VŠB – Technical University of Ostrava
Faculty of Electrical Engineering and Computer Science
Department of Applied Mathematics

Variational methods

Jiří Bouchala
Jan Zapletal

2018

The lecture notes ‘Variational methods’ are a translation of the Czech version ‘Variační metody’ by Jiří Bouchala available at `http://mi21.vsb.cz/modul/variacni-metody`.

October 29, 2018

Contents

1 Why to generalize the notion of a classical solution

Let us consider the heat equation modeling heat distribution (in a thin rod) with homogeneous Dirichlet boundary conditions, i.e., the boundary value problem

$$\begin{cases} -(p(x)y')' + q(x)y = f(x) \quad \text{in} \quad (0,2), \\ y(0) = y(2) = 0. \end{cases} \tag{1.1}$$

The solution, i.e., the function y, describes the stationary heat distribution in a thin rod ($y = y(x)$ is `the temperature in the cross section`) and

- p ... characterizes `the material of the bar` (related to the coefficient of thermal conductivity),
- q ... characterizes `heat exchange with the exterior` (related to the coefficient of heat transfer),
- f ... related to `the heat sources inside of the bar`, `the exterior temperature`, `the coefficient of heat transfer`.

Let us consider two – physically reasonable – cases, when the `the classical solution` to (1.1) `does not exist`:

1)
$$\begin{cases} -y'' = f(x) \quad \text{in} \quad (0,2), \\ y(0) = y(2) = 0; \end{cases} \qquad \text{where } f(x) := \begin{cases} 2 \quad \text{for} \quad x \in \langle 0,1), \\ -2 \quad \text{for} \quad x \in \langle 1,2 \rangle \end{cases}$$

(distribution of the sources is not continuous).

2)
$$\begin{cases} -(p(x)y')' = 1 \quad \text{in} \quad (0,2), \\ y(0) = y(2) = 0; \end{cases} \qquad \text{where } p(x) := \begin{cases} 1 \quad \text{for} \quad x \in \langle 0,1), \\ 2 \quad \text{for} \quad x \in \langle 1,2 \rangle \end{cases}$$

(the bar is made of two different materials).

In either of the two cases the classical solution (i.e., a smooth function satisfying the corresponding equation in every point of $(0,2)$ and vanishing in 0 and 2) does not exist, even though from the physical point of view they are reasonable. This brings up the question of how to solve such problems.

It seems reasonable to split the interval $(0,2)$ into two disjoint ones, namely $(0,1)$ and $(1,2)$. Let us illustrate such an approach on the former example (and leave the second one as an exercise to the reader)

$$\begin{cases} -y_1'' = 2 \quad \text{in} \quad (0,1), \\ y_1(0) = 0; \end{cases} \qquad \begin{cases} -y_2'' = -2 \quad \text{in} \quad (1,2), \\ y_2(2) = 0. \end{cases}$$

We can write

$$\begin{cases} y_1(x) = c_1 + c_2 x - x^2, \\ y_1(0) = c_1 = 0, \end{cases} \qquad \begin{cases} y_2(x) = d_1 + d_2 x + x^2, \\ y_2(2) = d_1 + 2d_2 + 4 = 0, \end{cases}$$

and thus

$$y_1(x) = c_2 x - x^2, \qquad y_2(x) = -4 - 2d_2 + d_2 x + x^2.$$

Now we can 'combine' the solutions on individual intervals:

$$y(x) := \begin{cases} y_1(x) = c_2 x - x^2 \quad \text{for} \quad x \in (0,1), \\ y_2(x) = -4 - 2d_2 + d_2 x + x^2 \quad \text{for} \quad x \in (1,2). \end{cases} \tag{1.2}$$

It remains to compute the constants c_2 a d_2.

The physically reasonable assumption of `the continuity of the solution` in $x = 1$, i.e., the condition

$$y(1-) = y(1+),$$

leads us to

$$y(1-) = y_1(1-) = \underline{c_2 - 1} = y(1+) = y_2(1+) = \underline{-4 - 2d_2 + d_2 + 1},$$

and

$$c_2 = -2 - d_2. \tag{1.3}$$

To fully define the solution `a transmission condition` has to be added:

$$p(1-)y'(1-) = p(1+)y'(1+).$$

From a physical point of view, this condition ensures that the amount of a substance 'flowing into' $x = 1$ is equal to the amount of a substance 'flowing out' of it.

In our case the condition reads

$$y'(1-) = y'(1+),$$

and thus

$$y'(1-) = \underline{c_2 - 2} = y'(1+) = \underline{d_2 + 2},$$

or

$$c_2 = 4 + d_2. \tag{1.4}$$

From (1.3) a (1.4) one can easily conclude that $c_2 = 1$ a $d_2 = -3$, and so (after the substitution into (1.2))

$$y(x) := \begin{cases} x - x^2 \quad \text{for} \quad x \in \langle 0,1 \rangle, \\ 2 - 3x + x^2 \quad \text{for} \quad x \in \langle 1,2 \rangle. \end{cases}$$

We have presented a situation, where the classical solution of a reasonable problem did not exist, yet it was possible to find a 'physically reasonable' solution by 'classical methods'. The aim is to generalize the notion of a solution to a differential equation (or the boundary value problem, to be more precise) also for much more 'exotic' cases and for situations where the 'classical methods' are not sufficient.

To get an idea of an alternative approach, let us consider the following example. Let us assume that u is a classical solution to the boundary value problem

$$\begin{cases} -u'' = f(x) \quad \text{in} \quad \langle 0, 2\rangle, \\ u(0) = u(2) = 0. \end{cases} \tag{1.5}$$

This means that

$$u \in C^2(\langle 0,2\rangle),\ u(0) = u(2) = 0,\ \forall x \in \langle 0,2\rangle\colon\ -u''(x) = f(x).$$

For every $v \in C^1(\langle 0,2\rangle)$ such that $v(0) = v(2) = 0$ it holds

$$\int_0^2 -u''(x)v(x)\,\mathrm{d}x = \int_0^2 f(x)v(x)\,\mathrm{d}x.$$

Applying integration by parts for the right-hand-side integral

$$\int_0^2 -u''(x)v(x)\,\mathrm{d}x = [-u'(x)v(x)]_0^2 + \int_0^2 u'(x)v'(x)\,\mathrm{d}x = \int_0^2 u'(x)v'(x)\,\mathrm{d}x$$

leads us to the fact that for every v of the qualities described above the (classical) solution u to the boundary value problem (1.5) has to satisfy the equality

$$\int_0^2 u'(x)v'(x)\,\mathrm{d}x = \int_0^2 f(x)v(x)\,\mathrm{d}x. \tag{1.6}$$

Note that this relation does not involve the second derivative of the solution u and one can use it for the definition of the so-called 'weak solution' to the corresponding boundary value problem.

Incidentally, physical reasoning (used to set up a mathematical model – an equation describing some phenomenon) often leads to similar 'integral equations'. The differential equation can be derived under additional assumptions on the smoothness of the solution (in our case: $u \in C^2(\langle 0,2\rangle)$). In a sense, it thus seems more natural to define the solution to a particular problem by the integral equation.

To support the above statement let us note that from (1.6) we immediately obtain the aforementioned transmission condition. Indeed, substituting

$$f(x) := \begin{cases} 2 \quad \text{for} \quad x \in \langle 0,1), \\ -2 \quad \text{for} \quad x \in \langle 1,2\rangle, \end{cases}$$

$$u(x) := y(x) = \begin{cases} c_2 x - x^2 \quad \text{for} \quad x \in \langle 0,1), \\ -4 - 2d_2 + d_2 x + x^2 \quad \text{for} \quad x \in \langle 1,2\rangle, \end{cases}$$

into (1.6) leads us to

$$\int_0^1 (c_2 - 2x)v'(x)\,\mathrm{d}x + \int_1^2 (d_2 + 2x)v'(x)\,\mathrm{d}x = \int_0^1 2v(x)\,\mathrm{d}x + \int_1^2 -2v(x)\,\mathrm{d}x, \tag{1.7}$$

and because

- $\int_0^1 2v(x)\,\mathrm{d}x = [2xv(x)]_0^1 - \int_0^1 2xv'(x)\,\mathrm{d}x = 2v(1) - \int_0^1 2xv'(x)\,\mathrm{d}x,$
- $\int_1^2 -2v(x)\,\mathrm{d}x = [-2xv(x)]_1^2 + \int_1^2 2xv'(x)\,\mathrm{d}x = 2v(1) + \int_1^2 2xv'(x)\,\mathrm{d}x,$

we obtain from (1.7) that

$$\int_0^1 (c_2 - 2x + 2x)v'(x)\,\mathrm{d}x + \int_1^2 (d_2 + 2x - 2x)v'(x)\,\mathrm{d}x = 4v(1),$$

$$c_2[v(x)]_0^1 + d_2[v(x)]_1^2 = 4v(1),$$

$$c_2 v(1) - d_2 v(1) = 4v(1).$$

Since we can choose v such that $v(1) \neq 0$, it must hold that

$$c_2 = 4 + d_2.$$

The integrals in (1.6) are all Riemann integrals. To be able to handle as general right-hand sides f (and other quantities) as possible, one first has to generalize the notion of the Riemann integral. There only exist ‘few’ Riemann integrable functions, it is non-trivial to handle taking limits, it limits us to closed intervals… The term function and its derivative will have to be generalized as well.

In the following, we will get to know functions, which

- are Lebesgue integrable,
- posses ‘a generalized’ derivative.

The first chapters of these notes will be devoted to the definitions and basic properties of these special function spaces.

2 Lebesgue measure, Lebesgue integral, $L^p(\Omega)$ spaces

2.1 Lebesgue measure

Motivation 2.1. Let us imagine a situation of having a handful of coins. To find out how much money we actually have we can follow two approaches:

1) taking a coin after coin and adding their value to the total sum
(€ 1 + € 2 + € 1 + € 5 + € 5 + € 2 + € 1 = € 17),
2) sorting the coins first with respect to their value, counting the number of coins in individual heaps and summing up to obtain the total value
(3 · € 1 + 2 · € 2 + 2 · € 5 = 17 €).

In a similar fashion, one can distinguish between the 'Riemann' and 'Lebesgue' approaches to the evaluation of an integral, e.g., $\int_a^b f(x)\,\mathrm{d}x$:

1) 'Riemann sums':
$$\int_a^b f(x)\,\mathrm{d}x \doteq \sum_k f(x_k) \underbrace{(x_{k+1}-x_k)}_{\substack{\text{the length (measure)}\\ \text{of the interval } \langle x_k, x_{k+1}\rangle}},$$
2) 'Lebesgue sums':
$$\int_a^b f(x)\,\mathrm{d}x \doteq \sum_k y_k \cdot \text{`measure } M_k\text{'},$$
where
$$M_k := \{x \in \langle a,b\rangle \colon f(x) \in \langle y_k, y_{k+1})\}.$$

Thus,

`we need to define a measure!`

Definition 2.2. Let X denote an arbitrary set and $\mathcal{P}(X)$ the system of all its subsets (the power set). $\mathcal{A} \subset \mathcal{P}(X)$ is called a σ-algebra (in X), if it holds

i) $X \in \mathcal{A}$,
ii) $\forall A \in \mathcal{A} \colon X \setminus A \in \mathcal{A}$,
iii) $[\forall n \in \mathbb{N} \colon A_n \in \mathcal{A}] \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

The elements of $\mathcal{A}$ are called measurable sets; the set X together with the σ-algebra $\mathcal{A}$ is called a measurable space and we denote it by $(X, \mathcal{A})$.

Exercise 2.3. *Let X denote an arbitrary set. Prove that the below given system of subsets $\mathcal{A}$ is a σ-algebra in X:*

1) $\mathcal{A} = \mathcal{P}(X)$,
2) $\mathcal{A} = \{\emptyset, X\}$,

3) $\mathcal{A} = \{A \subset X\colon A$ *is (at most) countable or* $X \setminus A$ *is a (at most) countable set*$\}$.

Observation 2.4. If $\mathcal{S} \subset \mathcal{P}(X)$, there exists a σ-algebra (in X) containing $\mathcal{S}$. This σ-algebra is said to be generated by the system $\mathcal{S}$.

Definition 2.5. Let $(X, \mathcal{A})$ be a measurable space. A function

$$\mu\colon \mathcal{A} \to \langle 0, +\infty) \cup \{+\infty\}$$

is called a measure (in $\mathcal{A}$), if it holds

i) μ is not identically equal to $+\infty$,

ii) μ is σ-additive, i.e.,

$$\left.\begin{array}{r} \forall n \in \mathbb{N}\colon A_n \in \mathcal{A}, \\ \forall i, j \in \mathbb{N}\colon [i \neq j \;\Rightarrow\; A_i \cap A_j = \emptyset] \end{array}\right\} \Rightarrow \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

The set X together with the σ-algebra $\mathcal{A}$ and the measure μ is called a measure space and is denoted by $(X, \mathcal{A}, \mu)$.

Example 2.6.

1) $\mathcal{A} = \mathcal{P}(X),\ a \in X,\ \mu(A) := \begin{cases} 1, & a \in A, \\ 0, & a \notin A, \end{cases}$

... *Dirac measure;*

2) $\mathcal{A} = \mathcal{P}(X),\ \mu(A) := \begin{cases} \textit{number of elements in } A, \textit{ if } A \textit{ is finite}, \\ +\infty, \textit{ if } A \textit{ is infinite}, \end{cases}$

... *arithmetic measure.*

Theorem 2.7 (Properties of a measure)**.** *Let* $(X, \mathcal{A}, \mu)$ *be a measure space. Then it holds*

i) $\mu(\emptyset) = 0$,

ii) $\forall A, B \in \mathcal{A}:\ [A \subset B \Rightarrow \mu(A) \leq \mu(B)]$,

iii) $\Big[(\forall n \in \mathbb{N}:\ A_n \in \mathcal{A}) \ \wedge\ (A_n \nearrow A)\Big] \Rightarrow \mu(A_n) \to \mu(A)$

$\left(\textit{by } A_n \nearrow A \textit{ we understand that } A_1 \subset A_2 \subset A_3 \subset \ldots \textit{ and } A = \bigcup_{n=1}^{\infty} A_n\right)$,

iv) $\Big[(\forall n \in \mathbb{N}:\ A_n \in \mathcal{A}) \ \wedge\ (A_n \searrow A) \ \wedge\ \big(\underline{\mu(A_1) < +\infty}\big)\Big] \Rightarrow \mu(A_n) \to \mu(A)$

$\left(\textit{by } A_n \searrow A \textit{ we understand that } A_1 \supset A_2 \supset A_3 \supset \ldots \textit{ and } A = \bigcap_{n=1}^{\infty} A_n\right)$.

Remark 2.8. Note that the fact that $A = \bigcap_{n=1}^{\infty} A_n$ is a measurable set is one of the statements of Theorem 2.7, part *iv).*

Definition 2.9. Let $(X, \mathcal{A}, \mu)$ denote a measure space. Then μ is a complete measure, if

$$\forall A, B \subset X\colon \big[\mu(A) = 0,\ B \subset A \ \Rightarrow\ B \in \mathcal{A}\ \ (\Rightarrow\ \mu(B) = 0)\big].$$

Theorem 2.10 (Completion of a measure)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $\mathcal{A}_0$ denote the system of all $E \subset X$ such that there exist sets $A, B \in \mathcal{A}$ satisfying*

$$A \subset E \subset B,\ \mu\left(B \setminus A\right) = 0.$$

Let us define for $E \in \mathcal{A}_0$:

$$\mu_0(E) := \mu(A).$$

Then $(X, \mathcal{A}_0, \mu_0)$ is a space with a complete measure.[1]

Observation 2.11. Obviously, $\mathcal{A} \subset \mathcal{A}_0$ and $\mu = \mu_0$ on $\mathcal{A}$.

Theorem 2.12. *Let us consider the measurable space $(\mathbb{R}^n, \mathcal{B})$, where $n \in \mathbb{N}$ and $\mathcal{B}$ is a σ-algebra generated by the system of all open subsets of $\mathbb{R}^n$.*[2]

Then there exists a unique measure λ on $\mathcal{B}$ such that for every set (an interval)

$$W := (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$$

$$(\forall i \in \{1, 2, \ldots, n\}\colon a_i, b_i \in \mathbb{R},\ a_i < b_i)$$

it holds

$$\lambda(W) = \prod_{i=1}^{n} (b_i - a_i).$$

Definition 2.13. The measure λ appearing in the previous theorem is called the Lebesgue–Borel measure. Let $(\mathbb{R}^n, \mathcal{B}_0, \lambda_0)$ denote the completion of $(\mathbb{R}^n, \mathcal{B}, \lambda)$. The measure λ_0 is called the Lebesgue measure and the system $\mathcal{B}_0$ consists of the Lebesgue measurable sets.

Convention 2.14. *In the following we denote by λ the Lebesgue measure.*

Remark 2.15. It can be shown that

$$\mathcal{B} \subsetneqq \mathcal{B}_0 \subsetneqq \mathcal{P}(\mathbb{R}^n).$$

Theorem 2.16 (Properties of Lebesgue measurable sets and the Lebesgue measure)**.**
It holds

i)

$$E \in \mathcal{B}_0 \ \Leftrightarrow\ \Big[\left(\forall \varepsilon \in \mathbb{R}^+\right) (\exists A, B \subset \mathbb{R}^n) : A \textit{ is closed, } B \textit{ is open,} \\ A \subset E \subset B,\ \lambda(B \setminus A) < \varepsilon \Big];$$

[1] $(X, \mathcal{A}_0, \mu_0)$ is called the completion of $(X, \mathcal{A}, \mu)$.
[2] $\mathcal{B}$ is a σ-algebra of Borel sets.

ii) λ *is* <u>*shift invariant*</u>,[3] *i.e.,*

$$(\forall E \in \mathcal{B}_0)\,(\forall x \in \mathbb{R}^n) : E + x := \{y + x :\ y \in E\} \in \mathcal{B}_0,\ \lambda(E + x) = \lambda(E);$$

iii)

$$\forall E \in \mathcal{B}_0 \colon \lambda(E) = \sup\,\{\lambda(K) \colon K \textit{ is compact, } K \subset E\} =$$
$$= \inf\,\{\lambda(G) \colon G \textit{ is open, } E \subset G\}$$

... <u>*regularity*</u> *of* λ*;*

iv) if $M \subset \mathbb{R}^n$ *is at most countable, then* $\lambda(M) = 0$*;*

v) if $T \colon \mathbb{R}^n \to \mathbb{R}^n$ *is a* `linear` *mapping, then*

$$\forall A \in \mathcal{B}_0 \colon \lambda\,(T(A)) = \lambda(A) \cdot |\det T|.$$

Remark 2.17. It can be shown that there does not exists a normalized measure in $\mathcal{P}(\mathbb{R}^n)$ invariant to shifting. This implies that if we want to work with a measure with such properties, there will exist subsets $M \subset \mathbb{R}^n$ that are not measurable.

One may thus ask, whether our requirements on a measure (see Definition 2.5) are not too ‘strict’. Is there at least a `finitely-additive normalized` set function defined in $\mathcal{P}(\mathbb{R}^n)$ which would be shift invariant? In 1920, it was shown by Stefan Banach that for $n = 1$ or $n = 2$ such a function exists, but it is not unique. In 1914 Felix Hausdorff showed that for $n \geq 3$ such a function does not exist at all.

2.2 Lebesgue integral

Definition 2.18. A function

$$f \colon \mathbb{R}^n \to \mathbb{R}^* := \mathbb{R} \cup \{-\infty, +\infty\}.$$

is <u>(Lebesgue) measurable</u> if the following conditions hold

- $Df \in \mathcal{B}_0$,
- $\forall \alpha \in \mathbb{R} \colon \{x \in \mathbb{R}^n \colon f(x) > \alpha\} \in \mathcal{B}_0$.

A function f is <u>(Lebesgue) measurable on a set</u> $M \subset Df$ if its restriction, i.e., $f_{|M}$, is measurable.

Definition 2.19. Let $M \subset \mathbb{R}^n$. The function

$$\chi_M(x) := \begin{cases} 1, & x \in M, \\ 0, & x \in \mathbb{R}^n \setminus M, \end{cases}$$

is called <u>the indicator function</u> of the set M.

[3] It can be shown that λ is the only <u>normalized</u> (i.e., $\lambda\left((0,1)^n\right) = 1$) <u>Borel</u> (i.e., defined on $\mathcal{B}$) measure in $\mathbb{R}^n$, which is shift invariant.

Observation 2.20. It holds that

$$M \in \mathcal{B}_0 \Leftrightarrow \chi_M \text{ is a measurable function.}$$

Definition 2.21. Any (finite) linear combination of indicator functions of sets in $\mathcal{B}_0$ is called <u>a step function</u> (in $\mathbb{R}^n$). In other words, a function $s\colon \mathbb{R}^n \to \mathbb{R}$ is a step function, if there exist numbers $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$ and sets $A_1, A_2, \dots, A_k \in \mathcal{B}_0$ such that

$$\forall x \in \mathbb{R}^n\colon s(x) = \sum_{i=1}^{k} \alpha_i \chi_{A_i}(x).$$

Observation 2.22. The above described step function s is thus `measurable` and $s(\mathbb{R}^n)$ is a `finite` subset of $\mathbb{R}$.

Definition 2.23. Let us define

$$0 \cdot \infty := 0, \ \infty \cdot 0 := 0.$$

Theorem 2.24 (Properties of measurable functions)**.**
It holds:

i) if the functions f, g, f_k (for every $k \in \mathbb{N}$) are measurable (in $\mathbb{R}^n$) and $\alpha \in \mathbb{R}$, then also the functions αf, $f + g$ (if the sum is well defined everywhere), $|f|$, fg, $\max(f, g)$, $\min(f, g)$, $\frac{f}{g}$ (if g does not vanish anywhere), $\sup f_k$, $\inf f_k$, $\limsup f_k$, $\liminf f_k$, $\lim f_k$ (if it exists) are measurable;

ii) if $f\colon \mathbb{R}^n \to \langle 0, +\infty) \cup \{+\infty\}$ is measurable in $\mathbb{R}^n$, there exists a sequence (s_k) of non-negative step functions such that $s_k \nearrow f$, i.e.,

$$(\forall k \in \mathbb{N}\colon s_{k+1} \geq s_k) \ \wedge \ (\forall x \in \mathbb{R}^n\colon \lim s_k(x) = f(x))$$

*(furthermore: if f is additionally **`bounded`**, the functions (s_k) can be chosen such that $s_k \rightrightarrows f$ in $\mathbb{R}^n$).*

Definition 2.25. Let

$$s\colon \mathbb{R}^n \to \langle 0, \infty)$$

denote a `step function`, $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}^+ \cup \{0\}$ its all `different` values and let (for every $i \in \{1, 2, \dots, k\}$)

$$A_i := \{x \in \mathbb{R}^n\colon s(x) = \alpha_i\},$$

i.e.,

$$s = \sum_{i=1}^{k} \alpha_i \chi_{A_i}.$$

For every $E \in \mathcal{B}_0$ we define

$$\int_E s \,\mathrm{d}\lambda := \sum_{i=1}^{k} \alpha_i \, \lambda(A_i \cap E).$$

Definition 2.26. Let $f\colon \mathbb{R}^n \to \langle 0, +\infty) \cup \{+\infty\}$ denote a `measurable function.` For every $E \in \mathcal{B}_0$, $E \subset Df$ we define

$$\int_E f \,\mathrm{d}\lambda := \sup \left\{ \int_E s \,\mathrm{d}\lambda \colon 0 \le s \le f \text{ in } E,\ s \text{ is a step function} \right\}.$$

Definition 2.27. If $f\colon\ \mathbb{R}^n \to \mathbb{R}^*$ denotes `a measurable function`, we define for every $E \in \mathcal{B}_0$, $E \subset Df$:

$$\int_E f \,\mathrm{d}\lambda := \int_E f^+ \,\mathrm{d}\lambda - \int_E f^- \,\mathrm{d}\lambda, \text{ if the right-hand side is meaningful}$$

$$\Big(f^+ := \max{(f, 0)},\ f^- := \max{(-f, 0)};\ \text{ i.e., } f = f^+ - f^-\Big).$$

Notation 2.28. *Sometimes we will also use the notation*

$$\int_E f \,\mathrm{d}\lambda := \int_E f(x) \,\mathrm{d}x.$$

Remark 2.29. It is useful also to define integrals for functions defined only almost everywhere (abbreviated to 'a.e.'), i.e., everywhere except for a set of measure zero – see the following definition.

Definition 2.30. If the function $f\colon \mathbb{R}^n \to \mathbb{R}^*$ is measurable, $E \in \mathcal{B}_0$, $E \setminus N \subset Df$, where $\lambda(N) = 0$, we define

$$\int_E f \,\mathrm{d}\lambda := \int_{E\setminus N} f \,\mathrm{d}\lambda.$$

Definition 2.31. Let $1 \le p < +\infty$, $E \in \mathcal{B}_0$. We define

$$\mathcal{L}^p(E) := \left\{ f\colon \mathbb{R}^n \to \mathbb{R}^* \colon f \text{ is measurable, } \int_E |f|^p \,\mathrm{d}\lambda < +\infty \right\}.$$

Theorem 2.32 (Levi – 'Lebesgue monotone convergence theorem')**.**
Assume that

- *for every $n \in \mathbb{N}$, f_n is a measurable function,*
- *$f_n \nearrow f$ almost everywhere in $E \in \mathcal{B}_0$,*
- *$\int_E f_1 \,\mathrm{d}\lambda > -\infty$.*

Then it holds

$$\int_E f_n \,\mathrm{d}\lambda \to \int_E f \,\mathrm{d}\lambda.$$

Example 2.33. *Consider*

$$f_n := -\frac{1}{n},\ f := 0.$$

Then every function f_n is measurable and $f_n \nearrow f$ everywhere in $\mathbb{R}$, but since

$$\forall n \in \mathbb{N} \colon \int_{\mathbb{R}} f_n \,\mathrm{d}\lambda = -\infty, \quad \int_{\mathbb{R}} f \,\mathrm{d}\lambda = 0,$$

it does not hold that

$$\int_{\mathbb{R}} f_n \,\mathrm{d}\lambda \;\to\; \int_{\mathbb{R}} f \,\mathrm{d}\lambda.$$

(The assumption $\int_{\mathbb{R}} f_1 \,\mathrm{d}\lambda > -\infty$ is not satisfied.)

Example 2.34. *Let (x_n) denote a sequence such that*

$$\mathbb{Q} \cap \langle 0, 1\rangle = \{x_n \colon n \in \mathbb{N}\},$$

and let

$$f_n := \chi_{\{x_1, x_2, \dots, x_n\}},\; f := \chi_{\mathbb{Q}\,\cap\,\langle 0,1\rangle}.$$

Then, since

$$0 \le f_n \nearrow f,$$

it holds

$$\int_{\langle 0,1\rangle} f_n \,\mathrm{d}\lambda = 0 \;\to\; \int_{\langle 0,1\rangle} f \,\mathrm{d}\lambda = 0.$$

(Note that the Riemann integral $\int_0^1 f(x)\,\mathrm{d}x$ does not exist.)

Theorem 2.35 (Lebesgue – 'Lebesgue dominated convergence theorem')**.**
Assume that

- *for every $n \in \mathbb{N}$, f_n is a measurable function,*
- *$f_n \to f$ almost everywhere in $E \in \mathcal{B}_0$,*
- *there exists a function $g \in \mathcal{L}^1(E)$ such that for every $n \in \mathbb{N}$ it holds that $|f_n| \le g$ almost everywhere in E.*

Then it holds

$$f \in \mathcal{L}^1(E),\; \int_E |f_n - f| \,\mathrm{d}\lambda \to 0,\; \int_E f_n \,\mathrm{d}\lambda \;\to\; \int_E f \,\mathrm{d}\lambda.$$

Example 2.36. *Let*

$$f_n := n\,\chi_{(0,\frac{1}{n})},\; f := 0.$$

Then every function f_n is measurable and $f_n \to f$ everywhere in $\mathbb{R}$, but since

$$\forall n \in \mathbb{N}\colon \int_{\mathbb{R}} f_n \,\mathrm{d}\lambda = 1, \quad \int_{\mathbb{R}} f \,\mathrm{d}\lambda = 0,$$

it does not hold that

$$\int_{\mathbb{R}} f_n \,\mathrm{d}\lambda \;\to\; \int_{\mathbb{R}} f \,\mathrm{d}\lambda$$

(the function g (a majorant) required in the Lebesgue theorem does not exist).

Theorem 2.37 (Fatou's lemma)**.**
Assume that

- *for every $n \in \mathbb{N}$, f_n is a measurable function,*
- *$g \in \mathcal{L}^1(E)$, $E \in \mathcal{B}_0$,*

- *for every* $n \in \mathbb{N}$ *je* $f_n \geq g$ *almost everywhere in* E.

Then it holds that

$$\int_E \liminf f_n \,\mathrm{d}\lambda \leq \liminf \int_E f_n \,\mathrm{d}\lambda.$$

Example 2.38. *Let*

$$f_n := \begin{cases} \chi_{(-1,0)}, & \text{for } n \texttt{ odd}, \\ \chi_{(0,1)}, & \text{for } n \texttt{ even}. \end{cases}$$

Then

$$\int_{\mathbb{R}} \liminf f_n \,\mathrm{d}\lambda = \int_{\mathbb{R}} 0 \,\mathrm{d}\lambda = 0$$

$$\wedge$$

$$\liminf \int_{\mathbb{R}} f_n \,\mathrm{d}\lambda = \lim \int_{\mathbb{R}} f_n \,\mathrm{d}\lambda = \lim 1 = 1.$$

Theorem 2.39 (Fubini)**.** *Let* $p, q \in \mathbb{N}$ *and assume that the integral (finite or infinite)*

$$\int_{\mathbb{R}^{p+q}} f \,\mathrm{d}\lambda$$

exists. Then

- *for almost all* $x \in \mathbb{R}^p$ *the integral*

$$\int_{\mathbb{R}^q} f_x \,\mathrm{d}\lambda := \varphi(x) \quad (f_x(y) := f(x,y))$$

  is well defined,

- *for almost* $y \in \mathbb{R}^q$ *the integral*

$$\int_{\mathbb{R}^p} f^y \,\mathrm{d}\lambda := \psi(y) \quad (f^y(x) := f(x,y))$$

  is well defined and it holds that

$$\int_{\mathbb{R}^{p+q}} f \,\mathrm{d}\lambda = \int_{\mathbb{R}^p} \varphi \,\mathrm{d}\lambda = \int_{\mathbb{R}^q} \psi \,\mathrm{d}\lambda,$$

  i.e., [4]

$$\int_{\mathbb{R}^{p+q}} f(x,y) \,\mathrm{d}x \,\mathrm{d}y = \int_{\mathbb{R}^p} \left(\int_{\mathbb{R}^q} f(x,y) \,\mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}^q} \left(\int_{\mathbb{R}^p} f(x,y) \,\mathrm{d}x \right) \mathrm{d}y.$$

Theorem 2.40 (substitution)**.**
Assume that

[4] We denote

$$\int_{\mathbb{R}^{p+q}} f \,\mathrm{d}\lambda := \int_{\mathbb{R}^{p+q}} f(x,y) \,\mathrm{d}x \,\mathrm{d}y.$$

- *$G \subset \mathbb{R}^n$ is an open set,*
- *$\varphi \colon \mathbb{R}^n \to \mathbb{R}^n$ is a mapping such that*
  - *φ is injective in G,*
  - *$\varphi \in C^1$ in G,*
  - *$\forall z \in G\colon\ \det \varphi'(z) \neq 0$,*
- *$f \colon \mathbb{R}^n \to \mathbb{R}^*$.*

Then it holds that

$$\int_{\varphi(G)} f(x)\,\mathrm{d}x = \int_G f(\varphi(t))\,|\det \varphi'(t)|\,\mathrm{d}t,$$

if `one` *of these integrals is well defined.*

Theorem 2.41 (connection between the Riemann and Lebesgue integrals)**.**
Let $f \colon \mathbb{R} \to \mathbb{R}$ denote a **`bounded`** *function in the interval $\langle a, b\rangle \subset \mathbb{R}$. Then it holds:*

i) if $(R)\displaystyle\int_a^b f(x)\,\mathrm{d}x$ is well defined, $(L)\displaystyle\int_a^b f(x)\,\mathrm{d}x$ is well defined too and the integrals are equal,

ii) $(R)\displaystyle\int_a^b f(x)\,\mathrm{d}x$ is well defined $\Leftrightarrow$ the function f is continuous almost everywhere $\langle a, b\rangle$.[5]

Definition 2.42.
Let us assume that

- $-\infty \leq a < b \leq +\infty$,
- $f \colon (a, b) \to \mathbb{R}$,
- $\forall x \in (a, b)\colon F'(x) = f(x)$.

If the limits

$$F(b-) := \lim_{x \to b-} F(x),\ \ F(a+) := \lim_{x \to a+} F(x)$$

exist and are finite, we call their difference a Newton integral of the function f in (a, b); i.e.,

$$(N)\int_a^b f(x)\,\mathrm{d}x = F(b-) - F(a+) =: \big[F(x)\big]_a^b \in \mathbb{R}.$$

Theorem 2.43 (connection between the Newton and Lebesgue integrals)**.**
It holds:

i) if $(N)\displaystyle\int_a^b f(x)\,\mathrm{d}x$ and $(L)\displaystyle\int_a^b f(x)\,\mathrm{d}x$ are well defined, we have

$$(N)\int_a^b f(x)\,\mathrm{d}x = (L)\int_a^b f(x)\,\mathrm{d}x,$$

[5] I.e.,

$$\exists N \subset \mathbb{R}\colon\ \Big[(\lambda(N) = 0)\ \wedge\ (\forall x \in \langle a, b\rangle \setminus N \colon f \text{ is continuous in } x)\Big].$$

ii)

$$\left.\begin{array}{c} f \in \mathcal{L}^1((a,b)) \\ f \text{ is continuous in } (a,b) \end{array}\right\} \Rightarrow (N) \int_a^b f(x)\,\mathrm{d}x \text{ is well defined,}$$

iii) if $(N) \displaystyle\int_a^b f(x)\,\mathrm{d}x$ *is well defined, then* f *is measurable. Furthermore, if* $(L) \displaystyle\int_a^b f(x)\,\mathrm{d}x$ *does not exist,* $(N) \displaystyle\int_a^b |f(x)|\,\mathrm{d}x$ *does not exist, either.*

Example 2.44.

$$(N) \int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}, \ (L) \int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x \text{ does not exist, } (N) \int_0^\infty \left|\frac{\sin x}{x}\right|\mathrm{d}x \text{ does not exist.}$$

2.3 $L^p(\Omega)$ spaces

Theorem 2.45 (properties of $\mathcal{L}^1(E)$)**.** *Let* $E \in \mathcal{B}_0$, *then it holds*

i) if $f \in \mathcal{L}^1(E)$, *then* f *is finite almost everywhere (in* E*),*

ii) if $f, g \in \mathcal{L}^1(E)$ *and* $\alpha, \beta \in \mathbb{R}$, *then* $\alpha f + \beta g \in \mathcal{L}^1(E)$ *and it holds*

$$\int_E (\alpha f + \beta g)\,\mathrm{d}\lambda = \alpha \int_E f\,\mathrm{d}\lambda + \beta \int_E g\,\mathrm{d}\lambda,$$

iii) if $f \in \mathcal{L}^1(E)$, *then* $|f| \in \mathcal{L}^1(E)$ *and it holds*

$$\left|\int_E f\,\mathrm{d}\lambda\right| \leq \int_E |f|\,\mathrm{d}\lambda$$

(i.e., the Lebesgue integral is <u>absolutely convergent</u>),

iv) if $f, g \in \mathcal{L}^1(E)$, *then also* $\max(f,g), \min(f,g) \in \mathcal{L}^1(E)$,

v) if f *is measurable and there exists* $g \in \mathcal{L}^1(E)$ *such that* $|f| \leq g$ *almost everywhere in* E, *then* $f \in \mathcal{L}^1(E)$.

Theorem 2.46 (Hölder and Minkowski inequalities)**.**
Let $E \in \mathcal{B}_0$; $p, q \in (1, +\infty)$;

$$\frac{1}{p} + \frac{1}{q} = 1 \quad (p, q \ \ldots \ \underline{\text{conjugate exponents}}).$$

Then it holds that

(H) if $f \in \mathcal{L}^p(E)$, $g \in \mathcal{L}^q(E)$, *then* $fg \in \mathcal{L}^1(E)$ *and it holds*

$$\int_E fg\,\mathrm{d}\lambda \leq \left(\int_E |f|^p\,\mathrm{d}\lambda\right)^{1/p} \left(\int_E |g|^q\,\mathrm{d}\lambda\right)^{1/q},$$

(M) if $f, g \in \mathcal{L}^p(E)$, then $f+g \in \mathcal{L}^p(E)$ and it holds

$$\left(\int_E |f+g|^p \,\mathrm{d}\lambda\right)^{1/p} \leq \left(\int_E |f|^p \,\mathrm{d}\lambda\right)^{1/p} + \left(\int_E |g|^p \,\mathrm{d}\lambda\right)^{1/p}.$$

Observation and definition 2.47. Let $E \in \mathcal{B}_0$, $1 \leq p < +\infty$. Although the properties of the space $\mathcal{L}^p(E)$ with the classically defined addition and multiplication and the functional

$$\|f\|_p := \left(\int_E |f|^p \,\mathrm{d}\lambda\right)^{1/p}$$

resemble the structure of a normed vector space, $\mathcal{L}^p(E)$ is not one.[6] To be able to 'utilize' the theory of normed vector spaces, we assign to each function $f \in \mathcal{L}^p(E)$ a class of functions

$$[f] := \{g \in \mathcal{L}^p(E) \colon g = f \text{ almost everywhere in } E\}$$

and define

$$L^p(E) := \{[f] \colon f \in \mathcal{L}^p(E)\}.$$

Then

$$\left(L^p(E), +, -, \|\cdot\|_{L^p(E)}\right),$$

where

$$[f]+[g] := [f+g],\ \alpha[f] := [\alpha f]\ (\alpha \in \mathbb{R}),\ \|[f]\|_{L^p(E)} := \|f\|_p,$$

is a normed vector space.[7]

In mathematical literature it is common not to distinguish between the `functions` and `classes of functions`. Intuitively, `we consider the functions differing on a set of measure zero identical`.

Let us generalize the spaces $L^p(E)$ also for the case of $p = \infty$.

Definition 2.48. Let $E \in \mathcal{B}_0$. We denote by $L^\infty(E)$ the set of all measurable functions f (more accurately: the classes of functions equal almost everywhere in E), for which there exists $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for a.a. $x \in E$ (i.e., $|f| \leq M$ almost everywhere in E).

The lowest constant M satisfying this property is called <u>the L^∞-norm</u> of the function f, i.e.,

$$\|f\|_{L^\infty(E)} := \operatorname*{ess\,sup}_E |f| := \inf_{\substack{N \subset E \\ \lambda(N)=0}} \sup_{x \in E \setminus N} |f(x)|.$$

Theorem 2.49 (properties of $L^p(\Omega)$)**.** *Let Ω denote a non-empty* `domain` *in $\mathbb{R}^n$ (i.e., an* `open` *and* `connected` *set). Then it holds*

---

[6] Think this through!

[7] The above operations $+, -, \|\cdot\|_{L^p(E)}$ are well defined as they are independent of the chosen representatives of the corresponding classes.

i) $\forall p \in \langle 1, +\infty) \cup \{+\infty\}$: $L^p(\Omega)$ *is a* `Banach space`,

ii) $L^2(\Omega)$ *with the inner product*

$$(u, v) := \int_\Omega uv \, \mathrm{d}\lambda = \int_\Omega u(x)v(x) \, \mathrm{d}x$$

is a `Hilbert space`*,*

iii) if $1 \leq p < +\infty$*, then* $C^\infty(\Omega) \cap L^p(\Omega)$ *is a* `dense subset` *of* $L^p(\Omega)$*,*

iv) if $1 < p_1 < p_2 < +\infty$ *and* $\lambda(\Omega) < +\infty$*, then*

$$L^\infty(\Omega) \subset L^{p_2}(\Omega) \subset L^{p_1}(\Omega) \subset L^1(\Omega).$$

Exercise 2.50. *Prove part iv) of Theorem 2.49.*

3 Generalized functions (distributions), generalized derivatives

Let Ω denote a `non-empty domain` in $\mathbb{R}^n$ ($n \in \mathbb{N}$).

We denote by $\mathcal{D}(\Omega)$ – the so-called space of testing functions – the set of all functions $\varphi \in C^\infty(\Omega)$, for which it holds

- $\operatorname{supp}\varphi := \overline{\{x \in \mathbb{R}^n : \varphi(x) \neq 0\}} \subset \Omega$
  ($\operatorname{supp}\varphi$ ... the so-called support of the function φ),
- $\operatorname{supp}\varphi$ is compact in $\mathbb{R}^n$
  (i.e., bounded and closed).

Note that if we consider addition and multiplication of functions in $\mathcal{D}(\Omega)$ defined in the standard way, i.e., for all $\varphi_1, \varphi_2 \in \mathcal{D}(\Omega)$ and $c \in \mathbb{R}$

$$(\varphi_1 + \varphi_2)(x) := \varphi_1(x) + \varphi_2(x), \quad (c\varphi_1)(x) := c\,\varphi_1(x),$$

then $\mathcal{D}(\Omega)$ is a vector space. In the text below we will use the following notation:

- $\alpha = (\alpha_1, \alpha_2, \ldots \alpha_n) \in \mathbb{R}^n,$ where $\alpha_i \in \mathbb{N} \cup \{0\}$,
  ... multiindex;
- $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_n$
  ... the length of a multiindex;
- $u\colon \mathbb{R}^n \to \mathbb{R}$, $u \in C^{|\alpha|}(\Omega)$:

$$D^\alpha u := \frac{\partial^{\alpha_1+\alpha_2+\cdots+\alpha_n}\, u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_n^{\alpha_n}} = \frac{\partial^{|\alpha|}\, u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_n^{\alpha_n}}.$$

Example 3.1. *For $n = 3$, $\alpha = (3, 0, 2)$ and $u \in C^5(\Omega)$ we have*

$$D^\alpha u = \frac{\partial^5 u}{\partial x^3 \partial z^2} = \frac{\partial^5 u}{\partial z^2 \partial x^3} = \frac{\partial^5 u}{\partial x \partial z \partial x^2 \partial z} = \ldots,$$

since the derivatives of the fifth order can be interchanged for

$$u \in C^{|\alpha|}(\Omega) = C^5(\Omega).$$

Definition 3.2. Let (φ_n) denote a sequence in $\mathcal{D}(\Omega)$ and let $\varphi \in \mathcal{D}(\Omega)$. We say that the sequence (φ_n) converges in $\mathcal{D}(\Omega)$ to a function φ, and we write

$$\underline{\varphi_n \to \varphi \text{ in } \mathcal{D}(\Omega)},$$

if there exists a `compact set` $K \subset \Omega$ such that:

- for all $n \in \mathbb{N}$:
$$\operatorname{supp}\varphi_n \subset K,\ \operatorname{supp}\varphi \subset K,$$
- for every multiindex α:
$$D^\alpha \varphi_n \rightrightarrows D^\alpha \varphi \text{ in } K.$$

Remark 3.3. It can be shown that the space $\mathcal{D}(\Omega)$ `is not metrizable`, i.e., that there `does not exist` any metric ϱ in $\mathcal{D}(\Omega)$ such that

$$\varphi_n \to \varphi \text{ in } \mathcal{D}(\Omega) \;\Leftrightarrow\; \varrho(\varphi_n, \varphi) \to 0 \text{ (in } \mathbb{R}).$$

Definition 3.4. We call the space of all continuous linear functionals acting on $\mathcal{D}(\Omega)$ the space of <u>distributions</u> in Ω and denote it by $\mathcal{D}^*(\Omega)$, i.e.,

$$F \in \mathcal{D}^*(\Omega) \;\Leftrightarrow\; \begin{cases} \bullet \;\; F\colon \mathcal{D}(\Omega) \to \mathbb{R}, \\ \bullet \;\; \varphi_n \to \varphi \text{ in } \mathcal{D}(\Omega) \;\Rightarrow\; F(\varphi_n) \to F(\varphi), \\ \bullet \;\; \left.\begin{matrix} \varphi_1, \varphi_2 \in \mathcal{D}(\Omega) \\ c_1, c_2 \in \mathbb{R} \end{matrix}\right\} \;\Rightarrow\; F(c_1\varphi_1 + c_2\varphi_2) = c_1F(\varphi_1) + c_2F(\varphi_2). \end{cases}$$

Notation 3.5. *If $F \in \mathcal{D}^*(\Omega)$ and $\varphi \in \mathcal{D}(\Omega)$, we will also use the notation*

$$\langle F, \varphi \rangle := F(\varphi).$$

Definition 3.6. We say that a function f belongs to $L^1_{loc}(\Omega)$ if

$$(\forall x \in \Omega)\,(\exists U(x) \subset \Omega) : f \in L^1(U(x)).$$

Remark 3.7. It can be shown that $f \in L^1_{loc}(\Omega)$ if and only if for every compact set $K \subset \Omega$ it holds that $f \in L^1(K)$.

Example 3.8. *Let us define f and g by*

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

Then

$$f \in L^1(0,1), \;\; \textit{and thus also } f \in L^1_{loc}(0,1),$$

$$g \in L^1_{loc}(0,1), \;\; \textit{but } g \notin L^1(0,1).$$

Remark 3.9. Let use denote

$$f \in L^1_{loc}(\Omega)$$

and define the mapping

$$F\colon \mathcal{D}(\Omega) \to \mathbb{R}$$

by[8]

$$F(\varphi) := \int_\Omega f(x)\varphi(x)\,\mathrm{d}x.$$

Then $F \in \mathcal{D}^*(\Omega)$.

[8] The (Lebesgue) integral exists under these conditions!

Let $G \in \mathcal{D}^*(\Omega)$ denote a distribution such that for some function $g \in L^1_{loc}(\Omega)$ it holds that

$$\forall \varphi \in \mathcal{D}(\Omega)\colon \langle G, \varphi \rangle = \int_\Omega g(x)\varphi(x)\,\mathrm{d}x.$$

Then

$$F = G \;\Leftrightarrow\; f = g \;\text{ a. e. in } \Omega.$$

We can thus identify $F \sim f$ to obtain the inclusion

$$L^1_{loc}(\Omega) \subset \mathcal{D}^*(\Omega).$$

Definition 3.10. A distribution $F \in \mathcal{D}^*(\Omega)$ is called <u>regular</u>, if there exists $f \in L^1_{loc}(\Omega)$ such that

$$\forall \varphi \in \mathcal{D}(\Omega)\colon \langle F, \varphi \rangle = \int_\Omega f(x)\varphi(x)\,\mathrm{d}x.$$

Example 3.11. *Let $a \in \Omega$. Let us define a distribution δ_a by*

$$\langle \delta_a, \varphi \rangle := \varphi(a).$$

Then δ_a (the so-called <u>Dirac function</u> or <u>Dirac distribution</u>) `is not a regular` *distribution. I.e.,*

$$\mathcal{D}(\Omega) \subsetneqq L^1_{loc}(\Omega) \subsetneqq \mathcal{D}^*(\Omega).$$

Observation 3.12. Let

$$f \in \mathcal{D}(\mathbb{R}) \subset \mathcal{D}^*(\mathbb{R})$$

$$\left(\forall \varphi \in \mathcal{D}(\mathbb{R})\colon \langle f, \varphi \rangle = \int_\mathbb{R} f(x)\varphi(x)\,\mathrm{d}x\right).$$

Then it also holds that

$$f' \in \mathcal{D}(\mathbb{R}) \subset \mathcal{D}^*(\mathbb{R})$$

$$\left(\forall \varphi \in \mathcal{D}(\mathbb{R})\colon \langle f', \varphi \rangle = \int_\mathbb{R} f'(x)\varphi(x)\,\mathrm{d}x\right).$$

Moreover,

$$(\forall \varphi \in \mathcal{D}(\mathbb{R}))\left(\exists a \in \mathbb{R}^+\right) : \operatorname{supp}\varphi \subset \langle -a, a \rangle,$$

and thus

$$\langle f', \varphi \rangle = \int_{-a}^{a} f'(x)\varphi(x)\,\mathrm{d}x = \underbrace{[f(x)\varphi(x)]_{-a}^{a}}_{=0} - \int_{-a}^{a} f(x)\varphi'(x)\,\mathrm{d}x = -\langle f, \varphi' \rangle.$$

In the following definition we generalize the concept of differentiation in $\mathcal{D}(\Omega)$ to $\mathcal{D}^*(\Omega)$.

Definition 3.13. Let

$$F \in \mathcal{D}^*(\Omega).$$

The (generalized) derivative of a distribution F with respect to the i-th variable is a distribution

$$\frac{\partial F}{\partial x_i} \in \mathcal{D}^*(\Omega)$$

defined by

$$\left\langle \frac{\partial F}{\partial x_i}, \varphi \right\rangle := -\left\langle F, \frac{\partial \varphi}{\partial x_i} \right\rangle.$$

For every multiindex α we define a distribution

$$D^\alpha F \in \mathcal{D}^*(\Omega)$$

by

$$\langle D^\alpha F, \varphi \rangle := (-1)^{|\alpha|} \langle F, D^\alpha \varphi \rangle.$$

Theorem 3.14. *Let*

$$f \in C^k(\Omega) \quad (\subset L^1_{loc}(\Omega) \subset \mathcal{D}^*(\Omega))$$

and α is a multiindex such that $|\alpha| = k$. Then for every $\varphi \in \mathcal{D}(\Omega)$ it holds that[9]

$$\langle D^\alpha f, \varphi \rangle = \int_\Omega (D^\alpha f)(x) \varphi(x) \,\mathrm{d}x.$$

Example 3.15. *Let us consider the Heaviside function*

$$\eta(x) := \begin{cases} 0, \ x < 0, \\ 1, \ x \geq 0. \end{cases}$$

Then

$$\eta \in L^1_{loc}(\mathbb{R}) \subset \mathcal{D}^*(\mathbb{R})$$

and for every $\varphi \in \mathcal{D}(\mathbb{R})$ it holds that

$$\langle \eta', \varphi \rangle = -\langle \eta, \varphi' \rangle = -\int_{\mathbb{R}} \eta(x) \varphi'(x) \,\mathrm{d}x = -\int_0^\infty \varphi'(x) \,\mathrm{d}x =$$

$$= -\left[\varphi(x)\right]_0^\infty = -(\varphi(\infty) - \varphi(0)) = \varphi(0) = \langle \delta_0, \varphi \rangle,$$

and thus[10]

$$\eta' = \delta_0.$$

[9] The so-called 'generalized' derivative = 'classical' derivative.

[10] By the derivative we understand the derivative `in the distributional sense` – see Definition 3.13.

4 Sobolev spaces

Assume that
$$\emptyset \neq \Omega \subset \mathbb{R}^n,$$
where $n \in \mathbb{N}$, is a `bounded` domain and
$$k \in \mathbb{N} \cup \{0\},\ 1 \leq p < \infty.$$
Let us consider the vector space $C^\infty(\overline{\Omega})$ equipped with a norm
$$\|u\|_{k,p} := \left(\sum_{|\alpha| \leq k} \int_\Omega |D^\alpha u(x)|^p \, \mathrm{d}x \right)^{1/p}.$$

Example 4.1.
$$\|u\|_{1,p} = \left(\int_\Omega \left(|u(x)|^p + \sum_{i=1}^n \left| \frac{\partial u}{\partial x_i}(x) \right|^p \right) \mathrm{d}x \right)^{1/p},$$
$$\|u\|_{2,p} = \left(\int_\Omega \left(|u(x)|^p + \sum_{i=1}^n \left| \frac{\partial u}{\partial x_i}(x) \right|^p + \sum_{i,j=1}^n \left| \frac{\partial^2 u}{\partial x_i \partial x_j}(x) \right|^p \right) \mathrm{d}x \right)^{1/p}.$$

Definition 4.2. We define the <u>Sobolev space</u>
$$W^{k,p}(\Omega)$$
as the completion of the space
$$\left(C^\infty(\overline{\Omega}), \| \cdot \|_{k,p} \right).$$

Observation 4.3.
$$L^2(\Omega) = W^{0,2}(\Omega) \supset W^{1,2}(\Omega) \supset W^{2,2}(\Omega) \supset \dots$$

Observation 4.4. Let $u \in W^{k,2}(\Omega)$. Then there exists a sequence (u_n) in $C^\infty(\overline{\Omega})$ such that $u_n \to u$ in $W^{k,2}(\Omega)$ and thus (u_n) is a Cauchy sequence in $W^{k,2}(\Omega)$. For an arbitrary multiindex α such that $|\alpha| \leq k$ it holds that
$$\|D^\alpha u_n - D^\alpha u_m\|_{L^2(\Omega)} \leq \|u_n - u_m\|_{k,2},$$
and so $(D^\alpha u_n)$ is `a Cauchy sequence` in the `(complete!)` space $L^2(\Omega)$. There thus exists a function $f_\alpha \in L^2(\Omega)$ such that
$$D^\alpha u_n \to f_\alpha \text{ in } L^2(\Omega).$$
It can be proven that for such functions f_α it holds that [11]
$$f_\alpha = D^\alpha f_{(0,0,\dots,0)} = D^\alpha u.$$

[11] Again, the derivatives have to be understood in the distributional sense.

We have found out that

$$W^{k,2}(\Omega) \subset \widetilde{W}^{k,2}(\Omega),$$

where $\widetilde{W}^{k,2}(\Omega)$ is defined as the set of functions $u \in L^2(\Omega)$ such that for every multiindex α, $|\alpha| \leq k$, the generalized (distributional) derivative $D^\alpha u$ is an element of $L^2(\Omega)$.

This can be further generalized to obtain

$$W^{k,p}(\Omega) \subset \widetilde{W}^{k,p}(\Omega) := \{u \in L^p(\Omega) \colon \; D^\alpha u \in L^p(\Omega) \text{ for every multiindex } \alpha, |\alpha| \leq k\}.$$

'Definition' 4.5. Let $1 < n \in \mathbb{N}$. We say that `a bounded` domain

$$\emptyset \neq \Omega \subset \mathbb{R}^n$$

is Lipschitz if there exists a `finite` cover of the boundary of Ω, i.e., a finite set of neighbourhoods such that for every such a neighbourhood $\mathcal{U}$ there exist

- a Cartesian system of coordinates $(\underbrace{y_1, y_2, \dots, y_{n-1}}_{=:\, y'}, y_n) =: (y', y_n)$,
- $\varepsilon, \delta \in \mathbb{R}^+$,
- a function $a \colon \; \mathbb{R}^{n-1} \to \mathbb{R}$,

so that

- $\Gamma := \mathcal{U} \cap \partial\Omega = \{(y', y_n) \colon \; \|y'\| < \delta, \; y_n = a(y')\}$,
- $\mathcal{U}^+ := \{(y', y_n) \colon \; \|y'\| < \delta, \; a(y') < y_n < a(y') + \varepsilon\} \subset \Omega$,
- $\mathcal{U}^- := \{(y', y_n) \colon \; \|y'\| < \delta, \; a(y') - \varepsilon < y_n < a(y')\} \subset \mathbb{R}^n \setminus \overline{\Omega}$,
- the function a is Lipschitz continuous in $\{y' \colon \; \|y'\| < \delta\}$, i.e., there exists $L \in \mathbb{R}^+$ such that for every $y', z' \in \mathbb{R}^{n-1}$ it holds that

$$(\|y'\| < \delta \;\wedge\; \|z'\| < \delta) \;\Rightarrow\; |a(y') - a(z')| \leq L \, \|y' - z'\|.$$

In $\mathbb{R}$ every bounded set (i.e., every bounded open interval) is Lipschitz.

Remark 4.6. It can be shown that for a Lipschitz domain Ω and 'almost all $x \in \partial\Omega$' there exists a unit exterior normal vector to $\partial\Omega$ in x,

$$\nu(x) = (\nu_1(x), \nu_2(x), \dots, \nu_n(x)) \in \mathbb{R}^n.$$

The coordinates

$$\nu_1, \nu_2, \dots, \nu_n : \; \mathbb{R}^n \to \mathbb{R}$$

of the `unit` exterior normal vectors are `bounded` '`measurable`' functions on $\partial\Omega$.

Theorem 4.7. *For a* `Lipschitz` *domain Ω it holds*

$$W^{k,p}(\Omega) = \widetilde{W}^{k,p}(\Omega).$$

Convention 4.8. *In the following we assume that $\Omega \subset \mathbb{R}^n$ denotes a Lipschitz domain.*

Theorem 4.9. *The space*

$$W^{k,2}(\Omega) = \widetilde{W}^{k,2}(\Omega)$$

equipped with the dot product

$$(u,v) := \sum_{|\alpha|\le k} \int_\Omega D^\alpha u(x)\, D^\alpha v(x)\,\mathrm{d}x$$

is ***a separable Hilbert*** *space.*

Lemma 4.10. *It holds that*

$$C_0^\infty(\overline{\Omega}) := \{\varphi \in C^\infty(\overline{\Omega})\colon\ \operatorname{supp}\varphi \subset \Omega\} \subsetneqq \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{k,p}} \subsetneqq W^{k,p}(\Omega).$$

Definition 4.11. We define *the Sobolev space*

$$W_0^{k,p}(\Omega)$$

as the closure of $C_0^\infty(\Omega)$ in the space $W^{k,p}(\Omega)$, i.e.,

$$\begin{aligned} W_0^{k,p}(\Omega) &:= \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{k,p}} \\ &= \Big\{u \in W^{k,p}(\Omega)\colon \text{there exists a sequence } (u_n) \text{ in } C_0^\infty(\Omega) \text{ such that } u_n \to u \text{ in } W^{k,p}(\Omega)\Big\}. \end{aligned}$$

Theorem 4.12 (Friedrichs)**.** *In $W_0^{k,p}(\Omega)$ the functional $\|\cdot\|_{k,p,0}$,*

$$\|u\|_{k,p,0} := \left(\sum_{|\alpha|=k} \int_\Omega |D^\alpha u(x)|^p\ \mathrm{d}x\right)^{1/p},$$

defines a norm ***equivalent*** *with the norm $\|\cdot\|_{k,p}$.*

Remark to the proof, or more precisely, to the equivalence of norms $\|\cdot\|_{k,p,0}$ and $\|\cdot\|_{k,p}$ in $W_0^{k,p}(\Omega)$.

Obviously,

$$\forall u \in W_0^{k,p}(\Omega)\colon\ \|u\|_{k,p,0} \le \|u\|_{k,p}.$$

Let us show, for simplicity only in a special case of

$$n = 1,\ \ k = 1,\ \ \Omega = (a,b),\ \text{ where } a,b \in \mathbb{R};\ \ a < b,$$

the validity of the so-called <u>Friedrichs inequality</u>

$$\Big(\exists s \in \mathbb{R}^+\Big)\ \Big(\forall u \in W_0^{k,p}(\Omega)\Big) : \|u\|_{k,p} \le s\|u\|_{k,p,0},$$

i.e.,

$$\Big(\exists s \in \mathbb{R}^+\Big)\ \Big(\forall u \in W_0^{1,p}(a,b)\Big) :\ \left(\int_a^b |u|^p\,\mathrm{d}x + \int_a^b |u'|^p\,\mathrm{d}x\right)^{1/p} \le s\left(\int_a^b |u'|^p\,\mathrm{d}x\right)^{1/p}.$$

Taking into account that

$$W_0^{1,p}(a,b) = \overline{C_0^\infty(a,b)}^{\|\cdot\|_{1,p}},$$

in remains to prove

$$\left(\exists c \in \mathbb{R}^+\right) (\forall \varphi \in C_0^\infty(\langle a,b\rangle)) : \underline{\int_a^b |\varphi(x)|^p \,\mathrm{d}x \le c \int_a^b |\varphi'(x)|^p \,\mathrm{d}x}.$$

Let $\varphi \in C_0^\infty(\langle a,b\rangle)$. Then for every $x \in (a,b)$ it holds that

$$\varphi(x) = \int_a^x \varphi'(t)\,\mathrm{d}t.$$

Using the Hölder inequality (see Theorem 2.46) we easily obtain the inequalities[12]

$$|\varphi(x)|^p = \left|\int_a^x \varphi'(t)\cdot 1 \,\mathrm{d}t\right|^p \le \int_a^x |\varphi'(t)|^p \,\mathrm{d}t \cdot \left(\int_a^x 1 \,\mathrm{d}t\right)^{\frac{p}{q}} \le \int_a^b |\varphi'(t)|^p \,\mathrm{d}t \cdot (b-a)^{\frac{p}{q}},$$

and thus

$$\underline{\int_a^b |\varphi(x)|^p \,\mathrm{d}x} \le \int_a^b |\varphi'(x)|^p \,\mathrm{d}x \cdot (b-a)^{\frac{p}{q}} \cdot \int_a^b 1 \,\mathrm{d}x = \underline{\underbrace{(b-a)^p}_{=:c} \int_a^b |\varphi'(x)|^p \,\mathrm{d}x}.$$

□

Example 4.13.

$$\|u\|_{1,2,0} = \left(\int_\Omega \sum_{i=1}^n \left(\frac{\partial u}{\partial x_i}(x)\right)^2 \,\mathrm{d}x\right)^{1/2} = \sqrt{\int_\Omega |\nabla u|^2 \,\mathrm{d}x}.$$

Remark 4.14. It makes no sense to think of the values of a function $u \in L^p(\Omega)$ at the boundary $u \in L^p(\Omega)$, since the n-dimensional Lebesgue measure of $\partial\Omega$ vanishes and functions only different in sets of measure zero represent the same element of $L^p(\Omega)$.

We will show that for $u \in W^{1,p}(\Omega)$ the situation is rather different: to every such an element one can `uniquely` and `reasonably` assign a function $f_u \in L^p(\partial\Omega)$ – the so-called <u>trace</u> of u – such that the mapping $u \mapsto f_u$ will be a natural extension of the restriction to the boundary.

'Definition' 4.15. Let $1 < n \in \mathbb{N}$, $1 \le p < \infty$ and recall that $\Omega \subset \mathbb{R}^n$ is a Lipschitz domain, i.e., there exist $m \in \mathbb{N}$ Cartesian systems of coordinates

$$(\underbrace{y_{1\,r}, y_{2\,r}, \dots y_{n-1\,r}}_{=:y'_r}, y_{n\,r}) := (y'_r, y_{n\,r}),$$

[12]Obviously, q is chosen such that

$$\frac{1}{p} + \frac{1}{q} = 1, \text{ i.e., } q := \frac{p}{p-1}.$$

numbers ε_r, δ_r and Lipschitz continuous functions a_r of corresponding qualities.[13] We say that the function

$$f\colon \partial\Omega \to \mathbb{R}$$

belongs to $\underline{L^p(\partial\Omega)}$ if for every $r \in \{1, 2, \dots, m\}$ the functions

$$y'_r \mapsto f(y'_r, a_r(y'_r))$$

belong to the space $L^p\left(\{y'_r \colon \|y'_r\| < \delta_r\}\right)$.

Moreover, in such a case we define

$$\|f\|_{L^p(\partial\Omega)} := \left(\sum_{r=1}^{m} \int\limits_{\{y'_r \colon \|y'_r\| < \delta_r\}} |f(y'_r, a_r(y'_r))|^p \, \mathrm{d}y'_r \right)^{1/p} .$$

Remark 4.16. Using the approach above, there exist infinitely many possibilities to describe the boundary $\partial\Omega$. However, it can be shown that the 'corresponding' spaces $L^p(\partial\Omega)$ are identical (contain the same set of functions) and the 'corresponding' norms $\|\cdot\|_{L^p(\partial\Omega)}$ are mutually equivalent.

Theorem 4.17 (traces)**.** *There exists a unique* `continuous` *and* `linear` *mapping*

$$T\colon W^{1,p}(\Omega) \to L^p(\partial\Omega)$$

such that for every $u \in C^\infty(\overline{\Omega})$ it holds

$$Tu = u_{|\partial\Omega} .$$

The element $Tu \in L^p(\partial\Omega)$ is called <u>*a trace*</u> *of $u \in W^{1,p}(\Omega)$.*

Observation 4.18. Let $u \in W^{k,p}(\Omega)$. Then for every multiindex $\alpha, |\alpha| \le k-1$ it holds

$$D^\alpha u \in W^{1,p}(\Omega),$$

and thus there exists $T(D^\alpha u)$. Moreover, it can be shown that

$$W_0^{k,p}(\Omega) = \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{k,p}} = \left\{ u \in W^{k,p}(\Omega) \colon \forall\alpha, |\alpha| \le k-1, \text{ it holds that } T(D^\alpha u) = 0 \right\};$$

in particular

$$W_0^{1,p}(\Omega) = \left\{ u \in W^{1,p}(\Omega) \colon T(u) = 0 \right\}.$$

The boundedness and Lipschitz regularity of $\emptyset \neq \Omega \subset \mathbb{R}^n$ allows us to define a 'surface' integral on $\partial\Omega$.

Lemma 4.19 (partition of unity)**.** *Let us denote by*

- *$F \subset \mathbb{R}^n$ a* `compact` *set,*

[13] See 'Definition' 4.5 of Lipschitz domains.

- *$G_1, G_2, \dots, G_m \subset \mathbb{R}^n$ open sets,*
- *$F \subset \bigcup_{r=1}^{m} G_r$. (The set $\{G_r\}_{r=1}^m$ is an open cover of the compact set F.)*

Then there exist functions $\varphi_1, \varphi_2, \dots, \varphi_m \colon \mathbb{R}^n \to \mathbb{R}$ such that

- *$(\forall x \in \mathbb{R}^n)\,(\forall r \in \{1, 2, \dots, m\}) : 0 \le \varphi_r(x) \le 1$,*
- *$\forall r \in \{1, 2, \dots, m\}\colon \ (\varphi_r \in C_0^\infty(\mathbb{R}^n),\ \operatorname{supp} \varphi_r \subset G_r)$,*
- *$\forall x \in F\colon \ \sum_{r=1}^{m} \varphi_r(x) = 1$.*

'Definition' 4.20. Let us recall the situation described in 'Definiton' 4.15 of the space $L^p(\partial\Omega)$ and let us define for every $r \in \{1, 2, \dots, m\}$

$$G_r := \left\{ (y_r', y_{n\,r})\colon \|y_r'\| < \delta_r,\ a_r(y_r') - \varepsilon_r < y_{n\,r} < a_r(y_r') + \varepsilon_r \right\}.$$

Let $\varphi_1, \varphi_2, \dots, \varphi_m$ denote a partition of unity related to the open cover $\{G_r\}_{r=1}^m$ of the boundary $\partial\Omega$. For a function $f \in L^1(\partial\Omega)$ we define

$$\int_{\partial\Omega} f(x)\,\mathrm{d}s := \sum_{r=1}^{m} \int_{\{y_r'\colon \|y_r'\| < \delta_r\}} f(y_r', a_r(y_r'))\, \varphi_r(y_r', a_r(y_r')) \sqrt{1 + \sum_{i=1}^{n-1} \left(\frac{\partial a_r}{\partial y_{i\,r}}(y_r') \right)^2}\, \mathrm{d}y_r'.$$

(It can be shown that the definition is independent both of the description of $\partial\Omega$ and the partition of unity on $\partial\Omega$.)

Theorem 4.21 (Green)**.**

i) For $1 < n \in \mathbb{N}$ let $\emptyset \ne \Omega \subset \mathbb{R}^n$ denote a Lipschitz domain. Then for every $u, v \in W^{1,2}(\Omega)$ and $i \in \{1, 2, \dots, n\}$ it holds that

$$\int_\Omega u \frac{\partial v}{\partial x_i}\,\mathrm{d}x = -\int_\Omega \frac{\partial u}{\partial x_i} v\,\mathrm{d}x + \int_{\partial\Omega} T(u)\, T(v)\, \nu_i\,\mathrm{d}s$$

$(\nu_i = \nu_i(x)$ *... i-th component of the exterior unit normal vector of $\partial\Omega$ in x).*

ii) Let $a, b \in \mathbb{R},\ a < b$. Then $W^{1,2}(a, b) \subset C(\langle a, b\rangle)$ and furthermore

$$\forall u, v \in W^{1,2}(a, b)\colon \int_a^b u v'\,\mathrm{d}x = -\int_a^b u' v\,\mathrm{d}x + \underbrace{u(b)v(b) - u(a)v(a)}_{=:\,[uv]_a^b}.$$

(Note that the derivatives in the Green theorem are understood in the distributional sense.)

Definition 4.22. A normed linear space X is continuously embedded in a normed linear space Y if the conditions

- $X \subset Y$,
- $(\exists c \in \mathbb{R}^+)\,(\forall x \in X) : \|x\|_Y \le c\|x\|_X$.

are met. We denote a continuous embedding by

$$X \hookrightarrow Y.$$

Theorem 4.23 (on continuous embedding)**.** *Let*

$$\emptyset \neq \Omega \subset \mathbb{R}^n$$

denote a bounded Lipschitz domain. Then it holds:

i) if either

$$kp < n,\ 1 \leq q \leq q^* := \frac{np}{n - pk},$$

or

$$kp = n,\ 1 \leq q < \infty,$$

we have

$$W^{k,p}(\Omega) \hookrightarrow L^q(\Omega),$$

$$i.e.,\ \left(\exists c \in \mathbb{R}^+\right) \left(\forall u \in W^{k,p}(\Omega)\right) : \|u\|_{L^q(\Omega)} \leq c\|u\|_{k,p};$$

ii) for

$$kp > n,$$

we have

$$W^{k,p}(\Omega) \hookrightarrow C(\overline{\Omega}),$$

$$i.e.,\ \left(\exists c \in \mathbb{R}^+\right) \left(\forall u \in W^{k,p}(\Omega)\right) : \|u\|_{C(\overline{\Omega})} := \max_{\overline{\Omega}} |u| \leq c\|u\|_{k,p}.$$

Definition 4.24. A normed linear space X is compactly embedded in a normed linear space Y if the conditions

- $X \subset Y$,
- from each sequence $(X, \|\cdot\|_X)$ one can extract a subsequence convergent in $(Y, \|\cdot\|_Y)$,

are met. We denote a compact embedding by

$$X \hookrightarrow\hookrightarrow Y.$$

Observation 4.25.

i) $X \hookrightarrow\hookrightarrow Y \ \Rightarrow\ X \hookrightarrow Y$,

ii) $X \hookrightarrow\hookrightarrow Y \ \Rightarrow\ \left[x_n \rightharpoonup x \text{ in } X \ \Rightarrow\ x_n \to x \text{ in } Y\right]$.

Theorem 4.26 (on compact embedding)**.** *Let*

$$\emptyset \neq \Omega \subset \mathbb{R}^n$$

denote a bounded Lipschitz domain. Then it holds *i) if either*

$$n > 1,\ kp < n,\ \frac{1}{q} > \frac{1}{p} - \frac{k}{n},$$

or

$$n > 1,\ kp = n,\ 1 < q < \infty,$$

we have

$$W^{k,p}(\Omega) \hookrightarrow\hookrightarrow L^q(\Omega);$$

ii) for

$$n \geq 1,\ kp > n,$$

we have

$$W^{k,p}(\Omega) \hookrightarrow\hookrightarrow C(\overline{\Omega}).$$

5 Weak solution to linear elliptic problems

Let us consider the Dirichlet problem for the Poisson equation

$$\begin{cases} -\Delta u = f & \text{v } \Omega, \\ u = 0 & \text{na } \partial\Omega, \end{cases}$$

where $\emptyset \neq \Omega \subset \mathbb{R}^n$ denotes a bounded Lipschitz domain, $f \in C(\overline{\Omega})$. In the following we assume that $u \in C^2(\Omega) \cap C(\overline{\Omega})$ is a (classical) solution to the problem.

For every $v \in C_0^\infty(\overline{\Omega}) := \{\varphi \in C^\infty(\overline{\Omega})\colon \operatorname{supp}\varphi \subset \Omega\}$ we have

$$\int_\Omega -\Delta u\, v \,\mathrm{d}x = \int_\Omega f\, v \,\mathrm{d}x.$$

Due to the Green theorem 4.21 we deduce [14]

$$\begin{aligned} \int_\Omega -\Delta u\, v \,\mathrm{d}x &= \sum_{i=1}^n \left(\int_\Omega -\frac{\partial^2 u}{\partial x_i^2} v \,\mathrm{d}x \right) \sum_{i=1}^n \left(\int_\Omega \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} \,\mathrm{d}x \right) + \sum_{i=1}^n \left(-\int_{\partial\Omega} \frac{\partial u}{\partial x_i} v\, \nu_i \,\mathrm{d}s \right) \\ &= \int_\Omega \nabla u \nabla v \,\mathrm{d}x - \underbrace{\int_{\partial\Omega} \frac{\partial u}{\partial \nu} v \,\mathrm{d}s}_{=0,\text{ since } v \in C_0^\infty(\overline{\Omega})} = \int_\Omega \nabla u \nabla v \,\mathrm{d}x \end{aligned}$$

to finally obtain

$$\int_\Omega \nabla u \nabla v \,\mathrm{d}x = \int_\Omega f\, v \,\mathrm{d}x.$$

Note that the formula is not only meaningful for the above specified u, v and f, but also for

$$u,\ v \in W^{1,2}(\Omega) \text{ and } f \in L^2(\Omega).$$

This leads us to the idea of a generalized 'solution' to the given problem. Moreover, considering only $u \in W_0^{1,2}(\Omega) = \{u \in W^{1,2}(\Omega)\colon Tu = 0\}$,[15] we can naturally generalize the boundary condition '$u = 0$ on $\partial\Omega$'.

Definition 5.1. A function $u \in W_0^{1,2}(\Omega)$ is a <u>weak solution to the Dirichlet problem</u>

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.1}$$

where $f \in L^2(\Omega)$ if

$$\int_\Omega \nabla u \nabla v \,\mathrm{d}x = \int_\Omega f\, v \,\mathrm{d}x \quad \text{for every } v \in W_0^{1,2}(\Omega).$$

Theorem 5.2. *The problem* (5.1) *is uniquely solvable.*

[14] Again, by $\nu = (\nu_1, \dots, \nu_n)$ we denote the exterior unit normal vector on $\partial\Omega$.
[15] T denotes the trace operator.

Proof. We already know that $W_0^{1,2}(\Omega)$ equipped with the inner product

$$(u,v) := \sum_{|\alpha|\le 1} \int_\Omega D^\alpha u\, D^\alpha v \,\mathrm{d}x = \int_\Omega u\, v \,\mathrm{d}x + \sum_{i=1}^n \int_\Omega \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}\,\mathrm{d}x =$$
$$= \int_\Omega u\, v\,\mathrm{d}x + \int_\Omega \nabla u \nabla v \,\mathrm{d}x$$

forms a Hilbert space.

Friedrichs theorem 4.12 ensures that the `inner product`

$$((u,v)) := \int_\Omega \nabla u \nabla v \,\mathrm{d}x$$

induces a norm $\|\cdot\|_{1,2,0}$ on $W_0^{1,2}(\Omega)$, which is equivalent with the norm $\|\cdot\|_{1,2}$.

Also note that the functional $F\colon W_0^{1,2}(\Omega) \to \mathbb{R}$ defined by

$$F(v) := \int_\Omega f\, v\,\mathrm{d}x$$

is:

- linear – obviously,
- continuous – here it suffices to show that it is bounded:[16]

$$|F(v)| = \left|\int_\Omega f\, v\,\mathrm{d}x\right| \le \sqrt{\int_\Omega f^2\,\mathrm{d}x}\sqrt{\int_\Omega v^2\,\mathrm{d}x} \le \|f\|_{L^2(\Omega)}\,\|v\|_{1,2} \le$$
$$\le \|f\|_{L^2(\Omega)}\, c\,\|v\|_{1,2,0}.$$

It follows that

$$F \in \left(W_0^{1,2}(\Omega)\right)^* = \left(W_0^{1,2}(\Omega), \|\cdot\|_{1,2,0}\right)^*.$$

From the Riesz representation theorem we finally conclude that

$$\left(\exists!\, u \in W_0^{1,2}(\Omega)\right)\left(\forall v \in W_0^{1,2}(\Omega)\right) : F(v) = ((u,v)).$$

□

Our goal now is to generalize the notion of 'a solution to a boundary value problem' for problems more complex than the Dirichlet problem for the Poisson equation. We will study problems of the type

$$\begin{cases} \mathcal{L}u = f \text{ v } \Omega, \\ \quad + \\ \text{boundary conditions on } \partial\Omega, \end{cases} \tag{5.2}$$

where

[16] The first inequality is the Hölder inequality (see 2.46), the last one follows from the Friedrichs inequality 4.12.

- $\emptyset \neq \Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain,
- $f \in L^2(\Omega)$,
- $\mathcal{L}$ is a differential operator of the $2k$-th order in the (so-called divergent) form:

$$(\mathcal{L}u)(x) := \sum_{|\alpha|,|\beta|\leq k} (-1)^{|\alpha|} D^\alpha \left(a_{\alpha,\beta}(x) D^\beta u(x)\right),$$

- $a_{\alpha,\beta} \in L^\infty(\Omega)$ for all multiindices α, β satisfying $|\alpha|, |\beta| \leq k$.

Remark 5.3. Note that in the case of the classical solution to (5.2) we require k-th order differentiability of the functions $a_{\alpha,\beta}$ and $2k$-th order differentiability of u. Using the 'trick' $\int_\Omega \mathcal{L}u\, v \,\mathrm{d}x = \int_\Omega f v \,\mathrm{d}x$ and applying the Green theorem we only require k-th order differentiability of u on the left-hand side of the equation.

To every operator $\mathcal{L}$ we can assign the bilinear form[17]

$$a(u,v) := \sum_{|\alpha|,|\beta|\leq k} \int_\Omega a_{\alpha,\beta}(x) D^\beta u(x) D^\alpha v(x) \,\mathrm{d}x. \tag{5.3}$$

It can be easily verified that the bilinear form (5.3) is well defined and continuous in

$$W^{k,2}(\Omega) \times W^{k,2}(\Omega);$$

as it holds (using the Hölder inequality 2.46)

$$\left| \int_\Omega a_{\alpha,\beta}(x) D^\beta u(x) D^\alpha v(x) \,\mathrm{d}x \right| \leq \|a_{\alpha,\beta}\|_{L^\infty(\Omega)} \int_\Omega \left|D^\beta u\right| |D^\alpha v| \,\mathrm{d}x \leq$$

$$\leq \|a_{\alpha,\beta}\|_{L^\infty(\Omega)} \sqrt{\int_\Omega |D^\beta u|^2 \,\mathrm{d}x} \sqrt{\int_\Omega |D^\alpha v|^2 \,\mathrm{d}x} \leq \|a_{\alpha,\beta}\|_{L^\infty(\Omega)} \|u\|_{k,2} \|v\|_{k,2},$$

and since there only exists a finite number of combinations of the coefficients α, β such that $|\alpha|, |\beta| \leq k$, there exists $c > 0$ satisfying

$$|a(u,v)| = \left| \sum_{|\alpha|,|\beta|\leq k} \int_\Omega a_{\alpha,\beta}(x) D^\beta u(x) D^\alpha v(x) \,\mathrm{d}x \right| \leq c \|u\|_{k,2} \|v\|_{k,2}.$$

Example 5.4.

$$\mathcal{L}u := -\operatorname{div}(\nabla u) = -\Delta u = \sum_{i=1}^n -\frac{\partial^2 u}{\partial x_i^2} = \sum_{i,j=1}^n (-1)^1 \frac{\partial}{\partial x_i} \left(a_{i,j}(x) \frac{\partial u}{\partial x_j} \right),$$

[17] Note that for $a_{\alpha,\beta} \in C^k(\overline{\Omega})$ and $u, v \in C_0^{2k}(\overline{\Omega})$ it holds

$$a(u,v) = \int_\Omega (\mathcal{L}u)(x) v(x) \,\mathrm{d}x.$$

where

$$a_{i,j}(x) := \begin{cases} 1, & \textit{for } i = j, \\ 0, & \textit{for } i \neq j; \end{cases}$$

$$a(u,v) = \int_\Omega \nabla u \, \nabla v \, \mathrm{d}x.$$

Example 5.5.

$$\mathcal{L}u := -\operatorname{div}\left(p(x)\nabla u\right) + q(x)u = \sum_{i=1}^{n} -\frac{\partial}{\partial x_i}\left(p(x)\frac{\partial u}{\partial x_i}\right) + q(x)u,$$

$$a(u,v) = \int_\Omega \left(p(x)\nabla u \, \nabla v + q(x) u \, v\right) \mathrm{d}x.$$

Let us further denote by

- $\mathcal{V}$ a linear subset $C^\infty(\overline{\Omega})$ such that

$$C_0^\infty(\overline{\Omega}) \subset \mathcal{V} \subset C^\infty(\overline{\Omega}),$$

- V the closure of $\mathcal{V}$ in the space $W^{k,2}(\Omega)$ (trivially, $W_0^{k,2}(\Omega) \subset V \subset W^{k,2}(\Omega)$),
- $b\colon W^{k,2}(\Omega) \times W^{k,2}(\Omega) \to \mathbb{R}$ a continuous bilinear form,
- $g \in V^*$ such $\forall v \in W_0^{k,2}(\Omega)\colon g(v) = 0$,
- $u_0 \in W^{k,2}(\Omega)$.

Below we will use the quantities V, b, g and u_0 to model boundary conditions on $\partial\Omega$.

Definition 5.6. Let

$$((u,v)) := a(u,v) + b(u,v).$$

A function $u \in W^{k,2}(\Omega)$ is a weak solution to the boundary value problem, if

i) $u - u_0 \in V$,

ii) $\forall v \in V\colon ((u,v)) = \int_\Omega f(x)v(x)\,\mathrm{d}x + g(v)$.

Again, in the following we assume that Ω denotes a bounded Lipschitz domain.

Example 5.7 (weak solution to the Dirichlet problem)**.**
Consider the Dirichlet problem

$$\begin{cases} -\operatorname{div}(p(x)\nabla u) + q(x)u &= f(x) \quad \textit{in } \Omega, \\ u &= h(x) \quad \textit{on } \partial\Omega, \end{cases} \tag{5.4}$$

and assume that it holds:

- $f \in L^2(\Omega)$,

- *p, $q \in L^\infty(\Omega)$,*
- *there exists $u_0 \in W^{1,2}(\Omega)$ such that $Tu_0 = h \in L^2(\partial\Omega)$.*

A function $u \in W^{1,2}(\Omega)$ is a weak solution to the Dirichlet problem (5.4)*, if*

- *$u - u_0 \in W_0^{1,2}(\Omega)$,*
- *$\forall v \in W_0^{1,2}(\Omega)\colon \int_\Omega (p(x)\nabla u\,\nabla v + q(x)u\,v)\,\mathrm{d}x = \int_\Omega f\,v\,\mathrm{d}x.$*

Here

$$V := W_0^{1,2}(\Omega),$$

$$a(u,v) := \int_\Omega (p(x)\nabla u\,\nabla v + q(x)u\,v)\ \mathrm{d}x,$$

$$b(u,v) := 0,\ \ g(v) := 0.$$

Observation 5.8. In case that the solution u and the 'data' of the problem (5.4) are 'smooth enough', we have

- $0 = T(u-u_0) = Tu - Tu_0 = u|_{\partial\Omega} - h\ \Rightarrow\ u(x) = h(x)$ for 'almost all' $x \in \partial\Omega$ (see the boundary conditions from (5.4)),
- $\forall v \in W_0^{1,2}(\Omega)\colon Tv = 0$, and thus also (again using the Green theorem 4.21)

$$\int_\Omega f\,v\,\mathrm{d}x = \int_\Omega (p(x)\nabla u\,\nabla v + q(x)u\,v)\ \mathrm{d}x = \int_\Omega (-\operatorname{div}(p(x)\nabla u)v + q(x)u\,v)\ \mathrm{d}x.$$

  Thus,

$$\forall v \in W_0^{1,2}(\Omega)\colon \int_\Omega \big(-\operatorname{div}(p(x)\nabla u) + q(x)u - f(x)\big)v\,\mathrm{d}x = 0$$

$$\Downarrow$$

$$-\operatorname{div}(p(x)\nabla u) + q(x)u = f(x) \text{ almost everywhere in } \Omega$$

  (see the differential equation from (5.4)).

`Conclusion:` In case that the weak solution u and the 'data' of the problem (5.4) are 'smooth enough' it holds

$$\text{weak solution} \equiv \text{classical solution.}$$

Example 5.9 (weak solution to the Neumann probem)**.**
Consider the Neumann problem

$$\begin{cases} -\operatorname{div}(p(x)\nabla u) + q(x)u &=& f(x) \quad \textit{in } \Omega, \\ p(x)\dfrac{\partial u}{\partial \nu} &=& h(x) \quad \textit{on } \partial\Omega, \end{cases} \tag{5.5}$$

where

- *$f \in L^2(\Omega)$,*

- $p,\, q \in L^\infty(\Omega)$,
- $h \in L^2(\partial\Omega)$.

The function $u \in W^{1,2}(\Omega)$ is <u>weak solution to the Neumann problem</u> (5.5), if

$$\forall v \in W^{1,2}(\Omega)\colon \int_\Omega \big(p(x)\nabla u\,\nabla v + q(x) u\, v\big)\,\mathrm{d}x = \int_\Omega f\, v\,\mathrm{d}x + \int_{\partial\Omega} h\, Tv\,\mathrm{d}s.$$

Here we use the notation

$$V := W^{1,2}(\Omega),$$
$$a(u,v) := \int_\Omega \big(p(x)\nabla u\,\nabla v + q(x) u\, v\big)\,\mathrm{d}x,$$
$$b(u,v) := 0,\;\; g(v) := \int_{\partial\Omega} h\, Tv\,\mathrm{d}s.$$

Note that it is not necessary to define $u_0 \in W^{1,2}(\Omega)$, the condition i) from Definition 5.6 is satisfied for all $u, u_0 \in W^{1,2}(\Omega)$.

Observation 5.10. In case that the weak solution u and the 'data' of the problem (5.5) are 'smooth enough' we have

$$\forall v \in W^{1,2}(\Omega)\colon \int_\Omega p(x) \sum_{i=1}^n \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}\,\mathrm{d}x + \int_\Omega q(x) u\, v\,\mathrm{d}x = \int_\Omega f\, v\,\mathrm{d}x + \int_{\partial\Omega} h\, Tv\,\mathrm{d}s,$$

since (see the Green theorem 4.21)

$$\int_\Omega p(x) \sum_{i=1}^n \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}\,\mathrm{d}x = \int_\Omega -\sum_{i=1}^n \frac{\partial}{\partial x_i}\left(p(x)\frac{\partial u}{\partial x_i}\right) v\,\mathrm{d}x + \int_{\partial\Omega} p(x) \underbrace{\sum_{i=1}^n \frac{\partial u}{\partial x_i}\nu_i}_{=\frac{\partial u}{\partial \nu}}\, Tv\,\mathrm{d}s.$$

Thus, for all $v \in W^{1,2}(\Omega)$ it holds that

$$\int_\Omega \big(-\operatorname{div}(p(x)\nabla u) + q(x)u - f(x)\big)v\,\mathrm{d}x + \int_{\partial\Omega}\left(p(x)\frac{\partial u}{\partial \nu} - h(x)\right) Tv\,\mathrm{d}s = 0.$$

It follows that

- $$\forall v \in C_0^\infty(\overline{\Omega})\colon \int_\Omega \big(-\operatorname{div}(p(x)\nabla u) + q(x)u - f(x)\big)v \,\mathrm{d}x = 0$$
$$\Downarrow$$
$$-\operatorname{div}(p(x)\nabla u) + q(x)u = f(x) \text{ almost everywhere in } \Omega$$
(see the differential equation from (5.5)),
- $$\forall v \in W^{1,2}(\Omega)\colon \int_{\partial\Omega} \big(p(x)\frac{\partial u}{\partial \nu} - h(x)\big)Tv \,\mathrm{d}s = 0$$
$$\Downarrow$$
$$p(x)\frac{\partial u}{\partial \nu} = h(x) \text{ `almost everywhere' on } \partial\Omega$$
(see the boundary condition from (5.5)).

Example 5.11 (weak solution to the Newton problem)**.**
Consider the Newton problem

$$\left\{ \begin{array}{rcll} -\operatorname{div}(p(x)\nabla u) + q(x)u &=& f(x) & \textit{in } \Omega, \\ \sigma(x)u + p(x)\dfrac{\partial u}{\partial \nu} &=& h(x) & \textit{on } \partial\Omega, \end{array} \right. \tag{5.6}$$

where

- $f \in L^2(\Omega)$,
- $p,\ q \in L^\infty(\Omega)$,
- $0 \neq \sigma \in L^\infty(\partial\Omega)$[18],
- $h \in L^2(\partial\Omega)$.

The function $u \in W^{1,2}(\Omega)$ is *a weak solution to the Newton problem* (5.6)*, if it holds*

$$\forall v \in W^{1,2}(\Omega)\colon \int_\Omega (p(x)\nabla u\,\nabla v + q(x)u\,v)\,\mathrm{d}x + \int_{\partial\Omega} \sigma(x)Tu\,Tv\,\mathrm{d}s = \int_\Omega f\,v\,\mathrm{d}x + \int_{\partial\Omega} h\,Tv\,\mathrm{d}s.$$

Here we use the notation

$$V := W^{1,2}(\Omega),$$

$$a(u,v) := \int_\Omega \big(p(x)\nabla u\,\nabla v + q(x)u\,v\big)\,\mathrm{d}x,$$

$$b(u,v) := \int_{\partial\Omega} \sigma(x)Tu\,Tv\,\mathrm{d}s,\ \ g(v) := \int_{\partial\Omega} h\,Tv\,\mathrm{d}s.$$

Again, it is not necessary to define $u_0 \in W^{1,2}(\Omega)$.

[18] $s \in L^\infty(\Omega) \Leftrightarrow s$ is bounded and ‘measurable’ on $\partial\Omega$

Observation 5.12. Similarly as in both previous cases one can show (using the Green theorem 4.21) that for a 'smooth enough' weak solution u and the 'data' of the Newton problem (5.6) it holds

$$\forall v \in W^{1,2}(\Omega):$$

$$\int_\Omega (-\operatorname{div}(p(x)\nabla u) + q(x)u - f(x))v \,\mathrm{d}x + \int_{\partial\Omega} \left(p(x)\frac{\partial u}{\partial \nu} + \sigma(x)Tu - h(x) \right) Tv \,\mathrm{d}s = 0,$$

and thus

$$-\operatorname{div}(p(x)\nabla u) + q(x)u = f(x) \text{ almost everywhere in } \Omega,$$

$$p(x)\frac{\partial u}{\partial \nu} + \sigma(x) \underbrace{Tu}_{=u} = h(x) \text{ `almost everywhere' on } \partial\Omega.$$

Example 5.13 (weak solution to the mixed problem)**.**
Consider the mixed problem

$$\begin{cases} -\operatorname{div}(p(x)\nabla u) + q(x)u &= f(x) & \text{in } \Omega, \\ u &= h_1(x) & \text{on } \Gamma_1, \\ p(x)\dfrac{\partial u}{\partial \nu} &= h_2(x) & \text{on } \Gamma_2, \\ \sigma(x)u + p(x)\dfrac{\partial u}{\partial \nu} &= h_3(x) & \text{on } \Gamma_3, \end{cases} \tag{5.7}$$

where the boundary $\partial\Omega$ is decomposed into pairwise disjoint 'measurable' components $\Gamma_1, \Gamma_2, \Gamma_3$ and it holds

- $f \in L^2(\Omega)$,
- $p,\, q \in L^\infty(\Omega)$,
- $0 \neq \sigma \in L^\infty(\Gamma_3)$,
- $u_0 \in W^{1,2}(\Omega)$ *satisfies* $Tu_0 = h_1$ *'almost everywhere' on* Γ_1,
- $h_2 \in L^2(\Gamma_2)$,
- $h_3 \in L^2(\Gamma_3)$.

Let

$$V := \left\{ v \in W^{1,2}(\Omega) \colon Tv = 0 \text{ na } \Gamma_1 \right\}.$$

A function $u \in W^{1,2}(\Omega)$ is <u>a weak solution to the mixed problem</u> (5.7)*, if*

- $u - u_0 \in V$,

- $\forall v \in V:$

$$\overbrace{\int_\Omega (p(x)\nabla u\,\nabla v + q(x)u\,v)\,\mathrm{d}x}^{=:\,a(u,v)} + \overbrace{\int_{\Gamma_3} \sigma(x)Tu\,Tv\,\mathrm{d}s}^{=:\,b(u,v)} =$$
$$= \int_\Omega f\,v\,\mathrm{d}x + \underbrace{\int_{\Gamma_2} h_2(x)Tv\,\mathrm{d}s + \int_{\Gamma_3} h_3(x)Tv\,\mathrm{d}s}_{=:\,g(v)}\,.$$

Observation 5.14. Note that the mixed problem is a generalization of the pure Dirichlet, Neumann, and Newton problems, since

- $\Gamma_1 = \partial\Omega$, $\Gamma_2 = \Gamma_3 = \emptyset$, $V = W_0^{1,2}(\Omega)\ldots$ problem (5.4),
- $\Gamma_2 = \partial\Omega$, $\Gamma_1 = \Gamma_3 = \emptyset$, $V = W^{1,2}(\Omega)\ldots$ problem (5.5),
- $\Gamma_3 = \partial\Omega$, $\Gamma_1 = \Gamma_2 = \emptyset$, $V = W^{1,2}(\Omega)\ldots$ problem (5.6).

Exercise 5.15. *Prove that for 'smooth enough' 'data' of* (5.7) *it holds*

$$\textit{weak solution} \equiv \textit{classical solution.}$$

Example 5.16 (weak solution to the transmission problem)**.**
Let us denote

- $\Omega = \{(x_1,\ldots,x_n) \in \mathbb{R}^n : \|(x_1,\ldots,x_n)\| < 1\}$,
- $\Omega^+ = \Omega \cap \{(x_1,\ldots,x_n) \in \mathbb{R}^n : x_n > 0\}$,
- $\Omega^- = \Omega \setminus \overline{\Omega^+}$,
- $\Gamma = \Omega \cap \{(x_1,\ldots,x_n) \in \mathbb{R}^n : x_n = 0\}$.

and consider the problem

$$\left\{\begin{array}{rcl} -\alpha\Delta u &=& f(x) \ \textit{v}\ \Omega^+, \\ -\beta\Delta u &=& f(x) \ \textit{v}\ \Omega^-, \\ u &=& h(x) \ \textit{na}\ \partial\Omega, \\ u && \textit{is continuous on } \Gamma, \\ \alpha\,\dfrac{\partial u}{\partial x_n}(x_1,\ldots,x_{n-1},0+) &=& \beta\,\dfrac{\partial u}{\partial x_n}(x_1,\ldots,x_{n-1},0-) \\ && \quad \textit{for all } (x_1,\ldots,x_{n-1},0) \in \Gamma, \end{array}\right. \tag{5.8}$$

where

- $\alpha,\ \beta \in \mathbb{R}^+$,
- $f \in L^2(\Omega)$,
- *there exists* $u_0 \in W^{1,2}(\Omega)$ *such that* $Tu_0 = h \in L^2(\partial\Omega)$.

A function $u \in W^{1,2}(\Omega)$ *is* <u>*a weak solution to the transmission problem*</u> *if*

- $u - u_0 \in W_0^{1,2}(\Omega)$,
- $\forall v \in W_0^{1,2}(\Omega)\colon \alpha \int_{\Omega+} \nabla u \,\nabla v \,\mathrm{d}x + \beta \int_{\Omega-} \nabla u \,\nabla v \,\mathrm{d}x = \int_{\Omega} f(x) v \,\mathrm{d}x.$

Here we made use of

$$V := W_0^{1,2}(\Omega),$$

$$a(u,v) := \alpha \int_{\Omega+} \nabla u \,\nabla v \,\mathrm{d}x + \beta \int_{\Omega-} \nabla u \,\nabla v \,\mathrm{d}x,$$

$$b(u,v) := 0,\ \ g(v) := 0.$$

Observation 5.17. For a 'smooth enough' weak solution u and the 'data' in (5.8) it holds

- $0 = T(u - u_0) = Tu - Tu_0 = u|_{\partial\Omega} - h \ \Rightarrow\ u(x) = h(x)$ for 'almost all' $x \in \partial\Omega$,
- $\forall v \in W_0^{1,2}(\Omega)$:

$$\alpha \int_{\Omega+} \nabla u \,\nabla v \,\mathrm{d}x + \beta \int_{\Omega-} \nabla u \,\nabla v \,\mathrm{d}x = \int_{\Omega+} f(x) v \,\mathrm{d}x + \int_{\Omega-} f(x) v \,\mathrm{d}x$$

$$\Downarrow \text{ see the Green theorem 4.21}$$

$$\begin{aligned} -\alpha \int_{\Omega^+} \Delta u \, v \,\mathrm{d}x + \alpha \int_{\partial\Omega^+} \frac{\partial u}{\partial \nu^+} Tv \,\mathrm{d}s - \beta \int_{\Omega^-} \Delta u \, v \,\mathrm{d}x + \beta \int_{\partial\Omega^-} \frac{\partial u}{\partial \nu^-} Tv \,\mathrm{d}s & \\ = \int_{\Omega^+} f(x) v \,\mathrm{d}x + \int_{\Omega^-} f(x) v \,\mathrm{d}x. & \end{aligned}$$

It follows that

- $\forall v \in C^\infty(\overline{\Omega})$, $\operatorname{supp} v \subset \Omega^+$:

$$\int_{\Omega^+} (-\alpha \Delta u - f(x)) v \,\mathrm{d}x = 0 \ \Rightarrow\ -\alpha \Delta u = f(x) \text{ almost everywhere in } \Omega^+,$$

- $\forall v \in C^\infty(\overline{\Omega})$, $\operatorname{supp} v \subset \Omega^-$:

$$\int_{\Omega^-} (-\beta \Delta u - f(x)) v \,\mathrm{d}x = 0 \ \Rightarrow\ -\beta \Delta u = f(x) \text{ almost everywhere in } \Omega^-,$$

and thus for all $v \in W_0^{1,2}(\Omega)$ it holds

$$\alpha \int_{\partial\Omega^+} \frac{\partial u}{\partial \nu^+} Tv \,\mathrm{d}s + \beta \int_{\partial\Omega^-} \frac{\partial u}{\partial \nu^-} Tv \,\mathrm{d}s = 0$$

$$\Downarrow$$

$$\int_{\Gamma} \left(\alpha \frac{\partial u}{\partial \nu^+} + \beta \frac{\partial u}{\partial \nu^-} \right) Tv \,\mathrm{d}s = 0,$$

or
$$\alpha \frac{\partial u}{\partial x_n}(x_1, \ldots, x_{n-1}, 0+) = \beta \frac{\partial u}{\partial x_n}(x_1, \ldots, x_{n-1}, 0-)$$
for 'almost all' $(x_1, \ldots, x_{n-1}, 0) \in \Gamma$. (Moreover, with the assumption of smooth u the transmission condition holds for every $(x_1, \ldots, x_{n-1}, 0) \in \Gamma$.)

Theorem 5.18 (Lax-Milgram lemma)**.**
Assume that

- *H denotes a Hilbert space endowed with an inner product $(\cdot, \cdot)$ inducing the norm*
$$\|\cdot\| = \sqrt{(\cdot, \cdot)},$$
- *$B\colon H \times H \to \mathbb{R}$ is a bilinear form that is*
  - *continuous, i.e., there exists $k > 0$ such that*
$$\forall u,\, v \in H\colon |B(u, v)| \le k\, \|u\|\, \|v\|,$$
  - *H-elliptic, i.e., there exists $c > 0$ such that*
$$\forall u \in H\colon B(u, u) \ge c\, \|u\|^2,$$
- *$F \in H^*$.*

Then there exists **`a unique element`** *$w \in H$ such that*
$$\forall v \in H\colon F(v) = B(w, v).$$

Furthermore, it holds
$$\|w\| \le \frac{1}{c}\, \|F\|_{H^*}.$$

Proof. Let us choose an arbitrary but fixed element $u \in H$ and consider the functional
$$G(v) := B(u, v).$$
Obviously, $G \in H^*$, and thus there exists (see the Riesz representation theorem) a unique element $Au \in H$ such that
$$\underline{\forall v \in H}\colon G(v) = \underline{B(u, v) = (Au, v)}.$$

For the mapping
$$A\colon H \to H$$
it holds

- $\mathcal{D}(A) = H,\ A(H) \subset H,$
- A is *linear*,

- A is *continuous* in H, since

$$\forall u \in H\colon \|Au\|^2 = (Au, Au) = B(u, Au) \le k\,\|u\|\,\|Au\|$$

(here we used the continuity of the bilinear form B), and thus

$$\forall u \in H\colon \|Au\| \le k\,\|u\|.$$

- A is an *injective* mapping (and thus its inversion $A^{-1}\colon A(H) \to H$) exists, because due to the linearity of A and H-ellipticity of the bilinear form B it holds

$$\underline{Au_1 = Au_2}$$
$$\Downarrow$$
$$A(u_1 - u_2) = 0$$
$$\Downarrow$$
$$c\|u_1 - u_2\|^2 \le B(u_1 - u_2, u_1 - u_2) = (A(u_1 - u_2), u_1 - u_2) = (0, u_1 - u_2) = 0$$
$$\Downarrow$$
$$\|u_1 - u_2\| = 0$$
$$\Downarrow$$
$$\underline{u_1 = u_2}.$$

- $\forall u \in H\colon c\,\|u\|^2 \le B(u,u) = (Au, u) \le \|Au\|\,\|u\|$, and thus[19]

$$\forall u \in H\colon c\,\|u\| \le \|Au\|. \tag{5.9}$$

- $A(H)$ is a *closed subset* of H. Let us prove this assertion. Consider a sequence $(u_n) \subset H$ and $y \in H$ such that $\underline{Au_n \to y}$. Then (Au_n) is a Cauchy sequence in H, and since (see (5.9))

$$c\,\|u_n - u_m\| \le \|A(u_n - u_m)\| = \|Au_n - Au_m\|,$$

also (u_n) is a Cauchy sequence in (`the complete space`) H. There thus exists $x \in H$ such that $u_n \to x$. The continuity of A then leads to $Au_n \to Ax$. Since we simultaneously assume that $Au_n \to y$, it obviously holds that $Ax = \underline{y \in A(H)}$.

- $A(H) = H$. We prove this assertion by contradiction. Assume that there exists an element $z \in H \setminus A(H)$ and denote $u = z - Pz$, where P is the orthogonal projector to the closed linear subspace $A(H)$. Then we have $\underline{u \ne 0}$ and $\forall y \in A(H)\colon (u, y) = 0$. In particular, for $y = Au$:

$$0 = (u, Au) = (Au, u) = B(u,u) \ge c\,\|u\|^2 \;\Rightarrow\; \underline{u = 0},$$

which contradicts our assumption.

[19]Note that we prove the continuity A^{-1}!

Returning back to the functional $F \in H^*$, from the Riesz representation theorem it follows that there exists a unique element $t \in H$ such that

$$\forall v \in H\colon F(v) = (t, v), \quad \|t\| = \|F\|_{H^*}.$$

Because $A\colon H \to H$ is both *injective* an *surjective*, there exists a unique $w \in H$ satisfying $t = Aw$. Thus,

$$\underline{\forall v \in H\colon F(v)} = (t, v) = (Aw, v) \underline{= B(w, v)}.$$

Moreover, (see (5.9))

$$\|w\| \leq \frac{1}{c}\|Aw\| = \frac{1}{c}\|t\| = \frac{1}{c}\|F\|_{H^*}.$$

□

Theorem 5.19 (existence and uniqueness of the weak solution).
Recall the setting from the definition of the weak solution to (5.6)*. If the bilinear from* $((u, v))$ *is also* V*-elliptic, i.e., there exists* $c > 0$ *such that*

$$\forall v \in V\colon ((v, v)) \geq c\,\|v\|_{k,2}^2,$$

then there exists a unique weak solution to the boundary value problem.

Moreover, there exists $j > 0$ *such that for the weak solution* $u \in W^{k,2}(\Omega)$ *it holds*[20]

$$\|u\|_{k,2} \leq j\left(\|f\|_{L^2(\Omega)} + \|g\|_{V^*} + \|u_0\|_{k,2}\right).$$

Proof. Choosing

i) $H := V$,
ii) $B(u, v) := ((u, v))$,
iii) $F(v) := \int_\Omega f\,v\,\mathrm{d}x + g(v) - ((u_0, v)) \in H^*$,

all assumptions of the Lax-Milgram lemma 5.19 are satisfied, and thus

$$(\exists!\, w \in V)(\forall v \in V)\colon B(w, v) = ((w, v)) = F(v),$$

$$\|w\|_{k,2} \leq \frac{1}{c}\,\|F\|_{V^*}.$$

Setting

$$u = u_0 + w$$

leads to

i) $u - u_0 = w \in V$,

[20] In other words, `the solution is continuously dependent on the right-hand side and the boundary conditions`.

ii) $\forall v \in V$:

$$((u, v)) = ((u_0 + w, v)) = ((u_0, v)) + ((w, v)) = ((u_0, v)) + F(v) = \int_\Omega f\, v \,\mathrm{d}x + g(v).$$

From the continuity of $((u, v))$ it follows that there exists $\ell > 0$ such that

$$\forall u, v \in V \colon |((u, v))| \leq \ell \,\|u\|_{k,2} \,\|v\|_{k,2},$$

and thus

$$\begin{aligned}\|u\|_{k,2} &= \|u_0 + w\|_{k,2} \leq \|u_0\|_{k,2} + \|w\|_{k,2} \leq \|u_0\|_{k,2} + \frac{1}{c}\|F\|_{V^*} \leq \\ &\leq \|u_0\|_{k,2} + \frac{1}{c}\left(\|f\|_{L^2(\Omega)} + \|g\|_{V^*} + \ell\|u_0\|_{k,2}\right) \leq j\left(\|f\|_{L^2(\Omega)} + \|g\|_{V^*} + \|u_0\|_{k,2}\right),\end{aligned}$$

where

$$j = \max\left\{\frac{1}{c}, 1 + \frac{\ell}{c}\right\} > 0.$$

□

Example 5.20. *Let us consider the above defined mixed problem* (5.7)*. One can show that*

♠ *if Γ_1 denotes a non-empty subset of $\partial\Omega$ with 'a positive measure', the functional*

$$\|v\|_{1,2,0} := \left(\int_\Omega |\nabla v|^2 \,\mathrm{d}x\right)^{\frac{1}{2}}$$

defines a norm in V equivalent with $\|\cdot\|_{1,2}$.[21]

♣ *if Γ_3 denotes a non-empty subset of $\partial\Omega$ with 'a positive measure', the following Friedrichs-type inequality holds: there exists $k > 0$ such that*

$$\forall v \in W^{1,2}(\Omega) \colon \|v\|_{1,2}^2 \leq k\left(\int_\Omega |\nabla v|^2 \,\mathrm{d}x + \int_{\Gamma_3} |Tv|^2 \,\mathrm{d}s\right).$$

Assume, in addition, that there exist constants p_0, $\sigma_0 > 0$ satisfying

- *$p(x) \geq p_0 > 0$ for almost all $x \in \Omega$,*
- *$\sigma(x) \geq \sigma_0 > 0$ for 'almost all' $x \in \Gamma_3$,*
- *$q(x) \geq 0$ for almost all $x \in \Omega$.*

Then it holds for every $v \in V$ that

$$\begin{aligned}((v, v)) &= \int_\Omega p(x)|\nabla v|^2 + q(x)v^2 \,\mathrm{d}x + \int_{\Gamma_3} \sigma(x)|Tv|^2 \,\mathrm{d}s \\ &\geq p_0 \int_\Omega |\nabla v|^2 \,\mathrm{d}x + \sigma_0 \int_{\Gamma_3} |Tv|^2 \,\mathrm{d}s \\ &\geq \min\{p_0, \sigma_0\}\left(\int_\Omega |\nabla v|^2 \,\mathrm{d}x + \int_{\Gamma_3} |Tv|^2 \,\mathrm{d}s\right),\end{aligned}$$

[21] For $\Gamma_1 = \partial\Omega$ one already has the Friedrichs theorem 4.12.

and thus under the assumption of either ♠ or ♣ there exists $c > 0$ such that

$$\forall v \in V\colon\ ((v,v)) \geq c\,\|v\|_{1,2}^2,$$

e.g., the bilinear form is V-elliptic.

To summarize, we have shown that under the assumptions of ♠ or ♣ the mixed problem (5.7) *admits a unique weak solution. In particular, the Dirichlet and Newton problems* (5.4) *and* (5.6), *respectively, admit a unique weak solution which is continuously dependent on the right-hand side and the boundary conditions.*

It remains to study the solvability of the Neumann problem (5.5). *For a special case of $q \equiv 0$, i.e., the boundary value problem*

$$\left\{\begin{array}{rcll} -\operatorname{div}(p(x)\nabla u) &=& f(x) & \text{in } \Omega, \\ p(x)\dfrac{\partial u}{\partial \nu} &=& h(x) & \text{on } \partial\Omega, \end{array}\right. \tag{5.10}$$

where p again satisfies

$$p(x) \geq p_0 > 0 \text{ for almost all } x \in \Omega,$$

the bilinear form

$$((u,v)) = \int_\Omega p(x)\nabla u\,\nabla v\,\mathrm{d}x$$

is not V-elliptic.[22] *(E.g., for $v \equiv 1 \in W^{1,2}(\Omega)$ it holds $\|v\|_{1,2} > 0$ and simultaneously $((v,v)) = 0$.) Moreover:*

- *$v \equiv 1 \in W^{1,2}(\Omega)$, and thus (should a weak solution exist) it must hold*

$$\int_\Omega f(x)\,\mathrm{d}x + \int_{\partial\Omega} h(x)\,\mathrm{d}s = 0;$$

- *if $u \in W^{1,2}(\Omega)$ is a weak solution to the Neumann problem* (5.10), *then also every*

$$u_c(x) := u(x) + c,$$

  with $c \in \mathbb{R}$ defines a weak solution to (5.10).

One can show that under the additional assumption

$$\int_\Omega f(x)\,\mathrm{d}x + \int_{\partial\Omega} h(x)\,\mathrm{d}s = 0,$$

there exist infinitely many weak solutions to the Neumann problem.

$$\left\{\begin{array}{rcll} -\operatorname{div}(p(x)\nabla u) &=& f(x) & \text{in } \Omega, \\ p(x)\dfrac{\partial u}{\partial \nu} &=& h(x) & \text{on } \partial\Omega. \end{array}\right.$$

[22] Here we have $V = W^{1,2}(\Omega)$.

Moreover, if u_1 *and* u_2 *solve the Neumann problem (in the weak sense), there exists* $c \in \mathbb{R}$ *such that*

$$u_1(x) = u_2(x) + c \textit{ for almost all } x \in \Omega.$$

Remark 5.21 (regularity of the weak solution). We have shown in several examples that under certain smoothness assumptions on the weak solution and the corresponding 'data', the weak solution is equivalent to the classical one. A natural question arises whether and how exactly the smoothness of the weak solution to the boundary value problem depends on the quality (smoothness) of the right-hand side, the boundary conditions, and the domain Ω itself.

In particular, which qualities should these functions and Ω possess for the weak solution to become a classical one? The problematic of the so-called regularity of weak solutions is very involved and there still exists a number of open problems. (By the way, the 19th and 20th problems from the famous list of problems assembled by David Hilbert are from this area of research.)

Let us briefly mention the known results for the boundary value problem

$$\begin{cases} -(p(x)u')' = f(x) \text{ in } (0,1), \\ \qquad u(0) = u(1) = 0, \end{cases} \tag{5.11}$$

where

- $f \in L^2(0,1)$,
- $p \in L^\infty(0,1)$,
- $\exists p_0 \in \mathbb{R}\colon 0 < p_0 \le p(x)$ for almost all $x \in (0,1)$.

we already know that there exists a unique weak solution to the problem – i.e., a function $u \in W_0^{1,2}(0,1)$ satisfying

$$\forall v \in W_0^{1,2}(0,1)\colon \int_0^1 p(x)u'(x)v'(x)\,\mathrm{d}x = \int_0^1 f(x)v(x)\,\mathrm{d}x.$$

We shall prove the following theorem.

Theorem 5.22. *If, in addition,*

- $p \in C^{k+1}(\langle 0,1\rangle)$,
- $f \in W^{k,2}(0,1)$,

where $k \in \{0,1,2,3,\dots\}$, *we have for the weak solution to* (5.11) *that*

$$u \in W^{k+2,2}(0,1).$$

Observation 5.23. If, in addition,

- $p \in C^2(\langle 0,1\rangle)$,
- $f \in W^{1,2}(0,1)$,

it holds (due to Theorem 5.22) that $u \in W^{3,2}(0,1)$, and thus (see Theorem 4.23)

$$u'' \in W^{1,2}(0,1) \subset C(\langle 0,1\rangle).$$

In other words, the weak solution is also the classical one.

Proof of Theorem 5.22. We will proceed in two steps.

1) First, assume that

$$p(x) \equiv 1.$$

For the weak solution $u \in W_0^{1,2}(0,1)$ to (5.11) it holds

$$\forall v \in W_0^{1,2}(0,1) = \overline{C_0^\infty(0,1)}\colon \int_0^1 u'(x)v'(x)\,\mathrm{d}x = \int_0^1 f(x)v(x)\,\mathrm{d}x.$$

This formula, in fact, states that

$$-u'' = f \text{ in the distributional sense.}$$

Thus, if[23]

$$f \in W^{k,2}(0,1) = \left\{v \in L^2(0,1)\colon v',\, v'',\, \ldots,\, v^{(k)} \in L^2(0,1)\right\},$$

we have

$$u \in W^{k+2,2}(0,1).$$

2) Let us now assume that

$$p \in C^{k+1}(\langle 0,1\rangle),$$

$$\forall x \in \langle 0,1\rangle\colon 0 < p_0 \le p(x),$$

$$f \in W^{k,2}(0,1).$$

Then

$$\forall w \in W_0^{1,2}(0,1)\colon\ v := \frac{w}{p} \in W_0^{1,2}(0,1),$$

and for the weak solution u to (5.11) it thus holds

$\underline{\forall w \in W_0^{1,2}(0,1)}$:

$$\begin{aligned}
\underline{\int_0^1 u'(x)w'(x)\,\mathrm{d}x} &= \int_0^1 u'(x)(p(x)v(x))'\,\mathrm{d}x \\
&= \int_0^1 p(x)u'(x)v'(x)\,\mathrm{d}x + \int_0^1 u'(x)p'(x)v(x)\,\mathrm{d}x \\
&= \int_0^1 f(x)v(x)\,\mathrm{d}x + \int_0^1 u'(x)p'(x)v(x)\,\mathrm{d}x \\
&= \int_0^1 (f(x) + u'(x)p'(x))\frac{w(x)}{p(x)}\,\mathrm{d}x = \underline{\int_0^1 \tilde{f}(x)w(x)\,\mathrm{d}x},
\end{aligned}$$

[23]In the following formula the derivatives are understood, obviously, in the distributional sense.

where
$$\tilde{f}(x) := \frac{1}{p(x)}\big(f(x) + u'(x)p'(x)\big).$$
From here we conclude that u solves the problem
$$\begin{cases} -u'' = \tilde{f}(x) \text{ v } (0,1), \\ u(0) = u(1) = 0, \end{cases}$$
in the weak sense and so (see 1)):
$$\tilde{f}(x) = \frac{1}{p(x)}\left(f(x) + u'(x)p'(x)\right) \in W^{k,2}(0,1) \;\Rightarrow\; u \in W^{k+2,2}(0,1).$$
Now it remains to think through the following 'chain' of implications (can be proven by induction):
$$\begin{aligned} u \in W_0^{1,2}(0,1) \;&\Rightarrow\; u' \in L^2(0,1) \;\Rightarrow\; \tilde{f} \in L^2(0,1) = W^{0,2}(0,1) \\ &\Rightarrow\; u \in W^{2,2}(0,1) \;\Rightarrow\; u' \in W^{1,2}(0,1) \;\Rightarrow\; \tilde{f} \in W^{1,2}(0,1) \\ &\Rightarrow\; u \in W^{3,2}(0,1) \;\Rightarrow\; \dots \;\Rightarrow\; u \in W^{k+2,2}(0,1). \end{aligned}$$
□

5.1 Relation to variational calculus

Recall the situation from the definition of the weak solution to the boundary value problem (see page 32) and let us assume `in addition` that

- $u_0 = 0$, $g = 0$ (i.e., homogeneous boundary conditions),
- the bilinear form $((u,v))$ is V-elliptic, i.e.,
$$(\exists c > 0)\,(\forall v \in V) : ((v,v)) \geq c\|v\|_{k,2}^2,$$
- the bilinear form $((u,v))$ is symmetric in V, i.e.,
$$\forall u,v \in V\colon ((u,v)) = ((v,u)).$$

Theorem 5.24. *Under the above specified assumptions the following statements are equivalent.*

1) *$u \in V$ is a weak solution to the boundary value problem, i.e.,*
$$\forall v \in V\colon ((u,v)) = \int_\Omega f\,v\,\mathrm{d}x;$$
2) *u is the minimizer of the functional*[24]
$$J(v) := ((v,v)) - 2\int_\Omega f\,v\,\mathrm{d}x$$

[24] The functional J is often called the quadratic functional or the energy functional.

in V, i.e.,

$$u \in V \ \wedge \ \min_{v \in V} J(v) = J(u).$$

Proof.

1) We will first prove the implication <u>*1) ⇒ 2)*</u>. Let $u \in V$ denote the weak solution, $v \in V$ an arbitrary but fixed element and
$$z := v - u.$$
Then
$$\begin{aligned} J(v) = J(u+z) &= ((u+z, u+z)) - 2\int_\Omega f(u+z)\,\mathrm{d}x = \\ &= ((u,u)) + 2((u,z)) + ((z,z)) - 2\int_\Omega f\,u\,\mathrm{d}x - 2\int_\Omega f\,z\,\mathrm{d}x = \\ &= ((u,u)) - 2\int_\Omega f\,u\,\mathrm{d}x + ((z,z)) = J(u) + ((z,z)) \geq \\ &\geq J(u) + c\|z\|_{k,2}^2 \geq J(u). \end{aligned}$$
Here we used the symmetry and V-ellipticity of the bilinear form $((\cdot,\cdot))$ and the fact that u solves the boundary value problem in the weak sense.
2) Now we prove the converse implication <u>*2) ⇒ 1)*</u>. Let $u \in V$ denote a function satisfying
$$J(u) = \min_{v \in V} J(v),$$
and $z \in V$ an arbitrary but fixed element. Then for all $t \in \mathbb{R}$ it holds
$$h(t) := J(u+tz) \geq J(u) = h(0),$$
i.e., the function h is minimized in $\mathbb{R}$ by 0, and thus: if $h'(0)$ exists, it must vanish. From the definition h it follows that for all $t \in \mathbb{R}$ it holds
$$\begin{aligned} h(t) &= ((u+tz, u+tz)) - 2\int_\Omega f(u+tz)\,\mathrm{d}x = \\ &= ((u,u)) + 2t((u,z)) + t^2((z,z)) - 2\int_\Omega f\,u\,\mathrm{d}x - 2t\int_\Omega f\,z\,\mathrm{d}x, \end{aligned}$$
and so h is a quadratic function. We can write
$$h'(0) = 0 = 2((u,z)) - 2\int_\Omega f\,z\,\mathrm{d}x.$$
We have found out that
$$\forall z \in V\colon ((u,z)) = \int_\Omega f\,z\,\mathrm{d}x,$$
and u thus defines a weak solution to the boundary value problem under consideration.

□

5.2 Discretization techniques

Recall the situation from Theorem 5.24 on page 46, i.e., let us assume that the bilinear form

$$((\cdot,\cdot))\colon W^{k,2}(\Omega)\times W^{k,2}(\Omega)\to\mathbb{R}$$

is

- continuous,
- symmetric,
- V-elliptic ($W_0^{k,2}(\Omega)\subset V=\overline{V}\subset W^{k,2}(\Omega)$).

For a given function $f\in L^2(\Omega)$ we seek $u\in V$ such that

$$\forall v\in V\colon ((u,v))=\int_\Omega f\,v\,\mathrm{d}x. \tag{5.12}$$

We already know that the problem is uniquely solvable in the weak sense.

First note that the bilinear form $((\cdot,\cdot))$ defines (under the given assumtpions) `an inner product` in the space V; moreover, from the V-ellipticity and continuity of the bilinear form $((\cdot,\cdot))$ we derive that

$$(\exists c_1,\, c_2>0)\,(\forall v\in V): c_1\|v\|_{k,2}^2\le((v,v))\le c_2\|v\|_{k,2}^2,$$

and thus the norms $\|\cdot\|:=\sqrt{((\cdot,\cdot))}$ and $\|\cdot\|_{k,2}$ are `equivalent` in V.

Using this result and the closedness of $(V,\|\cdot\|_{k,2})$ in $W^{k,2}(\Omega)$, we can deduce that

$(V,\|\cdot\|)$ is `a separable Hilbert space`.

Theorem 5.25. *Assume that* $\{e_1,e_2,\dots,e_j,\dots\}$ *defines a base of* V *orthonormal with respect to the inner product* $((\cdot,\cdot))$*. The solution to* (5.12) *can then be expressed as*

$$u=\sum_{j=1}^{\infty}\left(\int_\Omega f\,e_j\,\mathrm{d}x\right)e_j.$$

Proof. Let u denote the solution to (5.12) (it is well defined!). Then we have for every $j\in\mathbb{N}$ that

$$((u,e_j))=\int_\Omega f\,e_j\,\mathrm{d}x,$$

and thus (see the definition of an orthonormal basis in [2])

$$u=\sum_{j=1}^{\infty}((u,e_j))e_j=\sum_{j=1}^{\infty}\left(\int_\Omega f\,e_j\,\mathrm{d}x\right)e_j. \tag{5.13}$$

□

5.2.1 The Ritz method

We have already found out that $u \in V$ solves (5.12) if and only if

$$J(u) = \min_{v \in V} J(v),$$

where

$$J(v) := ((v,v)) - 2\int_\Omega f\, v \,\mathrm{d}x.$$

We can try to approximate u as follows: let $\{e_1, e_2, \ldots\}$ denote an orthonormal basis of V (with respect to $((\cdot,\cdot))$). We seek a function

$$u_m \in \operatorname{span}(e_1, \ldots, e_m) =: V_m$$

satisfying

$$J(u_m) = \min_{v \in V_m} J(v).$$

Since

$$u_m \in V_m \;\Leftrightarrow\; \left[\exists \alpha_1, \ldots, \alpha_m \in \mathbb{R} \colon u_m = \sum_{j=1}^m \alpha_j\, e_j\right],$$

we are looking for a minimizer of

$$h(\alpha_1, \ldots, \alpha_m) := J\left(\sum_{j=1}^m \alpha_j\, e_j\right) \text{ in } \mathbb{R}^m.$$

Note that

$$\begin{aligned} h(\alpha_1, \ldots, \alpha_m) &= \left(\left(\sum_{j=1}^m \alpha_j\, e_j, \sum_{j=1}^m \alpha_j\, e_j\right)\right) - 2\int_\Omega f \sum_{j=1}^m \alpha_j\, e_j \,\mathrm{d}x = \\ &= \sum_{j=1}^m \alpha_j^2 - 2\sum_{j=1}^m \alpha_j \int_\Omega f\, e_j \,\mathrm{d}x = \sum_{j=1}^m \underbrace{\left(\alpha_j^2 - 2\alpha_j \int_\Omega f\, e_j \,\mathrm{d}x\right)}_{\to\infty \text{ for } |\alpha_j|\to\infty}, \end{aligned}$$

and thus ($h \in C^\infty(\mathbb{R}^m)$) the minimum of h in $\mathbb{R}^m$ exists. The minimum is attained in the point $(\alpha_1, \ldots, \alpha_m)$, for which it holds

$$\forall j \in \{1, \ldots, m\} \colon \frac{\partial h}{\partial \alpha_j}(\alpha_1, \ldots, \alpha_m) = 2\alpha_j - 2\int_\Omega f\, e_j \,\mathrm{d}x = 0.$$

Thus, for the approximation u_m we can write

$$u_m = \sum_{j=1}^m \left(\int_\Omega f\, e_j \,\mathrm{d}x\right) e_j.$$

This means that $u_m \in V_m$ is in fact the sum of the first m elements in the Fourier series of u (see (5.13)), which among other things means

$$\|u - u_m\| = \inf_{z \in V_m} \|u - z\|,$$

$$u_m \to u \text{ in } (V, \|\cdot\|),$$

and thus – due to the equivalence of the norms $\|\cdot\|$ and $\|\cdot\|_{k,2}$ – also

$$u_m \to u \text{ in } (V, \|\cdot\|_{k,2}).$$

Also note another point: the functional J restricted to V_m is minimized by $u_m \in V_m$. This holds true `regardless` of the basis chosen for V_m.

Let

$$\{v_1, v_2, \ldots, v_j, \ldots\}$$

denote a basis (not necessarily orthogonal) of the space V, i.e.,

- $(\forall v \in V)\,(\exists (c_n) \subset \mathbb{R}) : v = \sum_{j=1}^{\infty} c_j\, v_j := \lim_{m\to\infty} \sum_{j=1}^{m} c_j\, v_j$,
- $\forall j \in \mathbb{N}$: $v_1, \ldots, v_j$ are linearly independent,

and denote

$$V_m = \operatorname{span}(v_1, \ldots, v_m).$$

Then (as we already know)

$$\exists!\, u_m \in V_m \colon J(u_m) = \min_{v \in V_m} J(v).$$

Similarly as in the previous case, let us show how to find the approximation

$$u_m = \sum_{j=1}^{m} c_j\, v_j \in V_m.$$

We define

$$\tilde{h}(c_1, \ldots, c_m) := J\left(\sum_{j=1}^{m} c_j\, v_j\right) = \sum_{i=1}^{m} c_i \left(\sum_{\ell=1}^{m} c_\ell((v_i, v_\ell))\right) - 2\sum_{j=1}^{m} c_j \int_\Omega f\, v_j \,\mathrm{d}x.$$

Then

$$\forall j \in \{1, \ldots, m\} \colon \frac{\partial \tilde{h}}{\partial c_j}(c_1, \ldots, c_m) = \sum_{\substack{i=1\\ i\neq j}}^{m} c_i((v_i, v_j)) + \frac{\partial}{\partial c_j}\left(c_j \sum_{\ell=1}^{m} c_\ell((v_j, v_\ell))\right) - 2\int_\Omega f\, v_j \,\mathrm{d}x$$

$$= \sum_{\substack{i=1\\ i\neq j}}^{m} c_i((v_i, v_j)) + \sum_{\ell=1}^{m} c_\ell((v_j, v_\ell)) + c_j((v_j, v_j)) - 2\int_\Omega f\, v_j \,\mathrm{d}x =$$

$$= 2\sum_{i=1}^{m} c_i((v_i, v_j)) - 2\int_\Omega f\, v_j \,\mathrm{d}x = 0$$

$$\Updownarrow$$

$$\forall j \in \{1,\ldots,m\}\colon\ \sum_{i=1}^{m} c_i((v_i,v_j)) = \int_\Omega f\, v_j \,\mathrm{d}x. \tag{5.14}$$

The solution of this system of linear equations (`uniquely!`) defines the sought coefficients $c_1,\ldots,c_m$. We have thus proven the following theorem.

Theorem 5.26. *Assume that $\{v_1,v_2,\ldots,v_m,\ldots\}$ denotes a basis of the space V and for every $m \in \mathbb{N}$*

$$u_m := \sum_{j=1}^{m} c_j\, v_j,$$

where $c_1,\ldots,c_m \in \mathbb{R}$ solve the system (5.14)*.*

Then

$$u := \lim u_m$$

solves the problem (5.12)*.*

In 'real-life' situations it is rather difficult to find a basis of the space V (defined in a bounded domain Ω) and thus also the finite dimensional subspaces V_m. This difficulty is partially overcome by the Galerkin method described in the following section. Differently from the Ritz method, the Galerkin approach can also be extended to problems with `non-symmetric` bilinear forms $((\cdot,\cdot))$.

5.2.2 The Galerkin method

Definition 5.27. Consider a class of finite dimensional subspaces of V denoted by

$$\{V_h\}_{h\in(0,1)}.$$

We say that

$$\underline{V_h \to V \text{ for } h \to 0+}\,,$$

if

$$\forall v \in V\colon\ \lim_{h\to 0+} (\mathrm{dist}(V_h, v)) = 0,$$

where

$$\mathrm{dist}(V_h, v) := \inf_{z\in V_h} \|v - z\|_{k,2}.$$

Example 5.28. *If $\{v_1,\ldots,v_m,\ldots\}$ denotes a basis of V, $V_m = \mathrm{span}(v_1,\ldots,v_m)$, we can construct a subspace $\{V_h\}_{h\in(0,1)}$ satisfying the property $V_h \to V$ pro $h \to 0+$ as*

$$V_h := V_m \ \textit{for}\ \frac{1}{m+1} \le h < \frac{1}{m}.$$

Theorem 5.29. *Assume that u solves* (5.12) *and $V_h \to V$ for $h \to 0+$. Then for all $h \in (0,1)$ there exists a unique $u_h \in V_h$ satisfying*

$$\forall v \in V_h \colon ((u_h, v)) = \int_\Omega f\, v \,\mathrm{d}x.$$

Moreover, it holds

$$\lim_{h\to 0+} u_h = u \ (\text{in } W^{k,2}(\Omega)),$$

i.e.,

$$(\forall \varepsilon > 0)\,(\exists \delta > 0)\,(\forall h \in (0,\delta)) : \|u - u_h\|_{k,2} < \varepsilon.$$

Proof. [25]

1) Choose $h \in (0,1)$ and $V_h = \operatorname{span}(v_1, v_2, \ldots, v_{m_h})$, where $v_1, \ldots, v_{m_h}$ are linearly independent. We will seek the function $u_h \in V_h$ in the form

$$u_h = \sum_{i=1}^{m_h} c_i\, v_i,$$

where $c_1, \ldots, c_{m_h} \in \mathbb{R}$. Obviously, it holds for u_h that

$$\forall v \in V_h \colon \ ((u_h, v)) = \int_\Omega f\, v \,\mathrm{d}x$$

if and only if (substitute $v = v_j$)

$$\forall j \in \{1, \ldots, m_h\} \colon \ \sum_{i=1}^{m_h} c_i((v_i, v_j)) = \int_\Omega f\, v_j \,\mathrm{d}x.$$

This results in the same system of linear equations (with `a unique solution` $c_1, \ldots, c_{m_h}$) as in the case of the Ritz method (see (5.14)).

2) It remains to prove that

$$\lim_{h\to 0+} u_h = u.$$

For an arbitrary but fixed $h \in (0,1))$ consider:

- $u \in V$ solving the problem (5.12), i.e.,

$$\forall v \in V \colon ((u, v)) = \int_\Omega f\, v \,\mathrm{d}x, \tag{5.15}$$

- $u_h \in V_h \subset V$ such that

$$\forall v \in V_h \colon ((u_h, v)) = \int_\Omega f\, v \,\mathrm{d}x, \tag{5.16}$$

(such u_h, as we already know, is uniquely defined!),

[25] Thanks, Dalibor!

- an arbitrary element $v_h \in V_h$.

Then there exist constants c_1, $c_2 > 0$ (independent of u, u_h, v_h!) such that

$$\begin{aligned} c_1\|u - u_h\|_{k,2}^2 &\leq ((u - u_h, u - u_h)) = ((u - u_h, u - u_h)) + \overbrace{((u - u_h, \underbrace{u_h - v_h}_{\in V_h \subset V}))}^{=0 \text{ (see (5.15) and (5.16))}} \\ &= ((u - u_h, u - u_h + u_h - v_h)) = ((u - u_h, u - v_h)) \\ &\leq c_2\|u - u_h\|_{k,2}\|u - v_h\|_{k,2}, \end{aligned}$$

and thus

$$\|u - u_h\|_{k,2} \leq \frac{c_2}{c_1}\|u - v_h\|_{k,2}\,.$$

Using this results and the property $V_h \to V$ for $h \to 0+$ finally leads to

$$\lim_{h\to 0+} u_h = u.$$

$\square$

Remark 5.30. A special choice of the spaces V_h (i.e., a special case of the Galerkin method) leads to the so-called <u>finite element method</u>. The key idea of this method is to define the elements of the basis $v_1, \dots, v_{m_h} \in V_h$ in such a way that the matrix

$$A = ((v_i, v_j))_{i,j=1}^{m_h}$$

is sparse (i.e., most of its entries vanish) and the non-zero entries are distributed close to the main diagonal. This is a quality crucial for practical (numerical) solution of 'large' systems (i.e., for 'large' m_h).

References

[1] R. Blaheta: *Matematické modelování a metoda konečných prvků*, http://mi21.vsb.cz/, 2012.

[2] J. Bouchala: *Úvod do funkcionální analýzy*, http://mi21.vsb.cz/, 2012. 48

[3] J. Franců: *The methods for solving operator equations with applications to mathematical physics*, Proceedings of the seminar IMAMM 94, ZČU Plzeň, 1994.

[4] S. Fučík, A. Kufner: *Nelineární diferenciální rovnice*, SNTL, Praha, 1978.

[5] S. Fučík, J. Milota: *Matematická analýza II*, skripta MFF UK, Praha, 1980.

[6] O. John, J. Nečas: *Rovnice matematické fyziky*, skripta MFF UK, Praha, 1977.

[7] J. Kačur: *Vybrané kapitoly z matematickej fyziky I*, skripta MFF Univerzity Komenského, Bratislava, 1988.

[8] J. Lukeš: *Zápisky z funkcionální analýzy*, skripta MFF UK, Praha, 1998.

[9] J. Lukeš, J. Malý: *Míra a integrál*, skripta MFF UK, Praha, 1993.

[10] S. Míka, A. Kufner: *Okrajové úlohy pro obyčejné diferenciální rovnice*, SNTL, Praha, 1981.

[11] S. Míka, A. Kufner: *Parciální diferenciální rovnice I, stacionární rovnice*, SNTL, Praha, 1983.

[12] J. Nečas: *Direct methods in the theory of elliptic equations*, Springer-Verlag, Berlin, 2012.

[13] K. Rektorys: *Variační metody v inženýrských problémech a v problémech matematické fyziky*, Academia, Praha, 1999.

[14] K. Rektorys: *Metoda časové diskretizace a parciální diferenciální rovnice*, SNTL, Praha, 1985.

[15] K. Rektorys: *Matematika V. Obyčejné a parciální diferenciální rovnice s okrajovými podmínkami*, skripta FSv ČVUT, Praha, 1989.

[16] M. Renardy, R.C. Rogers: *An introduction to partial differential equations*, Springer-Verlag, New York, 1993.

[17] E. Zeidler: *Applied functional analysis, applications to mathematical physics*, Springer-Verlag, New York, 1995.

[18] E. Zeidler: *Applied functional analysis, main principles and their applications*, Springer-Verlag, New York, 1995.

Index