A PRIMER ON FUNCTIONAL ANALYTIC METHODS FOR PDES

Lecture notes

HANNES MEINLSCHMIDT

FAU Erlangen-Nürnberg

Summer term 2021
Preliminary version
Last updated: January 7, 2022

Preface

These lecture notes cover topics suitable for an introductory 5 ECTS (2+1 weekly hours) primer on functional analytic methods in partial differential equations.

The goal of the course is to provide a solid foundation in the aspects and tools of functional analysis used in modern abstract PDE analysis. Of course such a course cannot comprehensively consider all aspects of either functional analysis or PDE applications within a 5 ECTS scope, so there will be shortcuts and several results without proofs. The proofs will be part of the exercises or can be found easily in the literature mentioned below.

The overarching idea is to establish essentially three fundamental ideas which are prevalent in modern PDE theory basing on functional analysis; these being: positivity (ellipticity), Fredholm theory and diagonalization. The theoretical foundation will therefor lead to bilinear forms, compactness, and spectral theory, particularly in Hilbert spaces. Due to their important compactness properties, weak topologies will also be considered. Sobolev spaces are then the natural environment to consider elliptic boundary value problems in. Due to time constraints, time-dependent problems (evolution equations) will merely be mentioned; however, the topics taught will transfer very well to this more advanced topic.

There are several very nice and comprehensive books about functional analysis methods in PDEs. The books I want to recommend most in relation to this lecture are the ones by Bressan [Bre13] and Brezis [Bre10]; the reader will surely notice several similarities in approaches. There are also two very concise and straight-to-the-point chapters with many common topics in the book by Hackbusch [Hac17]. The classical books by Alt [Alt16] (functional analysis side) and Evans [Eva98] (PDE side) may serve to delve further into the respective topics.

These lecture notes are written during the summer term 2021 at FAU Erlangen-Nürnberg. There will certainly be mistakes and inconsistencies, and maybe I will want to restructure some parts later on. In this sense, these notes are a **work in progress**.

I am grateful for any comments and suggestions and of course particularly so for any notification of errors, by eMail at `meinlschmidt[at]math.fau.de`.

Contents

1 Introduction

1.1 Physical motivation

We begin with probably the most classical example of a physical process modeled by a partial differential equation: the heat equation.

Suppose that we have an item at hand, made out of an isotropic heat-conducting material (such as steel), and there is a certain source of energy around will influence the temperature of the item. We want to understand the evolution of this temperature. Ideally, we imagine the item to occupy a certain volume $\Omega \subset \mathbb{R}^d$ in space ($d \in \mathbb{N}$), and we would like to have a value $\theta(t, \mathrm{x}) \in \mathbb{R}$ of the temperature at time t and in the spatial point $\mathrm{x} \in \Omega$. (We imagine to start investigating the whole process at time 0).

To keep things simple, let us consider an external heat source modeled by a function $f\colon \Omega \to \mathbb{R}$. (Such a heat source could e.g. occur in so-called *Joule heating*.) Quantities of interest are the internal heat energy normalized to volume $Q(t, \mathrm{x}) \in \mathbb{R}$, and, of course, the temperature $\theta(t, \mathrm{x}) \in \mathbb{R}$ at every point t in time and $\mathrm{x} \in \Omega$. Clearly, life tells us that there should be a particular relation between these two, and indeed,

$$\partial_t Q = c\rho\, \partial_t \theta,$$

where c and ρ are constants defined by the material of the item; c is the *specific heat capacity*, and ρ is the density of the material. The next ingredient is the famous **Fourier law**[1]. It says that if $q(t, \mathrm{x})$ is the (normalized, vector) *flow* of heat energy through a surface, then

$$q = -\kappa \nabla \theta,$$

where κ is the thermal conductivity coefficient of the material. (This may in general be a matrix.) According to the first law of thermodynamics, its relation to the internal heat energy Q is that the temporal evolution of Q—the *aggregation* of energy in a fixed point x—will be exactly the amount of heat energy "left behind at x" by the flow q plus the external energy source f, that is,

$$\partial_t Q = -\operatorname{div}(q) + f.$$

Putting the foregoing relations together, we obtain the **heat equation**

$$c\rho\, \partial_t \theta = \partial_t Q = -\operatorname{div}(q) + f = \operatorname{div}(\kappa \nabla \theta) + f. \tag{1.1}$$

Of course, this is just a very basic derivation. We need to mention the initial temperature distribution $\theta_0 = \theta(0, \cdot)\colon \Omega \to \mathbb{R}$ at time $t = 0$, and also talk about how the item interacts with its surroundings: is it insulated, is there maybe a heat source at the boundary, or will there be radiative heat/energy loss to the exterior? The latter question asks for **boundary conditions**, that is, a law for for $\theta(t, \cdot)$ on $\partial\Omega$.

[1]Joseph Fourier (1768–1830)

The first classical boundary condition is an energy source (or sink) at the boundary, so $\theta = g$ on $\partial\Omega$, called **Dirichlet**[2] boundary condition. The second one is heat flux over the boundary, so $\nu \cdot \nabla\theta = g$ on $\partial\Omega$ with the (unit) outer normal ν, called **Neumann**[3] or **natural** boundary condition, with the particular case $g = 0$ corresponding to insulation of the item. The subclass of natural boundary conditions where $g = \alpha(\theta_{\text{ext}} - \theta)$ for a modulation α is called **Robin**[4] boundary condition, and we can interpret it as an interaction of the temperature θ within Ω and the exterior temperature θ_{ext}. There might be more involved boundary conditions, e.g. in problems modeling radiation.

A time-dependent partial differential equation such as (1.1) is also called **evolution equation** because it describes the evolution of the function $t \mapsto \theta(t, \cdot)$ which is a *function of functions*. The natural setting to investigate such mappings is that of infinite-dimensional (normed) vector spaces, so, in general, Banach spaces. (Because we like completeness!) However, we will not go quite as deep into the subject to actually consider evolution equations. We will content ourselves with the **stationary** variant of (1.1):

$$-\operatorname{div}(\kappa\nabla u) = f \quad + \quad \text{boundary condition} \tag{1.2}$$

which we obtain from $\partial_t\theta = 0$ in (1.1). One may imagine that we have looked at the heat evolution for a *very* long time and it has reached an equilibrium state. (To distance ourselves from the evolution equation, we call the quantity of interest u now.) If κ is in fact a constant, then, for good measure, we also normalize $\kappa = 1$ and obtain **Poisson's equation**[5]

$$-\Delta u = f \quad + \quad \text{boundary condition} \tag{1.3}$$

with the **Laplace operator**[6] $\Delta := \sum_{i=1}^{d} \partial_{x_i}$. The **linear elliptic** partial differential equation (1.2) is exactly the type of partial differential equation that we will ultimately deal with in this lecture. The Poisson equation is the most simple problem in this class, but still quite prototypical.

Clearly, there arise quite some questions of interest: given f of a particular quality, does there exist a solution u to (1.2) or (1.3)? In which sense? Is it unique? And how does the quality of f transfer to a solution u? We will attempt to answer at least some of those. To this end, it is a good leitmotif to imagine such an elliptic linear PDE as above as a linear equation

$$Lu = f \tag{1.4}$$

in a vector space, in analogy to a system of linear equations for $x \in \mathbb{R}^n$

$$Ax = b \tag{1.5}$$

[2] Johann Peter Gustav Lejeune Dirichlet (1805–1859)
[3] Carl Gottfried Neumann (1832–1925)
[4] Victor Gustave Robin (1855–1897)
[5] Siméon Denis Poisson (1781–1840)
[6] Pierre-Simon, marquis de Laplace (1749–1827)

induced by a (quadratic) matrix $\mathbb{R}^{n\times n}$ and a vector $b \in \mathbb{R}^n$. Of course, as already hinted at above, the obvious but nevertheless crucial difference here is that u and f are functions on an infinite-dimensional set (a continuum in fact) and thus **infinite-dimensional objects** between which L acts. This is the single reason why we require so much foundations to consider even linear PDEs. Still, we can allow ourselves to be inspired by the various techniques and beautiful theory of linear algebra. In fact, it will turn out that one can transfer many ideas in an appropriate manners to the infinite-dimensional system (1.4).

1.2 Linear algebra inspirations

Positivity

Suppose that A is positive definite, that is, there is $\gamma > 0$ such that

$$(Ax, x) \geq \gamma |x|^2 \quad (x \in \mathbb{R}^n),$$

where $(\cdot,\cdot)$ is the Euclidean inner product and $|\cdot|$ the associated norm. Then A is invertible, A^{-1} is also positive definite, and (1.5) has a unique solution $x = A^{-1}b$ for every $b \in \mathbb{R}^n$.

We can rediscover this property in a very similar form for the infinite-dimensional problem (1.4). Indeed, if there is a constant $\alpha > 0$ such that

$$(L^{1/2}u, L^{1/2}u)_H \approx a(u,u) \geq \alpha \|u\|_V^2 \quad (u \in V), \tag{1.6}$$

then L is invertible and (1.4) has a unique solution $u = L^{-1}f$ for every $f \in V'$.

There are some things to unpack here. First, First, V and H are Hilbert spaces with $V \hookrightarrow H$. Then, a is a bilinear form on the Hilbert space V and V' is the dual space of V. The operator L then acts $V \to V'$. . We will give sense to all these notions throughout this course, although we will not directly encounter the square root $L^{1/2}$ of a differential operator. For now it is enough to note that the positivity condition (1.6) for L is formulated not for L directly but for a related object $L^{1/2}$ in the form of a bilinear form a. The idea behind this formulation involves two different infinite-dimensional spaces whose norms are in general **not equivalent**. This is very much related to infinite-dimensional spaces since on finite-dimensional spaces (with fixed dimension), all norms are equivalent!

Fredholm alternative

In linear algebra, the unique solvability of (1.5) for every $b \in \mathbb{R}^n$ is equivalent to unique (trivial) solvability of the *homogeneous equation* $Ax = 0$, that is, whether

$$Ax = 0 \quad \Longrightarrow \quad x = 0.$$

This is because the **rank-nullity theorem** implies that injectivity of (the mapping induced by) A is equivalent to its surjectivity. Unfortunately, the rank-nullity theorem is not true any more in infinite-dimensional spaces. That is, there are injective linear mappings which are not surjective, and vice-versa.

However, there is a (surprisingly large) class of operators for which a similar theory holds true; this is the class of so-called **Fredholm operators**[7] of **index** 0. For these, we also have the good property

$$L \text{ is injective} \quad \Longleftrightarrow \quad L \text{ is surjective.}$$

Hence, if L is a Fredholm operator of index 0, and

$$Lu = 0 \quad \Longrightarrow \quad u = 0,$$

then (1.4) has a unique solution u for every f. A particularly useful case is when $L = I - K$, where I is the identity operator and K is a **compact** operator. Of course, for this to make sense we need to understand what a compact operator is! (The titular **Fredholm alternative** says that *either* (1.4) is uniquely solvable for every f *or* $Lu = 0$ has a nonzero solution, which is a particular rephrasing of historical importance of the foregoing explanations.)

Diagonalization

Last but not least, let A be symmetric. Then the eigenvalues $\lambda_1, \dots, \lambda_n$ of A are real. Even (much) more, the **spectral theorem** says that there is an orthonormal basis $\{v_1, \dots, v_n\}$ of $\mathbb{R}^n$ given by eigenvectors of A. From expanding x and b with respect to this orthonormal basis, we obtain for (1.5)

$$x = \sum_{i=1}^{n} (x, v_i) v_i \quad \text{and} \quad \sum_{i=1}^{n} \lambda_i (x, v_i) v_i = Ax = b = \sum_{i=1}^{n} (b, v_i) v_i$$

Now, A is invertible if and only if all eigenvalues λ_i are nonzero, and in this case, explicitly,

$$x = \sum_{i=1}^{n} \lambda_i^{-1} (b, v_i) v_i.$$

Interestingly, we are able to transfer nearly all of this theory to the infinite-dimensional case (1.4), at least if L is a **compact symmetric** operator on a Hilbert space H. Then we indeed obtain an orthonormal basis $\{\phi_1, \phi_2, \dots\}$ of H—by definition, this basis cannot be finite!—consisting of eigenvectors of L corresponding to eigenvalues $\lambda_1, \lambda_2, \dots$, which now have the property that $\lambda_i \to \infty$ as $i \to \infty$. Again, if $\lambda_i \neq 0$ for all $i \in \mathbb{N}$, then

$$u = \sum_{i=1}^{\infty} \lambda_i^{-1} (f, \phi_i)_H \, \phi_i$$

is the unique solution to (1.4) for given f.

[7] Erik Ivar Fredholm (1866–1927)

2 Fundamentals

2.1 Normed vector spaces

In this section we collect several basic objects from the theory of normed vector spaces. In the following, X is always a vector space over the field $\mathbb{K}$ of real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$. The very most fundamental idea is that we want to have a notion of the *size* and the *distance* between elements x, y of X. This will be realized by a *norm*.

A mapping $\|\cdot\|\colon X \to \mathbb{R}_0^+$ is called a **norm**, if

a) for every $x \in X$ and $\lambda \in \mathbb{K}$, we have $\|\lambda x\| = |\lambda|\|x\|$ (*positive homogeneity*),

b) for every $x, y \in X$, we have $\|x + y\| \leq \|x\| + \|y\|$ (*triangle inequality*), and

c) we have $\|x\| = 0$ if and only if $x = 0$ (*definiteness*).

If from $\|x\| = 0$ it does *not* follow that $x = 0$, then $\|\cdot\|$ is called **seminorm**.

The pair $(X, \|\cdot\|)$ is called a **normed vector space**. If we mean a generic norm on X, or if there is a “standard” norm on X, then we just refer to X instead of $(X, \|\cdot\|)$ and indicate the norm meant by $\|\cdot\|_X$.

In a normed vector space, the norm $\|x\|_X$ measures the *size* of the element $x \in X$. Accordingly, we say that a set $M \subset X$ is **bounded** if there exists $C \geq 0$ such that $\|x\| \leq C$ for all $x \in M$.

Clearly, a size gives also rise to a notion of **distance** between points $x, y \in X$: the distance should be exactly the size of the “connection” vector $x - y \in X$! Indeed,

$$d(x, y) := \|x - y\| \qquad (x, y \in X)$$

defines a **metric** on X. A *normed* vector space is thus also always a *metric* vector space. In particular, it is a *topological* vector space and we can now talk about all sorts of exciting topological notions in X. (But a norm does much more than just induce a topology, as we will see soon.)

We begin with the basic notion of convergence. A sequence $(x_k) \subseteq X$ is said to **converge** to $x \in X$, in short, $x_k \to x$ or $\lim_{k\to\infty} x_k = x$ (in X), if

$$\lim_{k\to\infty} \|x_k - x\| = 0.$$

Then we can already say what we mean by continuity. Let $U \subseteq X$. A map $f\colon U \to Y$ is said to be **continuous** at $x \in U$ if

$$x_k \to x \quad \text{in } X \qquad \Longrightarrow \qquad f(x_k) \to f(x) \quad \text{in } Y \qquad ((x_k) \subseteq U).$$

As in elementary calculus, one shows that this is equivalent to the dreaded (ε, δ)-formulation: The map $f\colon U \to Y$ is continuous at $x \in U$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|y - x\|_X \le \delta \quad \Longrightarrow \quad \|f(y) - f(x)\|_Y \le \varepsilon \qquad (y \in U)$$

We say that $f\colon U \to Y$ is continuous if it is continuous at every point in U. If the choice of δ in dependence of ε can be done uniformly for every $x \in U$, then f is **uniformly continuous**. It is **Lipschitz continuous**[8] if there is a constant $L_f \ge 0$ such that

$$\|f(x) - f(y)\|_Y \le L_f \|x - y\|_X \qquad (x, y \in U).$$

Note that the norm on a normed vector space is always (Lipschitz) continuous. This can be seen by the *reverse triangle inequality*

$$\big|\|x\| - \|y\|\big| \le \|x - y\| \quad (x, y \in X).$$

In particular, from $x_k \to x$ it always follows that $\|x_k\| \to \|x\|$. (Clearly, the converse is false in general.)

A particular brand of sequences are the following: A sequence $(x_k) \subseteq X$ is called a **Cauchy sequence**[9] if for every $\varepsilon > 0$, there is an index $M \in \mathbb{N}$ (large) such that

$$\|x_k - x_\ell\| < \varepsilon \qquad (k, \ell \ge M).$$

It is easily seen that every convergent sequence is also a Cauchy sequence. But the converse is not always true. In fact, this is a fundamental property: we say that X is **complete** if every Cauchy sequence is also convergent. A complete normed vector space is a **Banach space**[10].

Now, let us turn to topological properties of sets. The **(open) ball** in X, centered at $x \in X$, with radius $r > 0$, is denoted by

$$B_X(x, r) := \{y \in X\colon \|x - y\| < r\}.$$

If the space X is clear from the context, we skip the subscript X in $B_X(x, r)$. The triangle inequality implies that (open) balls are **convex**: From $y, z \in B(x, r)$ it follows that $(1 - \lambda)y + \lambda z \in B(x, r)$ for every $\lambda \in [0, 1]$.

We say that a set $U \subseteq X$ is a **neighborhood** of $x \in U$ if there exists $r > 0$ such that $B(x, r) \subseteq U$. The set $U \subseteq X$ is **open** if it is a neighborhood of every one of its elements $x \in U$. Conversely, $V \subseteq X$ is **closed** if its complement $X \setminus V$ is open. (Recall the saying, though: *sets are not doors!*—they can be neither open nor closed.) The **closure** $\overline{U}$ of a set $U \subseteq X$ is the smallest closed set containing U. We say that a set $U \subseteq X$ is

[8] Rudolf Lipschitz (1832–1903)
[9] Louis Augustin Cauchy (1789–1857)
[10] Stefan Banach (1892–1945)

dense in X if $\overline{U} = X$. A normed vector space is **separable** if there exists a *countable* dense subset. Every subset of a separable space is again separable. The largest open set contained in U is called the **interior** of U and denoted by $\operatorname{int}(U)$. It is given by the union of all $x \in U$ such that there exists $r > 0$ such that $B(x,r) \subseteq U$.

Particularly nice sets are those where sequences cannot go fully wild in the following sense: A set $U \subseteq X$ is **compact** if every sequence $(x_k) \subseteq U$ admits a convergent subsequence (x_{k_ℓ}) whose limit is again in U. If the closure $\overline{U}$ of a set $U \subset X$ is compact, then we say that U is **relatively compact**. In a Banach space X, an equivalent notion is as follows: $U \subseteq X$ is relatively compact if and only if it is **totally bounded** (or **precompact**): for every $\varepsilon > 0$ there exists a finite number of point $u_1, u_2, \ldots, u_N$ such that

$$U \subseteq \bigcup_{k=1}^{N} B(u_k, \varepsilon),$$

that is, U can be covered by a finite number of balls of radius ε centered at points in U.

Some examples:

We use the opportunity to recall some examples of Banach spaces.

1) On the finite-dimensional space $\mathbb{K}^n$, the p-norms, for $1 \leq p \leq \infty$,

$$\|x\|_p := \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p}, \quad \|x\|_\infty := \max_{1 \leq i \leq n} |x_i|$$

all give rise to separable Banach spaces. The dense subset is derived from $\mathbb{Q}^n$.

2) Let $(\Omega, \mathcal{A}, \mu)$ be a σ-finite measure space. For $1 \leq p < \infty$, set

$$\|f\|_{L^p(\Omega,\mu)} := \Big(\int_\Omega |f(x)|^p \,\mathrm{d}\mu(x)\Big)^{1/p}$$

and

$$\|f\|_{L^\infty(\Omega,\mu)} := \mu\text{-}\operatorname*{ess\,sup}_{x \in \Omega} |f(x)| := \inf\Big\{M \geq 0 \colon \mu\big(\{x \in \Omega \colon |f(x)| > M\}\big) = 0\Big\}.$$

Then, for $1 \leq p \leq \infty$, we define the **Lebesgue spaces**[11] $L^p(\Omega, \mu)$ to consist of all (equivalence classes w.r.t. equality μ-almost everywhere of) μ-measurable functions $f \colon \Omega \to \mathbb{K}$ such that $\|f\|_{L^p(\Omega,\mu)} < \infty$. They are Banach spaces and separable if $1 \leq p < \infty$ due to approximation by step functions.

These spaces are prototypical function spaces. In the (standard) case of μ being the Lebesgue measure on $\mathbb{R}^n$, we write only $L^p(\Omega)$.

[11] Henri Léon Lebesgue (1875–1941)

3) Another prototypical Banach space is the space of continuous functions $C(E)$ on a compact set $E \subset \mathbb{R}^n$ with the **supremum norm**

$$\|f\|_\infty := \|f\|_{C(E)} := \max_{x \in E} |f(x)|.$$

It is also separable. However, the space of continuous functions with the L^1-type norm

$$\|f\|_1 := \int_E |f(x)| \, \mathrm{d}x$$

is **not** a Banach space.

2.2 Linear operators

We fix a field $\mathbb{K}$ underlying the considered vector spaces. Let X, Y be normed vector spaces.

A linear mapping $A\colon X \subseteq \operatorname{dom}(A) \to Y$ from a subspace $\operatorname{dom}(A) \subseteq X$ of X to Y is called a **linear operator** between X and Y with **domain** $\operatorname{dom}(A)$. We say that the operator A is **densely defined** if $\operatorname{dom}(A)$ is dense in X. The **range** of A is given by

$$\operatorname{ran}(A) := \big\{Ax\colon x \in \operatorname{dom}(A)\big\} \subseteq Y,$$

and the **kernel** of A by

$$\ker(A) := A^{-1}\big[\{0\}\big] := \big\{x \in \operatorname{dom}(A)\colon Ax = 0\big\} \subseteq X.$$

Of course, A is injective if and only if $\ker(A) = \{0\}$.

Linear operators have a particularly nice characterization of continuity. Indeed, consider the following notion: A mapping $\Lambda\colon X \to Y$ is said to be **bounded** if it maps bounded sets in X into bounded sets in Y. Now, a linear operator between X and Y with domain $\operatorname{dom}(A) = X$, so $A\colon X \to Y$, is bounded exactly when

$$\|A\|_{X \to Y} := \sup_{\|x\|_X \leq 1} \|Ax\|_Y < \infty. \tag{2.1}$$

We note that in fact

$$\sup_{\|x\|_X \leq 1} \|Ax\|_Y = \sup_{\|x\|_X = 1} \|Ax\|_Y = \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} \tag{2.2}$$

if any of these quantities is finite. Moreover, in this case,

$$\|Ax\|_Y \leq \|A\|_{X \to Y} \|x\|_X \qquad (x \in X). \tag{2.3}$$

It is thus both sufficient and necessary for A to be bounded that there exists a constant $C \geq 0$ such that

$$\|Ax\|_Y \leq C \, \|x\|_X \qquad (x \in X).$$

Now we can have our first theorem:

Theorem 2.1. *A linear operator $X \to Y$ is bounded if and only if it is continuous.*

Proof. If A is bounded, then by (2.3),

$$\|Ax - Ay\|_Y \leq \|A\|_{X\to Y}\|x - y\|_X \qquad (x, y \in X),$$

and from $y \to x$ in X it follows that $Ay \to Ax$ in Y. (In fact, A is Lipschitz continuous!)

Conversely, let A be continuous. We consider $x = 0$. For $\varepsilon = 1$, there exists $\delta > 0$ such that, for all $y \in X$, from $\|y\|_X \leq \delta$ it follows that $\|Ay\|_Y \leq \varepsilon = 1$. Now let $z \in X$ with $\|z\|_X = 1$ be arbitrary. Then $y := \delta z$ satisfies $\|y\|_X = \delta$, so

$$\|Az\|_Y = \delta^{-1}\|Ay\|_Y \leq \delta^{-1} \qquad (\|z\|_X = 1).$$

Taking the supremum, it follows that A is bounded with $\|A\|_{X\to Y} \leq \delta^{-1}$. $\square$

Remark 2.2. Note that we have actually proven that if A is continuous at the origin, then A is bounded, and then A is already continuous everywhere.

It is then natural to consider the set $\mathcal{L}(X \to Y)$ of all bounded linear operators A mapping X to Y, so with $\operatorname{dom}(A) = X$.

Theorem 2.3. *The space of bounded linear operators $\mathcal{L}(X \to Y)$ is a normed vector space with the norm defined by (2.1) or (2.2). If Y is a Banach space, so is $\mathcal{L}(X \to Y)$.*

If $X = Y$, then we just write $\mathcal{L}(X)$ instead of $\mathcal{L}(X \to X)$.

Note that if $A \in \mathcal{L}(X \to Y)$ and $B \in \mathcal{L}(Y \to Z)$, then $BA \in \mathcal{L}(X \to Z)$ and

$$\|BA\|_{X\to Z} \leq \|A\|_{X\to Y}\|B\|_{Y\to Z}.$$

In particular, $\mathcal{L}(X)$ is an algebra. (It even has a unit, the identity mapping given by $Ax := x$.)

Remark 2.4. A linear operator $A\colon X \supseteq \operatorname{dom}(A) \to Y$ with domain $\operatorname{dom}(A) \neq X$ is often called **unbounded** in order to make a distinction to bounded linear operators in $\mathcal{L}(X \to Y)$. It is in principle possible for an unbounded operator to be continuous, i.e., bounded, so this notion is not flawless. We will sometimes also use this notion to point out when we do not assume continuity (and $\operatorname{dom}(A) = X$) for the operator(s) considered.

We next consider a concept closely related to completion for a linear operator A between X and Y. Often, we understand an operator well on a subspace X_0 of X. We can recover the operator on the whole X in the following situation.

Proposition 2.5. *Let X be a normed vector space and let Y be a Banach space. Let X_0 be a dense subspace of X and let A_0 be a linear operator between X and Y with domain X_0. Suppose that*

$$\sup_{x \in X_0,\, \|x\|_X = 1} \|A_0 x\|_Y =: C_0 < \infty. \tag{2.4}$$

Then there exists a unique operator $A \in \mathcal{L}(X \to Y)$ with $\|A\|_{X \to Y} = C_0$ as in (2.4) such that $Ax = A_0 x$ for all $x \in X_0$. For $x \in X \setminus X_0$ and $(x_k) \subseteq X_0$ with $x_k \to x$ in X we have $Ax = \lim_{k\to\infty} A_0 x_k$.

Proof. Let $x \in X \setminus X_0$ and consider a sequence $X_0 \ni x_k \to x$ in X. We need to show that $Ax := \lim_{k\to\infty} A_0 x_k$ exists and that it is well defined, so independent of the choice of the sequence $(x_k) \subseteq X_0$. Since (x_k) is convergent, $(A_0 x_k)$ is a Cauchy sequence:

$$\|A_0 x_k - A_0 x_\ell\|_Y \le C_0 \|x_k - x_\ell\|_X.$$

We have assumed Y to be a Banach space, so $(A_0 x_k)$ is convergent and there is some $y \in Y$ such that $A_0 x_k \to y$ in Y. Now let $(\bar{x}_k) \subseteq X_0$ be some other sequence converging to x in X. By the analogous argument, there exists $\bar{y} \in Y$ such that $A_0 \bar{x}_k \to \bar{y}$ in Y. But

$$\|y - \bar{y}\|_Y = \lim_{k\to\infty} \|A_0 x_k - A_0 \bar{x}_k\|_X \le \lim_{k\to\infty} C_0 \|x_k - \bar{x}_k\|_X = 0,$$

so $y = \bar{y}$. Hence $Ax := y$ is well defined. Finally, A is bounded with norm C_0 because

$$C_0 \le \sup_{0 \ne x \in X} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{0 \ne x \in X} \lim_{k\to\infty} \frac{\|A_0 x_k\|_Y}{\|x_k\|_X} \le C_0$$

for any sequence $X_0 \ni x_k \to x$ in X, and then using (2.2). □

Lastly, we mention a type of bounded linear operator which occurs often in the context of functional analytic methods for PDEs. Given two Banach spaces X and Y, we call a bounded linear *injective* operator $A \in \mathcal{L}(X \to Y)$ an **embedding**. Usually, embeddings are considered in the context of $X \subseteq Y$ and the *identity mapping* $Ax = x$. Recall that the identity is continuous between X and Y if and only if there exists a constant $C \ge 0$ such that

$$\|x\|_Y \le C\,\|x\|_X \qquad (x \in X).$$

If there exists an embedding between two Banach spaces X and Y, then we say that X is **embedded** in Y and write $X \hookrightarrow Y$. For example, $L^\infty(0,1) \hookrightarrow L^p(0,1)$ for all $1 \le p \le \infty$ by the Hölder inequality.

3 Dual space, linear functionals and weak topology

3.1 Dual space and linear functionals

A manifestation of $\mathcal{L}(X \to Y)$ of particular interest is the case $Y = \mathbb{K}$. We in fact give this space its own notation: $X' := \mathcal{L}(X \to \mathbb{K})$, and we call it the **dual space** of X. A linear mapping $\phi\colon X \to \mathbb{R}$ will be called **(linear) functional**. (An element $\phi \in X'$ is thus a *bounded* linear functional.) Since $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$ which is each a Banach space, X' is always a Banach space, independent of X being a Banach space or not. This is already a first indicator that X' is an interesting object.

The most far reaching result about the dual space is the following (version of the) **Hahn-Banach extension theorem**[12]. It says that we can extend a bounded linear functional from a subspace $V \subseteq X$ of X to the whole X with preservation of the norm. However, the extension is *not unique* in general. This is to be compared with Proposition 2.5.

Theorem 3.1 (Hahn-Banach (extension))**.** *Let X be a normed vector space and let $V \subseteq X$ be a subspace. Let $\phi\colon (V, \|\cdot\|_X) \to \mathbb{K}$ be a bounded linear functional. Then ϕ can be extended to a bounded linear functional on X, i.e., there exists $\Phi\colon X \to \mathbb{K}$ such that*

$$\|\Phi\|_{X'} = \sup_{\|x\|_X \le 1} |\Phi(x)| = \sup_{\substack{x \in V \\ \|x\|_X \le 1}} |\phi(x)|.$$

The theorem in itself may, at first glance, not seem very spectacular. But note that the subspace V may be very small, for instance $V := \{\lambda x_0 \colon \lambda \in \mathbb{K}\}$ for some $x_0 \in X$. It is quite astonishing that it is possible to extend a bounded linear functional from such a small subspace to the whole space. (Of couse, the price we pay is that we lose uniqueness of the extension.) This particular one-dimensional subspace exapmple for V already leads to useful corollaries which, in their essence, say that X' is sufficiently rich to include interesting objects. Let X be a normed vector space for the following corollaries.

Corollary 3.2. *For every $x \in X$, there exists a bounded linear functional $\phi \in X'$ such that $\phi(x) = \|x\|_X$ and $\|\phi\|_{X'} = 1$.*

Proof. Let $x \in X$ be fixed and consider $V := \{\lambda x \colon \lambda \in \mathbb{K}\}$. Define $\phi(\lambda x) := \lambda\|x\|$ and use Theorem 3.1. $\square$

[12] Hans Hahn (1879–1934)

Corollary 3.3. *For every pair $x, y \in X$ with $x \neq y$, there exists a bounded linear functional $\phi \in X'$ such that $\phi(x) \neq \phi(y)$.*

Proof. Use Corollary 3.2 to obtain a functional $\phi \in X'$ such that

$$\phi(x) - \phi(y) = \phi(x - y) = \|x - y\| \neq 0. \qquad \square$$

Corollary 3.4. *Let $x \in X$. Then*

$$\|x\|_X = \sup_{\substack{\phi \in X' \\ \|\phi\|_{X'} \leq 1}} |\phi(x)| = \max_{\substack{\phi \in X' \\ \|\phi\|_{X'} \leq 1}} |\phi(x)|.$$

Proof. Clearly,

$$\sup_{\substack{\phi \in X' \\ \|\phi\|_{X'} \leq 1}} |\phi(x)| \leq \sup_{\substack{\phi \in X' \\ \|\phi\|_{X'} \leq 1}} \|\phi\|_{X'} \|x\|_X \leq \|x\|_X.$$

It remains to show that there is $\phi \in X'$ with $\|\phi\| \leq 1$ such that $|\phi(x)| = \|x\|$. But this is exactly the statement of Corollary 3.2. $\square$

It is imperative to compare the assertion of Corollary 3.4—which is a nontrivial *result*—with the *definition*, recall (2.1),

$$\|\phi\|_{X'} = \sup_{\|x\|_X \leq 1} |\phi(x)|.$$

We lastly mention a particular geometric form of the Hahn-Banach theorem. It is about the separation of convex sets. We use the notion of a **hyperplane** $[\phi = \alpha]$ given by

$$[\phi = \alpha] = \left\{ x \in X \colon \operatorname{Re} \phi(x) = \alpha \right\}$$

for a linear functional $\phi \colon X \to \mathbb{K}$ (not necessarily bounded!) and a number $\alpha \in \mathbb{R}$.

Theorem 3.5 (Hahn-Banach (geometric))**.** *Let X be a normed space and let $A, B \subseteq X$ be nonempty and convex with $A \cap B = \emptyset$.*

a) Suppose that A is open. Then there exists a bounded linear functional $\phi \in X'$ and a number $\alpha \in \mathbb{R}$ such that the hyperplane $[\phi = \alpha]$ separates A and B, that is:

$$\operatorname{Re} \phi(a) < \alpha \leq \operatorname{Re} \phi(b) \qquad (a \in A,\ b \in B).$$

b) Suppose that A is compact and B is closed. Then there exists a bounded linear functional $\phi \in X'$ and a number $\alpha \in \mathbb{R}$ such that the hyperplane $[\phi = \alpha]$ strictly *separates A and B, that is, there exists $\varepsilon > 0$ such that:*

$$\operatorname{Re}\phi(a) \leq \alpha - \varepsilon < \alpha \leq \operatorname{Re}\phi(b) \qquad (a \in A,\ b \in B).$$

Note that one can easily construct counterexamples (already in the plane $X = \mathbb{R}^2$) where the assertion of Theorem 3.5 is false if A and B are not convex.

3.2 Weak convergence

For the following, it will often be convenient to use the notion of a **duality pair**. We write

$$\langle \phi, x \rangle_{X',X} = \phi(x) \qquad (\phi \in X',\ x \in X).$$

As usual, if the pair X and X' is clear from context, we skip the index. It is useful to note that

$$X' \times X \ni (\phi, x) \mapsto \langle \phi, x \rangle \in \mathbb{K} \tag{3.1}$$

is a bilinear mapping.

We next introduce a new notion of convergence (and thus topology) on a normed vector space X. This will be **weak convergence**. A sequence $(x_k) \subseteq X$ converges weakly to $x \in X$ if

$$\langle \phi, x_k \rangle \to \langle \phi, x \rangle \quad \text{in } \mathbb{K} \qquad (\phi \in X').$$

We say that x is the **weak limit** of (x_k) and write $x_k \rightharpoonup x$. Note that the weak limit is unique and thus well defined by Corollary 3.3.

The name *weak* is appropriate, because weak convergence is indeed weaker than convergence with respect to the norm on X. In fact, every (norm) convergent sequence (x_k) is also weakly convergent with the same limit since

$$\big|\langle \phi, x_k \rangle - \langle \phi, x \rangle\big| \leq \|\phi\|_{X'} \|x_k - x\|_X \qquad (\phi \in X').$$

A set $U \subseteq X$ is called **weakly closed** if for every weakly convergent sequence, the (weak) limit is again an element of U. We have the following very useful consequence of the Hahn-Banach theorem:

Lemma 3.6. *Let X be a normed vector space and let $U \subseteq X$ be closed and convex. Then U is weakly closed.*

Proof. We use the Hahn-Banach theorem in its geometric form, Theorem 3.5. If $U = \emptyset$, there is nothing to prove. Otherwise let U be closed and convex and let $U \ni x_k \rightharpoonup x$ be a weakly convergent sequence with limit x. Suppose that $x \in X \setminus U$, i.e., that the sets $\{x\}$ and U are disjoint. Clearly, $\{x\}$ is a compact set in X. Hence the Hahn-Banach theorem says that there exists a hyperplane $[\phi = \alpha]$ induced by $\phi \in X'$ and $\alpha \in \mathbb{R}$ which strictly separates $\{x\}$ and U. In particular, there exists $\varepsilon > 0$ such that

$$\mathrm{Re}\langle \phi, x\rangle \le \alpha - \varepsilon < \alpha \le \mathrm{Re}\langle \phi, x_k\rangle \qquad (k \in \mathbb{N}).$$

But then

$$\left|\langle \phi, x_k\rangle - \langle \phi, x\rangle\right| \ge \varepsilon \qquad (k \in \mathbb{N})$$

which is a contradiction to $x_k \rightharpoonup x$. So $x \in U$ and U is weakly closed. □

Remark 3.7. The notion "weak" in weak convergence is not to be taken lightly. Note for instance how in Lemma 3.6, we had to require a strong structural property (convexity) together with "ordinary" closedness to obtain weak closedness. This is because weak closedness is a *stronger* property than regular closedness, because it makes an assertion about a *larger* class of sequences.

Also, one should by no means expect any kind of approximating behavior from weak convergence. We will soon learn to know generic families of sequences (x_k) such that $\|x_k\| = 1$, but $x_k \rightharpoonup 0$.

We now come to the possibly most useful property of the weak topology or convergence. To set the stage, we mention the following result:

Proposition 3.8. *Let X be a normed vector space. Then the closed unit ball $\overline{B(0,1)}$ is compact* if and only if *X is finite-dimensional.*

This proposition is of course bad news. It says that a compact set in an inifinite-dimensional space must be much "smaller" than a ball. In particular, a bounded and closed set in an infinite-dimensional normed vector space is not necessarily compact!

But we have just learned that there are much more weakly convergent sequences than norm convergent ones. So maybe a bounded sequence always has a weakly convergent subsequence? (We will later also see that every weakly convergent sequence is in fact bounded.) This is not true for *every* Banach space, but for many. And, conversely, the lack of this property for a Banach space X makes life much harder working with X.

To fix the notion, we say that a set $U \subseteq X$ is **weakly compact** if every sequence in U admits a weakly convergent subsequence whose limit is again in U. Unfortunately, we also need to introduce a somewhat technical property to make a meaningful statement. It is about the **bidual** $X'' := (X')'$ of X, that is, the space of all bounded linear functionals on X'. This may seem like a very abstract object, but it at least includes a subspace which is isomorphic to X itself: Indeed, given $x \in X$, the mapping $\phi \mapsto \langle \phi, x \rangle$ is a continuous linear functional on X', recall (3.1). Hence

$$J \colon x \mapsto \Big[\phi \mapsto \langle \phi, x \rangle \Big] \quad \in \quad \mathcal{L}(X \to X'').$$

We usually call J the **canonical injection**. The following relation is memorable:

$$\langle Jx, \phi \rangle_{X'',X'} = \langle \phi, x \rangle_{X',X} \qquad (\phi \in X',\ x \in X).$$

The canonical injection is in fact an isometry, so $\|Jx\|_{X''} = \|x\|_X$, and thus injective, but it may fail to be surjective. If it is surjective, then X and X'' are (isometrically) isomorphic via the canonical injection J and this warrants its own name.

Definition 3.9. Let X be a Banach space. We say that X is **reflexive** if the canonical injection $J \in \mathcal{L}(X \to X'')$ is surjective.

Note that X is necessarily a Banach space if it is reflexive, because it is then (isometrically) isomorphic to a dual space $X'' = (X')'$ which is always complete by Theorem 2.3. Now we can finally state the far-reaching result we have been aiming for:

Theorem 3.10. *Let X be a Banach space. Then X is reflexive* if and only if *the closed unit ball $\overline{B(0,1)}$ is weakly compact.*

Corollary 3.11. *Let X be a reflexive Banach space. Then every bounded, closed and convex set is weakly compact.*

It follows immediately that reflexivity is inherited to closed subspaces:

Lemma 3.12. *Let X be a reflexive Banach space and let M be a closed subspace. Then $(M, \|\cdot\|_X)$ is also a reflexive Banach space.*

Reflexivity in general and weak compactness of bounded, closed and convex sets is an extremely useful property and whole approaches to PDEs build upon this property. In fact, it is the backbone of the **Galerkin method**[13] which is already reason enough to introduce it here, although we will not cover this method in this lecture.

[13]Boris Grigorjewitsch Galjorkin (1871–1945)

3.3 Examples

1. In Euclidean space $X = \mathbb{K}^n$, we have $X' = \mathbb{K}^n$. Thus $\mathbb{K}^n$ is clearly reflexive. Also, $x_k \rightharpoonup x$ if and only if $x_k \to x$, so weak convergence and the weak topology are not very interesting here. All this is also true for any finite-dimensional Banach space X since these are isomorphic to $\mathbb{K}^n$ with n being the dimension of X.
2. Let $(\Omega, \mathcal{A}, \mu)$ be a σ-finite measure space. For $1 \le p < \infty$, we have
$$L^p(\Omega,\mu)' = L^q(\Omega,\mu) \qquad \Big(\frac{1}{p}+\frac{1}{q}=1\Big).$$
in the sense of (isometrically) isomorphic. The isomorphism is given by
$$\Psi\colon L^q(\Omega,\mu) \ni f \mapsto \Big[g \mapsto \int_\Omega f(x)g(x)\,\mathrm{d}\mu(x)\Big] \in L^p(\Omega,\mu)',$$
so
$$\langle \Psi f, g\rangle := \int_\Omega f(x)g(x)\,\mathrm{d}x \qquad \big(f \in L^q(\Omega,\mu),\ g \in L^p(\Omega,\mu)\big).$$
We usually identify $f \in L^q(\Omega,\mu)$ with $\Psi f \in L^p(\Omega,\mu)'$ without further notice. The dual space $L^\infty(\Omega,\mu)'$ of $L^\infty(\Omega,\mu)$ is in general—if Ω is not pathological—larger than $L^1(\Omega,\mu)$. In particular, $L^p(\Omega,\mu)$ is reflexive for $1 < p < \infty$ and $L^2(\Omega,\mu)$ is (isometrically) isomorphic to its own dual space. (This is a manifestation of higher powers—that $L^2(\Omega,\mu)$ is a Hilbert space—as we will see later.)
3. Let $E \subset \mathbb{R}^n$ be compact and consider the continuous functions $C(E)$ on E equipped with the supremum norm. Then $C(E)' = \mathcal{M}(E)$, where $\mathcal{M}(E)$ denotes the space of **Radon measures**[14] on E. The duality relation is given by
$$\langle \mu, f\rangle_{C(E)',C(E)} := \int_E f(x)\,\mathrm{d}\mu(x) \qquad (\mu \in \mathcal{M}(E),\ f \in C(E)).$$
Since $\mathcal{M}(E)' \neq C(E)$, the latter is not reflexive. This is a major drawback of the space of continuous functions. (In fact, this manifestation of non-reflexivity is related to the structure of the supremum norm and occurs, as a rule of thumb, whenever a similar norm is in play; as noted above, $L^\infty(\Omega,\mu)$ is also not reflexive!)

4 Linear operators in Banach spaces

4.1 Main theorems about linear operators in Banach spaces

We next give three fundamental theorems about bounded linear operators between Banach spaces and some of their consequences. The proofs of these fundamental the-

[14] Johann Radon (1887–1956)

orems are based on the important **Baire category theorem**[15]:

Theorem 4.1 (Baire Category Theorem). *Let X be a Banach space and let (U_k) be a sequence of closed subsets of X. If $\operatorname{int}(U_k) = \emptyset$ for all $k \in \mathbb{N}$, then*

$$\operatorname{int}\Big(\bigcup_{k=1}^{\infty} U_k\Big) = \emptyset.$$

It is in Theorem 4.1 where (a lot of) the "Banach space magic" happens and indeed, the following results which are proved from Theorem 4.1 are quite fundamental and also somewhat surprising.

The Uniform Boundedness Principle

The first one is the **uniform boundedness principle**, also known as the **Banach-Steinhaus theorem**.[16]

Theorem 4.2 (Uniform Boundedness Principle). *Let X, Y be Banach spaces and let $\mathcal{F} \subseteq \mathcal{L}(X \to Y)$ be a family of bounded linear operators between X and Y. Then* either *there exists a dense set $U \subseteq X$ such that*

$$\sup_{A \in \mathcal{F}} \|Ax\|_Y = \infty \qquad (x \in U),$$

or

$$\sup_{A \in \mathcal{F}} \|A\|_{X \to Y} < \infty.$$

Proof. If

$$C := \sup_{A \in \mathcal{F}} \|A\|_{X \to Y} < \infty,$$

then

$$\sup_{A \in \mathcal{F}} \|Ax\|_Y \leq \sup_{A \in \mathcal{F}} \|A\|_{X \to Y} \|x\|_X \leq C\|x\|_X < \infty \qquad (x \in X).$$

For the converse, suppose that for all dense sets $U \subseteq X$, we have

$$\sup_{A \in \mathcal{F}} \|Ax\|_Y < \infty \qquad (x \in U).$$

[15] René-Louis Baire (1874–1932)

[16] Władysław Hugo Dionizy Steinhaus (1887–1972)

Without loss of generality we can thus assume that $U = X$. Next, consider the sets

$$U_k := \left\{ x \in X \colon \|Ax\|_Y \leq k \text{ for all } A \in \mathcal{F} \right\} \qquad (k \in \mathbb{N}).$$

It is easily seen that each U_k is closed. Moreover, by assumption, $\cup_{k=1}^{\infty} U_k = X$. Since $\operatorname{int}(X) = X \neq \emptyset$, there must exist an index $n \in \mathbb{N}$ such that $\operatorname{int}(U_n) \neq \emptyset$. Otherwise the Baire category theorem (Theorem 4.1) would give a contradiction. Pick $x \in \operatorname{int}(U_n)$. By definition, there exists $r > 0$ such that $B(x,r) \subseteq U_n$. Without loss of generality, we can even assume that $\overline{B(x,r)} \subseteq U_n$. (Replace the original r by $r/2$.) We write $y \in \overline{B(x,r)}$ in the form $y = x + rz$ with $\|z\| \leq 1$ and obtain

$$\|Az\|_Y = \frac{\|A(x+rz) - Ax\|_Y}{r} \leq \frac{2n}{r} \qquad (z \in \overline{B(0,1)},\ A \in \mathcal{F}).$$

The claim follows. □

The uniform boundedness principle is a quite surprising result, because it says that if a family of operators $\mathcal{F}$ is bounded in a *pointwise* (local) sense, then it is already bounded in a *uniform* (global) sense in the operator norm. Moreover, if the family $\mathcal{F}$ is not uniformly bounded in the operator norm, then the pointwise unboundedness must already manifest on a dense set $U \subseteq X$. This is a quite extraordinary property of linear operators between Banach spaces.

Note that the uniform boundedness principle as in Theorem 4.2 does *not* imply that there exists some "limit operator"for $\mathcal{F}$. For this we need an extra assumption.

Lemma 4.3. *Let $(A_k) \subseteq \mathcal{L}(X \to Y)$ be a sequence of linear operators between Banach spaces X and Y. Suppose that for every $x \in X$, the limit $\lim_{k\to\infty} A_k x$ exists in Y. We call this limit Ax. Then $A \in \mathcal{L}(X \to Y)$ with*

$$\|A\|_{X\to Y} \leq \liminf_{k\to\infty} \|A_k\|_{X\to Y} \quad \text{and} \quad \sup_{k\in\mathbb{N}} \|A_k\|_{X\to Y} < \infty.$$

Another consequence of the uniform boundedness principle (Theorem 4.2) is that we can determine boundedness of a set $U \subseteq X$ by looking at U through linear bounded functionals:

Lemma 4.4. *Let X be a Banach space and $U \subseteq X$. Suppose that each of the sets*

$$U'_\phi := \left\{ \langle \phi, x \rangle \colon x \in U \right\} \subseteq \mathbb{K} \qquad (\phi \in X')$$

is bounded. Then U is bounded.

Remark 4.5. Lemma 4.4 is another manifestation of the leitmotif that the application $\langle \phi, x \rangle$ of bounded linear functionals $\phi \in X'$ to $x \in X$ is a means of replacing the concept of *coordinates* in finite-dimensional spaces $\mathbb{K}^n$. Indeed, a set $U \subseteq \mathbb{K}^n$ is bounded if and only if it is bounded in every coordinate. This can be reproduced by Lemma 4.4 by observing that the "k-coordinate extraction" $x = (x_1, \ldots, x_n) \mapsto x_k$ is a continuous linear functional on $\mathbb{K}^n$ for $k = 1, \ldots, n$ induced by the unit vector e_k which is 1 at coordinate k and 0 otherwise.

In this context, recall also Corollary 3.3 and how it tells us that we are always able to distinguish two elements $x \neq y \in X$ by bounded linear functionals.

It follows that weakly convergent sequences are bounded, as announced earlier.

Corollary 4.6. *Let X be a Banach space and let (x_k) be a weakly convergent sequence in X. Then (x_k) is bounded in X.*

The Open Mapping Theorem and the Closed Graph Theorem

We now give two further fundamental theorems which can be inferred from Theorem 4.1. The first one is the **open mapping theorem**:

Theorem 4.7 (Open Mapping Theorem)**.** *Let X, Y be Banach spaces and suppose that $A \in \mathcal{L}(X \to Y)$ is surjective. Then A is **open**, that is, if $U \subseteq X$ is open, then $AU \subseteq Y$ is also open. Equivalently, there exists an $r > 0$ such that*

$$B(0, r) \subseteq AB(0, 1). \tag{4.1}$$

Remark 4.8. Clearly, (4.1) is also a *sufficient* condition for A to be surjective. In this sense, we could formulate Theorem 4.7 also by saying that, for Banach spaces X and Y, a bounded linear operator $A \in \mathcal{L}(X \to Y)$ is surjective *if and only if* it is open.

If A is surjective, then it is trivial that $B(0, r) \subseteq AX = \operatorname{ran}(A) = Y$ for all $r > 0$. Thus, the assertion in Theorem 4.7 is that the "full dimensionality" of $AX = Y$ is already achieved by the image of any ball around 0 in X.

Recall further that a function $f \colon X \to Y$ is continuous *if and only if* for every open set $V \subseteq Y$, the preimage $f^{-1}[V] \subseteq X$ is also open. Thus we obtain the following very useful form of the open mapping theorem, the **bounded inverse theorem**:

Theorem 4.9 (Bounded Inverse Theorem)**.** *Let X, Y be Banach spaces and suppose that $A \in \mathcal{L}(X \to Y)$ is bijective. Then the inverse operator A^{-1} is also linear and continuous, i.e., $A^{-1} \in \mathcal{L}(Y \to X)$.*

Again, this is a most astonishing result. Of course, a bijective function always admits an inverse. But *a priori* this inverse function has no reason at all to be continuous! However, Theorem 4.9 tells us that it is indeed the case for bounded linear operators between Banach spaces.

The other main theorem is the **closed graph theorem**. We need to say what we mean by a closed graph first. Recall that the cartesian product $X \times Y$ is a Banach space if X, Y are Banach spaces and $X \times Y$ is equipped with the norm $\|(x,y)\| := \|x\|_X + \|y\|_Y$. (This will be a convenient choice.)

Definition 4.10 (Graph, closed operator)**.** Let X, Y be Banach spaces and let $A\colon X \supseteq \operatorname{dom}(A) \to Y$ be a linear (unbounded) operator between X and Y with domain $\operatorname{dom}(A)$. The **graph** of A is given by

$$\operatorname{graph}(A) := \Big\{(x, Ax) \in X \times Y \colon x \in \operatorname{dom}(A)\Big\} \subseteq X \times Y.$$

We say that A is **closed** if $\operatorname{graph}(A)$ is closed in $X \times Y$. The **graph norm** $\|\cdot\|_A$ on $\operatorname{dom}(A)$ is given by

$$\|x\|_A := \|x\|_X + \|Ax\|_Y \qquad (x \in \operatorname{dom}(A)).$$

Remark 4.11. By definition, $\operatorname{graph}(A)$ is closed in $X \times Y$ *if and only if* for every convergent sequence $\operatorname{dom}(A) \supseteq x_k \to x \in X$ for which (Ax_k) is also convergent in Y with limit y, we have $x \in \operatorname{dom}(A)$ and $Ax = y$. In this case, $\operatorname{graph}(A)$ is also a Banach space if equipped with the norm of $X \times Y$.

Via Remark 4.11, we easily see that the kernel $\ker(A)$ of a closed operator is closed in X. Moreover, a continuous function between X and Y is always closed. (In fact, this property does not rely on linearity at all.) The converse statement, for which linearity is however fundamental, is the **closed graph theorem**:

Theorem 4.12 (Closed Graph Theorem)**.** *Let X, Y be Banach spaces and let $A\colon X \to Y$ be a closed linear operator. Then A is continuous.*

Proof. Consider the graph norm $\|x\|_A$ on X. Since A is assumed to be closed, it is easy to see that $(X, \|\cdot\|_A)$ is a Banach space. Moreover, of course,

$$\|x\|_X \leq \|x\|_A \qquad (x \in X).$$

Hence the identity mapping is bijective and continuous as a linear operator $(X, \|\cdot\|_A) \to (X, \|\cdot\|_X)$. By the bounded inverse theorem (Theorem 4.9), it is also continuous and linear $(X, \|\cdot\|_X) \to (X, \|\cdot\|_A)$. In particular, there exists a constant $C \geq 0$ such that

$$\|Ax\|_Y \leq \|x\|_X + \|Ax\|_Y = \|x\|_A \leq C\|x\|_X \qquad (x \in X).$$

Hence, if $x_k \to x$ in X, then $Ax_k \to Ax$ in Y and A is continuous. □

Remark 4.13. We have proven the closed graph theorem by appealing to the open mapping theorem via the bounded inverse theorem. It is also possible to proceed inversely, i.e., prove the open mapping theorem via the closed graph theorem.

The concept of a closed operator within the class of unbounded operators can be confusing. We try to shed some more light on it. Suppose that A is an unbounded linear operator between X and Y with domain $\mathrm{dom}(A)$.

- If $\mathrm{dom}(A)$ is in fact a *closed* subspace of X, then $\mathrm{dom}(A)$ is also a Banach space with respect to $\|\cdot\|_X$, and we can consider A as a linear operator A_{dom} between $(\mathrm{dom}(A), \|\cdot\|_X)$ and Y. Per the closed graph theorem (Theorem 4.12), A_{dom} is continuous if and only if it is closed, and then it is also continuous as a linear operator $X \supseteq \mathrm{dom}(A) \to Y$. There is however no unique (continuous) extension to the whole X in general.
- The operator A is *always* continuous from $(\mathrm{dom}(A), \|\cdot\|_A)$ to Y. If A is closed, then $(\mathrm{dom}(A), \|\cdot\|_A)$ is *always* a Banach space. This may seem like conjuring good properties from thin air, but essentially, the graph norm is often not very useful if not used in the context of the closed graph theorem (Theorem 4.12).

In practice there are many closed operators which are not continuous. We just have seen that this can only occur when $\mathrm{dom}(A)$ is *not* a closed subspace of X. (And we do not consider the graph norm on $\mathrm{dom}(A)$.) In fact, many of the most interesting operators—such as the derivative—are closed and densely defined unbounded operators. It is thus of elevated interest to to conceptual reasons to study closed unbounded operators.

4.2 Adjoint operators

If A is a linear operator between $X = \mathbb{K}^n$ and $Y = \mathbb{K}^m$, then we know that it is represented by a matrix $A \in \mathbb{K}^{m\times n}$. The Hermitian (or transpose) $A^H \in \mathbb{K}^{n\times m}$ can then be considered as the representative of a linear operator $\mathbb{K}^m \to \mathbb{K}^n$. We consider the corresponding general construction for linear unbounded operators between Banach spaces X and Y. To fix the ideas, let $A \in \mathcal{L}(X \to Y)$. Then

$$\left[x \mapsto \langle \phi, Ax\rangle\right] \in X' \qquad (\phi \in Y').$$

In particular,

$$\phi \mapsto \left[x \mapsto \langle \phi, Ax\rangle\right]$$

is a continuous linear operator from Y' to X'. We denote this operator by $A' \in \mathcal{L}(Y' \to X')$. It is characterized by the memorable relation

$$\langle A'\phi, x\rangle_{X',X} = \langle \phi, Ax\rangle_{Y',Y} \qquad (\phi \in Y',\ x \in X)$$

from which we see its relation to the Hermitian/transpose of a matrix. (In fact, we note that A^H should be interpreted the representative of a linear mapping $(\mathbb{K}^m)' \to (\mathbb{K}^n)'$.)

If A is an unbounded operator, then the foregoing assertions are not so immediate and we need to be a bit more careful.

Definition 4.14 (Adjoint operator)**.** Let X, Y be normed vector spaces. Let further $A\colon X \supseteq \mathrm{dom}(A) \to Y$ be an unbounded operator between X and Y which is *densely defined*. Then we define the **adjoint operator** A' as an unbounded operator $Y' \supseteq \mathrm{dom}(A') \to X'$ as follows:

$$\mathrm{dom}(A') := \left\{\phi \in Y' \colon \exists x' \in X' \colon \langle \phi, Ax\rangle_{Y',Y} = \langle x', x\rangle_{X',X} \text{ for all } x \in \mathrm{dom}(A)\right\},$$
$$A'\phi := x'.$$

Note that x' in Definition 4.14 is unique, if it exists. This follows from the assumption that $\mathrm{dom}(A)$ is dense in X. Again we have the fundamental relation

$$\langle A'\phi, x\rangle_{X',X} = \langle \phi, Ax\rangle_{Y',Y} \qquad (\phi \in \mathrm{dom}(A'),\ x \in \mathrm{dom}(A)).$$

Some permanence principles for adjoint operators in the situation of Definition 4.14:

- Suppose that $B \in \mathcal{L}(X \to Y)$. Then

$$(A+B)' = A' + B' \quad \text{with} \quad \mathrm{dom}((A+B)') = \mathrm{dom}(A').$$

- If $C \in \mathcal{L}(Y)$, then consider the operator CA on X with $\mathrm{dom}(CA) = \mathrm{dom}(A)$. We have

$$(CA)' = A'C' \quad \text{with} \quad \mathrm{dom}((CA)') := \left\{\phi \in Y' \colon C'\phi \in \mathrm{dom}(A')\right\}.$$

Remark 4.15. The adjoint operator A' as in Definition 4.14 is always closed: Suppose that $\phi_k \to \phi$ in Y' and $A'\phi_k \to x'$ in X'. Then

$$\langle x', x\rangle_{X',X} = \lim_{k\to\infty} \langle A'\phi_k, x\rangle_{X',X} = \lim_{k\to\infty} \langle \phi_k, Ax\rangle_{Y',Y} = \langle \phi, Ax\rangle_{X',X} \quad (x \in \mathrm{dom}(A)),$$

hence $\phi \in \mathrm{dom}(A')$ and $x' = A'\phi$ in X'. But it may happen that $\mathrm{dom}(A')$ is not dense in Y'. We will see an unpleasant consequence of that below in Lemma 4.18. Still, if Y is reflexive, then $\mathrm{dom}(A')$ is dense in Y'.

With an argument similar to the one in Remark 4.15, we can prove the following useful property:

Lemma 4.16. *Let X, Y be normed vector spaces and let $A \in \mathcal{L}(X \to Y)$. If $x_k \rightharpoonup x$ in X, then $Ax_k \rightharpoonup Ax$ in Y. That is, a continuous linear operator is **weakly continuous.***

There is a natural relation between surjectivity of A (or A') and injectivity of A' (or A), respectively. This relies on a certain orthogonality type relation between $\ker(A)$ and $\mathrm{ran}(A')$ and $\ker(A')$ and $\mathrm{ran}(A)$, respectively. To formulate this properly, a definition:

Definition 4.17 (Annihilator)**.** Let X be a normed vector space and let $U \subseteq X$ and $V \subseteq X'$ be linear subspaces. Then

$$U^\perp := \Big\{\phi \in X' \colon \langle \phi, u\rangle = 0 \text{ for all } u \in U\Big\} \subseteq X'$$

and

$$V^\perp := \Big\{x \in X \colon \langle \phi, x\rangle = 0 \text{ for all } \phi \in V\Big\} \subseteq X$$

denotes the **annihilator** of U and V, respectively.

It is useful to imagine $U^\perp$ as *orthogonal* to U, and analogously for $V^\perp$ and V. Note further that in the situation of Definition 4.17, $U^\perp$ and $V^\perp$ are closed subspaces of X and X', respectively. Moreover, $\overline{U}^\perp = U^\perp$ and $\overline{V}^\perp$ as well as

$$(U^\perp)^\perp = \overline{U} \quad \text{and} \quad (V^\perp)^\perp \supseteq \overline{V}, \tag{4.2}$$

with equality in the latter if X is reflexive.

Lemma 4.18. *Let X, Y be normed vector spaces and let A be an unbounded closed linear operator between X and Y with dense domain* $\mathrm{dom}(A)$*. Then we have the following:*

a) $A \in \mathcal{L}(X \to Y)$ if and only if $A' \in \mathcal{L}(Y' \to X')$ *and* $\|A\|_{X\to Y} = \|A'\|_{Y'\to X'}$.

b) If A' is is surjective, then A is injective, and if A is surjective, then A' is injective:

$$\ker(A) = \mathrm{ran}(A')^\perp \quad \textit{and} \quad \ker(A') = \mathrm{ran}(A)^\perp.$$

c) If A' is injective, then A has dense range: $\ker(A')^\perp = \overline{\mathrm{ran}(A)}$.

d) *We have* $\ker(A)^{\perp} \supseteq \overline{\operatorname{ran}(A')}$*. Moreover, if* X *is reflexive:* $\ker(A)^{\perp} = \overline{\operatorname{ran}(A')}$*, so if* A *is injective, then* A' *has dense range.*

Proof. a) By Corollary 3.4:

$$\|A\|_{X\to Y} = \sup_{\|x\|\leq 1} \|Ax\|_Y = \sup_{\substack{\|x\|_X\leq 1\\ \|\phi\|_{Y'}\leq 1}} \langle \phi, Ax\rangle = \sup_{\substack{\|x\|_X\leq 1\\ \|\phi\|_{Y'}\leq 1}} \langle A'\phi, x\rangle$$
$$= \sup_{\|\phi\|_{Y'}\leq 1} \|A'\phi\|_{X'} = \|A'\|_{Y'\to X'}.$$

b) We prove $\ker(A') = \operatorname{ran}(A)^{\perp}$. We do not need that A is closed for this identity, but it is crucial that A is densely defined:

$$\begin{aligned}
\phi \in \ker(A') &\iff \langle A'\phi, x\rangle = 0\ (x \in X)\\
&\iff \langle A'\phi, x\rangle = 0\ (x \in \operatorname{dom}(A))\\
&\iff \langle \phi, Ax\rangle = 0\ (x \in \operatorname{dom}(A)) \iff \phi \in \operatorname{ran}(A)^{\perp}.
\end{aligned}$$

Now $\ker A = \operatorname{ran}(A')^{\perp}$. We try to do the same:

$$\begin{aligned}
x \in \ker(A) &\iff \langle \phi, Ax\rangle = 0\ (\phi \in Y')\\
&\implies \langle \phi, Ax\rangle = 0\ (\phi \in \operatorname{dom}(A'))\\
&\iff \langle A'\phi, x\rangle = 0\ (\phi \in \operatorname{dom}(A')) \iff x \in \operatorname{ran}(A')^{\perp}.
\end{aligned}$$

Thus we only find $\ker(A) \subseteq \operatorname{ran}(A')^{\perp}$. (The culprit here is $\operatorname{dom}(A')$ not necessarily being dense in Y'!) But so far we have not used that A is closed, so this will be the way to go. We want to show that $\operatorname{ran}(A')^{\perp} \subseteq \ker(A)$. Assume that there is $\bar{x} \in \operatorname{ran}(A')^{\perp}$ such that $\bar{x} \notin \ker(A)$, that is, either $\bar{x} \notin \operatorname{dom}(A)$ or $A\bar{x} \neq 0$. These conditions can be summarized to $(\bar{x}, 0) \notin \operatorname{graph}(A)$. By assumption, $\operatorname{graph}(A)$ is *closed*. Hence, the geometric Hahn-Banach theorem (Theorem 3.5) strikes, and there exist $(\phi, \psi) \in (X \times Y)' = X' \times Y'$ and $\alpha \in \mathbb{R}$ as well as $\varepsilon > 0$ such that

$$\operatorname{Re}\langle \phi, \bar{x}\rangle \leq \alpha - \varepsilon < \alpha \leq \operatorname{Re}\langle \phi, x\rangle + \operatorname{Re}\langle \psi, Ax\rangle \qquad (x \in \operatorname{dom}(A)).$$

Suppose that there exists $x \in \operatorname{dom}(A)$ such that the right-hand side in the foregoing inequality is nonzero. Then $\lambda x \in \operatorname{dom}(A)$ for all $\lambda \in \mathbb{R}$ and

$$\lambda = \frac{\alpha - 1}{\operatorname{Re}\langle \phi, x\rangle + \operatorname{Re}\langle \psi, Ax\rangle} \implies \operatorname{Re}\langle \phi, \lambda x\rangle + \operatorname{Re}\langle \psi, \lambda Ax\rangle = \alpha - 1,$$

which is a contradiction. Hence

$$\operatorname{Re}\langle \phi, x\rangle = -\operatorname{Re}\langle \psi, Ax\rangle \qquad (x \in \operatorname{dom}(A)).$$

It follows that also

$$\operatorname{Im}\langle \phi, x\rangle = -\operatorname{Re}\langle \phi, ix\rangle = \operatorname{Re}\langle \psi, A(ix)\rangle = -\operatorname{Im}\langle \psi, Ax\rangle \qquad (x \in \operatorname{dom}(A)),$$

so

$$\langle \phi, x\rangle = -\langle \psi, Ax\rangle \qquad (x \in \operatorname{dom}(A)).$$

In particular, $-\psi \in \operatorname{dom}(A')$ and $A'(-\psi) = \phi$. But then $\phi \in \operatorname{ran}(A')$, so that due to $\bar{x} \in \operatorname{ran}(A')^{\perp}$,

$$\operatorname{Re}\langle \phi, \bar{x}\rangle = 0 = \operatorname{Re}\langle \phi, x\rangle + \operatorname{Re}\langle \psi, Ax\rangle \qquad (x \in \operatorname{dom}(A)).$$

This is again a contradiction. Hence $(x, 0) \in \operatorname{graph}(A)$ and so $\operatorname{ran}(A')^{\perp} \subseteq \ker(A)$.

From the proof we also see that we can dispose of the assumption that A is closed if we assume that Y is reflexive instead because then the first attempt of the proof goes through; recall Remark 4.15.

c) Follows from (4.2).

d) Follows from (4.2) and the following remark.

□

Finally, an operator is continuously invertible if and only its adjoint is:

Proposition 4.19. *Let X, Y be normed vector spaces and let A be an unbounded closed linear operator between X and Y with dense domain $\operatorname{dom}(A)$. Then A is bijective and $A^{-1} \in \mathcal{L}(Y \to X)$ if and only if A' is bijective and $(A')^{-1} \in \mathcal{L}(Y' \to X')$. In either case,*

$$(A^{-1})' = (A')^{-1}.$$

4.3 Compact operators

We next consider operators inducing *compactness* in the following sense:

Definition 4.20 (Compact operator)**.** Let X, Y be normed vector spaces and let $A \in \mathcal{L}(X \to Y)$. We say that A is **compact** if $\overline{AB_X(0,1)}$ is compact in Y.

If $A\colon X \to Y$ is a linear operator and $\overline{AB_X(0,1)}$ is compact, then $AB_X(0,1)$ is relatively compact and thus bounded. In particular, A is bounded. This shows that the requirement for A to be bounded in Definition 4.20 is in fact necessary.

Remark 4.21. It is easy to see that the condition in Definition 4.20 is equivalent to the following: Let $(x_k) \subseteq X$ be a bounded sequence. Then $(Ax_k) \subseteq Y$ admits a convergent subsequence.

We will learn much more about compact operators later when we come to spectral theory. For now, we just collect some fundamental facts.

Proposition 4.22. *Let X, Y be normed vector spaces and let $A \in \mathcal{L}(X \to Y)$.*

a) If $\operatorname{ran}(A)$ is finite-dimensional, then A is compact.

b) Let (A_k) be a sequence of compact linear operators $X \to Y$. Suppose that Y is complete and that $A_k \to A$ in $\mathcal{L}(X \to Y)$. Then A is also compact.

Proof. a) If $\operatorname{ran}(A)$ is finite-dimensional, then by assumption, $\overline{AB(0,1)}$ is a closed and bounded subset of the finite dimensional normed vector space $(\operatorname{ran}(A), \|\cdot\|_Y)$ and thus compact (Proposition 3.8).

b) We argue using precompactness. Here it is important that Y is a Banach space. Let $\varepsilon > 0$ and choose a number $K \in \mathbb{N}$ such that $\|A - A_K\|_{X\to Y} < \varepsilon/2$. By the assumption on the sequence (A_k), we know that $A_K B(0,1)$ is precompact in Y. Hence, there exists $N \in \mathbb{N}$ and $y_1, \ldots, y_N \in A_K B(0,1)$ such that

$$A_K B(0,1) \subseteq \bigcup_{k=1}^{N} B(y_k, \varepsilon/2).$$

Now let $x \in X$ with $\|x\| \leq 1$. Then $\|Ax - A_K x\| < \varepsilon/2$ by choice of K. On the other hand, there is $i \in \{1, \ldots, N\}$ such that $\|A_K x - y_i\| < \varepsilon/2$. So, by the triangle inequality,

$$AB(0,1) \subseteq \bigcup_{k=1}^{N} B(y_k, \varepsilon).$$

and A is compact. □

Note that if X is finite-dimensional, then $\operatorname{ran}(A)$ also is. This is a particular case in Proposition 4.22. The next theorem shows that A and its adjoint A' are of the same quality.

Theorem 4.23 (Schauder[17]). *Let X, Y be Banach spaces and let $A \in \mathcal{L}(X \to Y)$. Then A is compact* if and only if *A' is compact $Y' \to X'$.*

[16]Juliusz Schauder (1899–1943)

In fact, compact operators translate weak convergence into strong convergence. This is an extremely useful property to have although we will not see very much of it in action in this lecture.

Lemma 4.24. *Let X, Y be Banach spaces and let $A \in \mathcal{L}(X \to Y)$ be compact. Suppose that $x_k \rightharpoonup x$ in X. Then $Ax_k \to Ax$ in Y.*

Proof. In the proof we use the following elementary fact:

Nitpicker lemma: A sequence $(x_k) \subseteq X$ converges to x *if and only if* every subsequence admits a subsequence which converges to x.

Now let $x_k \rightharpoonup x$ in X. Since X is a Banach space, the sequence (x_k) is bounded by Corollary 4.6. Thus, by compactness of A, the sequence (Ax_k) is relatively compact in Y. In particular, there exists $y \in Y$ and a subsequence (Ax_{k_ℓ}) such that $Ax_{k_\ell} \to y$ in Y. We want to show that $y = Ax$. To this end,

$$\langle \phi, y \rangle = \lim_{\ell\to\infty} \langle \phi, Ax_{k_\ell} \rangle = \lim_{\ell\to\infty} \langle A'\phi, x_{k_\ell} \rangle = \langle A'\phi, x \rangle = \langle \phi, Ax \rangle \qquad (\phi \in Y')$$

and $y = Ax$ follows. But so far we have only proven that a *subsequence* of (Ax_k) converges.

Now the Nitpicker lemma does its magic: Simply start the above proof with *any subsequence* (x_{k_m}) of (x_k) instead of (x_k) itself. Then we obtain that (Ax_{k_m}) of (Ax_k) admits a subsequence which converges to Ax. Hence (Ax_k) converges to Ax. $\square$

Remark 4.25. In fact, if X is reflexive and Y is a Banach space, then the converse to Lemma 4.24 is also true: The operator $A \in \mathcal{L}(X \to Y)$ is compact *if and only if* for every weakly convergent sequence $x_k \rightharpoonup x$ in X, we have $Ax_k \to Ax$ in Y. This follows easily via Corollary 3.11.

Fredholm alternative

We next consider a quite fundamental and interesting property of a class of operators for which compactness yields most useful properties. Recall that a continuous linear operator T on a finite-dimensional normed vector space X is injective ($\ker T = \{0\}$) if and only if it is surjective ($\operatorname{ran} T = X$). But in infinite-dimensional spaces this is not true at all, consider e.g. $X = L^2(\mathbb{R}_+)$ and T the right shift of length, say, 1:

$$(Tf)(x) := \begin{cases} f(x-1) & \text{if } x > 1, \\ 0 & \text{if } x \in (0,1). \end{cases}$$

Then $Tf = 0$ if and only $f = 0$, each in $L^2(\mathbb{R}_+)$, but of course T is not surjective. The following fundamental **Fredholm alternative** theorem says that this phenomenon does not occur for operators of the form $T = I - A$ with $A \in \mathcal{L}(X)$ *compact*:

Theorem 4.26 (Fredholm alternative). *Let X be a Banach space and let $A \in \mathcal{L}(X)$ be compact. Then we have the following properties:*

a) $\ker(I - A)$ *is finite-dimensional,*

b) $\operatorname{ran}(I - A)$ *is closed with* $\operatorname{ran}(I - A) = \ker(I - A')^{\perp}$,

c) $\ker(I - A) = \{0\}$ if and only if $\operatorname{ran}(I - A) = X$,

d) $\dim \ker(I - A) = \dim \ker(I - A')$.

Before we start with the proof, let us mention the dichotomy which give Theorem 4.26 its name. It says that *either* ...

- for every $f \in X$, the equation $x - Ax = f$ has a *unique* solution $x \in X$, *or*
- the homogeneous equation $x - Ax = 0$ admits N linearly independent solutions; in this case, the inhomogeneous equation $x - Ax = f$ admits a solution *if and only if* $f \in \ker(I - A')^{\perp}$. The latter in turn means that there are N basis vectors $\phi_1, \ldots, \phi_N$ of $\ker(I - A')$ such that $\langle \phi_k, f \rangle = 0$ for $k = 1, \ldots, N$, so N *orthogonality relations*.

This is the **Fredholm alternative**. We also note that if A is compact, then $\lambda^{-1}A$ is compact for every $\lambda \in \mathbb{K}$. Thus the Fredholm alternative in particular applies to the equations

$$\lambda x - Ax = f \qquad \Longleftrightarrow \qquad x - \lambda^{-1}Ax = \lambda^{-1}f.$$

Proof. a) Let $U = \ker(I - A)$. This is a closed subspace of the Banach space X and thus itself a Banach space. Let $B_U(0,1)$ be the unit ball in U. Then $AB_X(0,1) \supseteq AB_U(0,1) = B_U(0,1)$. But $\overline{AB_U(0,1)}$ is compact by assumption and thus so it $\overline{B_U(0,1)}$. By Proposition 3.8, this implies that U is finite-dimensional.

b) Let $(f_k) \subseteq \operatorname{ran}(I - A)$ such that $x_k - Ax_k = f_k \to f$ in X. We need to show that there is $x \in X$ such that $f = x - Ax$, so $f \in \operatorname{ran}(I - A)$. Set $d_k = \operatorname{dist}(x_k, \ker(I - A)) = \inf_{y \in \ker(I-A)} \|x_k - y\|$. We already know that $\ker(I - A)$ is finite dimensional, so there exist $y_k \in \ker(I - A)$ such that $d_k = \|x_k - y_k\|$. (See the exercises.) Then

$$f_k = (x_k - y_k) - A(x_k - y_k). \tag{4.3}$$

Note that

$$\operatorname{dist}(x_k - y_k, \ker(I - A)) = \operatorname{dist}(x_k, \ker(I - A)) = d_k = \|x_k - y_k\|. \tag{4.4}$$

We would like to extract convergent subsequences from $A(x_k - y_k)$ in order to have convergence in each term in (4.3). By compactness of A, this we are able to do if whenever the sequence $(x_k - y_k)$ is bounded.

Suppose it was not. Then there is a subsequence $(x_{k_\ell} - y_{k_\ell})$ such that $\|x_{k_\ell} - y_{k_\ell}\| \to \infty$. Let

$$w_{k_\ell} := \frac{x_{k_\ell} - y_{k_\ell}}{\|x_{k_\ell} - y_{k_\ell}\|}$$

be the normalized sequence. We have, recall (4.4), $\operatorname{dist}(w_{k_\ell}, \ker(I - A)) = 1$. On the other hand, again by compactness of A, there exists yet another subsequence (which we do not relabel) such that $Aw_{k_\ell} \to z \in X$. So, via (4.3) and convergence of f_k,

$$w_{k_\ell} - Aw_{k_\ell} = \frac{f_{k_\ell}}{\|x_{k_\ell} - y_{k_\ell}\|} \quad \longrightarrow \quad 0.$$

Hence $w_{k_\ell} \to z \in \ker(I - A)$. In particular, $\operatorname{dist}(w_{k_\ell}, \ker(I - A)) \to 0$. This is a contradiction. So $(x_k - y_k)$ is a bounded sequence.

Accordingly, the sequence $(A(x_k - y_k))$ admits a (new) subsequence such that $(A(x_{k_\ell} - y_{k_\ell})) \to g \in X$. Then, by (4.3), $x_{k_\ell} - y_{k_\ell} \to f + g$. Hence $A(f + g) = g$ and

$$f = f + g - g = f + g - A(f + g),$$

so $f \in \operatorname{ran}(I - A)$. Hence $I - A$ has closed range. Via Lemma 4.18 is then follows that

$$\operatorname{ran}(I - A) = \overline{\operatorname{ran}(I - A)} = \ker((I - A)')^\perp = \ker(I - A')^\perp.$$

c) Suppose first that $\ker(I - A) = \{0\}$. Suppose that $\operatorname{ran}(I - A) \neq X$. We already know that $\operatorname{ran}(I - A)$ is closed, so it is a Banach space itself. Let $X_0 = \operatorname{ran}(I - A)$. Suppose that $x = y - Ay \in X_0$ for some $y \in X$. Then $Ax = Ay - AAy$, so A maps X_0 into itself. In particular, if A_0 denotes A restricted to X_0, then $A_0 \in \mathcal{L}(X_0 \to X_0)$ is compact and $A_1 := \operatorname{ran}(I - A_0)$, a subspace of X_0, is again closed. Moreover, $X_1 \subsetneq X_0$: Pick $y \in X \setminus X_0$. Then $x := y - Ay \in X_0$. Suppose that $x \in X_1$. Then there is $y_0 \in X_0$ such that $x = y_0 - Ay_0$. But then $y - y_0 \in \ker(I - A) = \{0\}$ (assumption!), so $y = y_0$. But this is not possible since $y \in X \setminus X_0$ and $y_0 \in X_0$. So $X_1 \subsetneq X_0$.

Now set iteratively $X_k := \operatorname{ran}(I - A_k)$ to obtain a sequence of (strictly) decreasing closed subspaces of X. By the Riesz lemma (exercises), we find a sequence $(x_k) \subset X$ such that $x_k \in X_k$ with $\|x_k\| = 1$ for $k \in \mathbb{N}$ and $\operatorname{dist}(x_k, X_{k+1}) \geq 1/2$. Since A is compact, (Ax_k) must have a convergent subsequence. We lead this to a contradiction: Consider

$$Ax_k - Ax_\ell = \underbrace{-(x_k - Ax_k)}_{\in X_{k+1}} + \underbrace{(x_\ell - Ax_\ell)}_{\in X_{\ell+1}} + \underbrace{x_k}_{\in X_k} - \underbrace{x_\ell}_{\in X_\ell}.$$

Without loss of generality, let $k > \ell$. Then $X_{k+1} \subsetneq X_k \subseteq X_{\ell+1} \subsetneq X_\ell$ and

$$-(x_k - Ax_k) + (x_\ell - Ax_\ell) + x_k \in X_{\ell+1}$$

In particular,

$$\|Ax_k - Ax_\ell\| \geq \operatorname{dist}(x_\ell, X_{\ell+1}) \geq \frac{1}{2},$$

so (Ax_k) cannot have a convergent subsequence. This is the contradiction. So after all $X = X_0 = \operatorname{ran}(I - A)$.

Conversely, suppose that $\operatorname{ran}(I - A) = X$. From Lemma 4.18, we have that $\ker(I - A') = \operatorname{ran}(I - A)^\perp = \{0\}$. The Schauder theorem (Theorem 4.23) says that A' is also a compact operator. Hence from $\ker(I - A') = \{0\}$ it follows that $\operatorname{ran}(I - A') = X'$. But then Lemma 4.18 strikes again and gives $\ker(I - A) = \operatorname{ran}(I - A')^\perp = \{0\}$.

d) We skip this proof. □

The Fredholm alternative as in Theorem 4.26 is a particular case of the more general theory of **Fredholm operators**: If X, Y are Banach spaces and $T \in \mathcal{L}(X \to Y)$, then we say that T is a Fredholm operator (or **Noether operator**), if the following properties hold true:

- $\ker(T)$ is finite-dimensional, and
- $\operatorname{ran}(T)$ is closed and has finite co-dimension, that is: there is a subspace $Y_0 \subseteq Y$ with $\dim Y_0 < \infty$ such that $Y = Y_0 + \operatorname{ran}(T)$.

(In fact, one can show that $\operatorname{ran}(T)$ is always closed if it has finite co-dimension, so the latter is the actual requirement.) The **index** of a Fredholm operator is then

$$\operatorname{ind}(T) = \dim\ker(T) - \operatorname{co\,dim}\operatorname{ran}(T).$$

The Fredholm alternative (Theorem 4.26) says that $T = I - A$ with $A \in \mathcal{L}(X)$ compact is a Fredholm operator of index 0.

4.4 Examples

1. Consider $C([0,1])$ with the supremum norm and its subspace

$$C^1([0,1]) := \left\{ f \in C([0,1]) : f \text{ differentiable on } (0,1),\ f' \in C([0,1]) \right\}$$

of uniformly continuous and differentiable functions on $[0,1]$ with uniformly continuous derivative. Let $D : f \mapsto f'$ be the derivative operator with domain $\operatorname{dom}(D) = C^1([0,1])$. Clearly, D is not bounded on $C([0,1])$. (Construct a sequence $(f_k) \subset C([0,1])$ with $\|f_k\|_\infty = 1$ but $\|f_k'\|_\infty = k$.) But it is closed: If $f_k \to f$ and $f_k' \to g$, each in $C([0,1])$, then $f' = g$.

2. We consider again the derivative operator $D\colon f \mapsto f'$, this time however on the subspace
$$C_0([0,1]) := \Big\{ f \in C([0,1]) \colon f(0) = 0 \Big\}$$
with domain $\operatorname{dom}(D) = C_0^1([0,1])$ which is defined analogously to $C_0([0,1])$. The derivative D is clearly still a closed operator in this context. But it has an inverse which is continuous: Define $\mathcal{I}\colon C_0([0,1]) \to C_0([0,1])$ by
$$(\mathcal{I}f)(t) := \int_0^t f(\tau)\,\mathrm{d}\tau \qquad (t \in [0,1]).$$
Clearly, $\mathcal{I}f \in C_0^1([0,1]) \subseteq C_0([0,1])$ and $D\mathcal{I}f = f$ for every $f \in C_0([0,1])$. On the other hand, for every $f \in \operatorname{dom}(D) = C_0^1([0,1])$,
$$(\mathcal{I}Df)(t) = \int_0^t (Df)(\tau)\,\mathrm{d}\tau = \int_0^t f'(\tau)\,\mathrm{d}\tau = f(t).$$
Hence $\mathcal{I}$ is the inverse to D. Moreover,
$$\big|(\mathcal{I}f)(t)\big| \leq t\|f\|_\infty \leq \|f\|_\infty \qquad (f \in C_0([0,1]),\ t \in [0,1])$$
so $\mathcal{I} \in \mathcal{L}(C_0([0,1]))$.

3. Let X, Y be Banach spaces and suppose that $X \hookrightarrow Y$ densely with the embedding A. Then $A' \in \mathcal{L}(Y' \to X')$. Since A is an embedding, it is injective by definition. Hence, by Lemma 4.18, $\operatorname{ran}(A')^\perp = \{0\}$ from which it follows that $\operatorname{ran}(A')$ must be dense in X'. Moreover, $Y = \overline{\operatorname{ran}(A)} = \ker(A')^\perp$, so $\ker(A') = \{0\}$. Hence A' is also an embedding and realizes $Y' \hookrightarrow X'$ densely.

4. Consider $L^p(\mathbb{R})$ for $1 \leq p < \infty$ and the *shift operator* τ_h defined by $(\tau_h f)(x) := f(x+h)$ for some $h \in \mathbb{R}$. It is easy to see that τ_h is an isometry on $L^p(\mathbb{R})$. We compute its adjoint: Every $\phi \in L^p(\mathbb{R})'$ is given by $\phi = \Psi g$ for a function $g \in L^q(\mathbb{R})$ with $\frac{1}{p} + \frac{1}{q} = 1$, recall Section 3.3. Of course we use this representation:
$$\langle \phi, \tau_h f\rangle = \langle \Psi g, \tau_h f\rangle = \int_{\mathbb{R}} g(x) f(x+h)\,\mathrm{d}x = \int_{\mathbb{R}} g(x-h) f(x) = \langle \Psi\tau_{-h} g, f\rangle.$$
Hence $\tau_h' = \Psi\tau_{-h}\Psi^{-1}$. In particular, if $\phi \in L^p(\mathbb{R})'$ corresponds to $g \in L^q(\mathbb{R})$, then $\tau_h'\phi$ corresponds to $\tau_{-h}g$.

5. Let $E \subseteq \mathbb{R}^n$ be compact. For $0 < \alpha \leq 1$, consider the (α-)*Hölder spaces*[18] $C^{0,\alpha}(E)$ of (α-)Hölder continuous functions on E defined by
$$C^{0,\alpha}(E) := \Big\{ f \in C(E) \colon |f|_{C^{0,\alpha}(E)} < \infty \Big\}, \qquad \|f\|_{C^{0,\alpha}(E)} := \|f\|_\infty + |f|_{C^{0,\alpha}(E)},$$

[18] Otto Hölder (1859–1937)

with the (α-)Hölder-seminorm

$$|f|_{C^{0,\alpha}(E)} := \sup_{\substack{x,y\in E\\ x\neq y}} \frac{|f(x)-f(x)|}{|x-y|^\alpha}.$$

It is trivial that $C^{0,\alpha}(E) \hookrightarrow C(E)$ for all $0 < \alpha \le 1$. In fact, the embedding is *compact*, that is, if (f_k) is a sequence of Hölder continuous functions on E which is bounded in the $C^{0,\alpha}(E)$-norm, then this sequence has a subsequence which converges uniformly to some $f \in C(E)$. This follows from the **Arzelà-Ascoli theorem**[19]:

Theorem (Arzelà-Ascoli). Let $E \subseteq \mathbb{R}^n$ be compact. Then a set $U \subseteq C(E)$ is relatively compact in $C(E)$ *if and only* U is bounded and **equicontinuous**: For every $\varepsilon > 0$ there exists $\delta > 0$ such that *for all* $f \in U$:

$$|x-y| \le \delta \quad \Longrightarrow \quad |f(x)-f(y)| \le \varepsilon \qquad (x,y \in E).$$

5 Hilbert spaces

We now turn to spaces which admit an inner product. This will allow for all sorts of geometric intuition, in particular related to orthogonality. Let H be a vector space over $\mathbb{K}$. An **inner product** on H is a positive definite sesquilinear form $(\cdot,\cdot)_H\colon H\times H \to \mathbb{K}$, that is, for $x,y,z \in H$ and $\alpha \in \mathbb{K}$ we have

a) $(x,y) = \overline{(y,x)_H}$,

b) $(x+\alpha y, z)_H = (x,z)_H + \lambda(y,z)_H$,

c) $(x,x)_H \in \mathbb{R}_+$ with $(x,x)_H = 0 \iff x = 0$.

It is easy to see that the first properties also imply

$$(x, y+\alpha z)_H = (x,y)_H + \overline{\alpha}(x,z)_H.$$

If $\mathbb{K} = \mathbb{R}$, then we need not bother with complex conjugation and the inner product is merely a usual positive definite bilinear form on H. An inner product defines an **inner product norm** $|\cdot|_H$ on H by

$$|x|_H := \sqrt{(x,x)_H} \qquad (x \in H). \tag{5.1}$$

[19] Cesare Arzelà (1847–1912), Giulio Ascoli (1843–1896)

As usual, if the inner product space H is clear from context, then we do not refer to the index H and just write $(\cdot,\cdot)$ and $|\cdot|$. The fact that (5.1) indeed defines a norm on H follows from the fundamental **Cauchy-Schwarz inequality**[20]

$$\big|(x,y)\big| \leq |x|\,|y| \qquad (x,y \in H)$$

which in turn implies the **Minkowski inequality**[21]

$$|x+y| \leq |x| + |y| \qquad (x,y \in H).$$

An inner product space norm (5.1) satisfies the **parallelogram law**

$$|x+y|^2 + |x-y|^2 = 2\Big(|x|^2 + |y|^2\Big) \qquad (x,y \in H). \tag{5.2}$$

In fact, the parallelogram law *characterizes* normed vector spaces whose norm is given by an inner product; see the exercises.

Finally, if the inner product space H is complete with respect to the norm (5.1), then we say that H is a **Hilbert space**[22].

First and most important examples of Hilbert spaces and non-Hilbert spaces are the following:

a) Of course, $H = \mathbb{C}^n$ is a Hilbert space with the inner product $(x,y) := \sum_{i=1}^n x_i \overline{y_i}$.

b) We already mentioned earlier that if $(\Omega, \mathcal{A}, \mu)$ is a σ-finite measure space, then $L^2(\Omega,\mu)$ is a Hilbert space with the inner product

$$(f,g)_{L^2(\Omega,\mu)} := \int_\Omega f(x)\overline{g(x)}\,\mathrm{d}x \qquad (f,g \in L^2(\Omega,\mu)).$$

In particular, the standard $L^2(\Omega,\mu)$ norm coincides with the inner product norm induced by $(\cdot,\cdot)_{L^2(\Omega,\mu)}$. But if the measure space is not degenerated, then $L^p(\Omega,\mu)$ is not an inner product space for $p \in [1,\infty] \setminus \{2\}$.

c) The vector space of continuous functions on $[0,1]$ equipped with the inner product $(\cdot,\cdot)_{L^2(0,1)}$ is an inner product space. But it is not complete and thus not a Hilbert space.

A most fundamental property of Hilbert spaces is that for every closed convex set, we have a well defined **projection**:

[20] Baron Augustin-Louis Cauchy (1789–1857), Karl Hermann Amandus Schwarz (1843–1921)
[21] Hermann Minkowski (1864–1909)
[22] David Hilbert (1862–1943)

Theorem 5.1 (Projection). *Let $\emptyset \neq K \subseteq H$ be a closed and convex subset of the Hilbert space H. Then for every $f \in H$ there exists a unique $y \in K$ such that*

$$|f-y| = \min_{x\in K} |f-x| = \operatorname{dist}(f,K). \tag{5.3}$$

Moreover, the vector y satisfying (5.3) is characterized by

$$y \in K, \qquad \operatorname{Re}(f-y, y-x) \geq 0 \quad (x \in K). \tag{5.4}$$

Proof. Let $(x_k) \subseteq K$ be a sequence such that $|f-x_k| \to \inf_{x\in K} |f-x| = \operatorname{dist}(f,K)$. We show that (x_k) is a Cauchy sequence using the parallelogram law (5.2). Indeed, we have

$$\left|f - \frac{x_k+x_\ell}{2}\right|^2 + \left|\frac{x_k-x_\ell}{2}\right|^2 = \frac{|f-x_k|^2+|f-x_\ell|^2}{2} \qquad (k,\ell \in \mathbb{N}).$$

But $(x_k - x_\ell)/2 \in K$, so $|f-(x_k-x_\ell)/2| \geq \operatorname{dist}(f,K)$. Hence

$$\left|\frac{x_k-x_\ell}{2}\right|^2 \leq \frac{|f-x_k|^2+|f-x_\ell|^2}{2} - \operatorname{dist}(f,K)^2 \qquad (k,\ell \in \mathbb{N}).$$

This implies that (x_k) is a Cauchy sequence and, since we have assumed H to be a Hilbert space, convergent with limit y, so $x_k \to y$. Since K was assumed to be closed, $y \in K$, and by continuity of the norm, $|f-y| = \operatorname{dist}(f,K)$.

We next prove that (5.3) and (5.4) are equivalent. Suppose first that $y \in K$ satisfies (5.3) and let $x \in K$. Then $z = (1-t)y + tx \in K$ for every $t \in (0,1)$ by convexity of K, so

$$|f-y| \leq |f-z| = \big|(f-y)+t(x-y)\big|.$$

Squaring and expanding the right-hand side, we obtain

$$|f-y|^2 \leq \big|(f-y)+t(x-y)\big|^2 = |f-y|^2 + 2t\operatorname{Re}(f-y,x-y) + t^2|x-y|^2$$

and so

$$2\operatorname{Re}(f-y,x-y) \geq t|x-y|^2 \quad \xrightarrow{t\searrow 0} \quad 0.$$

This is (5.4). Conversely, suppose that (5.4) holds true. Then

$$|f-x|^2 - |f-y|^2 = |f-y+y-x|^2 - |f-y|^2 = 2\operatorname{Re}(f-y,y-x) + |x-y|^2 \geq 0$$

for all $x \in K$, and this implies (5.3).

It remains to show that $y \in K$ characterized by (5.3) and (5.4) is unique. Suppose that $y_1, y_2 \in K$ satisfy (5.4). Then

$$|y_1-y_2|^2 = (y_1-y_2, y_1-y_2) = (f-y_2, y_1-y_2) + (y_1-f, y_1-y_2)$$

Taking the real part and using (5.4) shows that $|y_1-y_2|^2 \leq 0$, hence $y_1 = y_2$. $\square$

Definition 5.2 (Projection)**.** In the setting of Theorem 5.1, we call $y \in K$ characterized by (5.3) and (5.4) the **projection** of f onto K and write $y = P_K f$.

We will assume that H is a Hilbert space from now on without further mentioning.

Of course, even in the general case with K convex and closed, the projection act linearly on K. (It is the identity there.) But in general, the projection P_K is *nonlinear*. However, it is still nonexpansive, so Lipschitz continuous with Lipschitz constant at most 1:

Lemma 5.3. *Let $\emptyset \neq K \subseteq H$ be closed and convex. Then we have*

$$\left|P_K f - P_K g\right| \leq |f - g| \qquad (f, g \in H).$$

Proof. If $P_K f = P_K g$, then there is nothing to prove. So let $P_K f \neq P_K g$. By characterization (5.4), we have

$$\operatorname{Re}\bigl(f - P_K f, P_K f - P_K g\bigr) \geq 0 \quad \text{and} \quad \operatorname{Re}\bigl(g - P_K g, P_K g - P_K f\bigr) \geq 0.$$

Adding these inequalities, we obtain

$$\operatorname{Re}\bigl(f - g - \bigl(P_K f - P_K g\bigr), P_K f - P_K g\bigr) \geq 0,$$

so by the Cauchy-Schwarz inequality

$$\left|P_K f - P_K g\right|^2 \leq \operatorname{Re}\bigl(f - g, P_K f - P_K g\bigr) \leq |f - g||P_K f - P_K g|.$$

Since $P_K f \neq P_K g$, this gives the claim. ☐

Corollary 5.4. *Suppose that $M \subseteq H$ is a closed linear subspace. Then $y = P_M f$ is characterized by*

$$y \in M, \qquad (f - y, x) = 0 \quad (x \in M). \tag{5.5}$$

*In particular, $f - P_M f$ is orthogonal to M. Moreover, P_M is in fact a bounded linear operator on H which we call the **orthogonal projection**.*

Proof. Let $y = P_M f$. Then, by (5.4), $\operatorname{Re}(f - y, y - x) \geq 0$ for all $x \in M$. In particular, since M is a linear space,

$$\operatorname{Re}(f - y, y - \alpha x) \geq 0 \quad \iff \quad \operatorname{Re}(f - y, y) \geq \operatorname{Re}(f - y, \alpha x) \qquad (x \in M,\ \alpha \in \mathbb{R}).$$

Assume there is $x \in M$ such that $(f - y, x) \neq 0$. Then we get a contradiction in the foregoing inequality by making the right-hand side arbitrarily large by appropriate choice of α.

Now suppose conversely that (5.5) is true. Since $y \in M$,

$$0 = -(f - y, x) = -(f - y, x) + (f - y, y) = (f - y, y - x) \qquad (x \in M).$$

This implies in particular (5.4), so $y = P_M f$.

The characterization (5.5) shows that P_M is linear, and we have already seen in Lemma 5.3 that it is (Lipschitz) continuous, so in particular bounded. □

5.1 Dual space of a Hilbert space

From the Cauchy-Schwarz inequality, we see that we can immediately write down bounded linear functionals on a Hilbert space: For every fixed $y \in H$, the mapping $x \mapsto (x, y)$ is in H'. The following remarkable Riesz-Fréchet[23] representation theorem says that in fact *every* bounded linear functional on a Hilbert space is of that form:

Theorem 5.5 (Riesz-Fréchet representation theorem). *Let $\phi \in H'$ be a bounded linear functional. Then there exists a unique $y \in H$ such that*

$$\langle \phi, x \rangle_{H',H} = (x, y)_H \qquad (x \in H).$$

Moreover, $\|\phi\|_{H'} = |y|_H$.

Proof. Set $M = \ker(\phi)$. Since ϕ is continuous, $M \subseteq H$ is a closed subspace. If $M = H$, then ϕ is the zero functional, and the assertion is trivial with the choice $y = 0$. So we can assume $M \subsetneq H$. Thus there exists $u \in H \setminus M$. Define

$$v := \frac{u - P_M u}{|u - P_M u|}.$$

Then $|v| = 1$ and $(v, x) = 0$ for all $x \in M$ due to the characterization (5.5). Now let $z \in H$. Set

$$w := z - \frac{\langle \phi, z \rangle}{\langle \phi, v \rangle} v.$$

We have

$$\langle \phi, w \rangle = \langle \phi, z \rangle - \frac{\langle \phi, z \rangle}{\langle \phi, v \rangle} \langle \phi, v \rangle = 0,$$

so $w \in \ker(\phi) = M$. Thus

$$0 = (w, v) = (z, v) - \frac{\langle \phi, z \rangle}{\langle \phi, v \rangle} (v, v)$$

[23] Frigyes Riesz (1880–1958), Maurice René Fréchet (1878–1973)

and we obtain
$$\left(z, \overline{\langle \phi, v\rangle} v\right) = \langle \phi, z\rangle.$$
Hence $y = \overline{\langle \phi, v\rangle} v$ satisfies the assertion.

Finally, from the Cauchy-Schwarz inequality we have $\|\phi\| \leq |y|$. On the other hand, $|y| = |\langle \phi, v\rangle| \leq \|\phi\|$ since $|v| = 1$. So $|y| = \|\phi\|$. □

The Riesz-Frechet Theorem 5.5 implies the following fundamental structural properties of Hilbert spaces:

Proposition 5.6 (Fundamental structural properties of Hilbert spaces)**.**

a) *Every Hilbert space H is (antilinear)* isometrically isomorphic *to its dual space H'. The (antilinear) isometry $R\colon H' \to H$ is defined by*
$$\langle \phi, x\rangle_{H',H} = \left(x, R\phi\right)_H \qquad (\phi \in H',\ x \in H).$$

b) *Every Hilbert space H is reflexive.*

Proof. The first assertion follows immediately from Theorem 5.5. (One also sees that the constructed y depends on ϕ in an antilinear manner in the proof of Theorem 5.5.) For reflexivity of H, note that
$$\left(\phi, \psi\right)_{H'} := \overline{\left(R\phi, R\psi\right)_H} \qquad (\phi, \psi \in H')$$
defines an inner product on H' with
$$\|\phi\|_{H'} = |R\phi|_H = \sqrt{\left(R\phi, R\phi\right)_H} = |\phi|_{H'},$$
so H' is also a Hilbert space and the dual norm agrees with the inner product norm. Accordingly, it also admits an (antilinear) Riesz-Fréchet-isometry $S\colon H'' \to H'$ such that
$$\left\langle \Psi, \phi\right\rangle_{H'',H'} = \left(\phi, S\Psi\right)_{H'} \qquad (\Psi \in H'',\ \phi \in H').$$
Thus
$$\begin{aligned}\left\langle \Psi, \phi\right\rangle_{H'',H'} = \left(\phi, S\Psi\right)_{H'} &= \overline{\left(R\phi, RS\Psi\right)_H} \\ &= \left(RS\Psi, R\phi\right)_H = \left\langle \phi, RS\Psi\right\rangle_{H',H} \qquad (\Psi \in H'',\ \phi \in H').\end{aligned}$$

Hence, by definition of the canonical injection $J\colon H \to H''$, recall Definition 3.9, $JRS\Psi = \Psi$; in particular, J is surjective and H is reflexive. □

Weak convergence

The Riesz-Fréchet theorem (Theorem 5.5) tells us that the functionals $\phi \in H'$ are fully represented by the inner product via R as in Proposition 5.6. Hence, it is not a surprise that weak convergence in Hilbert spaces is completely characterized by convergence in the inner product: Let $(x_k) \subseteq H$ and $x_k \rightharpoonup x$. Then

$$(x_k, y) = \langle R^{-1}y, x_k \rangle \quad \longrightarrow \quad \langle R^{-1}y, x \rangle = (x, y) \qquad (y \in H).$$

Conversely, if $(y, x_k) \to (y, x)$ for every $y \in H$, then

$$\langle \phi, x_k \rangle = (x_k, R\phi) \quad \longrightarrow \quad (x, R\phi) = \langle \phi, x \rangle \qquad (\phi \in H').$$

Hence $x_k \rightharpoonup x$ in H if and only if $(x_k, y) \to (x, y)$ for every $y \in H$.

The orthogonal complement

We already have investigated the annihilator $U^\perp$ of a subspace $U \subseteq H$, recall Definition 4.17. With the knowledge of Theorem 5.5, we can give the *orthogonality* intuition a literal foundation: We set

$$U_\perp := \Big\{ x \in H \colon (u, x) = 0 \text{ for all } u \in U \Big\}.$$

Then $U_\perp = RU^\perp$. Since now $U_\perp \subseteq H$, we can compare it with U, which is particularly effective when U is closed:

Proposition 5.7 (Complemented subspace)**.** *Let $U \subseteq H$ be a closed subspace. Then $U \cap U_\perp = \{0\}$ and U is **complemented**, written $H = U \oplus U_\perp$: for every $x \in H$ there exists a unique decomposition $x = u + v$ with $u \in U$ and $v \in U_\perp$. In fact, we have $|x|^2 = |u|^2 + |v|^2$ and $u = P_U x$ and $v = (I - P_U)x$.*

Proof. The intersection $U \cap U_\perp = \{0\}$ is obvious. Regarding the decomposition, we have $P_U x \in U$ and $(I - P_U)x \in U_\perp$ by Corollary 5.4, uniqueness was stated in Theorem 5.1. The norm equality follows from the *Pythagoras theorem*, since $(u, v) = 0$:

$$|x|^2 = |u + v|^2 = |u|^2 + |v|^2. \qquad \square$$

Note that Proposition 5.7 also implies that $P_{U_\perp} = I - P_U$.

The Hilbert space adjoint

The adjoint operator $A' \in \mathcal{L}(H')$ of $A \in \mathcal{L}(H)$ was given by

$$\langle \phi, Ax \rangle = \langle A'\phi, x \rangle \qquad (\phi \in H',\ x \in H).$$

Now that we know that H and H' are isometrically isomorphic in a canonical way, we are interested in the Hilbert space adjoint $A^\star$ which should do

$$(Ax, y) = (x, A^\star y) \qquad (x, y \in H).$$

In order to see how this operator should look like, we calculate:

$$(Ax, y) = (Ax, RR^{-1}y) = \langle R^{-1}y, Ax \rangle = \langle A'R^{-1}y, x \rangle = (x, RA'R^{-1}y) \qquad (x, y \in H),$$

so $A^\star := RA'R^{-1}$. The general definition is then as follows:

Definition 5.8 (Hilbert space adjoint)**.** Let $A\colon H \supseteq \operatorname{dom}(A) \to H$ be a densely defined unbounded linear operator in H. Then the **Hilbert space adjoint** $A^\star$ is a densely defined unbounded linear operator defined by

$$\begin{aligned} \operatorname{dom}(A^\star) &:= \Big\{ x \in H\colon R^{-1}x \in \operatorname{dom}(A') \Big\}, \\ A^\star x &:= RA'R^{-1}x. \end{aligned}$$

Note that density of $\operatorname{dom}(A^\star)$ in H follows from density of $\operatorname{dom}(A')$ in H', see Remark 4.15. The Hilbert space adjoint $A^\star$ is the unique operator which satisfies the fundamental relation

$$(Ax, y) = (x, A^\star y) \qquad (x \in \operatorname{dom}(A),\ y \in \operatorname{dom}(A^\star)).$$

In this sense it fully generalizes the Hermitian of a matrix. But, in this context, a **word of caution**: There holds

$$(\lambda A)' = \lambda A', \quad \text{but} \quad (\lambda A)^\star = \overline{\lambda} A^\star \qquad (\lambda \in \mathbb{C}),$$

due to the second component of an inner product being antilinear.

5.2 The Lax-Milgram lemma

We next come to one of the most important general results for elliptic PDEs, the Lax-Milgram lemma[24]. This can (but need not be) formulated in the language of forms. Recall that a form $a\colon H \times H \to \mathbb{K}$ is is sesquilinear if $u \mapsto a(u, v)$ is linear for every $v \in H$, and $v \mapsto \overline{a(u, v)}$ is linear for every $u \in H$.

[24] Peter Lax (1926–), Arthur Norton Milgram (1912–1961)

Definition 5.9 (Continuous coercive form)**.** Let $a\colon H \times H \to K$ be a sesquilinear form on H. Then:

a) We say that a is **continuous** if there is a constant $C \geq 0$ such that

$$|a(u,v)| \leq C\,|u||v| \qquad (u,v \in H), \tag{5.6}$$

b) We say that a is **coercive** if there is a constant $\alpha > 0$ such that

$$\operatorname{Re} a(u,u) \geq \alpha |u|^2 \qquad (u \in H). \tag{5.7}$$

The coercivity property (5.7) is very interesting and will prove to be extremely useful. Of course, the inner product $a(u,v) := (u,v)_H$ on a Hilbert space is a continuous coercive sesquilinear form as the Cauchy-Schwarz inequality implies. In this sense the following fundamental lemma is a generalization of the Riesz-Fréchet Theorem 5.5:

Theorem 5.10 (Lax-Milgram lemma)**.** *Let $a\colon H \times H \to \mathbb{K}$ be a continuous coercive sesquilinear form. Then, for every $\phi \in H'$, there is a unique $u \in H$ such that*

$$a(u,v) = \overline{\langle \phi, v\rangle}_{H',H} \qquad (v \in H). \tag{5.8}$$

Moreover $|u| \leq \alpha^{-1}\|\phi\|_{H'}$, where $\alpha > 0$ is a coercivity constant of a as in (5.7).

Before we give the proof, let us note a reformulation of Theorem 5.10: If $a\colon H \times H \to \mathbb{K}$ is a continuous sesquilinear form, then for every $u \in H$, the mapping $v \mapsto \overline{a(u,v)}$ defines a continuous linear functional ψ on H. By Theorem 5.5,

$$\overline{a(u,v)} = \langle \psi, v\rangle = (v, R\psi) \qquad (v \in H).$$

We define $A \in \mathcal{L}(H)$ by $Au := R\psi$. Then

$$\overline{a(u,v)} = (v, Au), \quad \text{so} \quad a(u,v) = (Au, v) \qquad (v \in H).$$

So, Theorem 5.10 becomes the question whether for every $\phi \in H'$ there exists $u \in H$ with $|u| \leq \alpha^{-1}\|\phi\|_{H'}$ such that

$$Au = R\phi.$$

That is, whether A is invertible and $\|A^{-1}\|_{H\to H} \leq \alpha^{-1}$ under the assumption that A is **positive definite**:

$$\operatorname{Re}(Au, u) \geq \alpha |u|^2. \tag{5.9}$$

We prove exactly this now.

Proof. The classical way to prove that A as constructed before is bijective, is as follows:

i) A is injective,

ii) $\operatorname{ran}(A)$ is dense,

iii) $\operatorname{ran}(A)$ is closed.

So:

i) This is an immediate consequence of coercivity (5.7). Let $Au = 0$. Then

$$0 = \operatorname{Re}(Au, u) = \operatorname{Re}\overline{(u, Au)} = \operatorname{Re} a(u, u) \geq \alpha|u|^2,$$

so $u = 0$.

ii) We show that $\operatorname{ran}(A)_\perp = \{0\}$. Let $u \in \operatorname{ran}(A)_\perp$. Then $(u, Av) = 0$ for all $v \in H$. In particular, $(u, Au) = 0$ and the foregoing coercivity argument applies.

iii) We estimate

$$\alpha|u|^2 \leq \operatorname{Re} a(u, u) = \operatorname{Re}(Au, u) \leq \big|(Au, u)\big| \leq |Au||u| \qquad (u \in H),$$

so

$$\alpha|u| \leq |Au| \qquad (u \in H). \tag{5.10}$$

Now consider a sequence $(u_k) \subseteq H$ such that $(Au_k) \to v$ in H. We need to show that $v \in \operatorname{ran}(A)$. The foregoing inequality (5.10) shows that (u_k) is a Cauchy sequence. Since H is a Hilbert space, it is convergent with limit $u \in H$. But then $v = Au$ by continuity of A, so $\operatorname{ran}(A)$ is closed.

So A is bijective. The bounded inverse theorem (Theorem 4.9) already implies that A^{-1} is continuous. The claimed norm bound follows again from (5.10):

$$\alpha\big|A^{-1}v\big| \leq \big|AA^{-1}v\big| = |v| \qquad (v \in H). \qquad \square$$

It is imperative to note how the coercivity property (5.7) is used in *every* part of the foregoing proof, whereas the construction of the operator A associated to a relied on the linearity and continuity properties of a.

5.3 Orthonormal basis

We give another structural, approximation-type result for Hilbert spaces. To this end, we define what we mean by an orthonormal basis.

Definition 5.11. Let $(e_k) \subseteq H$ be a sequence in H. We say that (e_k) is an **orthonormal basis** (ONB) of H (or a **Hilbert basis**, or a **complete orthonormal system**, or just a **basis**), if the following conditions are satisfied:

a) We have $|e_k| = 1$ for every $k \in \mathbb{N}$ and $(e_k, e_\ell) = 0$ for $k \neq \ell$.

b) The linear span of (e_k), so the set of all *finite* linear combinations of (e_k), is dense in H.

Then we have the following result:

Proposition 5.12. *Let $(e_k) \subset H$ be an orthonormal basis of H. Then, for every $u \in H$:*

$$u = \sum_{k=1}^{\infty} (u, e_k) e_k \quad \text{and} \quad |u|^2 = \sum_{k=1}^{\infty} \left|(u, e_k)\right|^2.$$

The infinite sums in Proposition 5.12 need be understood as convergence of the partial sums sequence (S_n) in H, where

$$S_n := \sum_{k=1}^{n} (u, e_k) e_k \qquad (n \in \mathbb{N})$$

in case of the first one.

Proposition 5.12 has several interesting consequences some of which will be explored in the exercises. For example, it implies that an orthonormal basis (e_n) is weakly convergent to zero, although it lies on the unit sphere!

Finally, we mention a structural result for the class of **separable** Hilbert spaces.

Proposition 5.13. *Every separable Hilbert space admits an ONB.*

Proof. Let $(x_k) \subseteq H$ be a countable dense subset of H. Define $V_n := \operatorname{span}(x_1, \dots, x_n)$ for $n \in \mathbb{N}$. Then $V_n \subseteq V_{n+1}$ and $\bigcup_{n=1}^{\infty} V_n$ is dense in H. Choose a unit vector $v_1 \in V_1$. If $V_2 \neq V_1$, then pick any vector $u_2 \in V_2 \setminus V_1$. Set

$$v_2 = \frac{u_2 - (v_1, u_2) v_1}{\|u_2 - (v_1, u_2) v_1\|}.$$

Then $|v_2| = 1$ and $(v_1, v_2) = 0$. Moreover, $V_2 = \operatorname{span}(v_1, v_2)$. Proceed iteratively to construct a sequence (v_k) which is then by construction an ONB. $\square$

We will learn to know a canonical way of constructing an orthonormal basis in Hilbert spaces soon, by looking at the spectrum of normal or self-adjoint operators.

Examples

1. Of course, the standard basis in $\mathbb{K}^n$ is an orthonormal basis in this Hilbert space for which we have $(x, e_k) = x_k$, the k-th coordinate, so
$$x = \sum_{k=1}^{n} x_k e_k.$$

2. More interestingly: The two most classical orthonormal basis in the Lebesgue Hilbert space $L^2(0, \pi)$ are given by
$$e_k(x) := \sqrt{\frac{2}{\pi}} \sin(kx) \qquad (k \geq 1)$$
and
$$e_k(x) := \sqrt{\frac{2}{\pi}} \cos(kx) \qquad (k \geq 0).$$
The verification that $(e_k, e_\ell) = \delta_{k=\ell}$, so 1 if $k = \ell$ and 0 otherwise, is an easy exercise for integration by parts. From this example, one immediately obtains an orthonormal basis family in $L^2(a, b)$ for $a, b \in \mathbb{R}$ with $a < b$ by rescaling.

6 Spectral theory

In Linear Algebra we learn that the behavior of a linear mapping $A\colon \mathbb{C}^n \to \mathbb{C}^n$ is fully described by its spectral values (eigenvalues) and the associated (generalized) eigenvectors. The particularly nice case is if A is even normal, so $AA^H = A^H A$. Then the basis of $\mathbb{C}^n$ made of eigenvectors of A can be chosen to be orthonormal and we have $A = UDU^H$, where D is a diagonal matrix made of eigenvalues of A and U is the matrix whose columns are given by the eigenvectors of A. Even more, if $A = A^H$, then all eigenvalues of A are real and so is D.

We will try and obtain similar results also in the infinite-dimensional case. Of course, things are slightly more involved, so we need some definitions first. We stay with the case of bounded linear operators for now, since we ultimately will consider compact operators. So, for the following, X is a fixed Banach space over $\mathbb{K}$ and we consider an unbounded *closed* operator $A\colon X \supseteq \operatorname{dom}(A) \to X$.

Definition 6.1 (Spectrum, resolvent set, eigenvalue)**.** The **resolvent set** $\rho(A)$ is defined by
$$\rho(A) := \Big\{\lambda \in \mathbb{K}\colon A - \lambda \text{ is bijective } \operatorname{dom}(A) \to X\Big\}.$$
The **spectrum** $\sigma(A)$ is given by the complement $\mathbb{K} \setminus \rho(A)$. We say that $\lambda \in \sigma(A)$

is an **eigenvalue** if $\ker(A-\lambda) \neq \{0\}$ and the set of eigenvalues is called the **point spectrum** $\sigma_P(A) \subseteq \sigma(A)$. For $\lambda \in \rho(A)$, we call the operator $R(\lambda, A) = (A - \lambda)^{-1} \in \mathcal{L}(X)$ the **resolvent** (operator).

Note that continuity of the resolvent follows automatically from the closed graph theorem (Theorem 4.12) because $A \in \mathcal{L}((\operatorname{dom}(A), \|\cdot\|_A) \to X)$. In contrast to the finite-dimensional case, the spectrum $\sigma(A)$ does not consist of eigenvectors only. (Consider again the right shift which we already had as a counterexample for an injective but not surjective linear operator right before the Fredholm alternative Theorem 4.26.)

Here are a few useful simple equations following straight from the definition:

$$AR(\lambda, A) = \lambda R(\lambda, A) + I \qquad (\lambda \in \rho(A)),$$

and the **resolvent identity**

$$R(\lambda, A) - R(\mu, A) = (\lambda - \mu) R(\lambda, A) R(\mu, A) \qquad (\lambda, \mu \in \rho(A)).$$

From this part we can already see that $\lambda \mapsto R(\lambda, A)$ will be differentiable. (Let $\lambda = \mu + h$ and send $h \to 0$.) In fact, much more is true as we can see in the proof of the next important result. This is the part where it matters whether $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$.

Proposition 6.2. *The spectrum $\sigma(A)$ is a* closed *set in $\mathbb{K}$. If $A \in \mathcal{L}(X)$, then $\sigma(A)$ is* compact *in $\mathbb{K}$ with*

$$\sigma(A) \subseteq \left\{ \lambda \in \mathbb{K} \colon |\lambda| \leq \|A\| \right\}.$$

If $A \in \mathcal{L}(X)$ and $\mathbb{K} = \mathbb{C}$, then the spectrum $\sigma(A)$ is nonempty.

Proof. We show that the resolvent set $\rho(A)$ is open. Let $\mu \in \rho(A)$ and $\lambda \in \mathbb{K}$ and write

$$A - \lambda = A - \mu + \mu - \lambda = \left[I + (\mu - \lambda) R(\mu, A) \right] (A - \mu). \tag{6.1}$$

Since $\mu \in \rho(A)$, the $A - \lambda$ is a bijective operator if $I + (\mu - \lambda) R(\mu, A)$ is bijective. But this is the case if $|\mu - \lambda| < 1/\|R(\mu, A)\|$ (exercise). Hence $\rho(A)$ is open and $\sigma(A)$ is closed.

Now let $A \in \mathcal{L}(X)$ and let $\lambda \in \mathbb{K}$ with $|\lambda| > \|A\|$. Then the infinite sum

$$-\sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}} = (A - \lambda)^{-1}$$

exists in $\mathcal{L}(X)$ (exercise) and we have $\lambda \in \rho(A)$. Hence, if $\mu \in \sigma(A)$, then we must have $|\mu| \leq \|A\|$. So $\sigma(A)$ is a bounded and closed subset of $\mathbb{K}$ and thus compact.

Finally, if $\mathbb{K} = \mathbb{C}$, then the spectrum is in fact nonempty. This follows from a clever application of Liouville's theorem from complex analysis but we will not prove it here. $\square$

Remark 6.3. In the proof of Proposition 6.2, from (6.1) we find that $\lambda \mapsto R(\lambda, A)$ is in fact (locally) *analytic*: if $\lambda, \mu \in \rho(A)$ such that $|\mu - \lambda| < 1/\|R(\mu, A)\|$, then again using the infinite sum expansion, we obtain

$$R(\lambda, A) = R(\mu, A)\big[I + (\mu - \lambda)R(\mu, A)\big]^{-1} = R(\mu, A)\sum_{k=0}^{\infty}(\lambda - \mu)^k R(\mu, A)^k$$

In particular,

$$\frac{\mathrm{d}^k}{\mathrm{d}\lambda^k}R(\lambda, A) = (-1)^k k!(\lambda - \mu)^k R(\mu, A)^k \qquad (k \in \mathbb{N}).$$

Examples

Consider $X = C([0,1])$ with the supremum norm. We pick up the derivative operator from Section 4.4.

1. Let $D\colon f \mapsto f'$ be the derivative operator in X with domain $\operatorname{dom}(D) = C^1([0,1])$. Then $\sigma(D) = \mathbb{K}$. Indeed, note that for every $\lambda \in \mathbb{K}$

$$(D - \lambda)f = 0 \qquad \Longleftrightarrow \qquad f' = \lambda f \qquad (f \in \operatorname{dom}(D)),$$

   so every solution to the ODE $f'(t) = \lambda f(t)$ is an eigenvector of D for the eigenvalue λ. But clearly $f_\lambda(t) := e^{\lambda t}$ is such a solution for every $\lambda \in \mathbb{K}$ and $f_\lambda \in \operatorname{dom}(D) = C^1([0,1])$. Hence $\sigma(D) = \sigma_P(D) = \mathbb{K}$.

2. Now consider D on X with domain $C_0([0,1])$. We had already seen that D was bijective with $D^{-1} = \mathcal{I}$ given by

$$(\mathcal{I}f) = \int_0^t f(\tau)\,\mathrm{d}\tau \qquad (t \in [0,1]).$$

   Now let $\lambda \in \mathbb{K}$ and let $\mathcal{I}_\lambda\colon C([0,1]) \to C_0([0,1])$ be defined by

$$(\mathcal{I}_\lambda f)(t) := \int_0^t e^{\lambda(t-\tau)} f(\tau)\,\mathrm{d}\tau \qquad (t \in [0,1]).$$

   Then

$$D\mathcal{I}_\lambda f = f + \lambda\mathcal{I}_\lambda, \quad \text{and} \quad \mathcal{I}_\lambda D f = f + \lambda\mathcal{I}_\lambda f.$$

   In particular, $\mathcal{I}_\lambda = R(\lambda, D)$. Hence $\sigma(D) = \emptyset$ in this case.

   Note how, as in the previous example, we again have the ODE $f'(t) = \lambda f(t)$ as a necessary condition for f to be an eigenvector for the eigenvalue $\lambda \in \mathbb{K}$, but this time with an initial condition $f(0) = 0$, which enforces $f = 0$!

3. Now let $\emptyset \neq \Omega \subseteq \mathbb{K}$ be any compact set and consider the multiplication operator $\mathcal{M}\colon C(\Omega) \to C(\Omega)$ given by
$$(\mathcal{M}f)(x) := x \cdot f(x) \qquad (x \in \Omega).$$
Then one can show that $\sigma(\mathcal{M}) = \Omega$. One can construct a similar example also for nonempty and closed sets Ω. In particular, the spectrum can be *any set* in $\mathbb{K}$.

Residual spectrum and the adjoint

Let $\lambda \in \sigma(A) \setminus \sigma_P(A)$. Then $A - \lambda$ is not surjective. Hence $\operatorname{ran}(A-\lambda)$ is not dense or not closed. (Of course both can happen at the same time.) It will thus be useful to have another sub-classification of $\sigma(A) \setminus \sigma_P(A)$:

Definition 6.4 (Residual spectrum)**.** Let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a closed unbounded operator. Then the **residual spectrum** $\sigma_R(A)$ is given by
$$\sigma_R(A) := \Big\{\lambda \in \sigma(A)\colon \operatorname{ran}(A-\lambda) \text{ is not dense in } X\Big\}.$$

The significance of the residual spectrum is the following characterization as the point spectrum of the adjoint:

Lemma 6.5. *Let A be a densely defined closed unbounded operator between Banach spaces X and Y. Then we have*
$$\sigma_R(A) = \sigma_P(A') \quad \text{and} \quad \sigma_P(A) \subseteq \sigma_R(A')$$
If X is reflexive, then also
$$\sigma_P(A) = \sigma_R(A').$$

Proof. We rely on Lemma 4.18 for all assertions. If $\lambda \in \sigma_R(A)$, then
$$X \neq \overline{\operatorname{ran}(A-\lambda)} = \ker\big((A-\lambda)'\big) = \ker(A'-\lambda)^{\perp}.$$
But then $\ker(A'-\lambda) \neq \{0\}$ and $\lambda \in \sigma_P(A')$, so $\sigma_R(A) \subseteq \sigma_P(A')$.

Conversely, if $\lambda \in \sigma_P(A')$, then
$$\{0\} \neq \ker(A'-\lambda) = \operatorname{ran}(A-\lambda)^{\perp},$$
hence $\operatorname{ran}(A-\lambda)$ cannot be dense in X. Thus $\sigma_P(A') \subseteq \sigma_R(A)$. The same argument works for $\sigma_P(A) \subseteq \sigma_R(A')$.

If X is reflexive, then $\ker(A-\lambda)^{\perp} = \overline{\operatorname{ran}(A'-\lambda)}$ and the first argument applies again. $\square$

6.1 Spectrum of compact operators

We now concentrate on the spectrum $\sigma(A)$ of a *compact* operator $A \in \mathcal{L}(X)$. It will turn out that this is of a particularly interesting structure as the following main theorem, sometimes called **Riesz-Schauder theory**:

Theorem 6.6 (Riesz-Schauder). *Let $A \in \mathcal{L}(X)$ be a compact operator and suppose that X is infinite-dimensional. Then we have the following:*

a) $0 \in \sigma(A)$,

b) $\sigma(A) \setminus \{0\} = \sigma_P(A) \setminus \{0\}$,

c) if $(\lambda_k) \subseteq \sigma(A)$ is a sequence of distinct *scalars such that $\lambda_k \to \lambda$, then $\lambda = 0$.*

Proof. a) If $0 \notin \sigma(A)$, then A is bijective; in particular, $I = AA^{-1}$ is compact because it is a composition of a continuous and a compact operator. But then $\overline{IB(0,1)} = \overline{B(0,1)}$ is compact and this is a contradiction to X being infinite-dimensional by Proposition 3.8.

b) Let $\lambda \in \sigma(A) \setminus \{0\}$. Suppose that λ was not eigenvalue, so $\lambda \notin \sigma_P(A)$. Then $\ker(A - \lambda) = \{0\}$. On the other hand, $A - \lambda = -\lambda(I - \lambda^{-1}A)$ and we have the Fredholm alternative for $I - \lambda^{-1}A$. Thus, by Theorem 4.26, $\operatorname{ran}(A - \lambda) = X$. But then $A - \lambda$ is bijective and $\lambda \in \rho(A)$, and this is a contradiction.

c) Without loss of generality, we can assume that $\lambda_k \neq 0$ for all $k \in \mathbb{N}$. Then $(\lambda_k) \subseteq \sigma_P(A)$. Pick eigenvectors (e_k) of (λ_k) and set $E_k := \operatorname{span}\{e_1, e_2, \dots, e_k\}$. We claim that $E_k \subsetneq E_{k+1}$ for each $k \in \mathbb{N}$.

We prove this by induction on k. Suppose that $(e_1, \dots, e_k)$ are linearly independent. Assume further that $e_{k+1} \in E_k$, so $E_{k+1} = E_k$. Then there are coefficients $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ such that $e_{k+1} = \sum_{i=1}^k \alpha_i e_i$. By construction,

$$Ae_{k+1} = \sum_{i=1}^{k} \alpha_i \lambda_i e_i \quad \text{and} \quad Ae_{k+1} = \lambda_{k+1} e_{k+1} = \sum_{i=1}^{k} \alpha_i \lambda_{k+1} e_i.$$

Hence $\alpha_i(\lambda_i - \lambda_{k+1}) = 0$, which implies $\alpha_i = 0$ for $i = 1, \dots, k$. This is a contradiction, so $E_k \subsetneq E_{k+1}$ for all $k \in \mathbb{N}$.

Now the Riesz lemma (exercises) strikes again: There is a sequence (x_k) such that $x_k \in E_k$ with $\|x_k\| = 1$ and $\operatorname{dist}(x_k, E_{k-1})$ for all $k \geq 2$. Then we have

$$E_{\ell-1} \subsetneq E_\ell \subseteq E_{k-1} \subsetneq E_k \qquad (2 \leq \ell < k).$$

Further $(A-\lambda_k)E_k \subseteq E_{k-1}$ and so

$$\left\|\frac{Ax_k}{\lambda_k} - \frac{Ax_\ell}{\lambda_\ell}\right\| = \left\|\frac{Ax_k - \lambda_k x_k}{\lambda_k} - \frac{Ax_\ell - \lambda_\ell u_\ell}{\lambda_\ell} + x_k - x_\ell\right\|$$
$$\geq \operatorname{dist}(x_k, E_{k-1}) \geq \frac{1}{2}$$

But (Ax_k) has a convergent subsequence because A is compact. Hence, if $\lambda_k \to \lambda \neq 0$, then the left-hand side converges to 0, which is a contradiction. $\square$

Remark 6.7. A common reformulation of the third assertion in Theorem 6.6 is that $A \in \mathcal{L}(X)$ is compact, then one of the following cases is true:

- $\sigma(A) = \{0\}$,
- $\sigma(A) \setminus \{0\}$ is a finite set,
- $\sigma(A) \setminus \{0\}$ is a sequence converging to 0.

Indeed, the sets $F_k := \sigma(A) \cap \{\lambda \in \mathbb{K} \colon |\lambda| \geq 1/k\}$ are either empty or finite: If they had infinitely many (distinct) elements, then there would be a sequence in F_k which converges to some element of F_k. (Recall that $\sigma(A)$ is compact by Proposition 6.2!) But this contradicts Theorem 6.6. Thus if there are infinitely many elements in $\sigma(A) \setminus \{0\}$ then we can arrange them to be a sequence going to 0. (For every k, pick up the finite number of elements in $F_k \setminus F_{k+1}$; this is a countable collection of finitely many elements, so countable!)

6.2 Spectral theorem for normal operators

We will now come to our final main result in spectral theory. It will be a "diagonalization" result for compact normal operators on Hilbert spaces. To this end we **fix a Hilbert space H for the rest of this section**.

First, we strengthen the inclusion of the spectrum of a bounded linear operator compared to Proposition 6.2:

Proposition 6.8 (Spectrum in numerical range)**.** *Let $A \in \mathcal{L}(H)$. Then $\sigma(A) \subseteq \overline{W(A)}$ where the **numerical range** $W(A)$ of A is defined by*

$$W(A) := \Big\{(Ax, x) \colon x \in H,\ |x| = 1\Big\}.$$

More precisely, if $\lambda \notin \overline{W(A)}$*, then* $\lambda \in \rho(A)$ *with*

$$\|R(\lambda, A)\| \leq \frac{1}{\operatorname{dist}(\lambda, W(A))}.$$

Proof. Suppose that $\lambda \notin \overline{W(A)}$ and let $\alpha = \operatorname{dist}(\lambda, W(A))$. Then we have

$$\alpha \leq \big|(Ax, x) - \lambda\big| = \big|(Ax - \lambda x, x)\big| \qquad (|x| = 1)$$

But this means that $A - \lambda$ is positive definite (recall (5.9)) and the Lax-Milgram lemma (Theorem 5.10) implies that $A - \lambda$ is bijective with $\|R(\lambda, A)\|_{H \to H} \leq \alpha^{-1}$. □

Remark 6.9. a) From the Cauchy-Schwarz inequality it is immediately clear that $W(A) \subseteq \overline{B(0, \|A\|)}$, so Proposition 6.8 is stronger than Proposition 6.2.

b) The numerical range $W(A)$ is an interesting object with many useful properties. We mention for instance the surprising fact that it is *convex*.

The abstract spectrum inclusions $\sigma(A) \subseteq W(A) \subseteq \overline{B(0, \|A\|)}$ of a bounded linear operator A are not sharp at all: Consider $H = \mathbb{C}^2$ and $A(x, y) := (y, 0)$, then $\sigma(A) = \{0\}$, but $W(A) = 1/2$ and $\|A\| = 1$. For a **complex** Hilbert space H and a **normal** operator A, the estimate turns out to be sharp. We call a bounded operator A **normal** if $AA^\star = A^\star A$.

Proposition 6.10. *Let* $A \in \mathcal{L}(H)$ *be a normal operator. Then*

$$\max\big\{|\lambda| \colon \lambda \in \sigma(A)\big\} = \|A\|.$$

Note that the foregoing Proposition 6.10 also implies that there exists $x \in H$ with $|x| = 1$ such that $(Ax, x) = \|A\|$. Having Proposition 6.10 at hand, we can now go for the diagonalization main theorem on a **separable** Hilbert space:

Theorem 6.11 (Spectral theorem for normal compact operators). *Let* H *be a separable Hilbert space over* $\mathbb{K} = \mathbb{C}$ *and let* A *be a compact normal operator on* H*. Then there exists an ONB of* H *composed of eigenvectors of* A*.*

Proof. We already know that $\sigma(A) \setminus \{0\} = \sigma_P(A)$. (Theorem 6.6!) Let (λ_k) be the (distinct) sequence of all nonzero eigenvalues of A. Moreover, set $\lambda_0 = 0$ and

$$X_0 := \ker(A), \qquad X_k := \ker(A - \lambda_k), \qquad (k \in \mathbb{N}).$$

From the Fredholm alternative theorem (Theorem 4.26) we have $0 < \dim X_k < \infty$. We claim that $X_k \perp X_\ell$ for $k \neq \ell$ and that $\operatorname{span}(\bigcup_{k=0}^{\infty} X_k)$ is dense in H.

To this end, note that if A is normal, then

$$\ker(A - \lambda) = \ker(A^\star - \overline{\lambda}) \qquad (\lambda \in \mathbb{C}). \tag{6.2}$$

(Exercise.) Hence, if $x_k \in X_k$ and $x_\ell \in x_\ell$, then

$$\lambda_k(x_k, x_\ell) = (Ax_k, x_\ell) = (x_k, A^\star x_\ell) = (x_k, \overline{\lambda_\ell}) = \lambda_\ell(x_k, x_\ell).$$

If $k \neq \ell$, then this can only be true if $(x_k, x_\ell) = 0$. So $X_k \perp X_\ell$.

Moreover, set $Y = \operatorname{span}(\bigcup_{k=0}^{\infty} X_k)$. We show that Y is dense in H. By construction, $AX_k \subseteq X_k$, and by (6.2), $A^\star X_k \subseteq X_k$. Thus $A^\star Y \subseteq Y$. But then we also have $AY_\perp \subseteq Y_\perp$: If $z \in Y_\perp$, then $(Az, y) = (z, A^\star y) = 0$ for all $y \in Y$, since $A^\star y \in Y$. Thus the restriction A to $Y_\perp$ is meaningful and still a normal compact operator. We call it A_0 and claim that $\sigma(A_0) = \{0\}$.

Suppose not. Then there exists an eigenvalue $\mu \in \sigma(A_0)$ with an eigenvector $z \in Y_\perp$. Clearly, μ must also be an eigenvalue of A, so $\mu = \lambda_k$ for some $k \in \mathbb{N}$. Thus $z \in X_k \cap Y_\perp \subseteq Y \cap Y_\perp = \{0\}$ and this is a contradiction.

So $\sigma(A_0) = \{0\}$. By Proposition 6.10 this means that $A_0 = 0$. In particular, $Y_\perp \subseteq X_0 \subseteq Y$ and by Proposition 5.7 this can only happen if $Y_\perp = \{0\}$. Thus Y is dense in H.

Finally, pick an ONB in each space X_k for $k \in \mathbb{N}_0$. For $k \geq 1$, this is immediate since $\dim X_k < \infty$. For X_0, we use that H and thus also X_0 is separable, hence it admits an ONB by Proposition 5.13. The union of all those ONBs is the object we want. □

Note that the eigenvalues in Theorem 6.11 will in general be complex! Since by definition $\sigma(A) \subseteq \mathbb{K}$, Theorem 6.11 is not true for $\mathbb{K} = \mathbb{R}$ without further ado. We enforce the real case by strengthening the structural assumption on A from normal to **selfadjoint**. Recall that $A \in \mathcal{L}(H)$ is selfadjoint if $A = A^\star$ which clearly implies that A is normal.

The reasoning is the following: If $A \in \mathcal{L}(H)$ is selfadjoint, then

$$(Ax, x) = (x, Ax) = \overline{(Ax, x)} \qquad (x \in H),$$

so $\sigma(A) \subseteq W(A) \subseteq \mathbb{R}$ by Proposition 6.8. We thus obtain the following real case of Theorem 6.11:

Theorem 6.12 (Spectral theorem for selfadjoint compact operators)**.** *Let H be a separable Hilbert space over $\mathbb{K} = \mathbb{R}$ and let A be a compact normal operator on H. Then there exists an ONB of H composed of eigenvectors of A.*

The proof is exactly the same as for Theorem 6.11.

Corollary 6.13 (Diagonalization). *In the situation of either Theorem 6.11 or Theorem 6.12, if (e_k) is the ONB of eigenvectors of A, then we have*

$$Ax = \sum_{k=0}^{\infty} \lambda_k (x, e_k)\, e_k \qquad (x \in H).$$

Proof. Recall (6.2) from the proof of Theorem 6.11, so $\ker(A - \lambda) = \ker(A^\star - \overline{\lambda})$ for any $\lambda \in \mathbb{K}$ since A is a normal operator. Thus,

$$Ax = \sum_{k=0}^{\infty} (Ax, e_k) e_k = \sum_{k=0}^{\infty} (x, A^\star e_k) e_k = \sum_{k=0}^{\infty} (x, \overline{\lambda_k} e_k) e_k = \sum_{k=0}^{\infty} \lambda_k (x, e_k) e_k \qquad (x \in H).$$

This is exactly the diagonalization of A. □

Let us finally note that there exist generalizations of the foregoing diagonalization theorems to *unbounded* operators. Already the spectrum of bounded operators is much more complicated than the one of compact operators, which, as we have seen, consists *essentially* of eigenvalues only. If one wants to consider unbounded operators, one in addition needs to very carefully define what is meant by **normal** or **selfadjoint** in this case, since in general $\mathrm{dom}(A)$ and $\mathrm{dom}(A^\star)$ need not agree at all. Unfortunately, this is out of scope for these lecture notes.

7 Sobolev spaces

In this penultimate section we establish a class of suitable function spaces to consider elliptic PDEs in. Fix an open set $\Omega \subseteq \mathbb{R}^d$.

In fact, there is a lot of classical theory for elliptic PDEs in Hölder spaces $C^{0,\alpha}(\overline{\Omega})$ which is very well rounded and useful in many areas. However, the Hölder spaces are not very nice from a functional-analytic point of view because they are not reflexive. (This is a shared problem by all function spaces involving a supremum-type norm.) This is a strong structural limitation since reflexive function spaces allow for *variational* approaches related to the very natural concept of energy minimization.

The first and fundamental concept towards a suitable function space framework is the following.

7.1 Weak derivative

Let $\Lambda \subseteq \mathbb{R}^d$ be arbitrary but nonempty.

First, we introduce a shorthand notion for mixed derivatives of a certain order. We say that $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ is a **multi-index** of order $|\alpha| := \sum_{i=1}^d \alpha_i$. Then we define the mixed differential operator

$$D^\alpha := \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} \cdots \left(\frac{\partial}{\partial x_d}\right)^{\alpha_d}.$$

It is immediate that $D^{\alpha+\beta} = D^\alpha D^\beta = D^\beta D^\alpha$ for two multi-indices α, β.

Let $k \in \mathbb{N}_0$ and denote the space of k times continuously differentiable functions by

$$C^k(\Lambda) := \Big\{ f\colon \Lambda \to \mathbb{K}\colon D^\alpha f \text{ is continuous on } \Lambda \text{ for all } |\alpha| \le k \Big\}.$$

We say that a function $f\colon \Lambda \to \mathbb{K}$ is **smooth** if $D^\alpha f$ is continuous on Λ for *any* multi-index α of arbitrary order $|\alpha| \in \mathbb{N}_0$. The associated function space is

$$C^\infty(\Lambda) := \bigcap_{k=0}^{\infty} C^k(\Lambda).$$

Recall further that $\operatorname{supp} f := \overline{\{x \in \Lambda\colon f(x) \neq 0\}}$ is the **support** of a function $\Lambda \to \mathbb{K}$. Now let the space of **test functions** be given by

$$C_c^\infty(\Lambda) := \Big\{ f \in C^\infty(\Lambda),\ \operatorname{supp} f \text{ compact in } \Lambda \Big\}.$$

Note that if Λ is compact itself, then $C_c^\infty(\Lambda) = C^\infty(\Lambda)$, but in general, this is not true. In particular, $C^\infty(\Omega) \neq C^\infty(\overline{\Omega})$ and the same for test functions. (Consider $\Omega = (0,1)$ and $f(x) = 1/x$.)

The first result making use of test functions is a variational principle of great importance. Let $L^1_{\mathrm{loc}}(\Omega)$ denote all measurable functions $f\colon \Omega \to \mathbb{K}$ such that $f \in L^1(K)$ for every *compact* set $K \subseteq \overline{\Omega}$. (With this notation, if Ω is bounded, then $L^1_{\mathrm{loc}}(\Omega)$ coincides with $L^1(\Omega)$.)

Lemma 7.1 (Fundamental lemma)**.** *Let $f \in L^1_{\mathrm{loc}}(\Omega)$ such that*

$$\int_\Omega f(x)\varphi(x)\,\mathrm{d}x = 0 \qquad (\varphi \in C_c^\infty(\Omega)).$$

Then $f = 0$ almost everywhere on Ω.

Now we define a weaker form of derivative.

Definition 7.2 (Weak derivative)**.** Let $f \in L^1_{\mathrm{loc}}(\Omega)$ and let α be a multi-index. Suppose that there exists a function $g \in L^1_{\mathrm{loc}}(\Omega)$ such that

$$\int_\Omega f(x)\, D^\alpha \varphi(x)\,\mathrm{d}x = (-1)^k \int_\Omega g(x)\,\varphi(x)\,\mathrm{d}x \qquad (\varphi \in C_c^\infty(\Omega)). \tag{7.1}$$

Then we say that g is the **weak (α)-derivative** of f and put $D^\alpha f := g$.

The fundamental lemma (Lemma 7.1) immediately implies that the weak derivative is unique up to equality almost everywhere on Ω.

The notion of *weak* derivative—as for weak convergence and associated concepts—suggests some sort of generalization of the classical derivative. It requires only that an ($|\alpha|$-fold) integration by parts formula holds true for $f \in L^1_{\mathrm{loc}}(\Omega)$ instead of a pointwise limit as in the classical derivative. (Note that $\varphi \in C_c^\infty(\Omega)$ implies that $\varphi\restriction_{\partial\Omega} = 0$ since the support of φ is compactly included in Ω.) More precisely, if $k = 1$ and $\alpha_i = 1$, then it is just the requirement that

$$\int_\Omega f(x)\frac{\partial\varphi(x)}{\partial x_i}\,\mathrm{d}x = -\int_\Omega \frac{\partial f(x)}{\partial x_i}\varphi(x)\,\mathrm{d}x \qquad (\varphi \in C_c^\infty(\Omega)).$$

In particular, we immediately see that if the (classical) derivatives $D^\alpha f$ exist and are continuous, then $D^\alpha f$ coincides with the weak derivative g, so it is appropriate to use the same notation also for the weak derivative $g = D^\alpha f$. But there are functions which are not classically differentiable but which admit a weak derivative, such as $f(x) := \max(0, x)$ considered on any open interval $\Omega = (a, b)$ with $a < 0 < b$. There we have

$$\int_\Omega f(x)\varphi'(x)\,\mathrm{d}x = \int_0^b x\varphi'(x)\,\mathrm{d}x = -\int_0^b \varphi(x)\,\mathrm{d}x = -\int_a^b g(x)\varphi(x)\,\mathrm{d}x \quad (\varphi \in C_c^\infty(\Omega))$$

with $g(x) = H(x)$, the Heaviside function

$$H(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0. \end{cases}$$

Analogously, the weak derivative of the absolute value function $x \mapsto |x|$ is given by $x \mapsto H(x) - H(-x)$. However, the Heaviside function itself does *not* admit a weak derivative due to its jump in 0.

Note that the foregoing functions are in fact smooth almost everywhere, in this case, the exceptional set is just the single point $\{0\}$. Indeed one can push this even further, for instance,

$$f(x) := \begin{cases} 0 & \text{if } x \in \mathbb{Q}, \\ 2 + \sin(x) & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

is nowhere smooth but admits the weak derivative $g(x) := \cos(x)$. (*Nota bene* $|\mathbb{Q}| = 0$ however!)

We say that a sequence (f_k) converges to f in $L^1_{\mathrm{loc}}(\Omega)$ if $f_k \to f$ in $L^1(K)$ for every compact subset $K \subseteq \overline{\Omega}$. With this notion, the weak derivative is closed in the following useful sense.

Lemma 7.3. *Let $(f_k) \subseteq L^1_{\mathrm{loc}}(\Omega)$ and suppose that for some multi-index α, the weak derivatives $(D^\alpha f_k)$ exist for all $k \in \mathbb{N}$. If $f_k \to f$ and $D^\alpha f_k \to g$ in $L^1_{\mathrm{loc}}(\Omega)$, then $g = D^\alpha f$.*

Proof. For every $\varphi \in C_c^\infty(\Omega)$, the support $\operatorname{supp}\varphi$ is compact in Ω by construction, and $\varphi, D^\alpha\varphi \in C(\operatorname{supp}\varphi)$. Thus, $f_k D^\alpha \varphi \to f D^\alpha \varphi$ and $D^\alpha f_k \varphi \to g\varphi$ in $L^1(\operatorname{supp}\varphi)$, and we find

$$\begin{aligned}\int_\Omega g(x)\varphi(x)\,\mathrm{d}x &= \int_{\operatorname{supp}\varphi} g(x)\varphi(x)\,\mathrm{d}x = \lim_{k\to\infty}\int_{\operatorname{supp}\varphi} D^\alpha f_k(x)\varphi(x)\,\mathrm{d}x \\ &= \lim_{k\to\infty}(-1)^{|\alpha|}\int_{\operatorname{supp}\varphi} f_k(x) D^\alpha\varphi(x)\,\mathrm{d}x = (-1)^{|\alpha|}\int_\Omega f(x)D^\alpha\varphi(x)\,\mathrm{d}x\end{aligned}$$

for every $\varphi \in C_c^\infty(\Omega)$. □

Just as for classical derivatives, a function with weak derivative zero almost everywhere is also constant almost everywhere.

Lemma 7.4. *Let Ω be connected and let $f \in L^1_{\mathrm{loc}}(\Omega)$. If $D^\alpha f = 0$ almost everywhere on Ω for any multi-index α with $|\alpha| = 1$, then f is constant almost everywhere.*

The proof uses the highly useful and very nice theory of **convolution** and **mollifiers**. Unfortunately, dealing with this topic properly is out of scope for this lecture. The interested reader is invited to consult the textbooks underlying these lecture notes.

7.2 Bascis in Sobolev spaces

Fix a number $1 \leq p \leq \infty$ and $k \in \mathbb{N}$. The fundamental definition of the function space of all functions with weak derivatives in $L^p(\Omega)$ is as follows:

Definition 7.5 (Sobolev space)**.** Define the **Sobolev space**

$$W^{k,p}(\Omega) := \left\{ f \in L^1_{\mathrm{loc}}(\Omega) \colon D^\alpha f \in L^p(\Omega) \text{ for all } |\alpha| \leq k \right\}$$

and equip it with the norm

$$\|f\|_{W^{k,p}(\Omega)} := \Big(\sum_{|\alpha| \le k} \int_\Omega \big|D^\alpha f(x)\big|^p \,\mathrm{d}x \Big)^{1/p} \qquad (1 \le p < \infty),$$

$$\|f\|_{W^{k,\infty}(\Omega)} := \sum_{|\alpha| \le k} \operatorname*{ess\,sup}_{x \in \Omega} \big|D^\alpha f(x)\big|.$$

Define further the subspaces

$$W_0^{k,p}(\Omega) := \overline{C_c^\infty(\Omega)}^{\|\cdot\|_{W^{k,p}(\Omega)}}.$$

It is obvious that $C_c^\infty(\Omega) \subseteq W^{k,p}(\Omega)$, so the definition of $W_0^{k,p}(\Omega)$ makes sense. We will see later that $W_0^{k,p}(\Omega)$ corresponds to the subspace of functions in $W^{k,p}(\Omega)$ for which $D^\alpha f\restriction_{\partial\Omega} = 0$ in a generalized sense, for all $|\alpha| \le k-1$. In particular, if $k=1$, then we will have $f\restriction_{\partial\Omega} = 0$.

For $p = 2$, the we give the Sobolev spaces a particular name:

Definition 7.6 (Sobolev-Hilbert space)**.** The **Sobolev-Hilbert space** $H^k(\Omega) := W^{k,2}(\Omega)$ is endowed with the inner product

$$(f,g)_{H^k(\Omega)} := \sum_{|\alpha| \le k} \int_\Omega D^\alpha f(x)\, \overline{D^\alpha g(x)}\,\mathrm{d}x \qquad (f,g \in H^k(\Omega)). \tag{7.2}$$

We also set $H_0^k(\Omega) := W_0^{k,2}(\Omega)$.

Theorem 7.7. *The Sobolev spaces $W^{k,p}(\Omega)$ are Banach spaces which are separable for $1 \le p < \infty$ and reflexive for $1 < p < \infty$. For $p = 2$, they are Hilbert spaces endowed with the inner product (7.2).*

Proof. It is elementary to prove that $W^{k,p}(\Omega)$ is a normed vector space since the norm as in Definition 7.5 is derived from the $L^p(\Omega)$ norms.

For completeness, we argue as follows. Let N be the number of multi-indices α of order $|\alpha| \le k$. Then

$$D\colon W^{k,p}(\Omega) \to L^p(\Omega)^N, \qquad f \mapsto \Big\{ D^\alpha f \colon |\alpha| \le k \Big\}$$

is a linear isometry from $W^{k,p}(\Omega)$ into $L^p(\Omega)^N$ if we equip $L^p(\Omega)^N$ with the norm

$$\|(g_1, \dots, g_N)\|_{L^p(\Omega)^N} := \Big(\sum_{k=1}^N \|g_k\|_{L^p(\Omega)}^p \Big)^{1/p}.$$

But by Lemma 7.3, $\operatorname{ran}(D)$ is a closed subspace of $L^p(\Omega)^N$ and thus a Banach space. Since D is an isometry between $W^{k,p}(\Omega)$ and $\operatorname{ran}(D)$, so is $W^{k,p}(\Omega)$. In particular, $\operatorname{ran}(D)$ and thus also $W^{k,p}(\Omega)$ is separable for $1 \leq p < \infty$ and reflexive for $1 < p < \infty$ (Lemma 3.12). $\square$

The fact that a Sobolev function admits weak derivatives which are again in $L^p(\Omega)$ has quite some intriguing consequences. In one space dimension $d = 1$, if $\Omega = (a, b)$ is an interval, then $f \in W^{1,1}(a, b)$ already implies that f is **absolutely continuous**. We use the convention that $C^{0,0}([a, b]) = C([a, b])$.

Lemma 7.8. *Let $f \in W^{1,1}(a, b)$ with $a < b$ and let $f' \in L^1(a, b)$ be its weak derivative. Then there exists an absolutely continuous function $F \in C([a, b])$ such that $F = f$ almost everywhere on (a, b) and*

$$f'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} \qquad (\text{a.a. } x \in (a, b)). \tag{7.3}$$

If $f \in W^{1,p}(a, b)$ with $1 < p \leq \infty$, then F is Hölder continuous with $F \in C^{0,1-\frac{1}{p}}([a, b])$; for $p = \infty$ this means that it is Lipschitz continuous. Moreover, for every $1 \leq p \leq \infty$ there exists a constant $C_p \geq 0$ such that

$$\|F\|_{C^{0,1-1/p}([a,b])} \leq C_p \|f\|_{W^{1,p}(a,b)} \qquad (f \in W^{1,p}(a, b)),$$

that is,

$$W^{1,p}(a, b) \hookrightarrow C^{0,1-\frac{1}{p}}([a, b]).$$

Recall that $C^{0,1-\frac{1}{p}}([a, b])$ was *compactly* embedded into $C^{0,\alpha}([a, b])$ for $0 \leq \alpha < 1 - 1/p$. The last statement in Lemma 7.8 thus shows that $W^{1,p}(\Omega)$ is also compactly embedded in these spaces. In particular, every bounded sequence in $W^{1,p}(\Omega)$ has a uniformly convergent subsequence!

Lemma 7.8 essentially describes the full situation for $d = 1$. In more space dimensions $d > 1$, the behavior of Sobolev functions is much more complicated and interesting. In the exercises we show that the function $f \colon B(0, 1) \to \mathbb{R}$ given by $f(0) = 0$ and

$$f(x) := |x|^{-\gamma} = \left(\sum_{k=1}^{d} |x_i|^2 \right)^{-\frac{\gamma}{2}} \qquad (0 < |x| < 1)$$

is in $W^{1,p}(B(0, 1))$ if and only if $p < d$ and $0 < \gamma < \frac{d-p}{p}$. (Compare the condition on p with the situation for $d = 1$ in Lemma 7.8.) This function is continuously differentiable on $B(0, 1) \setminus \{0\}$ but unbounded as $|x| \to 0$ so there is to no hope to recover good behavior there. This shows that we cannot even expect Sobolev functions to be bounded in general, not to speak of continuous. It will however turn out that this *is* the case if $p > d/k$.

Another result that we only mention but not state or prove formally is that Sobolev functions are continuous on almost every line parallel to the coordinate axes.

Also, while Sobolev functions in general might be singular, we can always approximate them by smooth functions.

Proposition 7.9 (Approximation by smooth functions). *Let $1 \leq p < \infty$. Then $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$, that is, for every function $f \in W^{k,p}(\Omega)$ and every $\varepsilon > 0$ there exists $g_\varepsilon \in C^\infty(\Omega)$ such that $\|f - g_\varepsilon\|_{W^{k,p}(\Omega)} < \varepsilon$.*

It is worthwhile to note that by definition of $C^\infty(\Omega)$, smooth functions from this space may also become unbounded as they approach the boundary $\partial\Omega$.

Finally, we mention an interesting estimate which will be of crucial importance for the treatment of elliptic PDEs in the next section. It concerns the question whether we can bound the $L^p(\Omega)$ norm of f by the norm of a (first order) weak derivative of f. Clearly this can only work for non-constant functions. We exclude them by considering $W_0^{1,p}(\Omega)$.

Proposition 7.10 (Poincaré inequality). *Let Ω be contained in a slab, that is, there exists a coordinate $i \in \{1, \ldots, d\}$ and $a, b \in \mathbb{R}$ such that $x_i \in (a, b)$ for every $x \in \Omega$. Consider $1 \leq p < \infty$. Then*

$$\|f\|_{L^p(\Omega)} \leq p(b-a)\|D^{e_i} f\|_{L^p(\Omega)} \qquad (f \in W_0^{1,p}(\Omega)).$$

Proof. Without loss of generality, we suppose that $i = 1$. We split $x = (x_1, x') \in \Omega$.

It is sufficient to prove the claim for $0 \neq f \in C_c^\infty(\Omega)$ due to the definition of $W_0^{1,p}(\Omega)$ and approximation. For $g(x) := |f(x)|^p$ we have

$$|f(x)|^p = |f(x_1, x')|^p = \int_a^{x_1} D^{e_1} g(s, x') \,\mathrm{d}s$$

For simplicity we consider f to be defined on $\mathbb{R}^d$ by extension by zero outside of Ω. So, using the Fubini theorem,

$$\begin{aligned} \|f\|_{L^p(\Omega)}^p = \int_\Omega |f(x)|^p \,\mathrm{d}x = \int_{\mathbb{R}^d} |f(x)|^p \,\mathrm{d}x &= \int_a^b \int_{\mathbb{R}^{d-1}} |f(t, x')|^p \,\mathrm{d}x' \,\mathrm{d}t \\ &= \int_{\mathbb{R}^{d-1}} \int_a^b \int_a^t D^{e_1} g(s, x') \,\mathrm{d}s \,\mathrm{d}t \,\mathrm{d}x'. \end{aligned}$$

Now we rely on a useful integration by parts:

$$\int_a^b 1 \cdot \int_a^t D^{e_1} g(s,x')\,\mathrm{d}s\,\mathrm{d}t = \Big[t \cdot \int_a^t D^{e_1} g(s,x')\,\mathrm{d}s\Big]_{t=a}^{t=b} - \int_a^b t \cdot D^{e_1} g(t,x')\,\mathrm{d}t$$

and

$$\Big[t \cdot \int_a^t D^{e_1} g(s,x')\,\mathrm{d}s\Big]_{t=a}^{t=b} = \int_a^b b \cdot D^{e_1} g(s,x')\,\mathrm{d}s.$$

Overall

$$\|f\|_{L^p(\Omega)}^p = \int_{\mathbb{R}^{d-1}} \int_a^b (b-t) \cdot D^{e_1} g(t,x')\,\mathrm{d}t\,\mathrm{d}x' \le (b-a) \int_{\mathbb{R}^d} \big|D^{e_1} g(x)\big|\,\mathrm{d}x.$$

It is time to insert the actual formula for $D^{e_1} g$:

$$D^{e_1} g(x) = p|f(x)|^{p-2}\,\mathrm{Re}\big(f(x)\overline{D^{e_1} f(x)}\big),$$

so we obtain, with Cauchy-Schwarz and the Hölder inequality ($\frac{1}{p} + \frac{p-1}{p} = 1$),

$$\begin{aligned} \|f\|_{L^p(\Omega)}^p &\le p(b-a) \int_{\mathbb{R}^d} |f(x)|^{p-1} \big|D^{e_1} f(x)\big|\,\mathrm{d}x \\ &\le p(b-a) \|f\|_{L^p(\Omega)}^{p-1} \big\|D^{e_1} f\big\|_{L^p(\Omega)}. \end{aligned}$$

This is the claim for $f \in C_c^\infty(\Omega)$. □

If Ω is in fact bounded in *all* directions, then from Proposition 7.10 we get an estimate of f in terms of the L^p norm of (the modulus of) its **weak gradient** $\nabla f = (D^{e_1} f, \ldots, D^{e_d} f)$.

Corollary 7.11 (Poincaré inequality I)**.** *Let Ω be bounded and $1 \le p < \infty$. Then there exists a constant C depending on p, d and Ω such that*

$$\|f\|_{L^p(\Omega)} \le C \|\nabla f\|_{L^p(\Omega)} \qquad (f \in W_0^{1,p}(\Omega)). \tag{7.4}$$

In particular,

$$|f|_{W_0^{1,p}(\Omega)} := \|\nabla f\|_{L^p(\Omega)} \qquad (f \in W_0^{1,p}(\Omega))$$

is an equivalent norm on $W_0^{1,p}(\Omega)$.

Proof. Suppose that $\Omega \subseteq [a_1, b_1] \times \cdots \times [a_d, b_d]$. We can apply Proposition 7.10 in all directions $i = 1, \ldots, d$. Write $f = \frac{1}{d} \sum_{k=1}^d f$ and set $\delta := \max_{i=1,\ldots,d} (b_i - a_i)/d$. Then there exist constants $C_{1,p}$ and $C_{p,2}$ such that

$$\begin{aligned}\|f\|_{L^p(\Omega)} \le \delta p \sum_{i=1}^d \|D^{e_i} f\|_{L^p(\Omega)} &\le \delta p\, C_{1,p} \Big(\sum_{i=1}^d \|D^{e_i} f\|_{L^p(\Omega)}^p \Big)^{\frac{1}{p}} \\ &= \delta p\, C_{1,p} \Big(\int_\Omega \sum_{i=1}^d |D^{e_i} f(x)|^p \,\mathrm{d}x \Big)^{\frac{1}{p}} \\ &\le \delta p\, C_{1,p} \Big(\int_\Omega C_{p,2}^p \Big[\sum_{i=1}^d |D^{e_i} f(x)|^2 \Big]^{\frac{p}{2}} \,\mathrm{d}x \Big)^{\frac{1}{p}} \\ &= \delta p\, C_{1,p} C_{p,2} \Big(\int_\Omega |\nabla f(x)|^p \,\mathrm{d}x \Big)^{\frac{1}{p}}.\end{aligned}$$

The constants $C_{1,p}$ and $C_{p,2}$ refer to the constants such that $\|\cdot\|_1 \le C_{1,p} \|\cdot\|_p$ and $\|\cdot\|_p \le C_{p,2} \|\cdot\|_2$ on $\mathbb{K}^d$. □

7.3 Extension, embedding and compactness

Lemma 7.8 established that the fact that Sobolev functions have weak derivatives in a Lebesgue space implies improved regularity properties for $d = 1$. On the other hand, we have also seen that there are unbounded and discontinuous functions in $W^{1,p}(\Omega)$ if $p < d$. The precise properties of Sobolev functions are highly important in several areas of research. The first and most important case is for $\Omega = \mathbb{R}^d$, the full space. We have the two following fundamental results:

Theorem 7.12 (Morrey's inequality). *Let $d < p < \infty$ and set $\gamma := 1 - d/p > 0$. Then there is a constant C which depends only on p and d such that*

$$\|f\|_{C_b^{0,\gamma}(\mathbb{R}^d)} \le C \|f\|_{W^{1,p}(\mathbb{R}^d)} \qquad (f \in C^1(\mathbb{R}^d) \cap W^{1,p}(\mathbb{R}^d)).$$

For the next result, we introduce the **Sobolev conjugate** $p^\star$ to $1 \le p < d$ by

$$p^\star := \frac{dp}{d-p} \qquad \iff \qquad \frac{1}{p^\star} = \frac{1}{p} - \frac{1}{d}.$$

Theorem 7.13 (Gagliardo-Nirenberg inequality). *Let $1 \le p < d$. Then there is a constant C which depends only on p and d such that*

$$\|f\|_{L^{p^\star}(\mathbb{R}^d)} \le C \|\nabla f\|_{L^p(\mathbb{R}^d)} \qquad (f \in W^{1,p}(\mathbb{R}^d)).$$

Note that $p^\star > p$. So even if $1 \leq p < d$, the existence of a weak derivative in $L^p(\mathbb{R}^d)$ of an $L^p(\mathbb{R}^d)$ function implies that f is in fact *more integrable*. The borderline case $p = d$ is a bit more involved and we skip it for now.

We would like to transfer these fundamental results also to open sets $\Omega \neq \mathbb{R}^d$. To this end, it is easy to see that if Λ is another open set and $\Omega \subseteq \Lambda$, then from $f \in W^{k,p}(\Lambda)$ it follows that $f \in W^{k,p}(\Omega)$. In particular, the restriction of a $W^{k,p}(\mathbb{R}^d)$ function will always be in $W^{k,p}(\Omega)$.

Suppose that we have a means to *reverse* this restriction procedure, that is: to **extend** a function $f \in W^{k,p}(\Omega)$ to a function $Ef \in W^{k,p}(\mathbb{R}^d)$ such that $Ef\restriction_\Omega = f$ almost everywhere on Ω. Then we could—now for $k = 1$—proceed as follows, with X either a Hölder space (Theorem 7.12) or a Lebesgue space (Theorem 7.13), depending on the magnitude of p compared to d:

$$f \in W^{1,p}(\Omega) \quad \to \quad Ef \in W^{1,p}(\mathbb{R}^d) \hookrightarrow X(\mathbb{R}^d) \quad \to \quad Ef\restriction_\Omega = f \in X(\Omega), \tag{7.5}$$

and if in fact $E \in \mathcal{L}(W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^d))$:

$$\|f\|_{X(\Omega)} \leq \|Ef\|_{X(\mathbb{R}^d)} \leq C\|Ef\|_{W^{1,p}(\mathbb{R}^d)} \leq C\|E\|_{W^{1,p}(\Omega)\to W^{1,p}(\mathbb{R}^d)}\|f\|_{W^{1,p}(\Omega)}.$$

In this way we obtain the analogues of Theorems 7.12 and 7.13 for Ω. However such an extension operator as claimed before does **not** exist for arbitrary open sets Ω. (Exercise.) A notable particular case is the following:

Lemma 7.14 (Zero extension)**.** *The extension by zero E_0 defines a bounded linear extension operator $W_0^{k,p}(\Omega) \to W^{k,p}(\mathbb{R}^d)$ for $1 \leq p \leq \infty$.*

Proof. Let $f \in C_c^\infty(\Omega)$. Then the $\operatorname{supp} f$ is compactly contained in Ω, in particular, there is $\varepsilon > 0$ such that $\operatorname{dist}(x, \partial\Omega) \geq \varepsilon$ for all $x \in \operatorname{supp} f$. Hence $f(y) = 0$ for all $y \in \Omega$ such that $\operatorname{dist}(y, \partial\Omega) < \varepsilon$ and $E_0 f \in C_c^\infty(\mathbb{R}^d) \subseteq W^{k,p}(\mathbb{R}^d)$ for all $1 \leq p \leq \infty$ with

$$\|E_0 f\|_{W^{k,p}(\mathbb{R}^d)} = \|f\|_{W^{k,p}(\Omega)}.$$

Since $C_c^\infty(\Omega)$ is by definition dense in $W_0^{k,p}(\Omega)$, the foregoing equality extends to all $f \in W_0^{k,p}(\Omega)$ by continuity. □

To have an extension operator for the full Sobolev space $W^{k,p}(\Omega)$, we will have to pose assumptions on Ω which are related to **boundary regularity**. To this end we only consider bounded sets in the following. (Unbounded sets can also be dealt with but have some cumbersome uniformity requirements.)

The idea is to require that, locally at its boundary, Ω can be transformed into a nice and smooth object; we choose the lower half ball. The quality of $\partial\Omega$ is then given by the regularity properties of the associated transformations.

Definition 7.15 (Boundary regularity)**.** Let Ω be **bounded**. We say that Ω has a C^1**-boundary**, or $\partial\Omega \in C^1$, if for every $x \in \partial\Omega$ there exists a radius r and a C^1-diffeomorphism $\Phi\colon B(x,r) \to B(0,1)$ such that $\Phi(B(x,r) \cap \Omega) = B(0,1) \cap (\mathbb{R}^{d-1} \times \mathbb{R}^-)$, the lower half ball.

The idea how to obtain an extension operator from such a boundary regularity property is the following: If Ω has a C^1-boundary, then one can consider an $W^{k,p}(\Omega)$ function f locally in $B(x,r) \cap \Omega$ for every boundary point $x \in \partial\Omega$. Using the transformation Φ, one obtains a $W^{k,p}$ function $f \circ \Phi$ on the lower half ball. (Here a certain minimum regularity of Φ is necessary.) We reflect the function over the mid-plate of $B(0,1)$ to construct a $W^{k,p}(B(0,1))$ function which can then be transferred back by Φ^{-1} to $B(x,r)$. This way we have extended f across $B(x,r) \cap \partial\Omega$ to the whole $B(x,r)$. In order to patch together all these local extensions, one needs to make use of a **partition of unity**. (Of course all of the foregoing can and needs to be argued properly.)

Proposition 7.16 (Extension for C^1-boundary)**.** *Let Ω be bounded with C^1-boundary. Let $\Lambda \subseteq \mathbb{R}^d$ be another bounded open set such that $\overline{\Omega} \subset \Lambda$. Then there exists a bounded linear extension operator $E \in \mathcal{L}(W^{k,p}(\Omega) \to W^{k,p}(\mathbb{R}^n))$, that is, such that $Ef\restriction_\Omega = f$ almost everywhere on Ω. Moreover, $\operatorname{supp} Ef \subset \Lambda$ for every $f \in W^{k,p}(\Omega)$.*

Now we can actually follow the strategy laid out in (7.5) to obtain the following most fundamental result:

Theorem 7.17 (Sobolev embedding)**.** *Let Ω be bounded with C^1-boundary.*

(a) If $1 \le p < d$, then

$$W^{1,p}(\Omega) \hookrightarrow L^{p^\star}(\Omega) \hookrightarrow L^q(\Omega) \qquad (q \in [1, p^\star]).$$

(b) If $p = d$, then

$$W^{1,p}(\Omega) \hookrightarrow L^q(\Omega) \qquad (q \in [1, \infty)).$$

(c) If $d < p \le \infty$, then

$$W^{1,p}(\Omega) \hookrightarrow C^{0,1-\frac{d}{p}}(\overline{\Omega}) \hookrightarrow C^{0,\alpha}(\overline{\Omega}) \qquad (\alpha \in [0, 1 - \tfrac{d}{p}]).$$

*The assertions stays true without **any** boundary regularity assumption for Ω if one replaces $W^{1,p}(\Omega)$ by $W_0^{1,p}(\Omega)$.*

Proof. With the strategy as in (7.5) we obtain the first embeddings for $p \neq d$ using Lemma 7.14 and Proposition 7.16 from Theorems 7.12 and 7.13. The second embeddings for $p \neq d$ follow from boundedness of Ω by the Hölder inequality and an elementary calculation with the Hölder seminorm. Finally, the Hölder inequality also shows that $W^{1,d}(\Omega) \hookrightarrow W^{1,d-\varepsilon}(\Omega)$ for any $\varepsilon > 0$. Since $(d-\varepsilon)^\star \nearrow \infty$ for $\varepsilon \searrow 0$, the assertion for $p = d$ follows from the case $1 \leq p < d$. $\square$

The existence of an extension operator in particular implies that $W^{k,p}(\Omega)$ coincides with the set of all restrictions $W^{k,p}(\mathbb{R}^d)\restriction_\Omega$. A consequence of this is that then smooth functions up to the boundary are dense in $W^{k,p}(\mathbb{R}^d)$ as well. (Compare this result to Proposition 7.9!)

Lemma 7.18. *Let Ω be bounded with C^1-boundary and let $1 \leq p < \infty$. Then $C^\infty(\overline{\Omega})$ is dense in $W^{k,p}(\mathbb{R}^d)$.*

Proof. Let $f \in W^{k,p}(\Omega)$. Then $Ef \in W^{k,p}(\mathbb{R}^d)$ and by Proposition 7.9 there exists a sequence $(G_k) \subseteq W^{k,p}(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d)$ such that $G_k \to Ef$ in $W^{k,p}(\mathbb{R}^d)$. Denote $(g_k) := (G_k\restriction_\Omega) \in C^\infty(\overline{\Omega}) \subset W^{k,p}(\Omega)$. Then

$$\|f - g_k\|_{W^{k,p}(\Omega)} \leq \|Ef - G_k\|_{W^{k,p}(\mathbb{R}^d)} \quad \longrightarrow \quad 0. \qquad \square$$

The Sobolev embeddings in Theorem 7.17 are stated just for first-order Sobolev spaces. In fact, a similar result holds true for $W^{k,p}(\Omega)$ with $k > 1$ which follows from a leapfrogging argument. (Exercise.) To this end, we consider the **net smoothness** $\alpha := k - \frac{d}{p}$. For $\alpha > 0$, we split α in its integer and fractional part $\alpha = m + \gamma$ with $m \in \mathbb{N}_0$ and $0 \leq \gamma < 1$. Then we have the following:

Lemma 7.19 (Higher order Sobolev embedding)**.** *Let Ω be bounded with C^1-boundary.*

(a) If $\alpha < 0$, then

$$W^{k,p}(\Omega) \hookrightarrow L^q(\Omega) \qquad \frac{1}{q} := \frac{1}{p} - \frac{k}{d}.$$

(b) If $\alpha = 0$, then

$$W^{k,p}(\Omega) \hookrightarrow L^q(\Omega) \qquad (q \in [1,\infty)).$$

(c) If $\alpha > 0$ and $\gamma > 0$, then

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

(d) If $\alpha > 0$ and $\gamma = 0$, then

$$W^{k,p}(\Omega) \hookrightarrow C^{m-1,\beta}(\overline{\Omega}) \qquad (\beta \in [0,1)).$$

The assertions stays true without ***any*** *boundary regularity assumption for Ω if one replaces $W^{k,p}(\Omega)$ by $W_0^{k,p}(\Omega)$.*

Let us go back to $k = 1$. From the Arzelà-Ascoli theorem combined with Sobolev embedding (Theorem 7.17) we already know that if $d < p \le \infty$, then $W^{1,p}(\Omega) \hookrightarrow C^{0,\alpha}(\overline{\Omega})$ **compactly** if $0 \le \alpha < 1 - d/p$ for Ω bounded with C^1-boundary. (Or for $W_0^{1,p}(\Omega)$.) In particular, $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ **compactly**. An analogous statement for $p < d$ follows from the following most important theorem.

Theorem 7.20 (Rellich-Kondrachov)**.** *Let Ω be bounded and let $1 \le p < d$. Then the Sobolev embedding $W_0^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $q \in [1, p^\star)$ is* ***compact****. If Ω has a C^1-boundary, then the same holds true for $W^{1,p}(\Omega)$.*

Corollary 7.21. *Let Ω be bounded and $1 \le p \le \infty$. Then $W_0^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$* ***compactly****. If Ω has a C^1-boundary, then the same holds true for $W^{1,p}(\Omega)$.*

A consequence of the Rellich-Kondrachov theorem is another version of the Poincaré inequality (Proposition 7.10). This time we exclude constant functions by enforcing mean zero over Ω:

Proposition 7.22 (Poincaré inequality II)**.** *Let Ω be bounded and connected with C^1-boundary and let $1 \le p \le \infty$. Then there exists a constant C depending on p, d and Ω such that*

$$\left\| f - \frac{1}{|\Omega|} \int_\Omega f(x)\,\mathrm{d}x \right\|_{L^p(\Omega)} \le C \|\nabla f\|_{L^p(\Omega)} \qquad (f \in W^{1,p}(\Omega)).$$

Proof. Suppose the assertion was false. Then there exists a sequence (f_k) in $W^{1,p}(\Omega)$ such that

$$\left\| f_k - \frac{1}{|\Omega|} \int_\Omega f_k(x)\,\mathrm{d}x \right\|_{L^p(\Omega)} > k \|\nabla f_k\|_{L^p(\Omega)} \qquad (k \in \mathbb{N}).$$

Put

$$g_k := \frac{f_k - \frac{1}{|\Omega|} \int_\Omega f_k(x)\,\mathrm{d}x}{\left\| f_k - \frac{1}{|\Omega|} \int_\Omega f_k(x)\,\mathrm{d}x \right\|_{L^p(\Omega)}} \qquad (k \in \mathbb{N}).$$

Then

$$\|g_k\|_{L^p(\Omega)} = 1, \quad \frac{1}{k} > \|\nabla g_k\|_{L^p(\Omega)} \quad \text{and} \quad \frac{1}{|\Omega|} \int_\Omega g_k(x)\,\mathrm{d}x = 0.$$

Hence (g_k) is a bounded sequence in $W^{1,p}(\Omega)$ with $|\nabla g_k| \to 0$ in $L^p(\Omega)$. By the Rellich-Kondrachov theorem and its corollary (Corollary 7.21), there is a subsequence of (g_k) which converges in $L^p(\Omega)$ with limit g. (We do not relabel subsequences for this proof.) So we have $g_k \to g$ and $\nabla g_k \to 0$ in $L^p(\Omega)$. From Lemma 7.3 it follows that $g \in W^{1,p}(\Omega)$ with $\nabla g = 0$. But then Lemma 7.4 says that $g = c \in \mathbb{K}$ must be constant almost everywhere. (Here we use that Ω is connected.) Indeed, it must be zero, since

$$0 = \lim_{k\to\infty} \frac{1}{|\Omega|} \int_\Omega g_k(x)\,\mathrm{d}x = \frac{1}{|\Omega|} \int_\Omega g(x)\,\mathrm{d}x = c.$$

But clearly this is a contradiction since

$$\|g\|_{L^p(\Omega)} = \lim_{k\to\infty} \|g_k\|_{L^p(\Omega)} = 1. \qquad \square$$

8 Linear elliptic partial differential equations

Fix a bounded open set $\Omega \subseteq \mathbb{R}^d$ for the following. We consider the **linear second-order partial differential operator** L defined formally by

$$Lu(x) := -\sum_{i,j}^{d} \frac{\partial}{\partial x_j}\Big(a_{ij}(x)\frac{\partial u(x)}{\partial x_i}\Big) + \sum_{i=1}^{d} b_i(x)\frac{\partial u(x)}{\partial x_i} + c(x)y(x) \qquad (x \in \Omega) \tag{8.1}$$

acting on a function $u\colon \Omega \to \mathbb{K}$, for given (measurable) functions $a_{ij}, b_i, c\colon \Omega \to \mathbb{R}$, where $i, j \in \{1, \dots, d\}$. (We only consider real coefficients.)

The associated (Dirichlet) **boundary value problem** is

$$\left.\begin{aligned} Lu(x) &= f(x) \qquad && (x \in \Omega), \\ u(x) &= 0 && (x \in \partial\Omega) \end{aligned}\right\} \tag{8.2}$$

with some measurable function $f\colon \Omega \to \mathbb{K}$. We need to include the boundary condition in (8.2) if there is to be any hope to obtain unique solutions for $Lu = f$; this is already the case for ordinary differential equations and the same principle applies here.

8.1 Weak formulation

If the coefficients in (8.1) and the right-hand side in (8.2) are regular, say, $a_{ij} \in C^1(\overline{\Omega})$ and $b_i, c, f \in C(\overline{\Omega})$, then one *could* hope for a regular solution $y \in C^2(\overline{\Omega})$ such that (8.1) and (8.2) would be well defined for every $x \in \overline{\Omega}$. Such $u \in C^2(\overline{\Omega})$ would be called a **classical** solution and there is a well developed theory for these situations.

However, many problems applications do not admit such regular data. We will thus only assume, **from now on**,

$$f \in L^2(\Omega) \quad \text{and} \quad a_{ij}, b_i, c \in L^\infty(\Omega; \mathbb{R}), \quad a_{ij} = a_{ji} \text{ a.e. on } \Omega.$$

Note that the matrix $A(x) = (a_{ij}(x)) \in \mathbb{R}^{d \times d}$ is assumed to be **real symmetric** for almost every $x \in \Omega$. Clearly, (a_{ij}) will in general not admit a derivative at all, so we need to give (8.1) and (8.2) a particular meaning, which is as follows.

Definition 8.1 (Weak solution)**.** We say that $u \in H_0^1(\Omega)$ is a *weak solution* to (8.2) if it is satisfies the *weak formulation* of (8.2)

$$\int_\Omega \left[\sum_{i,j}^d a_{ij}(x) \frac{\partial u(x)}{\partial x_i} \frac{\partial \overline{v(x)}}{\partial x_j} + \sum_{i=1}^d b_i(x) \frac{\partial u(x)}{\partial x_i} \overline{v(x)} + c(x) u(x) \overline{v(x)} \right] \mathrm{d}x$$
$$= \int_\Omega f(x) \overline{v(x)} \, \mathrm{d}x \qquad (v \in H_0^1(\Omega))$$

The weak formulation of (8.2) is formally obtained from multiplying (8.2) by $\overline{v}$ and integration by parts for the second-order part. The boundary integral drops out due to the zero boundary values of v.

Note how L is a *second-order* differential operator, but all terms in the weak formulation make perfect sense for $u, v \in H_0^1(\Omega)$ only. This is of course all in analogy to the definition of the weak derivative.

Remark 8.2. The Dirichlet boundary condition $u = 0$ on $\partial\Omega$ in (8.2) is incorporated in the notion of weak solution as in Definition 8.1 by requiring that $u \in H_0^1(\Omega)$ instead of merely $u \in H^1(\Omega)$.

A neat way to interpret the weak formulation of (8.2) is in the language of forms on the Hilbert space $H_0^1(\Omega)$. Define $a\colon H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{K}$ by

$$a(u,v) := \int_\Omega \left[\sum_{i,j}^d a_{ij}(x) \frac{\partial u(x)}{\partial x_i} \frac{\partial \overline{v(x)}}{\partial x_j} + \sum_{i=1}^d b_i(x) \frac{\partial u(x)}{\partial x_i} \overline{v(x)} + c(x) u(x) \overline{v(x)} \right] \mathrm{d}x.$$

Then it is a reformulation of Definition 8.1 to require that $u \in H_0^1(\Omega)$ satisfies

$$a(u,v) = (f,v)_{L^2(\Omega)} \qquad (v \in H_0^1(\Omega)). \tag{8.3}$$

This smells a lot like an opportunity to use the Lax-Milgram lemma (Theorem 5.10). To set the stage, we first consider a particular case which can be dealt with by its brother, the Riesz representation theorem (Theorem 5.5). Let

$$a_{ij} = \delta_{i=j}, \quad b_i = 0, \quad c = 1.$$

Then $A(x) = I$, the identity matrix, and (8.2) becomes

$$\left.\begin{aligned} -\Delta u(x) + u(x) &= f(x) \qquad && (x \in \Omega), \\ u(x) &= 0 && (x \in \partial\Omega) \end{aligned}\right\} \tag{8.4}$$

with the Laplacian $\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial_{x_i x_i}}$. In this case,

$$a(u,v) = (u,v)_{H^1(\Omega)} \qquad (u, v \in H_0^1(\Omega)).$$

Lemma 8.3. *For every $f \in L^2(\Omega)$ the Laplacian boundary value problem (8.4) admits a unique weak solution $u \in H_0^1(\Omega)$. Moreover, the solution map $L^{-1}\colon f \mapsto u$ is a* ***compact*** *operator $L^2(\Omega) \to H_0^1(\Omega)$ and thus also in $L^2(\Omega)$.*

Proof. Consider the embedding $\iota\colon H_0^1(\Omega) \hookrightarrow L^2(\Omega)$. By the Rellich-Kondrachov theorem (Corollary 7.21), this is a **compact** embedding. From the Schauder theorem (Theorem 4.23) and the example in Section 4.4 we obtain that $\iota'\colon L^2(\Omega)' \to H_0^1(\Omega)'$ is also a **compact** embedding.

By the Riesz representation theorem (Theorem 5.5), there is an *anti*linear isometry $R_2\colon L^2(\Omega)' \to L^2(\Omega)$ such that

$$(f, w)_{L^2(\Omega)} = \overline{\langle R_2^{-1} f, w\rangle}_{L^2(\Omega)', L^2(\Omega)} \qquad (w \in L^2(\Omega)).$$

Hence,

$$(f, v)_{L^2(\Omega)} = (f, \iota v)_{L^2(\Omega)} = \overline{\langle \iota' R_2^{-1} f, v\rangle}_{H_0^1(\Omega)', H_0^1(\Omega)} \qquad (v \in H_0^1(\Omega))$$

Again by the Riesz representation theorem (Theorem 5.5), this time for $H_0^1(\Omega)$, there is an *anti*linear isometry $R\colon H_0^1(\Omega)' \to H_0^1(\Omega)$ such that

$$\langle \iota' R_2^{-1} f, v\rangle_{H_0^1(\Omega)', H_0^1(\Omega)} = (v, R\iota' R_2^{-1} f)_{H_0^1(\Omega)} \qquad (v \in H_0^1(\Omega)).$$

But then $u := R\iota' R_2^{-1} f \in H_0^1(\Omega)$ is exactly the weak solution to (8.4) since

$$\begin{aligned} (R\iota' R_2^{-1} f, v)_{H_0^1(\Omega)} &= \overline{(v, R\iota' R_2^{-1} f)}_{H_0^1(\Omega)} \\ &= \overline{\langle \iota' R_2^{-1} f, v\rangle}_{H_0^1(\Omega)', H_0^1(\Omega)} = (f, v)_{L^2(\Omega)} \qquad (f \in L^2(\Omega)). \end{aligned}$$

Further, $L_0^{-1}\colon f \mapsto R\iota' R_2^{-1} f = u$ is linear and continuous $L^2(\Omega) \to H_0^1(\Omega)$. In fact, since R_2^{-1} and R are continuous and ι' is **compact**, so is $L_0^{-1}\colon L^2(\Omega) \to H_0^1(\Omega)$. $\square$

8.2 Existence and uniqueness for uniformly elliptic operators

Now in order to transfer Lemma 8.3 to the more general differential operator L as in (8.1), we need one more (crucial) assumption. Unlike the inner product as we had in Lemma 8.3, the form a as in (8.3) is not always coercive. Thus, we require that the differential operator L is **uniformly elliptic** in Ω: There exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^{d} a_{ij}(x)\xi_i\overline{\xi_j} \geq \theta|\xi|^2 \qquad \text{for all } x \in \Omega, \quad \xi \in \mathbb{K}^d.$$

This means that the matrix $A = (a_{ij}) \in \mathbb{R}^{d\times d}$ is symmetric and **uniformly positive definite**. In particular, $\xi^\top A\overline{\xi} \in \mathbb{R}$ for all $\xi \in \mathbb{K}^d$ and all eigenvalues of A are real and greater or equal to $\theta > 0$. Note that for the Laplacian as in (8.4), a_{ij} is just the identity matrix which is trivially uniformly elliptic with $\theta = 1$.

Let L_0 be the **principal part** of L, that is, the second order differential operator made of the second order terms in L:

$$L_0u(x) := -\sum_{i,j}^{d} \frac{\partial}{\partial x_j}\Big(a_{ij}(x)\frac{\partial u(x)}{\partial x_i}\Big) \qquad (x \in \Omega),$$

and consider the boundary value problem

$$\left.\begin{aligned} L_0u(x) &= f(x) \qquad && (x \in \Omega),\\ u(x) &= 0 && (x \in \partial\Omega). \end{aligned}\right\} \tag{8.5}$$

Theorem 8.4. *Let L_0 be uniformly elliptic. Then for every $f \in L^2(\Omega)$ the boundary value problem (8.5) admits a unique weak solution $u \in H_0^1(\Omega)$. Moreover, the solution map $L_0^{-1}\colon f \mapsto u$ is a* ***compact*** *linear operator $L^2(\Omega) \to H_0^1(\Omega)$ and thus also on $L^2(\Omega)$.*

Proof. Let $a_0\colon H_0^1(\Omega) \to H_0^1(\Omega)$ be the sesquilinear form associated to L_0. We check that a_0 satisfies the assumptions of the Lax-Milgram lemma (Theorem 5.10).

Clearly, it is a continuous sesquilinear form on $H_0^1(\Omega) \times H_0^1(\Omega)$, since

$$|a_0(u,v)| \leq \sum_{i,j=1}^{d} \|a_{ij}\|_{L^\infty(\Omega)}\left|\tfrac{\partial u}{\partial x_i}\right|_{L^2(\Omega)}\left|\tfrac{\partial v}{\partial x_j}\right|_{L^2(\Omega)} \leq C|\nabla u|_{L^2(\Omega)}|\nabla v|_{L^2(\Omega)}$$

for some constant $C > 0$ depending on $A = (a_{ij})$. Here it is crucial that $a_{ij} \in L^\infty(\Omega)$.

The critical question is whether a_0 is coercive. We have $a_0(u,u) \in \mathbb{R}$ for all $u \in H_0^1(\Omega)$. Since we have assumed Ω to be bounded, the Poincaré inequality (Corollary 7.11) says that there exists a constant $C_P > 0$ such that

$$|u|_{L^2(\Omega)} \leq C_P|\nabla u|_{L^2(\Omega)} \qquad (u \in H_0^1(\Omega)).$$

Morever, since L_0 is assumed to be uniformly elliptic:

$$a_0(u,u) = \int_\Omega \sum_{i,j}^d a_{ij}(x) \frac{\partial u(x)}{\partial x_i} \frac{\partial \overline{u(x)}}{\partial x_j} \,\mathrm{d}x \geq \theta \int_\Omega |\nabla u(x)|^2 \,\mathrm{d}x \qquad (u \in H_0^1(\Omega)).$$

So, for all $u \in H_0^1(\Omega)$:

$$|u|^2_{H^1(\Omega)} = |u|^2_{L^2(\Omega)} + \|\nabla u\|^2_{L^2(\Omega)} \leq (1 + C_P)|\nabla u|^2_{L^2(\Omega)} \leq \frac{C_P + 1}{\theta} a_0(u,u)$$

and a_0 is coercive on $H_0^1(\Omega)$ with coercivity constant $\alpha = \frac{\theta}{C_P+1}$.

Let ι and R_2 be as in the proof of Lemma 8.3. Then the weak formulation of (8.5) is

$$a_0(u,v) = (f,v)_{L^2(\Omega)} = \overline{\langle \iota' R_2^{-1} f, v \rangle}_{H_0^1(\Omega)', H_0^1(\Omega)} \qquad (v \in H_0^1(\Omega)).$$

The Lax-Milgram lemma (Theorem 5.10) says that there exists a unique $u \in H_0^1(\Omega)$ such that the foregoing weak formulation is satisfied, and we have

$$|u|_{H_0^1(\Omega)} \leq \alpha^{-1} \|\iota' R_2^{-1} f\|_{H_0^1(\Omega)'} \leq \frac{C_P + 1}{\theta} |f|_{L^2(\Omega)} \qquad (f \in L^2(\Omega)).$$

In particular, $L_0^{-1} \colon f \mapsto u \in \mathcal{L}(L^2(\Omega) \to H_0^1(\Omega))$ is **compact**, as in the proof of Lemma 8.3. $\square$

Eigenfunction expansion

The fact that the weak solution operator L_0^{-1} associated to (8.5) is **compact** in the Hilbert space $L^2(\Omega)$ allows to start the spectral theorem machinery from Section 6.2. We have not yet convinced ourselves that L_0^{-1} is normal (selfadjoint, in fact), but will do so in the following lemma. Intuitively, we need to use that we have assumed the coefficient matrix $A = (a_{ij})$ to be real and symmetric here.

Lemma 8.5. *Let L_0 be uniformly elliptic. Then the weak solution operator $L_0^{-1} \in \mathcal{L}(L^2(\Omega))$ to (8.5) from Theorem 8.4 is compact, injective, and selfadjoint.*

In particular, there is an ONB (ϕ_k) of $H_0^1(\Omega)$ consisting of eigenvectors of L_0^{-1}. The associated eigenvalues (λ_k) are real and strictly positive with $\lim_{k\to\infty} \lambda_k = 0$ and we have

$$L_0^{-1} f = \sum_{k=1}^{\infty} \lambda_k (f, \phi_k)_{L^2(\Omega)} \phi_k \qquad (f \in L^2(\Omega)).$$

Proof. We already know that $L_0^{-1} \in \mathcal{L}(L^2(\Omega))$ is in fact compact. Suppose that $u = L_0^{-1} f = 0$ for some $f \in L^2(\Omega)$. Then

$$0 = a_0(u, v) = (f, v)_{L^2(\Omega)} \qquad (v \in H_0^1(\Omega)).$$

In particular, the foregoing holds true for $v \in C_c^\infty(\Omega) \subset H_0^1(\Omega)$. But then the fundamental lemma (Lemma 7.1) implies that $f = 0$ in $L^2(\Omega)$. Hence L_0^{-1} is injective.

We next show that L_0^{-1} is selfadjoint. Let $f, g \in L^2(\Omega)$ and set $u = L_0^{-1} f$ and $v = L_0^{-1} g$. Then, because we have assumed $A = (a_{ij})$ to be **symmetric**, and $u, v \in H_0^1(\Omega)$ are admissible test functions for each other's weak formulation:

$$\begin{aligned}(g, L_0^{-1} f)_{L^2(\Omega)} = (g, u)_{L^2(\Omega)} &= a_0(v, u) \\ &= \overline{a_0(u, v)} = \overline{(f, v)}_{L^2(\Omega)} = (v, f)_{L^2(\Omega)} = (L_0^{-1} g, f)_{L^2(\Omega)}.\end{aligned}$$

Hence L_0^{-1} is selfadjoint on $L^2(\Omega)$.

Now the spectral theorem (Theorem 6.11) shows that there is an ONB of $L^2(\Omega)$ made of eigenvectors $(\phi_k) \subset H_0^1(\Omega)$. The associated eigenvalues (λ_k) are real since L_0^{-1} is selfadjoint and $\lim_{k\to\infty} \lambda_k = 0$ by the Riesz-Schauder theorem (Theorem 6.6) and Remark 6.7. Finally, we have

$$1 = (\phi_k, \phi_k)_{L^2(\Omega)} = (L_0^{-1} \phi_k, \phi_k)_{L^2(\Omega)} = a_0(L_0^{-1} \phi_k, \phi_k) = \lambda_k a_0(\phi_k, \phi_k),$$

so $\lambda_k = a_0(\phi_k, \phi_k)^{-1} > 0$ since a_0 was coercive. $\square$

Example: Let $\Omega = (0, \pi) \subset \mathbb{R}$ and $(L_0 u)(x) = -u''(x)$. Then (8.5) becomes

$$-u''(x) = f(x) \qquad (x \in (0, \pi)), \qquad u(0) = u(\pi) = 0.$$

One can compute the eigenvalues and associated eigenfunctions of L_0 to be $\mu_k = k^2$ and $\phi_k(x) := \sqrt{2/\pi} \sin(kx)$. Clearly, $\lambda_k := \mu_k^{-1} = 1/k^2$ are then the eigenvalues of L_0^{-1} with eigenfunctions ϕ_k. Per Lemma 8.5, (ϕ_k) gives rise to an ONB of $L^2(\Omega)$. (This is exactly the example mentioned at the end of Section 5.3.) Moreover, for $f \in L^2(\Omega)$ we obtain the **Fourier sine series representation** for the weak solution $u = L_0^{-1} f$:

$$u(x) = \sum_{k=1}^\infty \lambda_k (\phi_k, f)_{L^2(0,\pi)} \phi_k(x) = \sum_{k=1}^\infty \frac{2}{\pi k^2} \Big(\int_0^\pi f(y) \sin(ky)\, dy \Big) \sin(kx)$$

for $x \in (0, \pi)$.

8.3 General linear elliptic operators

So far we have a satisfying theory for the principal part L_0 of the second order differential operator L. The crucial assumption was uniform ellipticity via a condition on the coefficient matrix $A = (a_{ij})$. Going back to L, unfortunately it turns out that the lower order terms induced by b_i and c may interfere with ellipticity. For example, if c is negative and sufficiently large, then one can find functions $u \in H_0^1(\Omega)$ such that $a(u,u) < 0$ even if L_0 is uniformly elliptic. So, in full generality we cannot expect an analogue of Theorem 8.4 for L. But we can make good use of the properties of L_0 for L uniformly elliptic by leveraging the Fredholm alternative. This we do as follows:

Put $L_1 u := Lu - L_0 u$. Then L_1 consists of the lower-order terms of L and we have $L_1 \in \mathcal{L}(H_0^1(\Omega) \to L^2(\Omega))$ since $b_i, c \in L^\infty(\Omega)$. In particular, the weak form of (8.2) is given by

$$a_0(u,v) = (f,v)_{L^2(\Omega)} - (L_1 u, v)_{L^2(\Omega)} \qquad (v \in H_0^1(\Omega)),$$

and $u \in H_0^1(\Omega)$ is a weak solution to (8.2) exactly when

$$u = L_0^{-1}[f - L_1 u] \quad \iff \quad (I + L_0^{-1} L_1)u = L_0^{-1} f \quad \text{in } H_0^1(\Omega).$$

Set $K := L_0^{-1} L_1$. Then $K \in \mathcal{L}(H_0^1(\Omega))$ and it is **compact** by Theorem 8.4. Hence the **Fredholm alternative** (Theorem 4.26) applies and says that either $I + K$ is bijective or there exists $0 \neq u \in H_0^1(\Omega)$ such that $(I+K)u = 0$. This yields the following full characterization of weak solutions to (8.2).

Theorem 8.6. *Let L be uniformly elliptic. Then* (8.2) *admits a unique weak solution* $u \in H_0^1(\Omega)$ *for every* $f \in L^2(\Omega)$ if and only if *we have*

$$\Big[a(w,v) = 0 \quad (v \in H_0^1(\Omega))\Big] \quad \Longrightarrow \quad w = 0 \qquad (w \in H_0^1(\Omega)).$$

References

[Alt16] H. W. Alt. *Linear Functional Analysis*. Springer London, 2016. DOI: 10.1007/978-1-4471-7280-2.

[Bre10] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer New York, 2010. DOI: 10.1007/978-0-387-70914-7.

[Bre13] A. Bressan. *Lecture Notes on Functional Analysis: With Applications to Linear Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2013.

[Eva98] L. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 1998.

[Hac17] W. Hackbusch. *Elliptic Differential Equations*. Springer Berlin Heidelberg, 2017. DOI: 10.1007/978-3-662-54961-2.