $\mathcal{X}$ is said to be* ***dense*** *in $\mathcal{X}$ if its closure is equal to $\mathcal{X}$.*

This property means that every element $x \in \mathcal{X}$ may be approximated as closely as we like by some element $v \in \mathcal{V}$, that is, for any $x \in \mathcal{X}$ and any $\varepsilon > 0$ there exists an element $v \in \mathcal{V}$ such that $\|x - v\| < \varepsilon$.

All normed linear spaces have dense subsets, but they need not be obligatory countable subsets.

Definition 18.2 *A normed linear space $\mathcal{X}$ is said to be **separable** if it contains at least one dense subset which is countable.*

The separable spaces have special properties that are important in different applications. In particular, denoting the elements of such countable subset by $\{e_i\}_{i=1,...}$ it is possible to represent each element $x \in \mathcal{X}$ as the convergent series

$$\boxed{x = \sum_{i=1}^{\infty} \xi_i e_i} \tag{18.1a}$$

where the scalars $\xi_i \in \mathbb{R}$ are called the coordinates of the element x in the *basis* $\{e_i\}_{i=1,...}$.

18.1 Linear and normed spaces of functions

Below we will introduce the examples of some functional spaces with the corresponding norm within. The lineality and main properties of a norm (metric) can be easily verified that's why we leave this for a reader as an exercise.

18.1.1 Space m_n of all bounded complex numbers

Let us consider a set m of sequences $x := \{x_i\}_{i=1}^{\infty}$ such that

$$\boxed{x_i \in \mathbb{C}^n \text{ and } \sup_i \|x_i\| < \infty} \tag{18.2}$$

where $\|x_i\| := \sqrt{\sum_{s=1}^{n} x_{is}\bar{x}_{is}}$ and introduce the norm in m as

$$\boxed{\|x\| := \sup_i \|x_i\|} \tag{18.3}$$

18.1.2 Space l_p^n of all summable complex sequences

By the definition

$$\boxed{l_p^n := \left\{ x = \{x_i\}_{i=1}^{\infty} \mid x_i \in \mathbb{C}^n,\ \|x\|_{l_p^n} := \left(\sum_{i=1}^{\infty} \|x_i\|^p \right)^{1/p} \right\} < \infty} \tag{18.4}$$

18.1.3 Space $C[a,b]$ of continuous functions

It is defined as follows

$$\boxed{\begin{array}{c} C[a,b] := \{ f(t) \mid f \text{ is continuous for all } t \in [a,b], \\ \|f\|_{C[a,b]} := \max\limits_{t \in [a,b]} |f(t)| < \infty \} \end{array}} \tag{18.5}$$

18.1.4 Space $C^k[a,b]$ of continuously differentiable functions

It contains all functions which are k-times differentiable and the k-th derivative is continuous, that is

$$\boxed{\begin{array}{c} C^k[a,b] := \{ f(t) \mid f^{(k)} \text{ exists and continuous} \\ \text{for all } t \in [a,b],\ \|f\|_{C^k[a,b]} := \sum\limits_{i=0}^{k} \max\limits_{t \in [a,b]} \left| f^{(i)}(t) \right| < \infty \} \end{array}} \tag{18.6}$$

18.1.5 Lebesgue spaces $L_p[a,b]$ $(1 \le p < \infty)$

For each $1 \le p < \infty$ it is defined by the following way:

$$\boxed{\begin{array}{c} L_p[a,b] := \left\{ f(t) : [a,b] \to \mathbb{C} \mid \int\limits_{t=a}^{b} |f(t)|^p \, dt < \infty \right. \\ \text{(here the integral is understood in the Lebesgue sense)}, \\ \left. \|f\|_p := \left(\int\limits_{t=a}^{b} |f(t)|^p \, dt \right)^{1/p} \right\} \end{array}} \tag{18.7}$$

Remark 18.1 *Sure, here functions $f(t)$ are not obligatory continuous.*

18.1.6 Lebesgue spaces $L_\infty[a,b]$

It contains all measurable function from $[a,b]$ to $\mathbb{C}$, namely,

$$\boxed{\left.\begin{array}{c} L_\infty[a,b] := \{f(t) : [a,b] \to \mathbb{C} \mid \\ \|f\|_\infty := \underset{t\in[a,b]}{ess\sup} |f(t)| < \infty \end{array}\right\}} \tag{18.8}$$

18.1.7 Sobolev spaces $S_p^l(G)$

It consists of all functions (for the simplicity, real valued) $f(t)$ defined on G which have p-integrable continuous derivatives $f^{(i)}(t)$ $(i{=}1,...,l)$, that is,

$$\boxed{\left.\begin{array}{c} S_p^l(G) := \{f(t) : G \to \mathbb{R} \mid < \infty \ \ (i=1,...,l) \\ \text{(the integral is understood in the Lebesgue sense),} \\ \|f\|_{S_p^l(G)} := \left(\int\limits_{t\in G} |f(t)|^p \, dt + \sum\limits_{i=1}^{l} \int\limits_{t\in G} \left|f^{(i)}(t)\right|^p dt \right)^{1/p} \end{array}\right\}} \tag{18.9}$$

More exactly, the Sobolev space is the completion (see definition below) of (18.9).

18.1.8 Frequency domain spaces $\mathbb{L}_p^{m\times k}$, $\mathbb{RL}_p^{m\times k}$, $\mathbb{L}_\infty^{m\times k}$ and $\mathbb{RL}_\infty^{m\times k}$

By the definition

1) The Lebesgue space $\mathbb{L}_p^{m\times k}$ is the space of all p-integrable complex matrices, i.e.,

$$\left.\begin{array}{c} \mathbb{L}_p^{m\times k} := \{F : \mathbb{C} \to \mathbb{C}^{m\times k} \mid \\ \|F\|_{\mathbb{L}_p^{m\times k}} := \left(\frac{1}{2\pi} \int\limits_{\omega=-\infty}^{\infty} \left(\mathrm{tr}\,\{F(j\omega)\, F^{\sim}(j\omega)\}\right)^{p-1} d\omega \right)^{1/p} < \infty \\ \text{(here } F^{\sim}(j\omega) := F^{\intercal}(-j\omega)\text{)} \end{array}\right\} \tag{18.10}$$

2) The Lebesgue space $\mathbb{RL}_p^{m\times k}$ is the subspace of $\mathbb{L}_p^{m\times k}$ containing only complex matrices with rational elements, i.e., in

$$F = \|F_{i,j}(s)\|_{i=\overline{1,m};\ j=\overline{1,k}}$$

each element $F_{i,j}(s)$ represents the polynomial ratio

$$\boxed{\begin{array}{c} F_{i,j}(s) = \dfrac{a_{i,j}^0 + a_{i,j}^1 s + \ldots + a_{i,j}^{p_{i,j}} s_{i,j}^{p_{i,j}}}{b_{i,j}^0 + b_{i,j}^1 s + \ldots + b_{i,j}^{q_{i,j}} s_{i,j}^{q_{i,j}}} \\ p_{i,j} \text{ and } q_{i,j} \text{ are positive integer} \end{array}} \tag{18.11}$$

Remark 18.2 *If $p_{i,j} \le q_{i,j}$ for each element F_{ij} of , then $F(s)$ can be interpreted as a matrix* ***transfer function*** *of a linear (finite-dimensional) system.*

3) The Lebesgue space $\mathbb{L}_\infty^{m\times k}$ is the space of all complex matrices with bounded (almost everywhere) on the imaginary axis elements, i.e.

$$\boxed{\begin{array}{c} \mathbb{L}_\infty^{m\times k} := \Big\{ F : \mathbb{C} \to \mathbb{C}^{m\times k} \mid \\ \|F\|_{\mathbb{L}_\infty^{m\times k}} := \underset{s:\operatorname{Re} s>0}{ess\sup}\ \lambda_{\max}^{1/2} \left\{ F(s) F^{\sim}(s) \right\} < \infty \Big\} \\ = \underset{\omega\in(-\infty,\infty)}{ess\sup}\ \lambda_{\max}^{1/2} \left\{ F(j\omega) F^{\sim}(j\omega) \right\} < \infty \end{array}} \tag{18.12}$$

(the last equality may be regarded to as a the generalization of the Maximum Modulus Principle 17.10 for matrix functions).

4) The Lebesgue space $\mathbb{RL}_\infty^{m\times k}$ is the subspace of $\mathbb{L}_\infty^{m\times k}$ containing only complex matrices with rational elements given in the form (18.11).

18.1.9 Hardy spaces $\mathbb{H}_p^{m\times k}$, $\mathbb{RH}_p^{m\times k}$, $\mathbb{H}_\infty^{m\times k}$ and $\mathbb{RH}_\infty^{m\times k}$

The Hardy spaces $\mathbb{H}_p^{m\times k}$, $\mathbb{RH}_p^{m\times k}$, $\mathbb{H}_\infty^{m\times k}$ and $\mathbb{RH}_\infty^{m\times k}$ are subspaces of the corresponding Lebesgue spaces $\mathbb{L}_p^{m\times k}$, $\mathbb{RL}_p^{m\times k}$, $\mathbb{L}_\infty^{m\times k}$ and $\mathbb{RL}_\infty^{m\times k}$ containing complex matrices with only regular (holomorphic) (see Definition 17.2) elements on the open half-plane $\operatorname{Re} s > 0$.

Remark 18.3 *If $p_{i,j} \leq q_{i,j}$ for each element F_{ij} of , then $F(s) \in \mathbb{RH}_p^{m\times k}$ can be interpreted as a matrix transfer function of a* ***stable*** *linear (finite-dimensional) system.*

Example 18.1

$$\boxed{\begin{array}{ll} \dfrac{1}{2-s} \in \mathbb{RL}_2 := \mathbb{RL}_2^{1\times 1}, & \dfrac{1-s}{2-s} \in \mathbb{RL}_\infty := \mathbb{RL}_\infty^{1\times 1} \\ \dfrac{1}{2+s} \in \mathbb{HL}_2 := \mathbb{HL}_2^{1\times 1}, & \dfrac{1-s}{2+s} \in \mathbb{RH}_\infty := \mathbb{RH}_\infty^{1\times 1} \\ \dfrac{e^{-s}}{2-s} \in \mathbb{L}_2 := \mathbb{L}_2^{1\times 1}, & \dfrac{e^{-s}}{2+s} \in \mathbb{H}_2 := \mathbb{H}_2^{1\times 1} \\ e^{-s}\dfrac{1-s}{2-s} \in \mathbb{L}_\infty := \mathbb{L}_\infty^{1\times 1}, & e^{-s}\dfrac{1-s}{2+s} \in \mathbb{H}_\infty := \mathbb{H}_\infty^{1\times 1} \end{array}}$$

18.2 Banach spaces

18.2.1 Basic definition

Remember that a linear normed (topological) space $\mathcal{X}$ is said to be complete (see Definition 14.14) if every Cauchy (fundamental) sequence has a limit in the same space $\mathcal{X}$. The concept of a complete space is very import since even without evaluating the limit one can determine whether a sequence is convergent or not. So, if a metric (topological) space is not a complete it is impossible talk about a convergence, limits, differentiation and so on.

Definition 18.3 *A linear, normed and complete space is called a* ***Banach space****.*

18.2.2 Examples of incomplete metric spaces

Sure that not all linear normed (metric) spaces are complete. The example given below illustrates this fact.

Example 18.2 (of a noncomplete normed space) *Let us consider the space $CL[0,1]$ of all continuous functions $f : [0,1] \to R$ which are* *absolutely integrable (in this case, in the Riemann sense) on* $[0,1]$*, that is, for which*

$$\|f\|_{CL[0,1]} := \int_{t=-1}^{1} |f(t)|\, dt < \infty \tag{18.13}$$

Consider the sequence $\{f_n\}$ *of the continuous functions*

$$f_n := \begin{cases} nt & \text{if } t \in [0, 1/n] \\ 1 & \text{if } t \in [1/n, 1] \end{cases}$$

Then for $n > m$

$$\|f_n - f_m\|_{CL[0,1]} = \int_{t=0}^{1} |f_n(t) - f_m(t)|\, dt =$$
$$\int_{t=0}^{1/n} |nt - mt|\, dt + \int_{t=1/n}^{1/m} |1 - mt|\, dt + \int_{t=1/m}^{1} |1 - 1|\, dt$$
$$\frac{(n-m)}{2n^2} + \frac{(1 - m/n)^2}{2m} = \frac{1}{2}\left(\frac{1}{m} - \frac{1}{n}\right) \to 0$$

as $n, m \to \infty$*. So,* $\{f_n\}$ *is a Cauchy sequence. However, its pointwise limit is*

$$f_n(t) \to \begin{cases} 1 & \text{if } 0 < t \le 1 \\ 0 & \text{if } t = 0 \end{cases}$$

In other words, the limit is a discontinuous function and, hence, it is not in $CL[0,1]$*. This means that the functional space* $CL[0,1]$ *is not complete.*

Example 18.3 *By the same reason, the spaces* $CL_p[0,1]$ *(the space of continuous and p-integrable functions) are not complete.*

18.2.3 Completion of metric spaces

There exist two possibilities to correct the situation and to provide the completeness property for a linear normed space if initially bit is not a complete:

- try to change the definition of a norm;
- try to extend the class of considered functions (it was suggested by Cauchy).

Changing of a norm

To illustrate the first approach related to changing of a norm let us consider again the space of all functions *continuous* at the interval $[0,1]$, but instead of the Lebesgue norm (18.13) we consider the Chebyshev's type norm $\|f\|_{C[a,b]}$ as in (18.5). This means that instead of the space $CL\,[0,1]$ we will consider the space $C\,[a,b]$ (18.5). Evidently, that this space is complete, since it is known that uniform convergent sequences of continuous functions converges to a continuous function. Hence, $C\,[a,b]$ is a *Banach space* under this norm.

Claim 18.1 *By the same reasons it is not difficult to show that all spaces $C^k\,[a,b]$ (18.6) are Banach.*

Claim 18.2 *The spaces $L_p\,[a,b]$ $(1 \leq p < \infty)$ (18.7), $L_\infty\,[a,b]$ (18.8), $\mathbb{L}_p^{m\times k}$, (23.19) and $\mathbb{L}_\infty^{m\times k}$ (18.12) are Banach too.*

Completion

Theorem 18.1 *Any linear normed space $\mathcal{X}$ with a norm $\|x\|_{\mathcal{X}}$ can be considered as a linear manifold which is complete in some Banach space $\hat{\mathcal{X}}$. This space $\hat{\mathcal{X}}$ is called the* ***completion*** *of $\mathcal{X}$.*

Proof. Consider two fundamental sequences $\{x_n\}$ and $\{x_n'\}$ with elements from $\mathcal{X}$. We say that they are *equivalent* if $\|x_n - x_n'\| \to 0$ as $n \to \infty$ and we will write $\{x_n\} \sim \{x_n'\}$. The set of all fundamental sequences may be separated (*factorized*) at non crossed classes: $\{x_n\}$ and $\{x_n'\}$ are included in the same class if and only if $\{x_n\} \sim \{x_n'\}$. The set of all such classes $\mathcal{X}_i$ we denoted by $\hat{\mathcal{X}}$. So,

$$\hat{\mathcal{X}} := \bigcup_i \mathcal{X}_i, \ \mathcal{X}_i \underset{i\neq j}{\cap} \mathcal{X}_j = \varnothing$$

Let us make the space $\hat{\mathcal{X}}$ a normed space. To do that, define the *operation of summing of the classes $\mathcal{X}_i$* by the following manner: if $\{x_n\} \in \mathcal{X}_i$ and $\{y_n\} \in \mathcal{X}_j$ then class $(\mathcal{X}_i + \mathcal{X}_j)$ may be defined as the class containing $\{x_n + y_n\}$. The *operation of the multiplication by a constant* may be introduced as follows: we denoted by $\lambda\mathcal{X}_i$ the class containing $\{\lambda x_n\}$ if $\{x_n\} \in \mathcal{X}_i$. It is evident that $\hat{\mathcal{X}}$ is *a linear space.* Define now the norm in $\hat{\mathcal{X}}$ as

$$\|\mathcal{X}_i\| := \lim_{n\to\infty} \|x_n\|_{\mathcal{X}} \quad (\{x_n\} \in \mathcal{X}_i)$$

It easy to check the norm axioms for such norm and to show that

a) $\mathcal{X}$ may be considered as a linear manifold in $\hat{\mathcal{X}}$;

b) $\mathcal{X}$ is dense in $\hat{\mathcal{X}}$, i.e., there exists $\{x_n\} \in \mathcal{X}$ such that $\|x_n - \mathcal{X}_i\|_{\mathcal{X}} \to 0$ as as $n \to \infty$ for some $\mathcal{X}_i \in \mathcal{X}$;

c) $\hat{\mathcal{X}}$ is complete (Banach).

This complete the proof. ■

This theorem can be interpreted as the following statement.

Corollary 18.1 *For any linear norm space $\mathcal{X}$ there exists a Banach space $\hat{\mathcal{X}}$ and a linear, injective map $T : \mathcal{X} \to \hat{\mathcal{X}}$ such that $T(\mathcal{X})$ is dense in $\hat{\mathcal{X}}$ and for all $x \in \mathcal{X}$*

$$\boxed{\|Tx\|_{\hat{\mathcal{X}}} = \|x\|_{\mathcal{X}}}$$

18.3 Hilbert spaces

18.3.1 Definition and examples

Definition 18.4 *A **Hilbert space** $\mathcal{H}$ is an inner (scalar) product space that is complete as a linear normed space under the **induced norm***

$$\boxed{\|z\|_{\mathcal{H}} := \sqrt{\langle z, z\rangle}} \tag{18.14}$$

Example 18.4 *The following spaces are Hilbert*

1. *The **space** l_2^n **of all summable complex sequences** (see (18.4) for $p = 2$) under the inner product*

$$\boxed{\langle x, y\rangle_{l_2^n} := \sum_{i=1}^{\infty} x_i \bar{y}_i} \tag{18.15}$$

2. *The **Lebesgue space** $L_2[a, b]$ of all integrable (in Lebesgue sense) complex functions (see (18.7) for $p = 2$) under the inner product*

$$\boxed{\langle x, y\rangle_{L_2[a,b]} := \int_{t=a}^{b} x(t)\bar{y}(t)dt} \tag{18.16}$$

3. *The **Sobolev's space** $S_2^l(G)$ **of** all l times differentiable on G quadratically integrable (in Lebesgue sense) complex functions (see (18.9) for $p = 2$) under the inner product*

$$\boxed{\langle x, y\rangle_{S_p^l(G)} := \sum_{i=0}^{l} \left\langle \frac{d^i}{dt^i}x, \frac{d^i}{dt^i}y \right\rangle_{L_2[a,b]}} \tag{18.17}$$

4. *The **frequency domain space** $\mathbb{L}_2^{m\times k}$ of all p-integrable complex matrices (23.19) under the inner product*

$$\boxed{\langle x, y\rangle_{\mathbb{L}_p^{m\times k}} := \int_{\omega=-\infty}^{\infty} tr\{X(j\omega)\, Y^{\sim}(j\omega)\}\, d\omega} \tag{18.18}$$

5. *The **Hardy spaces** $\mathbb{H}_2^{m\times k}$ (the subspace of $\mathbb{L}_2^{m\times k}$ containing only holomorphic in the right-hand semi-plan $\mathbb{C}^+ := \{s \in \mathbb{C} \mid \operatorname{Re} s > 0\}$ functions) under the inner product (18.18).*

18.3.2 Orthogonal complement

Definition 18.5 *Let $\mathcal{M}$ be a subset of a Hilbert space $\mathcal{H}$, i.e., $\mathcal{M} \subset \mathcal{H}$. Then the **distance** between a point $x \in \mathcal{H}$ and $\mathcal{M}$ is defined by*

$$\boxed{\rho(x, \mathcal{M}) := \inf_{y\in\mathcal{M}} \|x - y\|} \tag{18.19}$$

The following claim seems to be evident.

Claim 18.3 *If $x \in \mathcal{M}$, then $\rho(x, \mathcal{M}) = 0$. If $x \notin \mathcal{M}$ and $\mathcal{M}$ is closed set (see Definition 14.7), then $\rho(x, \mathcal{M}) > 0$.*

Corollary 18.2 *If $\mathcal{M} \subset \mathcal{H}$ is closed convex set and $x \notin \mathcal{M}$, then there exists a unique element $y \in \mathcal{M}$ such that $\rho(x, \mathcal{M}) = \|x - y\|$.*

Proof. Indeed, suppose that there exists another element $y^* \in \mathcal{M}$ such that

$$\rho(x, \mathcal{M}) = \|x - y\| = \|x - y^*\| := d$$

Then

$$4d^2 = 2\|x-y\|^2 + 2\|x-y^*\|^2 = \|x-y^*\|^2 + 4\left\|x - \frac{y+y^*}{2}\right\|^2 \geq \|x-y^*\|^2 + 4\inf_{y\in\mathcal{M}} \|x-y\|^2 \geq \|x-y^*\|^2 + 4d^2$$

that gives $\|x-y^*\|^2 \leq 0$, or equivalently, $y = y^*$. ■

Corollary 18.3 *If $\mathcal{M} \subset \mathcal{H}$ is a subspace of $\mathcal{H}$ (this means that it is closed convex linear manifold in $\mathcal{H}$) then for any $x \in \mathcal{H}$ there exists a unique element $x_{\mathcal{M}} \in \mathcal{M}$ such that*

$$\boxed{\rho(x, \mathcal{M}) := \inf_{y\in\mathcal{M}} \|x-y\| = \|x - x_{\mathcal{M}}\|} \tag{18.20}$$

This element $x_{\mathcal{M}} \in \mathcal{M}$ is called the ***orthogonal projection*** *of the element $x \in \mathcal{H}$ onto the subspace $\mathcal{M} \subset \mathcal{H}$.*

Lemma 18.1 *Let $\rho(x, \mathcal{M}) = \|x - x_{\mathcal{M}}\|$ where $\mathcal{M}$ is a subspace of a Hilbert space $\mathcal{H}$ with the inner product $\langle x, y\rangle_{\mathcal{H}}$. Then $(x - x_{\mathcal{M}}) \perp \mathcal{M}$, that is, for any $y \in \mathcal{M}$*

$$\boxed{\langle x - x_{\mathcal{M}}, y\rangle_{\mathcal{H}} = 0} \tag{18.21}$$

Proof. By the definition (18.20) for any $\lambda \in \mathbb{C}$ (here $x_{\mathcal{M}} + \lambda y \in \mathcal{M}$) we have

$$\|x - (x_{\mathcal{M}} + \lambda y)\| \geq \|x - x_{\mathcal{M}}\|$$

that implies

$$\lambda \langle x - x_{\mathcal{M}}, y\rangle_{\mathcal{H}} + \bar{\lambda}\langle y, x - x_{\mathcal{M}}\rangle_{\mathcal{H}} + \lambda\bar{\lambda}\|y\|^2 \geq 0$$

Taking $\lambda = -\dfrac{\langle x - x_{\mathcal{M}}, y\rangle_{\mathcal{H}}}{\|y\|^2}$ one has $-\dfrac{|\langle x - x_{\mathcal{M}}, y\rangle|^2}{\|y\|^2} \geq 0$ that leads to the equality $\langle x - x_{\mathcal{M}}, y\rangle_{\mathcal{H}} = 0$. Lemma is proven. ■

Definition 18.6 *If $\mathcal{M}$ is a subspace of a Hilbert space $\mathcal{H}$ then the* ***orthogonal complement*** *$\mathcal{M}^{\perp}$ is defined by*

$$\boxed{\mathcal{M}^{\perp} := \{x \in \mathcal{H} \mid \langle x, y\rangle_{\mathcal{H}} = 0 \text{ for all } y \in \mathcal{M}\}} \tag{18.22}$$

It is easy to show that $\mathcal{M}^{\perp}$ is a closed linear subspace of and that $\mathcal{H}$ can be uniquely decomposed as the direct sum

$$\boxed{\mathcal{H} = \bar{\mathcal{M}} \oplus \mathcal{M}^{\perp}} \tag{18.23}$$

This means that any element $x \in \mathcal{H}$ has the unique representation

$$\boxed{x = x_{\bar{\mathcal{M}}} + x_{\mathcal{M}^{\perp}}} \tag{18.24}$$

where $x_{\mathcal{M}} \in \bar{\mathcal{M}}$ and $x_{\mathcal{M}^{\perp}} \in \mathcal{M}^{\perp}$ such that $\|x\|^2 = \|x_{\mathcal{M}}\|^2 + \|x_{\mathcal{M}^{\perp}}\|^2$.

Theorem 18.2 *Let If $\mathcal{M}$ is a subspace of a Hilbert space $\mathcal{H}$. $\mathcal{M}$ is dense in $\mathcal{H}$ if and only if $\mathcal{M}^{\perp} = \{0\}$.*

Proof. *a) Necessity.* Let $\mathcal{M}$ is dense in $\mathcal{H}$. This means that $\bar{\mathcal{M}} = \mathcal{H}$. Assume that there exists $x_0 \in \mathcal{H}$ such that $x_0 \perp \mathcal{M}$. Let $\{y_n\} \subset \mathcal{M}$ and $y_n \to y \in \mathcal{H}$. Then $0 = \langle y_n, x_0 \rangle \to \langle y, x_0 \rangle = 0$ since $\mathcal{M}$ is dense in $\mathcal{H}$. Taking $y = x_0$ we get that $\langle x_0, x_0 \rangle = 0$ that gives $x_0 = 0$. *b) Sufficiency.* Let $\mathcal{M}^{\perp} = \{0\}$, that is, if $\langle y, x_0 \rangle = 0$ for any $y \in \mathcal{M}$, then $x_0 = 0$. Suppose that $\mathcal{M}$ is not dense in $\mathcal{H}$. This means that there exists $x_0 \notin \bar{\mathcal{M}}$. Then by the orthogonal decomposition $x_0 = y_0 + z_0$ where $y_0 \in \bar{\mathcal{M}}$ and $z_0 \in \left(\bar{\mathcal{M}}\right)^{\perp} = \mathcal{M}^{\perp}$. Here $z_0 \neq 0$ for which $\langle z_0, y \rangle_{\mathcal{H}} = 0$ for any $y \in \bar{\mathcal{M}}$. By the assumption such element $z_0 = 0$. We get the contradiction. Theorem is proven. ■

18.3.3 Fourier series in Hilbert spaces

Definition 18.7 *An* ***orthonormal system (set)*** *$\{\phi_n\}$ of functions in a Hilbert space $\mathcal{H}$ is a nonempty subset $\{\phi_n \mid n \geq 1\}$ of $\mathcal{H}$ such that*

$$\boxed{\langle \phi_n, \phi_m \rangle_{\mathcal{H}} = \delta_{n,m} = \begin{cases} 1 & if \quad n = m \\ 0 & if \quad n \neq m \end{cases}} \tag{18.25}$$

1. *The series $\sum_{n=1}^{\infty} \alpha_k \phi_n$ is called the series in $\mathcal{H}$ with respect of the system $\{\phi_n\}$ (18.25);*

2. *For any $x \in \mathcal{H}$ the representation (if it exists)*

$$\boxed{x(t) = \sum_{n=1}^{\infty} \alpha_n \phi_n(t)} \tag{18.26}$$

is called the ***Fourier expansion*** *of x with respect to $\{\phi_n\}$.*

Lemma 18.2 *In (18.26)*

$$\boxed{\alpha_k = \langle x, \phi_k \rangle_{\mathcal{H}}} \tag{18.27}$$

Proof. Premultiplying (18.26) by ϕ_k and using (18.25) we find

$$\langle x, \phi_k \rangle_{\mathcal{H}} = \sum_{n=1}^{\infty} \alpha_n \langle \phi_k, \phi_n \rangle_{\mathcal{H}} = \sum_{k=1}^{\infty} \alpha_k \delta_{k,n} = \alpha_k$$

that proves (18.27). Lemma is proven. ■

Corollary 18.4 (the Parseval equality)

$$\boxed{\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, \phi_n \rangle_{\mathcal{H}}|^2} \tag{18.28}$$

Proof. It follows from the relation

$$\langle x, y \rangle_{\mathcal{H}} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \langle x, \phi_n \rangle_{\mathcal{H}} \overline{\langle x, \phi_m \rangle_{\mathcal{H}}} \langle \phi_n, \phi_m \rangle_{\mathcal{H}} =$$
$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \langle x, \phi_n \rangle_{\mathcal{H}} \overline{\langle x, \phi_m \rangle_{\mathcal{H}}} \delta_{n.m} = \sum_{n=1}^{\infty} |\langle x, \phi_n \rangle_{\mathcal{H}}|^2$$

■

Example 18.5

1) ***Classical Fourier expansion.*** *In $\mathcal{H} = L_2[0,1]$ the corresponding orthogonal basis $\{\phi_n\}$ is*

$$\boxed{\{\phi_n\} = \left\{1,\ \sqrt{2}\sin(2\pi nt),\ \sqrt{2}\cos(2\pi nt),\ n \geq 1\right\}}$$

that implies

$$x(t) = a_0 + \sqrt{2}\sum_{n=1}^{\infty} a_n \sin(2\pi nt) + \sqrt{2}\sum_{n=1}^{\infty} b_n \cos(2\pi nt)$$

where

$$a_0 = \int_{t=0}^{1} x(t)\,dt,\ a_n = \int_{t=0}^{1} x(t)\sqrt{2}\cos(2\pi nt)\,dt$$
$$b_n = \int_{t=0}^{1} x(t)\sqrt{2}\sin(2\pi nt)\,dt$$

*2) **Legandre expansion.** In* $\mathcal{H} = L_2[0,1]$ *the corresponding orthogonal basis* $\{\phi_n\}$ *is* $\{\phi_n\} = \{p_n\}$ *where*

$$\boxed{p_k := \frac{1}{2^k k!} \frac{d^k}{dt^k} \left[\left(t^2 - 1\right)^k \right], \ k \geq 1}$$

18.3.4 Linear n-manifold approximation

Definition 18.8 *The collection of the elements*

$$\boxed{u_n := \sum_{k=1}^{n} c_k \phi_k \in \mathcal{H}, \ c_k \in \mathbb{C} \ (k \geq 1)} \tag{18.29}$$

*is called the **linear** n**-manifold** generated by the system of functions* $\{\phi_k\}_{k=\overline{1,n}}$.

Theorem 18.3 *The best* L_2*-approximation of any elements* $x \in \mathcal{H}$ *by the element* u_n *from the n-manifold (18.29) is given by the Fourier coefficients* $c_k = \alpha_k$ *(18.27), namely,*

$$\boxed{\inf_{c_k : k=\overline{1,n}} \left\| x - \sum_{k=1}^{n} c_k \phi_k \right\|_{L_2}^2 = \left\| x - \sum_{k=1}^{n} \alpha_k \phi_k \right\|_{L_2}^2} \tag{18.30}$$

Proof. It follows from the identity

$$\|x - u_n\|_{L_2}^2 = \left\| x - \sum_{k=1}^{n} c_k \phi_k \right\|_{L_2}^2 =$$
$$\left\| \sum_{n=1}^{\infty} \alpha_n \phi_n(t) - \sum_{k=1}^{n} c_k \phi_k \right\|_{L_2}^2 =$$
$$\sum_{k=n+1}^{\infty} |\alpha_k|^2 \|\phi_k\|_{L_2}^2 + \sum_{k=1}^{n} |\alpha_k - c_k|^2 \|\phi_k\|_{L_2}^2$$

that reaches the minimum if $c_k = \alpha_k \ \left(c_k : k = \overline{1,n}\right)$. Theorem is proven. ■

18.4 Linear operators and functionals in Banach spaces

18.4.1 Operators and functionals

Definition 18.9

1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be be linear normed spaces (usually, either Banach or Hilbert spaces) and $T : \mathcal{D} \to \mathcal{Y}$ be a* ***transformation*** *(or* ***operator****) from a subset $\mathcal{D} \subset \mathcal{X}$ to $\mathcal{Y}$. $\mathcal{D} = \mathcal{D}(T)$ is called the* ***domain*** *(image) of the operator T and values $T(\mathcal{D})$ constitute the* ***range*** *(the set of possible values) $\mathcal{R}(T)$ of T. If the range of the operator T is finite-dimensional the we say that the operator has the* ***finite range****.*

2. *If $\mathcal{Y}$ is a scalar field $\mathcal{F}$ (usually $\mathbb{R}$) then the transformations T are called* ***functionals****.*

3. *A functional T is* ***linear*** *if it is additive, i.e., for any $x, y \in \mathcal{D}$*

$$\boxed{T(x+y) = Tx + Ty}$$

and homogeneous, i.e., for any $x \in \mathcal{D}$ and any $\lambda \in \mathcal{F}$

$$\boxed{T(\lambda x) = \lambda Tx}$$

4. *Operators for which the domain $\mathcal{D}$ and the range $T(\mathcal{D})$ are in one-to-one correspondence are called* ***invertible****. The* ***inverse operator*** *is denoted by $T^{-1} : T(\mathcal{D}) \to \mathcal{D}$, so that*

$$\boxed{\mathcal{D} \supseteq T^{-1}(T(\mathcal{D}))}$$

Example 18.6

1. *The* ***shift operator*** *$T_{sh} : l_p^n \to l_p^n$ defined by*

$$\boxed{T_{sh} x_i = x_{i+1}}$$

for any $i = 1, 2, \ldots$.

2. *The* ***integral operator*** $T_g : L_2[a,b] \to \mathbb{R}$ *defined by*

$$\boxed{T_g f := \int_{t=a}^{b} f(t)\, g(t)\, dt}$$

for any $f, g \in L_2[a,b]$.

3. *The* ***differential operator*** $T_d : \mathcal{D}(T) = C^1[a,b] \to C[a,b]$ *defined by*

$$\boxed{T_d f := \frac{d}{dt} f(t)}$$

for any $f \in C^1[a,b]$ *and any* $t \in [a,b]$.

It is evident the following statement.

Claim 18.4

1. T *is invertible if and only if it is* ***injective****, that is,* $Tx = 0$ *implies* $x = 0$*. The set* $\{x \in \mathcal{D} \mid Tx = 0\}$ *is called the kernel of the operator and denoted by*

$$\boxed{\ker T := \{x \in \mathcal{D} \mid Tx = 0\}}$$

So, T *is injective if and only if* $\ker T = \{0\}$.

2. *If* T *is linear and invertible then* T^{-1} *is also linear.*

18.4.2 Continuity and boundedness

Continuity

Definition 18.10

1. *Let* $T : \mathcal{D}(T) \to \mathcal{Y}$ *be a map* (**operator**) *between two linear normed spaces* $\mathcal{X}$ *(with a norm* $\|\cdot\|_{\mathcal{X}}$*) and* $\mathcal{Y}$ *(with a norm* $\|\cdot\|_{\mathcal{Y}}$*). It is said to be a* **continuous at** $x_0 \in \mathcal{X}$ *if, given* $\varepsilon > 0$*, there exists a* $\delta = \delta(\varepsilon) > 0$ *such that* $\|T(x) - T(x_0)\|_{\mathcal{Y}} < \varepsilon$*, whenever* $\|x - x_0\|_{\mathcal{X}} < \delta$.

2. *T is **semi-continuous** at a point $x_0 \in \mathcal{X}$ if it transforms any convergent sequence $\{x_n\} \subset \mathcal{D}(T)$, $x_n \to x_0$, $n \to \infty$ in to a sequence $\{T(x_n)\} \subset \mathcal{R}(T)$ weakly convergent to $T(x_0)$, i.e., $\|T(x_n) \to T(x_0)\| \to 0$ when $n \to \infty$.*

3. *T is **continuous** (or **semi-continuous**) **on** $\mathcal{D}(T)$ if it is continuous (or semi-continuous) at every point in $\mathcal{D}(T)$.*

Lemma 18.3 *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces and A be a linear operator defined at $\mathcal{X}$. If A is **continuous** at the point $0 \in \mathcal{X}$, then A is continuous at any point $x_0 \in \mathcal{X}$.*

Proof. This result follows from the identity $Ax - Ax_0 = A(x - x_0)$. If $x \to x_0$, then $z := x - x_0 \to 0$. By continuity at zero $Az \to 0$ that implies $Ax - Ax_0 \to 0$. Lemma is proven. ■

So, a linear operator A may be called *continuous*, if it is continuous at the point $x_0 = 0$.

Boundedness

Definition 18.11

1. *A linear operator $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{Y}$ between two linear normed spaces $\mathcal{X}$ (with a norm $\|\cdot\|_{\mathcal{X}}$) and $\mathcal{Y}$ (with a norm $\|\cdot\|_{\mathcal{Y}}$) is said to be **bounded** if there exists a real number $c > 0$ such that for all $x \in \mathcal{D}(A)$*
$$\boxed{\|Ax\|_{\mathcal{Y}} \leq c\,\|x\|_{\mathcal{X}}} \tag{18.31}$$
*The **set of all bounded linear operators** $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{Y}$ is usually denoted by $\mathcal{L}(\mathcal{X}, \mathcal{Y})$.*

2. *A linear operator $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{Y}$ is called a **compact operator** if it maps any bounded subset of $\mathcal{X}$ onto a compact set of $\mathcal{Y}$.*

3. *The **induced norm** of a linear bounded operator $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{Y}$ may be introduced as follows*
$$\boxed{\|A\| := \sup_{x \in \mathcal{D}(A),\, x \neq 0} \frac{\|Ax\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}} = \sup_{x \in \mathcal{D}(A),\, \|x\|_{\mathcal{X}} = 1} \|Ax\|_{\mathcal{Y}}} \tag{18.32}$$

(here it is assumed that if $\mathcal{D}(A) = \{0\}$ then by the definition $\|A\| = 0$ since $A0 = 0$).

It seems to be evident that the continuity and boundedness for linear operators are equivalent concepts.

Claim 18.5 *A linear operator $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{Y}$ is continuous if and only if it is bounded.*

Example 18.7

1. *If $\beta := \left(\sum_{i,j=1}^{\infty} |a_{ij}|^q \right)^{1/q} < \infty$ $(q > 1)$, then the **"weighting" operator** A defined by*

$$\boxed{y = Ax_i := \sum_{j=1}^{\infty} a_{ij} x_j} \tag{18.33}$$

making from l_p to l_q $(p^{-1} + q^{-1} = 1)$ is linear and bounded since by the Hölder inequality (16.134)

$$\begin{aligned} \|Ax_i\|_{l_q}^q &= \sum_{i=1}^{\infty} \left| \sum_{j=1}^{\infty} a_{ij} x_j \right|^q \le \sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} |a_{ij}|^q \right) \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{q/p} \\ &= \beta \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{q/p} = \beta \|x\|_{l_p}^q \end{aligned}$$

2. *If $\beta := \int_{x=a}^{b} \int_{s=a}^{b} |K(x,s)|^q \, ds dx < \infty$, then the **integral operator** $A : \mathcal{X} = L_p[a,b] \to \mathcal{Y} = L_q[a,b] = \mathcal{Y}$ $(p^{-1} + q^{-1} = 1)$ defined by*

$$\boxed{y = Af := \int_{s=a}^{b} K(x,s) f(s) \, ds} \tag{18.34}$$

is linear and bounded since by the Hölder inequality (16.134)

$$\|Af\|_{L_q[a,b]}^q := \int_{x=a}^{b} \left| \int_{s=a}^{b} K(x,s) f(s) \, ds \right|^q dx \le$$

$$\int_{x=a}^{b} \left(\int_{s=a}^{b} |K(x,s)|^q \, ds \right) \left(\int_{s=a}^{b} |f(s)|^p \, ds \right)^{q/p} dx = \beta \|f\|_{L_p[a,b]}^q$$

3. *If* $\beta := \max_{t\in\bar{\mathcal{D}}} \sum_{\alpha=0}^{l} |a_\alpha (t)| < \infty$*, then the* ***differential operator*** $A : \mathcal{D} \subset \mathcal{X} = C^k [a,b] \to \mathcal{Y} = C[a,b] = \mathcal{Y}$ *defined by*

$$\boxed{y = Af := \sum_{\alpha=0}^{l} a_\alpha (t) f^{(\alpha)} (t)} \tag{18.35}$$

is linear and bounded since

$$\|Af\|_{C[a,b]} := \max_{t\in\bar{\mathcal{D}}} \left| \sum_{\alpha=0}^{l} a_\alpha (t) f^{(\alpha)} (t) \right| \leq$$
$$\max_{t\in\bar{\mathcal{D}}} \left(\sum_{\alpha=0}^{l} |a_\alpha (t)| \sum_{\alpha=0}^{l} \left| f^{(\alpha)} (t) \right| \right) \leq \beta \|f\|_{C^l[a,b]}$$

Sequence of linear operators and uniform convergence

It is possible to introduce several different notions of a convergence in the space of linear bounded operators $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ acting from $\mathcal{X}$ to $\mathcal{Y}$.

Definition 18.12 *Let* $\{A_n\} \subset \mathcal{L}(\mathcal{X}, \mathcal{Y})$ *be a sequence of operators.*

1. *We say that*

   - A_n ***uniformly converges*** *to* $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ *if* $\|A_n - A\| \to 0$ *whenever* $n \to \infty$*. Here the norm* $\|A_n - A\|$ *is understood as in (18.32);*

   - A_n ***strongly converges*** *to* $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ *if* $\|A_n f - Af\|_{\mathcal{Y}} \to 0$ *whenever* $n \to \infty$ *for any* $f \in \mathcal{X}$*.*

2. *If the operator* A *is dependent on the parameter* $\alpha \in \mathcal{A}$*, then*

   - $A(\alpha)$ *is* ***uniformly continuous*** *at* $\alpha_0 \in \mathcal{A}$*, if*

   $$\boxed{\|A(\alpha) - A(\alpha_0)\| \to 0 \text{ as } \alpha \to \alpha_0}$$

   - $A(\alpha)$ *is* ***strongly continuous*** *at* $\alpha_0 \in \mathcal{A}$*, if for all* $f \in \mathcal{X}$

   $$\boxed{\|A(\alpha) f - A(\alpha_0) f\|_{\mathcal{Y}} \to 0 \text{ as } \alpha \to \alpha_0}$$

In view of this definition the following claim seems to be evident.

Claim 18.6 *A_n uniformly converges to $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ if and only if $A_n f \to Af$ uniformly on $f \in \mathcal{X}$ in the ball $\|f\|_{\mathcal{X}} \leq 1$.*

Theorem 18.4 *If $\mathcal{X}$ is a linear normed space and $\mathcal{Y}$ is a Banach space, then $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ is a Banach space too.*

Proof. Let $\{A_n\}$ be a fundamental sequence in the metric of $\mathcal{L}(\mathcal{X}, \mathcal{Y})$, that is, for any $\varepsilon > 0$ there exists a number $n_0 = n_0(\varepsilon)$ such that for any $n > n_0$ and any natural p we have $\|A_{n+p} - A\| < \varepsilon$. Then the sequence $\{A_n f\}$ is also fundamental. But $\mathcal{Y}$ is complete, and hence, $\{A_n f\}$ converges. Denote $y := \lim_{n\to\infty} A_n f$. By this formula any element $f \in \mathcal{X}$ is mapped into an element of $\mathcal{Y}$, and, hence, it defines the operator $y = Af$. Let us prove that the linear operator A is bounded (continuous). First, notice that $\{\|A_n\|\}$ is also fundamental. This follows from the inequality $|\|A_{n+p}\| - \|A_n\|| \leq \|A_{n+p} - A_n\|$. But it means that $\{\|A_n\|\}$ is bounded, that is, there exists $c > 0$ such that $\|A_n\| \leq c$ for every $n \geq 1$. Hence, $\|A_n f\| \leq c\|f\|$. Taking the limit in the right-hand side we obtain $\|Af\| \leq c\|f\|$ that shows that A is bounded. Theorem is proven. ■

Extension of linear bounded operators

Bounded linear operators that map into a Banach space always have a unique extension to the closure of their domain without changing of its norm value.

Theorem 18.5 *Let $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{Y}$ be a linear bounded operator (functional) mapping the linear normed space $\mathcal{X}$ into a Banach space $\mathcal{Y}$. Then it has a unique bounded extension $\tilde{A} : \overline{\mathcal{D}(A)} \to \mathcal{Y}$ such that*

1. *$\tilde{A}f = Af$ for any $f \in \mathcal{D}(A)$;*
2. *$\left\|\tilde{A}\right\| = \|A\|$.*

Proof. If $f \in \mathcal{D}(A)$, put $\tilde{A}f = Af$. Let $f \in \mathcal{X}$, but $x \notin \mathcal{D}(A)$. By the density of $\mathcal{D}(A)$ in $\mathcal{X}$, there exists the sequence $\{f_n\} \subset \mathcal{D}(A)$ converging to x. Put $\tilde{A}f = \lim_{n\to\infty} Af_n$. Let us show that this definition is correct, namely, that the limit exists and it does not depend on the selection of the convergent sequence $\{f_n\}$. The existence follows from the completeness property of $\mathcal{Y}$ since $\|Af_n - Af_m\| \leq \|A\| \|f_n - f_m\|_{\mathcal{X}}$. Hence, $\lim_{n\to\infty} Af_n$ exists. Supposing that there exists another sequences $\{f'_n\} \subset \mathcal{D}(A)$ converging to f we may denote $a := \lim_{n\to\infty} Af_n$ and $b := \lim_{n\to\infty} Af'_n$. Then we get

$$\|a - b\| \leq \|a - Af_n\| + \|Af_n - Af'_n\| + \|Af'_n - b\| \underset{n\to\infty}{\to} 0$$

But $\|Af_n\| \leq \|A\| \|f_n\|$ that for $n \to \infty$ implies $\left\|\tilde{A}f\right\| \leq \|A\| \|f\|$, or equivalently, $\left\|\tilde{A}\right\| \leq \|A\|$. We also have $\left\|\tilde{A}\right\| := \sup_{\|f\|_{\mathcal{X}}\leq 1} \left\|\tilde{A}f\right\| \geq \sup_{f\in\mathcal{D}(A),\, \|f\|_{\mathcal{X}}\leq 1} \left\|\tilde{A}f\right\| = \|A\|$. So, we have $\left\|\tilde{A}\right\| = \|A\|$. The linearity property of $\tilde{A}$ follows from the linearity of A. Theorem is proven. ■

Definition 18.13 *The operator $\tilde{A}$ constructed in the theorem 18.5 is called the* ***extension*** *of A to the closure $\overline{\mathcal{D}(A)}$ of its domain $\mathcal{D}(A)$ without increasing its norm.*

The principally more complex case arises when $\overline{\mathcal{D}(A)} = \mathcal{X}$. The following important theorem says that any linear bounded functional (operator) can be extended to the *whole* space $\mathcal{X}$ without increasing into norm. A consequence of this result is the existence of nontrivial linear bounded functionals on any normed linear space.

Theorem 18.6 (The Hahn-Banach theorem) *Any linear bounded functional $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{Y}$ defined on a linear subspace $\mathcal{D}(A)$ of a linear normed space $\mathcal{X}$ can be extended to a linear bounded functional $\tilde{A}$ defined on the whole $\mathcal{X}$ with the preservation of the norm, i.e., $\tilde{A}f = Af$ for any $f \in \mathcal{D}(A)$ such that $\left\|\tilde{A}\right\| = \|A\|$.*

Proof. Here we present only the main idea of the proof.

a) If $\mathcal{X}$ is separable, then the proof is based on the theorem 18.5 using the following lemma.

Lemma 18.4 *Let $\mathcal{X}$ is a real normed space and $\mathcal{L}$ is a linear manifold in $\mathcal{X}$ where there is defined a linear functional A. If $f_0 \notin \mathcal{L}$ and $\mathcal{L}_1 := \{f + tf_0 \mid f \in \mathcal{L},\ t \in \mathbb{R}\}$ is a linear manifold containing all elements $f + tf_0$, then there exists a linear bounded functional A_1 defined on $\mathcal{L}_1$ such that it coincides with A on $\mathcal{L}$ and preserving the norm on $\mathcal{L}_1$, namely, $\|A_1\| = \|A\|$.*

Then, since $\mathcal{X}$ is separable, there exists a basis $\{f_n\}_{n\geq 1}$ such that we can construct the sequence of s-manifolds

$$\mathcal{L}_{s\geq 1} := \left\{ \sum_{i=1}^{s} \lambda_i f_i \mid f_i \in \mathcal{X},\ \lambda_i \in \mathbb{R} \right\}$$

connected by $\mathcal{L}_{s+1} = \mathcal{L}_s + \{f_{n+1}\}$, $\mathcal{L}_0 := \varnothing$. Then we make the extension of A to each of the subspaces $\mathcal{L}_{s\geq 1}$ based on the lemma above. Finally we apply the theorem 18.5 to the space $\mathcal{X} = \underset{s\geq 1}{\cup} \mathcal{L}_s$ using the density property of $\mathcal{X}$.

b) In general case, the proof is based on the Zorn's lemma (see (Yoshida 1979)). ■

Corollary 18.5 *Let $\mathcal{X}$ be a normed (topological) space and $x \in \mathcal{X}$, $x \neq 0$. Then there exists a linear bounded functional f, defined on $\mathcal{X}$, such that its value at any point x is equal to*

$$\boxed{f(x) := \langle x, f \rangle = \|x\|} \tag{18.36}$$

and

$$\boxed{\|f\| := \sup_{x \in D(f), \|x\| \leq 1} \langle x, f \rangle = 1} \tag{18.37}$$

Proof. Consider the linear manifold $\mathcal{L} := \{tx\}$, $t \in \mathbb{R}$ where we define f as follows: $\langle tx, f \rangle = t\|x\|$. So, we have $\langle x, f \rangle = \|x\|$. Then for any $y = tx$ it follows $|\langle y, f \rangle| = |t| \cdot \|x\| = \|tx\| = \|y\|$. This means that $\|f\| = 1$ and completes the proof. ■

Corollary 18.6 *Let in a normed space $\mathcal{X}$ there is defined a linear manifold $\mathcal{L}$ and the element $x_0 \notin \mathcal{L}$ having the distance d up to this manifold, that is, $d := \inf_{x \in \mathcal{L}} \|x - x_0\|$. Then there exists a linear functional f defined on the whole $\mathcal{X}$ such that*

1. $\langle x, f \rangle = 0$ *for any* $x \in \mathcal{L}$
2. $\langle x_0, f \rangle = 1$
3. $\|f\| = 1/d$

Proof. Take $\mathcal{L}_1 := \mathcal{L} + \{x_0\}$. Then any element $y \in \mathcal{L}_1$ is uniquely defined by $y = x + tx_0$ where $x \in \mathcal{L}$ and $t \in \mathbb{R}$. Define on $\mathcal{L}_1$ the functional $f := t$. Now, if $y \in \mathcal{L}$, then $t = 0$ and $\langle y, f \rangle = 0$. So, the statement 1 holds. If $y = x_0$, then $t = 1$ and, hence, $\langle x_0, f \rangle = 1$ that verifies the statement 2. Finally,

$$|\langle y, f \rangle| = |t| = \frac{|t| \cdot \|y\|}{\|y\|} = \frac{\|y\|}{\left\| \frac{x}{t} + x_0 \right\|} \leq \frac{\|y\|}{d}$$

that gives $\|f\| \leq 1/d$. On the other hand, by the "inf" definition, there exists a sequence $\{x_n\} \in \mathcal{L}$ such that $d = \lim_{n \to \infty} \|x_n - x_0\|$. This implies

$$1 = \langle x_0 - x_n, f \rangle \leq \|x_n - x_0\| \cdot \|f\|$$

Taking limit in the last inequality we obtain $1 \leq d \|f\|$ that gives $\|f\| \geq 1/d$. Combining both inequalities we conclude the statement 3. Corollary is proven. ■

Corollary 18.7 *A linear manifold $\mathcal{L}$ is not **dense** in a Banach space $\mathcal{X}$ if and only if there exists a linear bounded functional $f \neq 0$ such that $\langle x, f \rangle = 0$ for any $x \in \mathcal{L}$.*

Proof. *a) Necessity.* Let $\bar{\mathcal{L}} \neq \mathcal{X}$. Then there exists a point $x_0 \in \mathcal{X}$ such that the distance between x_0 and $\mathcal{L}$ is positive, namely, $\rho(x_0, \mathcal{L}) = d > 0$. By the Corollary 18.6 there exists f such that $\langle x_0, f \rangle = 1$, that is, $f \neq 0$ but $\langle x, f \rangle = 0$ for any $x \in \mathcal{L}$. *b) Sufficiency.* Let now $\bar{\mathcal{L}} = \mathcal{X}$. Then for any $x \in \mathcal{X}$, in view of the density property, there exists $\{x_n\} \in \mathcal{L}$ such that $x_n \to x$ when $n \to \infty$. By the conditions that there exists $f \neq 0$, $= 0$ for any $x \in \mathcal{L}$, we have $\langle x, f \rangle = \lim_{n \to \infty} \langle x_n, f \rangle = 0$. Since x is arbitrary, it follows that $f = 0$. Contradiction. Corollary is proven. ■

Corollary 18.8 *Let $\{x_k\}_1^n$ be a system of linearly independent elements in a normed space $\mathcal{X}$. Then there exists a system of linear bounded functionals $\{f_l\}_1^n$, defined on the whole $\mathcal{X}$, such that*

$$\boxed{\langle x_k, f_l \rangle = \delta_{kl} \ (k, l = 1, ..., n)} \tag{18.38}$$

*These two systems $\{x_k\}_1^n$ and $\{f_l\}_1^n$ are called **bi-orthogonal**.*

Proof. Take x_1 and denote by L_1 the linear span of the elements $x_2, ..., x_n$. By the linear independency, it follows that $\rho(x_1, L_1) > 0$. By the Corollary 18.6 we can find the linear bounded functional f_1 such that $\langle x_1, f_1 \rangle = 1$, $\langle x, f_1 \rangle = 0$ on L_1. Iterating this process we construct the desired system $\{f_l\}_1^n$. ■

18.4.3 Compact operators

In this subsection we will consider a special subclass of bounded linear operators having properties rather similar to those enjoyed by operators on finite-dimensional spaces.

Definition 18.14 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces. An operator $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is said to be a* ***compact operator*** *if A maps bounded set of $\mathcal{X}$ onto* ***relative compact sets*** *of $\mathcal{Y}$, that is, A is linear and for any bounded sequence $\{x_n\}$ in $\mathcal{X}$ the sequence $\{Ax_n\}$ has a convergence subsequence in $\mathcal{Y}$.*

Claim 18.7 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces and $A : \mathcal{X} \to \mathcal{Y}$ be a linear operator. Then the following assertions holds:*

a) *If A is bounded, that is, $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and $\dim(Ax) < \infty$, then the operator A is compact.*

b) *If $\dim(\mathcal{X}) < \infty$, then A is compact.*

c) *The range of A is separable if A is compact.*

d) *If $\{A_n\}$ is a sequence of compact operators from $\mathcal{X}$ to Banach space $\mathcal{Y}$ that converges uniformly to A, then A is a compact operator.*

e) *The identity operator I on the Banach space $\mathcal{X}$ is compact if and only if $\dim(\mathcal{X}) < \infty$.*

f) *If A is a compact operator in $\mathcal{L}(, \mathcal{Y})$ whose range is closed subspace of $\mathcal{Y}$, then the range of A is finite-dimensional.*

Proof. It can be found in (Rudin 1976) and (Yoshida 1979). ■

Example 18.8

1. *Let* $\mathcal{X} = l_2$ *and* $A : l_2 \to l_2$ *is defined by* $Ax := \left(x_1, \frac{x_2}{2}, \frac{x_3}{3}, ...\right)$. *Then* A *is compact. Indeed, defining* A_n *by*

$$A_n x := \left(x_1, \frac{x_2}{2}, \frac{x_3}{3}, ..., \frac{x_n}{n}, 0, 0, ...\right)$$

we have

$$\|Ax - A_n x\|^2 = \sum_{k=n+1}^{\infty} \frac{1}{k^2} |x_i|^2 \leq \frac{\|x\|^2}{(n+1)^2}$$

and, hence, $\|A - A_n\| \leq (n+1)^{-1}$. *This means that* A_n *converges uniformly to* A *and, by the previous claim (d),* A *is compact.*

2. *Let* $k(t, s) \in L_2([a, b] \times [a, b])$. *Then the integral operator* $K : L_2([a, b]) \to L_2([a, b])$ *defined by* $(Ku)(t) := \int\limits_{s=a}^{b} k(t, s) u(s) ds$ *is a compact operator (see (Yoshida 1979)).*

Theorem 18.7 (Approximation theorem) *Let* $\Phi : \mathcal{M} \subset \mathcal{X} \to \mathcal{Y}$ *be a compact operator where* $\mathcal{X}$,$\mathcal{Y}$ *are Banach spaces and* $\mathcal{M}$ *is a bounded nonempty subset of* $\mathcal{X}$. *Then for every* $n = 1, 2, ...$ *there exists a continuous operator* $\Phi_n : \mathcal{M} \to \mathcal{Y}$ *such that*

$$\boxed{\sup_{x \in \mathcal{M}} \|\Phi(x) - \Phi_n(x)\| \leq n^{-1} \text{ and } \dim(\text{span}\Phi_n(\mathcal{M})) < \infty} \tag{18.39}$$

as well as $\Phi_n(\mathcal{M}) \subseteq \text{co}\Phi(\mathcal{M})$ *- convex hull of* $\Phi(\mathcal{M})$.

Proof (see (Zeidler 1995))**.** For every n there exists a finite $(2n)^{-1}$-net for $A(\mathcal{M})$ and elements $u_j \in \Phi(\mathcal{M})$ $(j = 1, ..., J)$ such that for all $x \in \mathcal{M}$

$$\min_{1 \leq j \leq J} \|\Phi(x) - u_j\| \leq (2n)^{-1}$$

Define for all $x \in \mathcal{M}$ the, so-called, *Schauder operator* A_n by

$$\Phi_n(x) := \sum_{j=1}^{J} a_j(x) u_j \left(\sum_{j=1}^{J} a_j(x)\right)^{-1}$$

where

$$a_j(x) := \max\left\{n^{-1} - \|\Phi(x) - u_j\|; 0\right\}$$

are continuous functions. In view of this A_n is also continuous and, moreover,

$$\begin{aligned}\|\Phi(x)\text{-}\Phi_n(x)\| = \Phi_n(x) &= \left\|\sum_{j=1}^{J} a_j(x)\left(u_j\text{-}\Phi_n(x)\right)\right\| \left(\sum_{j=1}^{J} a_j(x)\right)^{-1} \\ &\le \sum_{j=1}^{J} a_j(x)\|(u_j - \Phi_n(x))\| \left(\sum_{j=1}^{J} a_j(x)\right)^{-1} \\ &\le \sum_{j=1}^{J} a_j(x)\, n^{-1} \left(\sum_{j=1}^{J} a_j(x)\right)^{-1} = n^{-1}\end{aligned}$$

Theorem is proven. ■

18.4.4 Inverse operators

Many problems in Theory of Ordinary and Partial Differential equations may be presented as a linear equation $Ax = y$ given in functional spaces $\mathcal{X}$ and $\mathcal{Y}$ where $A : \mathcal{X} \to \mathcal{Y}$ is a linear operator. If there exists the inverse operator $A^{-1} : \mathcal{R}(A) \to \mathcal{D}(A)$, then the solution of this linear equation may be formally represented as $x = A^{-1}y$. So, it seems to be very important to notice under which conditions the inverse operator exists.

Set of nulls and isomorphic operators

Let $A : \mathcal{X} \to \mathcal{Y}$ be a linear operator where $\mathcal{X}$ and $\mathcal{Y}$ are linear spaces such that $\mathcal{D}(A) \subseteq \mathcal{X}$ and $\mathcal{R}(A) \subseteq \mathcal{Y}$.

Definition 18.15 *The subset $\mathcal{N}(A) \subseteq \mathcal{D}(A)$ defined by*

$$\boxed{\mathcal{N}(A) := \{x \in \mathcal{D}(A) \mid Ax = 0\}} \tag{18.40}$$

*is called the **null space** of the operator A.*

Notice that

1. $\mathcal{N}(A) \neq \varnothing$ since $0 \in \mathcal{N}(A)$.

2. $\mathcal{N}(A)$ is a linear subspace (manifold).

Theorem 18.8 *An operator A is **isomorphic** (it transforms each point $x \in \mathcal{D}(A)$ only into unique point $y \in \mathcal{R}(A)$) if and only if $\mathcal{N}(A) = \{0\}$, that is, when the set of nulls consists only of the single 0-element.*

Proof. *a) Necessity.* Let A be isomorphic. Suppose that $\mathcal{N}(A) \neq \{0\}$. Take $z \in \mathcal{N}(A)$ such that $z \neq 0$. Let also $y \in \mathcal{R}(A)$. Then the equation $Ax = y$ has a solution . Consider a point $x^* + z$. By lineality of A it follows $A(x^* + z) = y$. So, the element y has at least two different image x^* and $x^* + z$. We have obtained the contradiction to isomorphic property assumption. *b) Sufficiency.* Let $\mathcal{N}(A) = \{0\}$. But assume that there exist at least two $x_1, x_2 \in \mathcal{D}(A)$ such that $Ax_1 = Ax_2 = y$ and $x_1 \neq x_2$. The last implies $A(x_1 - x_2) = 0$. But this means that $(x_1 - x_2) \in \mathcal{N}(A) = \{0\}$, or, equivalently, $x_1 = x_2$. Contradiction. ■

Claim 18.8 *Evidently that*

- *if a linear operator A is isomorphic then there exists the inverse operator A^{-1}.*

- *the operator A^{-1} is a linear operator too.*

Bounded inverse operators

Theorem 18.9 *An operator A^{-1} **exists** and, simultaneously, is **bounded** if and only if the following inequality holds*

$$\boxed{\|Ax\| \geq m \|x\|} \tag{18.41}$$

for all $x \in \mathcal{D}(A)$ and some $m > 0$.

Proof. *a) Necessity.* Let A^{-1} exists and bounded on $\mathcal{D}(A^{-1}) = \mathcal{R}(A)$. This means that there exists $c > 0$ such that for any $y \in \mathcal{R}(A)$ we have $\|A^{-1}y\| \leq c\|y\|$. Taking $y = Ax$ in the last inequality, we obtain (18.41). *b) Sufficiency.* Let now (18.41) holds. Then if $Ax = 0$

then by (18.41) we find that $x = 0$. This means that $\mathcal{N}(A) = \{0\}$ and by Theorem 18.8 it follows that A^{-1} exists. Then taking in (18.41) $x = A^{-1}y$ we get $\|A^{-1}y\| \leq m^{-1}\|y\|$ for all $y \in \mathcal{R}(A)$ that proofs the boundedness of A^{-1}. ■

Definition 18.16 *A linear operator $A : \mathcal{X} \to \mathcal{Y}$ is said to be **continuously invertible** if $\mathcal{R}(A) = \mathcal{Y}$, A is invertible and $A^{-1} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ (that is, it is bounded).*

Theorem 18.9 may be reformulated in the following manner.

Theorem 18.10 *An operator A is **continuously invertible** if and only if $\mathcal{R}(A) = \mathcal{Y}$ and for some constant $m > 0$ the inequality (18.41) holds.*

It is not so difficult to prove the following result.

Theorem 18.11 (Banach) *If $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ (that is, A is linear bounded), $\mathcal{R}(A) = \mathcal{Y}$ and A is invertible, then it is continuously invertible.*

Example 18.9 *Let us consider in $C[0,1]$ the following simplest integral equation*

$$(Ax)(t) := x(t) - \int_{s=0}^{1} tsx(s)\,ds = y(t) \tag{18.42}$$

The linear operator $A : C[0,1] \to C[0,1]$ is defined by the left-hand side of (18.42). Notice that $x(t) = y(t) + ct$, where $c = \int_{s=0}^{1} sx(s)\,ds$. Integrating the equality $sx(s) = sy(s) + cs^2$ on $[0,1]$, we obtain $c = \frac{3}{2}\int_{s=0}^{1} sy(s)\,ds$. Hence, for any $y(t)$ in the right-hand side of (18.42) the solution is $x(t) = y(t) + \frac{3}{2}\int_{s=0}^{1} tsy(s)\,ds := (A^{-1}y)(t)$. Notice that A^{-1} is bounded, but this means by the definition that the operator A is continuously invertible.

Example 18.10 *Let $y(t)$ and $a_i(t)$ $(i = 1, ..., n)$ are continuous on $[0, T]$. Consider the following linear ordinary differential equation (ODE)*

$$(Ax)(t) := x^{(n)}(t) + a_1(t)\, x^{(n-1)}(t) + ... + a_n(t)\, x(t) = y(t) \quad (18.43)$$

under the initial conditions $x(0) = x'(0) = ... = x^{(n-1)}(0) = 0$ and define the operator A as the left-hand side of (18.43) which is, evidently, linear with $\mathcal{D}(A)$ consisting of all functions which are n-times continuously differentiable, i.e., $x(t) \in C^n[0, T]$. We will solve the Cauchy problem finding the corresponding $x(t)$. Let $x_1(t)$, $x_2(t)$, ..., $x_n(t)$ be the system of n linearly independent solutions of (18.43) when $y(t) \equiv 0$. Construct the, so-called, Wronsky's determinant

$$W(t) := \begin{vmatrix} x_1(t) & \cdots & x_n(t) \\ x_1'(t) & \cdots & x_n'(t) \\ \vdots & \vdots & \vdots \\ x_1^{(n-1)}(t) & & x_n^{(n-1)}(t) \end{vmatrix}$$

It is well known (see, for example (El'sgol'ts 1961)) that $W(t) \neq 0$ for all $t \in [0, T]$. According to the Lagrange approach dealing with the variation of arbitrary constants we may find the solution of (18.43) for any $y(t)$ in the form

$$x(t) = c_1(t)\, x_1(t) + c_2(t)\, x_2(t) + ... + c_n(t)\, x_n(t)$$

that leads to the following ODE-system for $c_i(t)$ $(i = 1, ..., n)$:

$$\left.\begin{array}{c} c_1'(t)\, x_1(t) + c_2'(t)\, x_2(t) + ... + c_n'(t)\, x_n(t) = 0 \\ c_1'(t)\, x_1'(t) + c_2'(t)\, x_2'(t) + ... + c_n'(t)\, x_n'(t) = 0 \\ \cdots \\ c_1'(t)\, x_1^{(n-1)}(t) + c_2'(t)\, x_2^{(n-1)}(t) + ... + c_n'(t)\, x_n^{(n-1)}(t) = y(y) \end{array}\right\}$$

Resolving this system by the Cramer's rule we derive $c_k'(t) = \dfrac{w_k(t)}{W(t)} y(t)$ $(k = 1, ..., n)$ where $w_k(t)$ is the algebraic complement of k-th element of the last n-th row. Taking into account the initial conditions we conclude that

$$x(t) = \sum_{k=1}^{n} x_k(t) \int_{s=0}^{t} \frac{w_k(s)}{W(s)} y(s)\, ds := \left(A^{-1} y\right)(t)$$

that implies the following estimate $\|x\|_{C[0,T]} \leq c\,\|y\|_{C[0,T]}$ *with* $c = \max_{t\in[0,T]} \sum_{k=1}^{n} |x_k(t)| \int_{s=0}^{t} \left|\frac{w_k(s)}{W(s)}\right| ds$ *that proofs that the operator* A *is continuously invertible.*

Bounds for $\|A^{-1}\|$

Theorem 18.12 *Let* $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ *be a linear bounded operator such that* $\|I - A\| < 1$ *where* I *is the identical operator (which is, obviously, continuously invertible). Then* A *is continuously invertible and the following bounds holds:*

$$\boxed{\|A^{-1}\| \leq \frac{1}{1 - \|I - A\|}} \tag{18.44}$$

$$\boxed{\|I - A^{-1}\| \leq \frac{\|I - A\|}{1 - \|I - A\|}} \tag{18.45}$$

Proof. Consider in $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ the series $(I + C + C^2 + ...)$ where $C := I - A$. Since $\left\|C^k\right\| \leq \|C\|^k$ this series uniformly converges (by the Weierstrass rule), i.e.,

$$S_n := I + C + C^2 + ...C^n \underset{n\to\infty}{\rightarrow} S$$

It is easy to check that

$$\begin{aligned} (I - C)\,S_n &= I - C^{n+1} \\ S_n\,(I - C) &= I - C^{n+1} \\ C^{n+1} &\underset{n\to\infty}{\rightarrow} 0 \end{aligned}$$

Taking the limits in the last identities we obtain

$$(I - C)\,S = I,\ \ S\,(I - C) = I$$

that shows that the operator S is invertible and $S^{-1} = I - C = A$. So, $S = A^{-1}$ and

$$\begin{aligned} \|S_n\| &\leq \|I\| + \|C\| + \|C\|^2 + ...\,\|C\|^n = \frac{1 - \|C\|^{n+1}}{1 - \|C\|} \\ \|I - S_n\| &\leq \|C\| + \|C\|^2 + ...\,\|C\|^n = \frac{\|C\| - \|C\|^{n+1}}{1 - \|C\|} \end{aligned}$$

Taking $n \to \infty$ we obtain (18.44) and (18.45). ■

18.5 Duality

Let $\mathcal{X}$ be a linear normed space and $\mathcal{F}$ be the real axis $\mathbb{R}$, if $\mathcal{X}$ is real, and be the complex plane $\mathbb{C}$, if $\mathcal{X}$ is complex.

18.5.1 Dual spaces

Definition 18.17 *Consider the space $\mathcal{L}(\mathcal{X}, \mathcal{F})$ of all linear bounded functional defined on $\mathcal{X}$. This space is called* ***dual to*** $\mathcal{X}$ *and is denoted by $\mathcal{X}^*$, so that*

$$\boxed{\mathcal{X}^* := \mathcal{L}(\mathcal{X}, \mathcal{F})} \tag{18.46}$$

The value of linear functional $f \in \mathcal{X}^*$ on the element $x \in \mathcal{X}$ we will denote by $f(x)$, or $\langle x, f \rangle$, that is,

$$\boxed{f(x) = \langle x, f \rangle} \tag{18.47}$$

The notation $\langle x, f \rangle$ is analogous to the usual scalar product and turns out to be very useful in concrete calculations. In particular, the lineality of $\mathcal{X}$ and $\mathcal{X}^*$ implies the following identities (for any scalars $\alpha_1, \alpha_2, \beta_1, \beta_2$, any elements $x_1, x_2 \in \mathcal{X}$ and any functionals $f, f_1, f_2 \in \mathcal{X}^*$):

$$\boxed{\begin{array}{c} \langle \alpha_1 x_1 + \alpha_2 x_2, f \rangle = \alpha_1 \langle x_1, f \rangle + \alpha_2 \langle x_2, f \rangle \\ \langle x, \beta_1 f_1 + \beta_2 f_2 \rangle = \bar{\beta}_1 \langle x, f_1 \rangle + \bar{\beta}_2 \langle x, f_2 \rangle \end{array}} \tag{18.48}$$

($\bar{\beta}$ means the complex conjugated value to β. In real case $\bar{\beta} = \beta$). If $\langle x, f \rangle = 0$ for any $x \in \mathcal{X}$, then $f = 0$. This property can be considered as the definition of the "*null*"-functional. Less trivial seems to be the next property.

Lemma 18.5 *If $\langle x, f \rangle = 0$ for any $f \in \mathcal{X}^*$, then $x = 0$.*

Proof. It is based on the Corollary 18.5 of the Hahn-Banach theorem 18.6. Assuming the existence of $x \neq 0$, we can find $f \in \mathcal{X}^*$ such that $f \neq 0$ and $\langle x, f \rangle = \|x\| \neq 0$ that contradicts to the identity $\langle x, f \rangle = 0$ valid for any $f \in \mathcal{X}^*$. So, $x = 0$. ■

Definition 18.18 *In $\mathcal{X}^*$ one can introduce two types of convergence.*

- ***Strong convergence*** *(on the norm in $\mathcal{X}^*$):*

  $f_n \underset{n\to\infty}{\to} f \quad (f_n, f \in \mathcal{X}^*)$, *if* $\|f_n - f\| \underset{n\to\infty}{\to} 0$.

- ***Weak convergence*** *(in the functional sense) :*

  $f_n \underset{n\to\infty}{\overset{*}{\to}} f \quad (f_n, f \in \mathcal{X}^*)$, *if for any* $x \in \mathcal{X}$ *one has*

  $\langle x, f_n\rangle \underset{n\to\infty}{\to} \langle x, f\rangle$.

Remark 18.4

1. *Notice that the **strong convergence** of a functional sequence $\{f_n\}$ **implies** its **weak convergence**.*

2. *(**Banach-Shteingauss**): $f_n \underset{n\to\infty}{\overset{*}{\to}} f$ if and only if*

   a) $\{\|f_n\|\}$ is bounded;

   b) $\langle x, f_n\rangle \underset{n\to\infty}{\to} \langle x, f\rangle$ on some dense linear manifold in $\mathcal{X}$.

Claim 18.9 *Independently of the fact whether the original topological space $\mathcal{X}$ is Banach or not, **the space $\mathcal{X}^* = \mathcal{L}(\mathcal{X}, \mathcal{F})$ of all linear bounded functional is always Banach**.*

Proof. It can be easily seen from the definition 18.3. ■
More exactly this statement can be formulated as follows.

Lemma 18.6 *$\mathcal{X}^*$ is a Banach space with the norm*

$$\boxed{\|f\| = \|f\|_{\mathcal{X}^*} := \sup_{x\in\mathcal{X},\ \|x\|_{\mathcal{X}}\le 1} |f(x)|} \tag{18.49}$$

Furthermore, the following duality between two norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{X}^*}$ takes place:

$$\boxed{\|x\|_{\mathcal{X}} = \sup_{f\in\mathcal{X}^*,\ \|f\|_{\mathcal{X}^*}\le 1} |f(x)|} \tag{18.50}$$

Proof. The details of the proof can be found in (Yoshida 1979).
■

Example 18.11 *The spaces $L_p[a,b]$ and $L_q[a,b]$ are dual, that is,*

$$\boxed{L_p^*[a,b] = L_q[a,b]} \tag{18.51}$$

where $p^{-1} + q^{-1} = 1$, $1 < p < \infty$. Indeed, if $x(t) \in L_p[a,b]$ and $y(t) \in L_q[a,b]$, then the functional

$$\boxed{f(x) = \int_{t=a}^{b} x(t)\, y(t)\, dt} \tag{18.52}$$

is evidently linear, and boundedness follows from the Hölder inequality (16.137).

Since the dual space of a linear normed space is always a Banach space, one can consider the bounded linear functionals on $\mathcal{X}^*$, which we shall denote by $\mathcal{X}^{**}$. Moreover, each element $x \in \mathcal{X}$ gives rise to a bounded linear functional f^* in $\mathcal{X}^*$ by $f^*(f) = f(x)$, $f \in \mathcal{X}^*$. It can be shown that $\mathcal{X} \subset \mathcal{X}^{**}$, that called the natural embedding of $\mathcal{X}$ into $\mathcal{X}^{**}$. Sometimes it happens that these spaces coincide. Notice that this is possible if $\mathcal{X}$ is Banach space (since $\mathcal{X}^{**}$ is always Banach).

Definition 18.19 *If $\mathcal{X} = \mathcal{X}^{**}$, the the Banach space $\mathcal{X}$ is called **reflexive**.*

Such spaces play important role in different applications since they possess many properties resembling ones in Hilbert spaces.

Claim 18.10 *Reflexive space are all Hilbert spaces, $\mathbb{R}^n$, l_p^n, and $L_{p>1}(\bar{G})$.*

Theorem 18.13 *The Banach space $\mathcal{X}$ is reflexive if and only if any bounded (by a norm) sequence of its elements contains a subsequence which weakly converges to some point in $\mathcal{X}$.*

Proof. See (Trenogin 1980) the section 17.5, and (Yoshida 1979) (p.264, the Eberlein-Shmulyan theorem) ■

18.5.2 Adjoint (dual) and self-adjoint operators

Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ where $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces. Construct the linear functional $\varphi(x) = \langle x, \varphi \rangle := \langle Ax, f \rangle$ where $x \in \mathcal{X}$ and $f \in \mathcal{Y}^*$.

Lemma 18.7 *1) $\mathcal{D}(\varphi) = \mathcal{X}$, 2) φ is linear operator, 3) φ is bounded.*

Proof. 1) is evident. 2) is valid since

$$\varphi(\alpha_1 x_1 + \alpha_2 x_2) = \langle A(\alpha_1 x_1 + \alpha_2 x_2), f \rangle =$$

$$\alpha_1 \langle A(x_1), f \rangle + \alpha_2 \langle A(x_2), f \rangle = \alpha_1 \varphi(x_1) + \alpha_2 \varphi(x_2)$$

And 3) holds since $|\varphi(x)| = |\langle Ax, f \rangle| \leq \|Ax\| \|f\| \leq \|A\| \|f\| \|x\|$. ■

From this lemma it follows that $\varphi \in \mathcal{X}^*$. So, there is correctly defined the linear continuous operator $\varphi = A^* f$.

Definition 18.20 *The operator $A^* \in \mathcal{L}(\mathcal{Y}^*, \mathcal{X}^*)$ defined by*

$$\boxed{\langle x, A^* f \rangle := \langle Ax, f \rangle} \tag{18.53}$$

is called ***adjoint*** *(or* ***dual****)* ***operator*** *of A.*

Lemma 18.8 *The representation $\langle Ax, f \rangle = \langle x, \varphi \rangle$ is unique $(\varphi \in \mathcal{X}^*)$ for any $x \in \mathcal{D}(A)$ if and only if $\overline{\mathcal{D}(A)} = \mathcal{X}$.*

Proof. *a) Necessity.* Suppose $\overline{\mathcal{D}(A)} \neq \mathcal{X}$. Then by the corollary 18.7 from the Hahn-Banach theorem 18.6 there exists $\varphi_0 \in \mathcal{X}^*$, $\varphi_0 \neq 0$ such that $\langle x, \varphi_0 \rangle = 0$ for all $x \in \overline{\mathcal{D}(A)}$. But then $\langle Ax, f \rangle = \langle x, \varphi + \varphi_0 \rangle = 0$ for all $x \in \mathcal{D}(A)$ that contradicts with the assumption of the uniqueness of the presentation.

b) *Sufficiency.* Let $\overline{\mathcal{D}(A)} = \mathcal{X}$. If $\langle Ax, f \rangle = \langle x, \varphi_1 \rangle = \langle x, \varphi_2 \rangle$ then $\langle x, \varphi_1 - \varphi_2 \rangle = 0$ and by the same corollary 18.7 it follows that $\varphi_1 - \varphi_2 = 0$ that means that the representation is unique. ■

Lemma 18.9 *If $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ where $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, then $\|A^*\| = \|A\|$.*

Proof. By the property 3) of the previous lemma we have $\|\varphi\| \leq \|A\| \, \|f\|$, i.e., $\|A^*\| \leq \|A\|$. But, by the corollary 18.5 from the Hahn-Banach theorem 18.6, for any x_0 such that $Ax_0 \neq 0$ there exists a functional $f_0 \in \mathcal{Y}^*$ such that $\|f_0\| = 1$ and $|\langle Ax_0, f_0\rangle| = \|Ax_0\|$ that leads to the following estimate:

$$\|Ax_0\| = |\langle Ax_0, f_0\rangle| = |\langle x_0, A^* f_0\rangle| \leq \|A^*\| \, \|f_0\| \, \|x_0\| = \|A^*\| \, \|x_0\|$$

So, $\|A^*\| \geq \|A\|$ and, hence, $\|A^*\| = \|A\|$ that proves the lemma. ■

Example 18.12 *Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}^n$ be n-dimensional Euclidian spaces. Consider the linear operator*

$$\boxed{y = Ax \quad \left(y_i := \sum_{k=1}^{n} a_{ik} x_k, \; i = 1, ..., n \right)} \tag{18.54}$$

Let $z \in (\mathbb{R}^n)^* = \mathbb{R}^n$. Since in Euclidian spaces the action of an operator is the corresponding scalar product, then $\langle Ax, z\rangle = (Ax, z) = (x, A^\intercal z) = \langle x, A^* z\rangle$. So,

$$\boxed{A^* = A^\intercal} \tag{18.55}$$

Example 18.13 *Let $\mathcal{X} = \mathcal{Y} = L_2\,[a, b]$. Let us consider the integral operator $y = Kx$ given by*

$$\boxed{y\,(t) = \int_{s=a}^{b} k\,(t, s)\, x\,(s)\, ds} \tag{18.56}$$

with the kernel $k\,(t, s)$ which is continuous on $[a, b] \times [a, b]$. We will consider the case when all variables are real. Then we have

$$\begin{gathered} \langle Kx, z\rangle = \int_{t=a}^{b} \left(\int_{s=a}^{b} k\,(t, s)\, x\,(s)\, ds \right) z\,(t)\, dt = \\ \int_{s=a}^{b} \left(\int_{t=a}^{b} k\,(t, s)\, z\,(t)\, dt \right) x\,(s)\, ds = \\ \int_{t=a}^{b} \left(\int_{s=a}^{b} k\,(s, t)\, z\,(s)\, ds \right) x\,(t)\, dt = \langle x, K^* z\rangle \end{gathered}$$

that shows that the operator K^* $(\omega = K^* z)$ is defined by

$$\boxed{\omega\,(t) = \int_{s=a}^{b} k\,(s, t)\, z\,(s)\, ds} \tag{18.57}$$

that is, K^* is also integral with the kernel $k(s,t)$ which is inverse to the kernel $k(t,s)$ of K.

Definition 18.21 *The operator $A \in \mathcal{L}(\mathcal{X},\mathcal{Y})$, where $\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces, is said to be* ***self-adjoint*** *(or* ***Hermitian****) if $A^* = A$, that is, if it coincides with its adjoint (dual) form.*

Remark 18.5 *Evidently that for self-adjoint operators $\mathcal{D}(A) = \mathcal{D}(A^*)$.*

Example 18.14

1. *In $\mathbb{R}^n$, where any linear operator A is a matrix transformation, it will be self-adjoint if it is symmetric, i.e., $A = A^\intercal$, or, equivalently, $a_{ij} = a_{ji}$.*
2. *In $\mathbb{C}^n$, where any linear operator A is a complex matrix transformation, it will be self-adjoint if it is Hermitian, i.e., $A = A^*$, or, equivalently, $a_{ij} = \bar{a}_{ji}$.*
3. *The integral operator in the example 18.13 the integral operator K is self-adjoint in $L_2[a,b]$ if its kernel is symmetric, namely, if $k(t,s) = k(s,t)$.*

It is easy to check the following simple properties of self-adjoint operators.

Proposition 18.1 *Let A and B be self-adjoint operators. Then*

1. *$(\alpha A + \beta B)$ is also self-adjoint for any real α and β.*
2. *(AB) is self-adjoint if an only if these two operators commute, i.e., if $AB = BA$. Indeed, $(ABx,f) = (Bx,Af) = (x,BAf)$.*
3. *The value (Ax,x) is always real for any $x \in \mathcal{F}$ (real or complex).*
4. *For any self-adjoint operator A we have*

$$\boxed{\|A\| = \sup_{\|x\| \leq 1} |(Ax,x)|} \tag{18.58}$$

18.5.3 Riesz representation theorem for Hilbert spaces

Theorem 18.14 (F. Riesz) *If $\mathcal{H}$ is a Hilbert space (complex or real) with a scalar product $(\cdot,\cdot)$, then for any linear bounded functional f, defined on $\mathcal{H}$, there exists the unique element $y \in \mathcal{H}$ such that for all $x \in \mathcal{H}$ one has*

$$\boxed{f(x) = \langle x, f\rangle = (x, y)} \tag{18.59}$$

and, furthermore, $\|f\| = \|y\|$.

Proof. Let L be a subspace of $\mathcal{H}$. If $L = \mathcal{H}$, then for $f = 0$ one can take $y = 0$ and the theorem is proven. If $L \neq \mathcal{H}$, there there exists $z_0 \perp L$, $z_0 \neq 0$ (it is sufficient to consider the case $f(z_0) = \langle z_0, f\rangle = 1$; if not, instead of z_0 we can consider $z_0/\langle z_0, f\rangle$). Let now $x \in \mathcal{H}$. Then $x - \langle x, f\rangle z_0 \in L$, since

$$\langle x - \langle x, f\rangle z_0, f\rangle = \langle x, f\rangle - \langle x, f\rangle \langle z_0, f\rangle = \langle x, f\rangle - \langle x, f\rangle = 0$$

Hence, $[x - \langle x, f\rangle z_0] \perp z_0$ that implies

$$0 = (x - \langle x, f\rangle z_0, z_0) = (x, z_0) - \langle x, f\rangle \|z_0\|^2$$

or, equivalently, $\langle x, f\rangle = \left(x, z_0/\|z_0\|^2\right)$. So, we can take $y = z_0/\|z_0\|^2$. Show now the uniqueness of y. If $\langle x, f\rangle = (x, y) = (x, \tilde{y})$, then $(x, y - \tilde{y}) = 0$ for any $x \in \mathcal{H}$. Taking $x = y - \tilde{y}$ we obtain $\|y - \tilde{y}\|^2 = 0$ that proves the identity $y = \tilde{y}$. To complete the proof we need to prove that $\|f\| = \|y\|$. By the Cauchy-Bounyakovski-Schwartz inequality $|\langle x, f\rangle| = |(x, y)| \leq \|f\| \|y\|$. By the definition of the norm $\|f\|$ it follows that $\|f\| \leq \|y\|$. On the other hand, $\langle y, f\rangle = (y, y) \leq \|f\| \|y\|$ that leads to the inverse inequality $\|y\| \leq \|f\|$. So, $\|f\| = \|y\|$. Theorem is proven. ■

Different application of this theorem can be found in (Riesz & Nagy 1978 (original in French, 1955)).

18.5.4 Orthogonal projection operators in Hilbert spaces

Let M be a subspace of a Hilbert space $\mathcal{H}$.

Definition 18.22 *The operator $P \in \mathcal{L}(\mathcal{H}, M)$ $(y = Px)$, acting in $\mathcal{H}$ such that*

$$y := \arg\min_{z \in M} \|x - z\| \tag{18.60}$$

is called the ***orthogonal projection operator*** *to the subspace M.*

Lemma 18.10 *The element $y = Px$ is unique and $(x - y, z) = 0$ for any $z \in M$.*

Proof. See Subsection 18.3.2. ■

The following evident properties of the projection operator hold.

Proposition 18.2

1. *$x \in M$ if and only if $Px = x$.*

2. *Let $M^{\perp}$ be the orthogonal complement to M, that is,*

$$\boxed{M^{\perp} := \{z \in \mathcal{H} \mid z \perp M\}} \tag{18.61}$$

Then any $x \in \mathcal{H}$ can be represented as $x = y + z$ where $y \in M$ and $z \perp M$. Then the operator $P^{\perp} \in \mathcal{L}(\mathcal{H}, M^{\perp})$, defining the orthogonal projection any point from $\mathcal{H}$ to $M^{\perp}$, has the following representation:

$$\boxed{P^{\perp} = I - P} \tag{18.62}$$

3. *$x \in M^{\perp}$ if and only if $Px = 0$.*

4. *P is linear operator, i.e., for any real α and β one has*

$$\boxed{P(\alpha x_1 + \beta x_2) = \alpha P(x_1) + \beta P(x_2)} \tag{18.63}$$

5.

$$\boxed{\|P\| = 1} \tag{18.64}$$

Indeed, $\|x\|^2 = \|Px + (I - P)x\|^2 = \|Px\|^2 + \|(I - P)x\|^2$ that implies $\|Px\|^2 \leq \|x\|^2$ and thus $\|P\| \leq 1$. On the other hand, if $M \neq \{0\}$, take $x_0 \in M$ with $\|x_0\| = 1$. Then $1 = \|x_0\| = \|Px_0\|$ $\leq \|P\| \|x_0\| = \|P\|$. The inequalities $\|P\| \leq 1$ and $\|P\| \geq 1$ give (18.64).

6.
$$\boxed{P^2 = P} \tag{18.65}$$
since for any $x \in M$ *we have* $P^2(Px) = Px$.

7. P *is self-adjoint, that is,*
$$\boxed{P^* = P} \tag{18.66}$$

8. *For any* $x \in \mathcal{H}$
$$\boxed{(Px, x) = (P^2x, x) = (Px, Px) = \|Px\|^2} \tag{18.67}$$
that implies
$$\boxed{(Px, x) \geq 0} \tag{18.68}$$

9. $\|Px\| = \|x\|$ *if and only if* $x \in M$.

10. *For any* $x \in \mathcal{H}$
$$\boxed{(Px, x) \leq \|x\|^2} \tag{18.69}$$
that follows from (18.67), the Cauchy - Bounyakovsky - Schwartz inequality and (18.64).

11. *Let* $A = A^* \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ *and* $A^2 = A$. *Then* A *is obligatory an orthogonal projection operator to some subspace* $M = \{x \in \mathcal{H} \,|Ax = x\} \subset \mathcal{H}$. *Indeed, since* $x = Ax + (I - A)x$ *it follows that* $Ax = A^2x = A(Ax) \in M$ *and* $(I - A)x \in M^{\perp}$.

The following lemma can be easily verified.

Lemma 18.11 *Let* P_1 *be the orthogonal projector to a subspace* M_1 *and* P_2 *be the orthogonal projector to a subspace* M_2. *Then following 5 statements are equivalent:*

1.
$$\boxed{M_2 \subset M_1}$$

2.
$$\boxed{P_1P_2 = P_2P_1 = P_2}$$

3. $\|P_2x\| \leq \|P_1x\|$ *for any* $x \in \mathcal{H}$.

4. $(P_2x, x) \leq (P_1x, x)$ *for any* $x \in \mathcal{H}$.

Corollary 18.9

1. $M_2 \perp M_1$ *if and only if* $P_1P_2 = 0$.

2. P_1P_2 *is a projector if and only if* $P_1P_2 = P_2P_1$.

3. *Let* P_i $(1 = 1, ..., N)$ *be a projection operators. Then* $\sum_{i=1}^{N} P_i$ *is a projection operator too if and only if*

$$\boxed{P_iP_k = \delta_{ik}P_i}$$

4. $P_1 - P_2$ *is a projection operator if and only if* $P_1P_2 = P_2$, *or equivalently, when* $P_1 \geq P_2$.

18.6 Monotonic, nonnegative and coercive operators

Remember the following elementary lemma from Real Analysis.

Lemma 18.12 *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a continuous function such that*

$$(x - y)\left[f(x) - f(y)\right] \geq 0 \tag{18.70}$$

for any $x, y \in \mathbb{R}$ *and*

$$xf(x) \to \infty \text{ when } |x| \to \infty \tag{18.71}$$

Then the equation $f(x) = 0$ *has a solution. If (18.70) holds in the strong sense, i.e.,*

$$(x - y)\left[f(x) - f(y)\right] > 0 \text{ when } x \neq y \tag{18.72}$$

then the equation $f(x) = 0$ *has a unique solution.*

Proof. For $x < y$ from (18.70) it follows that $f(x)$ is non-decreasing function and, in view of (18.71), there exist numbers a and b such that $a < b$, $f(a) < 0$ and $f(b) < 0$. Then, considering $f(x)$ on $[a, b]$, by the theorem on intermediate values, there exists a point $\xi \in [a, b]$ such that $f(\xi) = 0$. If (18.72) is fulfilled, then $f(x)$ is monotonically increasing function and the root of the function $f(x)$ is unique. ■

The following definitions and theorems represent the generalization of this lemma to functional spaces and nonlinear operators.

18.6.1 Basic definitions and properties

Let $\mathcal{X}$ be a real separable normed space and $\mathcal{X}^*$ be a space dual to $\mathcal{X}$. Consider a nonlinear operator $T : \mathcal{X} \to \mathcal{X}^*$ ($\mathcal{D}(T) = \mathcal{X}$, $\mathcal{R}(T) \subset \mathcal{X}^*$) and, as before, denoted by $f(x) = \langle x, f \rangle$ the value of the linear functional $f \in \mathcal{X}^*$ on the element $x \in \mathcal{X}$.

Definition 18.23

1. *An operator T is said to be* ***monotone*** *if for any $x, y \in \mathcal{D}(T)$*

$$\boxed{\langle x - y, T(x) - T(y) \rangle \geq 0} \tag{18.73}$$

2. *It is called* ***strictly monotone*** *if for any $x \neq y$*

$$\boxed{\langle x - y, T(x) - T(y) \rangle > 0} \tag{18.74}$$

   and the equality is possible only if $x = y$.

3. *It is called* ***strongly monotone*** *if for any $x, y \in \mathcal{D}(T)$*

$$\boxed{\langle x - y, T(x) - T(y) \rangle \geq \alpha(\|x - y\|)\, \|x - y\|} \tag{18.75}$$

   where the nonnegative function $\alpha(t)$, defined at $t \geq 0$, satisfies the condition $\alpha(0) = 0$ and $\alpha(t) \to \infty$ when $t \to \infty$.

4. *An operator T is called* ***nonnegative*** *if for all $x \in \mathcal{D}(T)$*

$$\boxed{\langle x, T(x) \rangle \geq 0} \tag{18.76}$$

5. *An operator T is* ***positive*** *if for all $x \in \mathcal{D}(T)$*

$$\boxed{\langle x, T(x) \rangle > 0} \tag{18.77}$$

6. *An operator T is called* ***coercive*** *(or,* ***strongly positive****) if for all $x \in \mathcal{D}(T)$*

$$\boxed{\langle x, T(x) \rangle \geq \gamma(\|x\|)} \tag{18.78}$$

where function $\gamma(t)$, defined at $t \geq 0$, satisfies the condition $\gamma(t) \to \infty$ when $t \to \infty$.

Example 18.15 *The function $f(x) = x^3 + x - 1$ is the strictly monotone operator in $\mathbb{R}$.*

The following lemma installs the relation between monotonicity and coercivity properties.

Lemma 18.13 *If an operator $T : \mathcal{X} \to \mathcal{X}^*$ is* ***strongly monotone*** *then it is* ***coercive*** *with*

$$\boxed{\gamma(\|t\|) = \alpha(t) - \|T(0)\|} \tag{18.79}$$

Proof. By the definition (when $y = 0$) it follows that

$$\langle x, T(x) - T(0) \rangle \geq \alpha(\|x\|)\|x\|$$

This implies

$$\langle x, T(x) \rangle \geq \langle x, T(0) \rangle + \alpha(\|x\|)\|x\| \geq \alpha(\|x\|)\|x\| - \\ \|x\| \|T(0)\| = [\alpha(\|x\|) - \|T(0)\|]\|x\|$$

that proves the lemma. ■

Remark 18.6 *Notice that an operator $T : \mathcal{X} \to \mathcal{X}^*$ is coercive then $\|T(x)\| \to \infty$ when $\|x\| \to \infty$. This follows from the inequalities*

$$\|T(x)\| \|x\| \geq \langle x, T(x) \rangle \geq \gamma(\|x\|)\|x\|$$

or, equivalently, from $\|T(x)\| \geq \gamma(\|x\|) \to \infty$ when $\|x\| \to \infty$.

Next theorem generalizes Lemma 18.12 to the nonlinear vector-function case.

Theorem 18.15 ((Trenogin 1980)) *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a nonlinear operator (a vector function) which is continuous everywhere in $\mathbb{R}^n$ and such that for any $x, y \in \mathcal{X}$*

$$\boxed{\langle x - y, T(x) - T(y)\rangle \geq c\,\|x - y\|^2,\ c > 0} \tag{18.80}$$

(i.e., in (18.75) $\alpha(t) = ct$). Then the system of nonlinear equations

$$\boxed{T(x) = 0} \tag{18.81}$$

has a unique solution $x^* \in \mathbb{R}^n$.

Proof. Let us apply the induction method. For $n = 1$ the result is true by Lemma 18.12. Let it be true in $\mathbb{R}^{n-1}$ $(n \geq 2)$. Show that this results holds in $\mathbb{R}^n$. Consider in $\mathbb{R}^n$ a standard orthonormal basis $\{e_i\}_{i=1}^n$ $(e_i = (\delta_{ik})_{k=1}^n)$. Then $T(x)$ can be represented as $T(x) = \{T_i(x)\}_{i=1}^n$, $x = \sum_{j=1}^{n} x_j e_j$. For some fixed $t \in \mathbb{R}$ define the operator T_t by $T_t : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ for all $x = \sum_{j=1}^{n-1} x_j e_j$ acting as $T_t(x) := \{T_i(x + te_n)\}_{i=1}^{n-1}$. Evidently, $T_t(x)$ is continuos on $\mathbb{R}^{n-1}$ and, by the induction supposition, for any $x, y \in \mathbb{R}^{n-1}$ it satisfies the following inequality

$$\begin{gathered}\langle x - y, T_t(x) - T_t(y)\rangle = (t - t)\,[T_n(x + te_n) - T_n(y + te_n)] + \\ \sum_{t=1}^{n-1} (x_i - y_i)\,[T_i(x + te_n) - T_i(y + te_n)] \geq c\,\|x - y\|^2\end{gathered}$$

This means that the operator T_t also satisfies (18.80). By the induction supposition the system of nonlinear equations

$$T_i(x + te_n) = 0,\ i = 1, ..., n - 1 \tag{18.82}$$

has a unique solution $\hat{x} \in \mathbb{R}^{n-1}$. This exactly means that there exists a vector-function $\hat{x} = \sum_{j=1}^{n-1} x_j e_j : \mathbb{R} \to \mathbb{R}^{n-1}$ which solves the system of nonlinear equations $T_t(x) = 0$. It is not difficult to check that the function $\hat{x} = \hat{x}(t)$ is continuous. Consider then the function $\psi(t)$: $T_n(x + te_n)$. It is also not difficult to check that this function satisfies all conditions of Lemma 18.12. Hence, there exists such $\tau \in \mathbb{R}$ that $\psi(\tau) = 0$. This exactly means that the equation (18.81) has a unique solution. ■

It seems to be useful the following proposition.

Theorem 18.16 ((Trenogin 1980)) *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous monotonic operator such that for all $x \in \mathbb{R}^n$ with $\|x\| > \lambda$ the following inequality holds:*

$$\boxed{(x, T(x)) \geq 0} \tag{18.83}$$

Then the equation $T(x) = 0$ has a solution x^* such that $\|x^*\| \leq \lambda$.

Proof. Consider the sequence $\{\varepsilon_k\}$, $0 < \varepsilon_k \underset{k\to\infty}{\to} 0$ and the associated sequence $\{T_k\}$, $T : \mathbb{R}^n \to \mathbb{R}^n$ of the operators defined by $T_k(x) := \varepsilon_k x + T(x)$. Then, in view of monotonicity of T, we have for all $x, y \in \mathbb{R}^n$

$$\begin{gathered}\langle x - y, T_k(x) - T_k(y)\rangle = (x - y, T_k(x) - T_k(y)) = \\ \varepsilon_k (x - y, (x - y) + (x - y, T(x) - T(y)) \geq \varepsilon_k \|x - y\|^2\end{gathered}$$

Hence, by Theorem 18.15 it follows that the equation $T_k(x) = 0$ has the unique solution x_k^* such that $\|x_k^*\| \leq \lambda$. Indeed, if not, we obtain the contradiction: $0 = (x_k^*, T_k(x_k^*)) \geq \varepsilon_k \|x_k^*\|^2 > 0$. Therefore, the sequence $\{x_k^*\} \subset \mathbb{R}^n$ is bounded. By the Bolzano-Weierstrass theorem there exists a subsequence $\{x_{k_t}^*\}$ convergent to some point $\bar{x} \subset \mathbb{R}^n$ when $k_t \to \infty$. This implies $0 = T_{k_t}(x_{k_t}^*) = \varepsilon_{x_{k_t}^*} x_{k_t}^* + T(x_{k_t}^*)$. Since $T(x)$ is continuous then when $k_t \to \infty$ we obtain $T(\bar{x})$. Theorem is proven. ■

18.6.2 Galerkin method for equations with monotone operators

The technique given below presents the constructive method for finding an approximative solution of the operator equation $T(x) = 0$ where $T : \mathcal{X} \to \mathcal{X}^*$ ($\mathcal{D}(T) = \mathcal{X}$, $\mathcal{R}(T) \subset \mathcal{X}^*$). Let $\{\varphi_k\}_{k=1}^{\infty}$ be a complete sequence of linearly independent elements from $\mathcal{X}$, and $\mathcal{X}_n$ be a subspace spanned on $\varphi_1, ..., \varphi_n$.

Definition 18.24 *The element $x_n \in \mathcal{X}$ having the construction*

$$\boxed{x_n = \sum_{l=1}^{n} c_l \varphi_l} \tag{18.84}$$

is said to be ***the Galerkin approximation*** *to the solution of the equation $T(x) = 0$ with the monotone operator T if it satisfies the following system of equations*

$$\boxed{\langle \varphi_k, T(x_n) \rangle = 0, k = 1, ..., n} \tag{18.85}$$

or, equivalently,

$$\boxed{\sum_{l=1}^{n} \langle \varphi_l, T(x_n) \rangle \varphi_l = 0} \tag{18.86}$$

Remark 18.7

1. *It is easy to prove that x_n is a solution of (18.85) if and only if $\langle u, T(x_n) \rangle = 0$ for any $u \in \mathcal{X}_n$.*

2. *The system (18.85) can be represented in the operator form $J_n \bar{c}_n = x_n$ where the operator J_n is defined by (18.85) with $\bar{c}_n := (c_1, ..., c_n)$. Notice that $\|J_n\| \leq \sqrt{\sum_{l=1}^{n} \|\varphi_l\|^2}$. In view of this, the equation (18.85) can be rewritten in the standard basis as*

$$\boxed{\langle \varphi_k, T(J_n \bar{c}_n) \rangle = 0, k = 1, ..., n} \tag{18.87}$$

Lemma 18.14 *If an operator $T : \mathcal{X} \to \mathcal{X}^*$ ($\mathcal{D}(T) = \mathcal{X}$, $\mathcal{R}(T) \subset \mathcal{X}^*$) is strictly monotone (18.74) then*

1. *The equation (18.81) has a unique solution.*

2. *For any n the system (18.85) has a unique solution.*

Proof. If u and v are two solutions of (18.81), then $T(u) = T(v) = 0$, and, hence, $\langle x - y, T(u) - T(v) \rangle = 0$ that, in view of (18.74), takes place if and only if $u = v$. Again, if x'_n and x''_n are two solutions of (18.85) then $\langle x'_n, T(x''_n) \rangle = \langle x''_n, T(x'_n) \rangle = \langle x'_n, T(x'_n) \rangle = \langle x''_n, T(x''_n) \rangle = 0$, or, equivalently,

$$\langle x'_n - x''_n, T(x'_n) - T(x''_n) \rangle = 0$$

that, by (18.74), is possible if and only if $x'_n = x''_n$. ■

Lemma 18.15 ((Trenogin 1980)) *Let an operator $T : \mathcal{X} \to \mathcal{X}^*$ ($\mathcal{D}(T) = \mathcal{X}$, $\mathcal{R}(T) \subset \mathcal{X}^*$) is monotone and semi-continuous, and there exists a constant $\lambda > 0$ such that for all $x \in \mathcal{X}$ with $\|x\| > \lambda$ we have $\langle x, T(x) \rangle > 0$. Then for any n the system (18.85) has the solution $x_n \in \mathcal{X}$ such that $\|x_n\| \leq \lambda$.*

Proof. It is sufficient to introduce in $\mathbb{R}^n$ the operator T_n defined by

$$T_n(\bar{c}_n) := \{\langle \varphi_k, T(J_n \bar{c}_n) \rangle\}_{k=1}^n$$

and to check that it satisfies all condition of Theorem 18.16. ■

Based on these two lemmas it is possible to prove the following *main result on the Galerkin approximations.*

Proposition 18.3 ((Trenogin 1980)) *Let the conditions of Lemma 18.15 be fulfilled and $\{x_n\}$ is the sequence of solutions of the system (18.85). Then the sequence $\{T(x_n)\}$ weakly converges to zero.*

18.6.3 Main theorems on the existence of solutions for equations with monotone operators

Theorem 18.17 ((Trenogin 1980)) *Let $T : \mathcal{X} \to \mathcal{X}^*$ ($\mathcal{D}(T) = \mathcal{X}$, $\mathcal{R}(T) \subset \mathcal{X}^*$) be an operator, acting from a real separable reflexive Banach space $\mathcal{X}$ in to its dual space $\mathcal{X}^*$, which is monotone and semi-continuous. Let also there exists a constant $\lambda > 0$ such that for all $x \in \mathcal{X}$ with $\|x\| > \lambda$ we have $\langle x, T(x) \rangle > 0$. Then the equation $T(x) = 0$ has the solution x^* such that $\|x^*\| \leq \lambda$.*

Proof. By Lemma 18.15 for any n the Galerkin system (18.85) has the solution x_n such that $\|x_n\| \leq \lambda$. By reflexivity, from any sequence $\{x_n\}$ one can take out the subsequence $\{x_{n'}\}$ weakly convergent to some $x_0 \in \mathcal{X}$ such that $\|x_0\| \leq \lambda$. Then, by monotonicity of T, it follows that

$$S_{n'} := \langle x - x_{n'}, T(x) - T(x_{n'}) \rangle \geq 0$$

But $S_{n'} = \langle x - x_{n'}, T(x) \rangle - \langle x, T(x_{n'}) \rangle$, and, by Proposition 18.3, $\langle x, T(x_{n'}) \rangle \to 0$ weakly if $n' \to \infty$. Hence, $S_{n'} \to \langle x - x_0, T(x) \rangle$, and, therefore for all $x \in \mathcal{X}$

$$\langle x - x_0, T(x) \rangle \geq 0 \tag{18.88}$$

If $T(x_0) = 0$ then the theorem is proven. Let now $T(x_0) \neq 0$. Then, by the Corollary 18.5 from the Hahn-Banach theorem 18.6 (for the case $\mathcal{X} = \mathcal{X}^{**}$), it follows the existence of the element $z_0 \in \mathcal{X}$ such that $\langle z_0, T(x_0) \rangle = \|T(x_0)\|$. Substitution of $x := x_0 - tz_0$ $(t > 0)$ in to (18.88) implies $\langle z_0, T(x_0 - tz_0) \rangle \leq 0$, that for $t \to +0$ gives $\langle z_0, T(x_0) \rangle = \|T(x_0)\| \leq 0$. This is equivalent to the identity $T(x_0) = 0$. So, the assumption that $T(x_0) \neq 0$ is incorrect. Theorem is proven. ■

Corollary 18.10 *Let an operator T be, additionally, coercive. Then the equation*

$$\boxed{T(x) = y} \tag{18.89}$$

has a solution for any $y \in \mathcal{X}^*$.

Proof. For any fixed $y \in \mathcal{X}^*$ define the operator $F(x) : \mathcal{X} \to \mathcal{X}^*$ acting as $F(x) := T(x) - y$. It is monotone and semi-continuous too. So, we have

$$\langle x, F(x) \rangle = \langle x, T(x) \rangle - \langle x, y \rangle \geq \gamma(\|x\|) \|x\| - \|y\| \|x\| = \\ [\gamma(\|x\|) - \|y\|] \|x\|$$

and, therefore, there exists $\lambda > 0$ such that for all $x \in \mathcal{X}$ with $\|x\| > \lambda$ one has $\langle x, F(x) \rangle > 0$. Hence, the conditions of Theorem 18.17 hold that implies the existence of the solution for the equation $F(x) = 0$. ■

Corollary 18.11 *If in Corollary 18.10 the operator is strictly monotone, then the solution of (18.89) is unique, i.e., there exists the operator T^{-1} inverse to T.*

Example 18.16 ***(Existence of the unique solution for ODE boundary problem)*** *Consider the following ODE boundary problem*

$$\left.\begin{array}{c} \mathcal{D}x(t) - f(t,x) = 0,\ t \in (a,b) \\ \mathcal{D}x(t) := \sum\limits_{l=1}^{m} (-1)^l D^l \left\{ P_l(x) D^l x(t) \right\}, \\ D := \dfrac{d}{dt} \text{ is the differentiation operator} \\ D^k x(a) = D^k x(b) = 0,\ 0 \le k \le m-1 \end{array}\right\} \tag{18.90}$$

in the Sobolev space $S_2^m(a,b)$ (18.9). Suppose that $f(t,x)$ for all x_1 and x_2 satisfies the condition

$$[f(t,x_1) - f(t,x_2)](x_1 - x_2) \ge 0$$

Let for the functions $P_l(x)$ the following additional condition is fulfilled for some $\alpha > 0$:

$$\int\limits_{t=a}^{b} \left(\sum_{l=1}^{m} P_l(x) \left[D^l x(t) \right]^2 \right) dt \ge \alpha \|x\|_{S_2^m(a,b)}$$

Consider now in $S_2^m(a,b)$ the bilinear form

$$b(x,z) := \int\limits_{t=a}^{b} \sum_{l=1}^{m} P_l(x) \left[D^l x(t) \right] \left[D^l z(t) \right] dt + \int\limits_{t=a}^{b} f(t, x(t)) z(t)\, dt$$

defining in $S_2^m(a,b)$ the nonlinear operator

$$(T(x), z)_{S_2^m(a,b)} = b(x,z)$$

which is continuos and strongly monotone since

$$b(x_1,z) - b(x_2,z) \ge \alpha \|x_1 - x_2\|_{S_2^m(a,b)}$$

Then by Theorem 18.17 and Corollary 18.10 it follows that the problem (18.90) has the unique solution.

18.7 Differentiation of Nonlinear Operators

Consider a nonlinear operator $\Phi : \mathcal{X} \to \mathcal{Y}$ acting from a Banach space $\mathcal{X}$ to another Banach space $\mathcal{Y}$ and having a domain $\mathcal{D}(\Phi) \subset \mathcal{X}$ and a range $\mathcal{R}(T) \subset \mathcal{Y}$.

18.7.1 Fréchet derivative

Definition 18.25 *We say that an operator $\Phi : \mathcal{X} \to \mathcal{Y}$ ($\mathcal{D}(\Phi) \subset \mathcal{X}$,$\mathcal{R}(\Phi) \subset \mathcal{Y}$) acting in Banach spaces is* ***Fréchet-differentiable*** *in a point $x_0 \in \mathcal{D}(\Phi)$, if there exists a linear bounded operator $\Phi'(x_0) \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ such that*

$$\boxed{\begin{array}{c} \Phi(x) - \Phi(x_0) = \Phi'(x_0)(x - x_0) + \omega(x - x_0) \\ \|\omega(x - x_0)\| = o(\|x - x_0\|) \end{array}} \tag{18.91}$$

or, equivalently,

$$\boxed{\lim_{x \to x_0} \frac{\Phi(x) - \Phi(x_0) - \langle x - x_0, \Phi'(x_0)\rangle}{\|x - x_0\|} = 0} \tag{18.92}$$

Definition 18.26 *If the operator $\Phi : \mathcal{X} \to \mathcal{Y}$ ($\mathcal{D}(\Phi) \subset \mathcal{X}$,$\mathcal{R}(\Phi) \subset \mathcal{Y}$), acting in Banach spaces, is* ***Fréchet-differentiable*** *in a point $x_0 \in \mathcal{D}(\Phi)$ the expression*

$$\boxed{d\Phi(x_0 \mid h) := \langle h, \Phi'(x_0)\rangle} \tag{18.93}$$

is called the ***Fréchet differential of the operator*** *Φ in the point $x_0 \in \mathcal{D}(\Phi)$ under the variation $h \in \mathcal{X}$, that is, the Fréchet-differential of Φ in x_0 is, nothing more, then the value of the operator $\Phi'(x_0)$ at the element $h \in \mathcal{X}$.*

Remark 18.8 *If originally $\Phi(x)$ is a linear operator, namely, if $\Phi(x) = Ax$ where $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, then $\Phi'(x_0) = A$ in any point $x_0 \in \mathcal{D}(A)$.*

Several simple propositions follows from these definitions.

Proposition 18.4

1. *If $F, G : \mathcal{X} \to \mathcal{Y}$ and both operators are Fréchet-differentiable in $x_0 \in \mathcal{X}$ then*

$$\boxed{(F+G)'(x_0) = F'(x_0) + G'(x_0)} \tag{18.94}$$

and for any scalar α

$$\boxed{(\alpha F)'(x_0) = \alpha F'(x_0)} \tag{18.95}$$

2. *If $F : \mathcal{X} \to \mathcal{Y}$ is Fréchet-differentiable in $x_0 \in \mathcal{D}(A)$ and $G : \mathcal{Z} \to \mathcal{X}$ is Fréchet-differentiable in $z_0 \in \mathcal{D}(G)$ such that $G(z_0) = x_0$ then is well-defined and continuous in the point z_0 the superposition $(F \circ G)$ of the operators F and G, namely,*

$$\boxed{F(G(z)) := (F \circ G)(z)} \tag{18.96}$$

and

$$\boxed{(F \circ G)'(z_0) = F'(x_0)\, G'(z_0)} \tag{18.97}$$

Example 18.17 *In finite-dimensional spaces $F : \mathcal{X} = \mathbb{R}^k \to \mathcal{Y} = \mathbb{R}^l$ and $G : \mathcal{Z} = \mathbb{R}^m \to \mathcal{X} = \mathbb{R}^k$ we have the systems of two algebraic nonlinear equations*

$$y = F(x), \ x = G(x)$$

and, moreover,

$$F'(x_0) = A := \left\| \frac{\partial f_i(x_0)}{\partial x_j} \right\|_{i=1,\ldots,k;j=1,\ldots,l}$$

*where A is called the **Jacobi-matrix**. Additionally, (18.97) is converted in to the following representation:*

$$(F \circ G)'(z_0) = \left\| \sum_{j=1}^{k} \frac{\partial f_i(x_0)}{\partial x_j} \frac{\partial g_j(z_0)}{\partial z_s} \right\|_{i=1,\ldots,l;s=1,\ldots,m}$$

Example 18.18 *If F is the nonlinear integral operator acting in $C[a,b]$ and is defined by*

$$F(u) := u(x) - \int_{t=a}^{b} f(x,t,u(t))\,dt$$

then $F'(u_0)$ exists in any point $u_0 \in C[a,b]$ such that

$$F'(u_0)h = h(x) - \int_{t=a}^{b} \frac{\partial f(x,t,u_0(t))}{\partial u} h(t)dt$$

18.7.2 Gáteaux derivative

Definition 18.27 *If for any $h \in \mathcal{X}$ there exists the limit*

$$\boxed{\lim_{t \to +0} \frac{\Phi(x_0 + th) - \Phi(x_0)}{t} = \delta\Phi(x_0 \mid h)} \tag{18.98}$$

*then the nonlinear operator $\delta\Phi(x_0 \mid h)$ is called the **first-variation of the operator** $\Phi(x)$ in the point $x_0 \in \mathcal{X}$ at the direction h.*

Definition 18.28 *If in (18.98)*

$$\boxed{\delta\Phi(x_0 \mid h) = A_{x_0}(h) = \langle h, A_{x_0} \rangle} \tag{18.99}$$

*where $A_{x_0} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is a linear bounded operator then Φ is **Gáteaux-differentiable** in a point $x_0 \in \mathcal{D}(\Phi)$ and the operator $A_{x_0} := \Phi'(x_0)$ is called the **Gáteaux derivative** of Φ in the point x_0 (independently on h). Moreover, the value*

$$\boxed{d\Phi(x_0 \mid h) := \langle h, A_{x_0} \rangle} \tag{18.100}$$

*is known as the **Gáteaux differential of** Φ in the point x_0 at the direction h.*

It is easy to check the following connections between the Gáteaux and Fréchet differentiability.

Proposition 18.5

1. *The Fréchet-differentiability* ***implies*** *the Gáteaux-differentiability.*

2. *The Gáteaux-differentiability* ***does not guarantee*** *the Fréchet-differentiability. Indeed, for the function*

$$f(x,y) = \begin{cases} 1 & \text{if} \quad y = x^2 \\ 0 & \text{if} \quad y \neq x^2 \end{cases}$$

which is, evidently, is not differentiable in the point $(0,0)$ *in the Fréchet sense, the Gáteaux differential in the point* $(0,0)$ *exists and equal to zero since, in view of the properties* $f(0,0) = 0$ *and* $f(th,tg) = 0$ *for any* (h,g)*, one has* $\dfrac{f(th,tg) - f(0,0)}{t} = 0.$

3. *The existence of the first variation* ***does not imply*** *the existence of the Gáteaux differential.*

18.7.3 Relation with "Variation Principle"

The main justification of the concept of differentiability is related with the optimization (or, optimal control) theory in Banach spaces and is closely connected with the, so-called, *Variation Principle* which allows us to replace a minimization problem by an equivalent problem in which the loss function is linear.

Theorem 18.18 ((Aubin 1979)) *Let* $\Phi : \mathcal{U} \to \mathcal{Y}$ *be a functional Gáteaux-differentiable on a convex subset* $\mathcal{X}$ *of a topological space* $\mathcal{U}$*. If* $x^* \in \mathcal{X}$ *minimizes* $\Phi(x)$ *on* $\mathcal{X}$ *then*

$$\boxed{\langle x^*, \Phi'(x^*)\rangle = \min_{x \in \mathcal{X}} \langle x, \Phi'(x^*)\rangle} \tag{18.101}$$

In particular, if x^* *is an interior point of* $\mathcal{X}$*, i.e.,* $x^* \in \operatorname{int}\mathcal{X}$*, then this condition implies*

$$\boxed{\Phi'(x^*) = 0} \tag{18.102}$$

Proof. Since $\mathcal{X}$ is convex then $\tilde{y} = x^* + \lambda(x - x^*) \in \mathcal{X}$ for any $\lambda \in (0,1]$ whenever $x \in \mathcal{X}$. Therefore, since x^* is a minimizer of $\Phi(x)$ on $\mathcal{X}$, we have $\dfrac{\Phi(\tilde{y}) - \Phi(x^*)}{\lambda} \geq 0$. Taking the limit $\lambda \to +0$ we deduce from the Gáteaux-differentiability of $\Phi(x)$ on $\mathcal{X}$ that $\langle x - x^*, \Phi'(x^*)\rangle \geq 0$ for any $x \in \mathcal{X}$. In particular if $x^* \in \mathrm{int}\mathcal{X}$ then for any $y \in \mathcal{X}$ there exists $\varepsilon > 0$ such that $x = x^* + \varepsilon y \in \mathcal{X}$, and, hence, $\langle x - x^*, \Phi'(x^*)\rangle = \varepsilon \langle y, \Phi'(x^*)\rangle \geq 0$ that is possible for any any $y \in \mathcal{X}$ if $\Phi'(x^*) = 0$. Theorem is proven. ■

18.8 Fixed-point Theorems

This section deals with the most important topics of Functional Analysis related with

- The existence principle
- The convergence analysis

18.8.1 Fixed-points of a nonlinear operator

In this section we follows ((Trenogin 1980)) and ((Zeidler 1995)).

Let an operator $\Phi : \mathcal{X} \to \mathcal{Y}$ $(\mathcal{D}(\Phi) \subset \mathcal{X}, \mathcal{R}(\Phi) \subset \mathcal{Y})$ acts in Banach space $\mathcal{X}$. Suppose that the set $\mathcal{M}_\Phi := \mathcal{D}(\Phi) \cap \mathcal{R}(\Phi)$ is not empty.

Definition 18.29 *The point $x^* \in \mathcal{M}_\Phi$ is called a **fixed point** of the operator Φ if it satisfies the equality*

$$\boxed{\Phi(x^*) = x^*} \tag{18.103}$$

Remark 18.9 *Any operator equation (18.81): $T(x) = 0$ can be transformed to the form (18.103). Indeed, one has*

$$\tilde{T}(x) := T(x) + x = x$$

That's why any results, concerning the existence of the solution to the operator equation (18.81), can be considered as ones but with respect to the equation $\tilde{T}(x) = x$. The inverse statement is also true.

Example 18.19 *The fixed points of the operator $\Phi(x) = x^3$ are $\{0, -1, 1\}$ that follows from the relation $0 = x^3 - x = x(x^2 - 1) = x(x-1)(x+1)$.*

Example 18.20 *Let try to find the fixed-points of the operator*

$$\Phi(x) := \int_{s=0}^{1} x(t)\,x(s)\,ds + f(t) \tag{18.104}$$

assuming that it acts in $C[0,1]$ (which is real) and that $\int_{t=0}^{1} f(t)\,dt \leq 1/4$. *By the definition (18.103) we have* $x(t)\int_{s=0}^{1} x(s)\,ds + f(t) = x(t)$. *Integrating this equations leads to the following:*

$$\left(\int_{t=0}^{1} x(t)\,dt\right)^2 + \int_{t=0}^{1} f(t)\,dt = \int_{t=0}^{1} x(t)\,dt$$

that gives

$$\int_{t=0}^{1} x(t)\,dt = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \int_{t=0}^{1} f(t)\,dt} \tag{18.105}$$

So, any function $x(t) \in C[0,1]$ satisfying (18.105) is a fixed point of the operator (18.104).

The main results related to the existence of the solution of the operator equation

$$\boxed{\Phi(x) = x} \tag{18.106}$$

are as follows:

- The **contraction principle** (see (14.17)) or **the Banach theorem** (1920) which state that *if the operator $\Phi : X \to X$ (X is a compact) is k-contractive, i.e., for all $x, x' \in X$*

  $$\|\Phi(x) - \Phi(x')\| \leq k\,\|x - x'\|, k \in [0,1)$$

  then

  a) *the solution of (18.106) exists and unique;*

  b) *the iterative method $x_{n+1} = \Phi(x_n)$ exponentially converges to this solution.*

- The **Brouwer fixed-point theorem** for finite-dimensional Banach space.
- The **Schauder fixed-point theorem** for infinite-dimensional Banach space.
- **The Leray-Schauder principle** which state that *a priory estimates yield existence.*

There are known many others versions of these fixed-point theorem such as Kakutani, Ky-Fan and etc. related with some generalizations of the theorems mentioned above. For details see (Aubin 1979) and (Zeidler 1986).

18.8.2 Brouwer fixed-point theorem

To deal correctly with the Brouwer fixed-point theorem we need the preparations considered below.

The Sperner lemma

Let

$$\boxed{\begin{array}{c} S_N\left(x_0,...,x_N\right) := \\ \left\{x \in \mathcal{X} \mid x = \sum_{i=0}^{N} \lambda_i x_i,\ \lambda_i \geq 0,\ \sum_{i=0}^{N} \lambda_i = 1\right\} \end{array}} \tag{18.107}$$

be an N-simplex in a finite-dimensional normed space $\mathcal{X}$ and $\{S_1, ..., S_J\}$ be a *triangulation* of S_N consisting of N-simplices S_j $(j = 1, ..., J)$ (see Fig.18.1) such that

a) $S_N = \bigcup_{j=1}^{J} S_j$;

b) if $j \neq k$, then the intersection $S_j \bigcap_{j \neq k} S_k$ or empty or a common face of dimension less than N.

Let one of the numbers $(0, 1, ..., N)$ be associated with each vertex v of the simplex S_j. So, suppose that if $v \in S_j := S_j\left(x_{i_0}, ..., x_{i_N}\right)$, then one of numbers i_0, ...,i_N is associated with v.

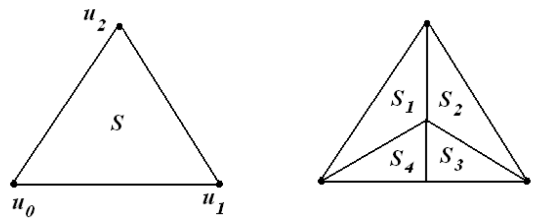

Figure 18.1: N-simplex and its triangulation.

Definition 18.30 *S_j is called a* ***Sperner simplex*** *if and only if all of its vertices carry different numbers, i.e., the vertices of S_N carry different numbers* 0, 1, ..., N.

Lemma 18.16 (Sperner) *The number of Sperner simplices is always odd.*

Proof. It can be easily proven by induction if note that for $N = 1$ each S_j is a 1-simplex (segment). In this case a 0-face (vertex) of S_j is called distinguished if and only if it carries the number 0. So, one has exactly two possibilities (see Fig.18.2 *a*): i) S_j has precisely one distinguished $(N-1)$-face, i.e., S_j is a Sperner simplex; ii) S_j has precisely two or more distinguished $(N-1)$-face, i.e., S_j is not a Sperner simplex. But since the distinguished 0-face occur twice in the interior and once on the boundary, the total number of distinguished 0-faces is odd. Hence, the number of Sperner simplices is odd. Let now $N = 2$ (see Fig.18.2 *b*). Then S_j is 2-simplex and a 1-face (segment) of S_j is called distinguished if and only if it carries the numbers $0, 1$. The the conditions i) and ii) given above are satisfied for $N = 2$. The distinguished 1-faces occur twice in the interior and, by the case $N = 1$, it follows that the number of the distinguished 1-faces is odd. Therefore, the number of the Sperner simplices is odd. Now let $N \geq 3$. Supposing that the lemma is true for $(N-1)$, as in the case $N = 2$, we easily obtain the result. ■

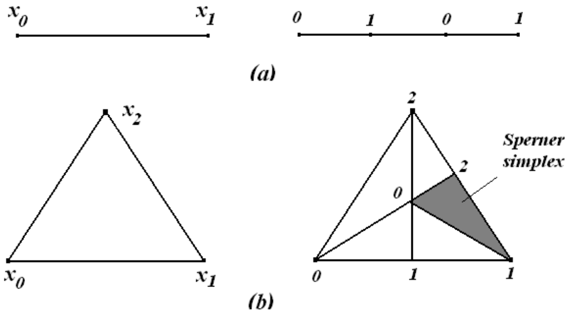

Figure 18.2: The Sperner simplex.

The Knaster-Kuratowski-Mazurkiewicz (KKM) lemma

Lemma 18.17 (Knaster-Kuratowski-Mazurkiewicz) *Let $S_N(x_0, ..., x_N)$ be a N-simplex in a finite-dimensional normed space $\mathcal{X}$. Suppose we are given closed sets $\{C_i\}_{i=1}^N$ in $\mathcal{X}$ such that*

$$\boxed{S_N(x_0, ..., x_N) \subseteq \bigcup_{m=0}^{k} C_{i_m}} \tag{18.108}$$

for all possible systems of indices $\{i_0, ..., i_k\}$ and all $k = 0, ..., N$. Then there exists a point $v \in S_N(x_0, ..., x_N)$ such that $v \in C_j$ for all $j = 0, ..., N$.

Proof. Since for $N = 0$ the set $S_0(x_0)$ consists of a single point x_0, and the statements looks trivial. Let $N \geq 1$. Let v be any vertex of S_j $(j = 0, ..., N)$ (for a triangulation S_1, ..., S_N) such that $v \in S_j(x_{i_0}, ..., x_{i_N})$. By the assumptions of this lemma there exists a set C_k such that $v \in C_k$. We may associate the index k with the vertex v. By the Sperner lemma 18.16 it follows that there exists a Sperner simplex S_j whose vertices carry the numbers $0, ..., N$. Hence the vertices v_0, ..., v_N satisfy the condition $v_k \in C_k$ $(k = 0, ..., N)$. Consider now a sequence of triangulations of simplex $S_N(x_0, ..., x_N)$ such that the diameters of the simplices of the triangulations tend to zero (selecting, for example, a sequence barycentric subdivisions of S). So, there are points $v_k^{(n)} \in C_k$ $(k = 0, ..., N;\ n = 1, 2, ...)$ such that

$\lim_{n\to\infty} \text{diam} S_N \left(v_0^{(n)}, ..., v_N^{(n)}\right) = 0$. Since the simplex $S_N(x_0, ..., x_N)$ is a compact, there exists a subsequence $\left\{v_k^{(s)}\right\}$ such that $v_k^{(s)} \underset{s\to\infty}{\to} v \in S_N(x_0, ..., x_N)$ for all $k = 0, ..., N$. And since the set C_k is closed, this implies $v \in C_k$ for all $k = 0, ..., N$. Lemma is proven. ■

Now we are ready to formulate the main result of this section

The Brouwer theorem

Theorem 18.19 (Brouwer, 1912) *The continuos operator $\Phi : \mathcal{M} \to \mathcal{M}$ has at least one fixed point when $\mathcal{M}$ is a compact, convex, non-empty set in a finite-dimensional normed space over the field $\mathcal{F}$ (real or complex).*

Proof. *a)* Consider this operator when $\mathcal{M} = S_N$ and demonstrate that the continuos operator $\Phi : S_N \to S_N$ $(N = 0, 1, ...)$ has at least one fixed point when $S_N = S_N(x_0, ..., x_N)$ is a N-simplex in a finite-dimensional normed space $\mathcal{X}$. For $N = 0$ the set S_0 consists of a single point and the the statement is trivial. For $N = 1$ the statement is also trivial. Let now $N = 2$. Then $S_2 = S_2(x_0, x_1, x_2)$ and any point x in S_2 can be represented as

$$x = \sum_{i=0}^{2} \lambda_i(x) x_i, \lambda_i \geq 0, \sum_{i=0}^{2} \lambda_i = 1 \tag{18.109}$$

We set

$$C_i := \{x \in S_N \mid \lambda_i(\Phi x) \leq \lambda_i(x), \ i = 0, 1, 2\}$$

Since $\lambda_i(x)$ and Φ are continuous on S_N, the sets C_i are closed and the condition (18.108) of Lemma 18.17 is fulfilled, that is, $S_N \in \overset{k}{\underset{m=0}{\cup}} C_{i_m}$ $(k = 0, 1, 2)$. Indeed, if it is not true, then there exists a point $x \in S_2(x_{i_0}, x_{i_1}, x_{i_2})$ such that $x \notin \overset{k}{\underset{m=0}{\cup}} C_{i_m}$, i.e., $\lambda_{i_m}(\Phi x) > \lambda_{i_m}(x)$ for all $m = 0, ..., k$. But this is in the contradiction to the representation (18.109). Then by Lemma 18.17 there is a point $y \in S_2$ such that $y \in C_j$ $(j = 0, 1, 2)$. This implies $\lambda_i(\Phi y) \leq \lambda_i(y)$ for all $j = 0, 1, 2$. Since also $\Phi y \in S_2$ we have

$$\sum_{i=0}^{2} \lambda_i(y) = \sum_{i=0}^{2} \lambda_i(\Phi y) = 1$$

and, hence, $\lambda_i(\Phi y) = \lambda_i(y)$ for all $j = 0, 1, 2$ that is equivalent to the expression $\Phi y = y$. So, y is the desired fixed-point of Φ in the case $N = 2$. In $N \geq 3$ one can use the same arguments as for $N = 2$.

b) Now, when $\mathcal{M}$ is a compact, convex, non-empty set in a finite-dimensional normed space, it is easy to show that $\mathcal{M}$ is homeomorphic to some N-simplex ($N = 0, 1, 2, ..$). This means that there exist homeomorphisms $\Phi : \mathcal{M} \to \mathcal{B}$ and $C : S_N \to \mathcal{B}$ such that the map

$$C^{-1} \circ \Phi : \mathcal{M} \xrightarrow{\Phi} \mathcal{B} \xrightarrow{C^{-1}} S_N$$

is the desired homeomorphism from the given set $\mathcal{M}$ onto the simplex S_N. Using now this fact shows that each continuos operator $\Phi : \mathcal{M} \to \mathcal{M}$ has at least one fixed point. This completes the proof. ■

Corollary 18.12 *The continuous operator $B : \mathcal{K} \to \mathcal{K}$ has at least fixed point when $\mathcal{K}$ is a subset of a normed space that is homeomorphic to a set $\mathcal{M}$ as it is considered in Theorem 18.19.*

Proof. Let $C : \mathcal{M} \to \mathcal{K}$ be a homeomorphism. Then the operator

$$C^{-1} \circ B \circ C : \mathcal{M} \xrightarrow{C} \mathcal{K} \xrightarrow{B} \mathcal{K} \xrightarrow{C^{-1}} \mathcal{M}$$

is continuous. By Theorem 18.19 there exists a fixed point x^* of the operator $\Phi := C^{-1} \circ B \circ C$, i.e., $C^{-1}(B(Cx^*)) = x^*$. Let $y = Cx$. Then $By = y$, $y \in \mathcal{K}$. Therefore B has a fixed point. Corollary is proven. ■

18.8.3 Schauder fixed-point theorem

This result represents the extension of the Brouwer fixed-point theorem 18.19 to a infinite-dimensional Banach space.

Theorem 18.20 (Schauder, 1930) *The compact operator $\Phi : \mathcal{M} \to \mathcal{M}$ has at least one fixed-point when $\mathcal{M}$ is a bounded, closed convex, nonempty subset of a Banach space $\mathcal{X}$ over the field $\mathcal{F}$ (real or complex).*

Proof ((Zeidler 1995)). Let $x \in \mathcal{M}$. Replacing x with $x - x_0$, if necessary, one may assume that $0 \in \mathcal{M}$. By Theorem 18.7 on the approximation of compact operators it follows that for every $n = 1, 2, ...$ there exists a finite-dimensional subspace $\mathcal{X}_n$ of $\mathcal{X}$ and a continuous operator $\Phi_n : \mathcal{M} \to \mathcal{X}_n$ such that $\|\Phi_n(x) - \Phi_n(x)\| \leq n^{-1}$ for any $x \in \mathcal{M}$. Define $\mathcal{M}_n := \mathcal{M} \cap \mathcal{X}_n$. Then $\mathcal{M}_n$ is a bounded, closed, convex subset of $\mathcal{X}_n$ with $0 \in \mathcal{M}_n$ and $\Phi_n(\mathcal{M}) \subseteq \mathcal{M}$ since $\mathcal{M}$ is convex. By the Brouwer fixed-point theorem 18.19 the operator $\Phi_n : \mathcal{M}_n \to \mathcal{M}_n$ has a fixed point, say x_n, that is, for all $n = 1, 2, ...$ we have $\Phi_n(x_n) = x_n \in \mathcal{M}_n$. Moreover, $\|\Phi(x_n) - x_n\| \leq n^{-1}$. Since $\mathcal{M}_n \subseteq \mathcal{M}$, the sequence $\{x_n\}$ is bounded. The compactness of $\Phi : \mathcal{M} \to \mathcal{M}$ implies the existence of a sequence $\{\tilde{x}_n\}$ such that $\Phi(x_n) \to v$ when $n \to \infty$. By the previous estimate

$$\begin{gathered} \|v - x_n\| = \|[v - \Phi(x_n)] + [\Phi(x_n) - x_n]\| \leq \\ \|[v - \Phi(x_n)]\| + \|\Phi(x_n) - x_n\| \underset{n\to\infty}{\to} 0 \end{gathered}$$

So, $x_n \to v$. Since $\Phi(x_n) \in \mathcal{M}$ and the set $\mathcal{M}$ is closed, we get that $v \in \mathcal{M}$. And, finally, since the operator $\Phi : \mathcal{M} \to \mathcal{M}$ is continuous, it follows that $\Phi(x) = x \in \mathcal{M}$. Theorem is proven. ■

Example 18.21 (Existence of solution for integral equations) *Let solve the following integral equation*

$$\boxed{\begin{gathered} u(t) = \lambda \int_{y=a}^{b} F(t, y, u(y))\, dy \\ -\infty < a < b < \infty, t \in [a, b], \lambda \in \mathbb{R} \end{gathered}} \tag{18.110}$$

Define

$$Q_r := \left\{(t, y, u) \in \mathbb{R}^3 \mid t, y \in [a, b],\ \ |u| \leq r\right\}$$

Proposition 18.6 ((Zeidler 1995)) *Assume that*

a) The function $F : Q_r \to \mathbb{R}$ is continuous;

b) $|\lambda|\, \mu \leq r$, $\mu := \dfrac{1}{b-a} \max_{(t,x,u) \in Q_r} |F(t, x, u)|$;

Setting $\mathcal{X} := C[a, b]$ and $\mathcal{M} := \{u \in \mathcal{X} \mid \|u\| \leq r\}$, it follows that the integral equation (18.110) has at least one solution $u \in \mathcal{M}$.

Proof. For all $t \in [a, b]$ define the operator

$$(Au)(t) := \lambda \int_{y=a}^{b} F(t, y, u(y))\, dy$$

Then the integral equation (18.110) corresponds to the following fixed-point problem $Au = u \in \mathcal{M}$. Notice that the operator $A : \mathcal{M} \to \mathcal{M}$ is compact and for all $u \in \mathcal{M}$

$$\|Au\| \leq |\lambda| \max_{t \in [a,b]} \left| \int_{y=a}^{b} F(t, y, u(y))\, dy \right| \leq |\lambda| \mu \leq r$$

Hence, $A(\mathcal{M}) \subseteq \mathcal{M}$. Thus, by the Schauder fixed-point theorem 18.20 it follows that the equation (18.110) has a solution. ■

18.8.4 The Leray-Schauder principle and a priory estimates

In this subsection we will again concern the solution of the operator equation

$$\boxed{\Phi(x) = x \in \mathcal{X}} \tag{18.111}$$

using the properties of the associated parametrized equation

$$\boxed{t\Phi(x) = x \in \mathcal{X},\ t \in [0, 1)} \tag{18.112}$$

For $t = 0$ the equation (18.112) has the trivial solution $x = 0$, and for $t = 1$ coincides with (18.111). Assume that the following conditions holds:

(A) There is a number $r > 0$ such that if x is a solution of (18.112), then

$$\boxed{\|x\| \leq r} \tag{18.113}$$

Remark 18.10 *Here we do not assume that (18.112) has a solution and, evidently, that the assumption (A) is trivially satisfied if the set $\Phi(\mathcal{X})$ is bounded since $\|\Phi(x)\| \leq r$ for all $x \in \mathcal{X}$.*

Theorem 18.21 (Leray-Schauder, 1934) *If the compact operator $\Phi : \mathcal{X} \to \mathcal{X}$ given on the Banach space $\mathcal{X}$ over the field $\mathcal{F}$ (real or complex) satisfies the assumption (A), then the original equation (18.111) has a solution (non obligatory unique).*

Proof ((Zeidler 1995)). Define the subset

$$\mathcal{M} := \{x \in \mathcal{X} \mid \|x\| \leq 2r\}$$

and the operator

$$B(x) := \begin{cases} \Phi(x) & \text{if } \|\Phi(x)\| \leq 2r \\ 2r\dfrac{\Phi(x)}{\|\Phi(x)\|} & \text{if } \|\Phi(x)\| > 2r \end{cases}$$

Obviously, $\|B(x)\| \leq 2r$ for all $x \in \mathcal{X}$ that implies $B(\mathcal{M}) \subseteq \mathcal{M}$. Show that $B : \mathcal{M} \to \mathcal{M}$ is a compact operator. First, notice that B is continuous because of the continuity of Φ. Then consider the sequences $\{u_n\} \in \mathcal{M}$ and $\{v_n\}$ such that a) $\{v_n\} \in \mathcal{M}$ or b) $\{v_n\} \notin \mathcal{M}$. In the case a) the boundedness of $\mathcal{M}$ and the compactness of Φ imply that there is a subsequence $\{v_{n_k}\}$ such that $B(v_{n_k}) = \Phi(v_{n_k}) \to z$ as $n \to \infty$. In the case b) we may choose this subsequence so that $1/\|\Phi(v_{n_k})\| \to \alpha$ and $\Phi(v_{n_k}) \to z$. Hence, $B(v_{n_k}) \to 2r\alpha z$. So, B is compact. The Schauder fixed-point theorem 18.20 being applied to the compact operator $B : \mathcal{M} \to \mathcal{M}$ provides us with a point $x \in \mathcal{M}$ such that $x = B(x)$. So, if $\|\Phi(x)\| \leq 2r$, then $B(x) = \Phi(x) = x$ and we obtain the solution of the original problem. Another case $\|\Phi(x)\| > 2r$ is impossibly by the assumption (A). Indeed, suppose $\Phi(x) = x$ for $\|\Phi(x)\| > 2r$. Then $x = Bx = t\Phi(x)$ with $t := 2r/\|\Phi(x)\| < 1$. This forces $\|x\| = t\|\Phi(x)\| = 2r$ that contradicts with the assumption (A). Theorem is proven. ■

Remark 18.11 *Theorem 18.21 turns out to be very useful for the justification of the existence of solution for different types of partial differential equations (such as the famous Navier-Stokes equations for viscous fluids, quasi-linear elliptic and etc.).*