A Short Course on the Hilbert Space Theory of Evolutionary Equations.

R. Picard
TU Dresden

24th Jyväskylä Summer School Course 2014

Version 1.3

This course presents the Hilbert space background and the solution theory of a class of operator equations involving time-differentiation (evolutionary equations). This class is comprehensive in the sense that it covers all typical models of mathematical physics such as acoustics, thermodynamics, visco-elastics, electrodynamics, quantum dynamics as well as systems describing various couplings of such model equations in a unified Hilbert space framework. It extends the typical problem classes accessible via a classical evolution equation approach to so-called differential-algebraic systems. In order to make the theoretical framework easily accessible, the course will to a large extent be reviewing the needed results from Hilbert space theory on which our approach to evolutionary equations is based. In later parts of the course we shall focus on specific applications of the approach to problems of mathematical physics and engineering.

Contents

Index

Nomenclature

A^*	adjoint relation or mapping to a relation or mapping A
$\bigwedge$	for all or for every, as in $\bigwedge_{x\in M} x \in N$, all quantifier
$\bigvee$	there is or there exists, as in $\bigvee_{x\in M} x \in N$, existence quantifier
$\bar{A}$	closure of A
$\mathring{A}$	closure of a closable operator A restricted to elements in $\mathring{C}_\infty(\Omega)$ for some open subset in $\mathbb{R}^{n+1}$, $n \in \mathbb{N}$
z^*	complex conjugate of a complex number z
$\mathring{C}_\infty(\Omega))$	set or space of smooth functions with compact support contained in the open subset in $\mathbb{R}^{n+1}$, $n \in \mathbb{N}$
D	the selfadjoint differentiation operator $D := \frac{1}{2\pi \mathrm{i}}\partial$
$\oplus$	direct sum, orthogonal sum
$\bigoplus$	direct summation sign or orthogonal summation sign as in $\bigoplus_{t\in M} H_t$
Div	divergence of $(0,2)$-tensor fields
$\in$	element sign as in $x \in C$
$\in$	element function as in $x = \in (\{x\})$ giving the element of a set containing only one element
$\mathcal{L}_\nu$	Fourier-Laplace transform with parameter $\nu \in \mathbb{R}^{n+1}$
$\mathbb{L}_\nu$	temporal Fourier-Laplace transform with parameter $\nu \in \mathbb{R}$
Grad	(0,2)-tensor obtained as symmetrized covariant derivative of a 1-form/(0,1)-tensor
$\Rightarrow$	if ... then ..., as in $x > 1 \Rightarrow x > 2$
$\iff$	... if and only if ... , as in $1 = 1 \iff 2 = 2$
$\langle \cdot \vert \cdot \rangle_X$	inner product of the inner product space X
$E^{-1/2}[X]$	inner product space derived from the inner product space X by modifying the inner product $\langle \cdot \vert \cdot \rangle_X$ to $\langle \cdot \vert E \cdot \rangle_X$, where $E : X \to X$ is continuous, linear, symmetric and strictly positive definit
$\lfloor r \rfloor$	largest integer less or equal to the real number r, integer part
$\bigcap$	big intersection symbol, as in $\bigcap M = \bigcap \{X \vert X \in M\} = \big\{y \vert \bigwedge_{X\in M} y \in X\big\}$ or $\bigcap_{X\in M} X$
$\ell^2(M)$	space of square summable complex-valued functions with at most countable arguments of non-zero values defined on M, i.e. functions $f \in \mathbb{C}^M$ with $M \setminus ([\{0\}]f)$ countable, such that $\sum_{m\in M} \vert f(m)\vert^2 = \sum_{m\in M\setminus([\{0\}]f)} \vert f(m)\vert^2 < \infty$
$L^2(\Omega)$	space of (equivalence classes of) square integrable functions on Ω, (equivalence relation is equality 'almost' everywhere)
Δ	(spatial) Laplacian
$\to$	mapping, as in $A \to B$
$\to$	strong convergence, as in $x_n \overset{n\to\infty}{\to} x_\infty$ for the convergence of a sequence $(x_n)_{n\in\mathbb{N}}$ to its limit x_∞ in the norm topology
Lin	linear hull, as in $\mathrm{Lin}_{\mathbb{K}} A$, the smallest linear space over the field $\mathbb{K}$ containing A
$\wedge$	logical *and*
$\neg$	logical negation, *not*
$\vee$	logical non-exclusive *or*
$\{\cdots \vert \cdots\}$	the set of all $\cdots$ such that $\cdots$, as in $\{x \in \mathbb{R} \vert x > 2\}$
$\bigwedge_{\cdots}$	for all $\cdots$ / for every $\cdots$
$\bigvee_{\cdots}$	there is $\cdots$ / there exists $\cdots$
$\mapsto$	maps to, as in $x \mapsto x^2$
$-A$	the set of all ordered pairs $(a, -b)$ for (a, b) in a relation $A \subseteq X \times Y$

$-[A]$	the set of all negatives of elements in the set A
$[\{0\}]f$, $N(f)$	null space or kernel of a mapping or function f
$\mathbb{C}$	field or set of complex numbers
$\mathfrak{Re}$	real part
$\mathfrak{Im}$	imaginary part
$\mathbb{K}$	field or set of numbers (either $\mathbb{K}=\mathbb{R}$ or $\mathbb{K}=\mathbb{C}$)
$\mathbb{N}$	monoid or set of natural numbers $\{0,1,2,,\ldots\}$
$\mathbb{Z}$	group or set of integers
$w\,(A)$	numerical range of A
$\perp$	orthogonal, as in $x \perp y$
$M^{\perp}$	ortho-complement of M,
2^B	the set of (left-total) mappings from B into $2=0,1A$, i.e. the power set of B
A^B	the set of (left-total) mappings from B into A
$\times$	Cartesian product as in $X \times Y$ or Cartesian multiplication sign as in $\underset{s\in M}{\times} H_x$, where X,Y,M,H_s are sets, $t \in M$
$\times$	vector product in $\mathbb{R}^3$ as in $x \times y$, where $x,y \in \mathbb{R}^3$
P_C	orthogonal projector onto the closed subspace C
$\varrho(C)$	resolvent set of operator C
$A\big\vert_M$	A restricted to M for a mapping $A: D(A) \subseteq X \to Y$, i.e. the mapping $A\big\vert_M : D(A)\cap M \subseteq X \to Y$ where $x \mapsto A(x)$
R_H	Riesz mapping, which unitarily maps H^* onto H
$\sigma(C)$	spectrum of operator C
$C\sigma(C)$	continuous spectrum of C
$P\sigma(C)$	point spectrum of C
$R\sigma(C)$	residual spectrum of C
supp	support
supp_{ν_0}	support in direction ν_0
supp_0	temporal support
$\bigotimes$	tensor product sign, as in $\bigotimes_{n\in M} W_n$
$\otimes$	tensor product
σ_α	re-scaling transformation, $\alpha \in \mathbb{R}\setminus\{0\}$
$\bigcup$	big union symbol, as in $\bigcup M = \bigcup\{X \vert X \in M\} = \big\{y \vert \bigvee_{X\in M} y \in X\big\}$ or $\bigcup_{X\in M} X$
grad	vectoranalytic differential operator grad, gradient
curl	vectoranalytic differential operator curl, curl
div	vectoranalytic differential operator div, divergence

Part 1

Some Basics of Hilbert Space Theory

CHAPTER 0

Some Useful "Well-Known" Elementary Preliminaries

The foundations of mathematics are logic & set theory. **All** mathematical objects are classes (i.e. of the form

$$\{x|P(x)\} \equiv \{x|P(x) = 1\},$$

where P is a predicate in the sense of predicate logic, $0, 1$ truth values, 1 for true and 0 for false). Usually even only "good classes" are of interest, i.e. **sets**, as defined in the axioms of set theory).

In particular (following John von Neumann)

$$\mathbb{N} := \bigcap \{M|0 \in M,\ n \in M \implies n \cup \{n\} \in M\}$$

is a set, although $\{M|0 \in M,\ n \in M \implies n \cup \{n\} \in M\}$ is merely a class and also the members of this class may be classes. In other words, the smallest class containing the empty set 0 and with every element n also its successor $n \cup \{n\}$ is (by axiom) a set, called the set of natural numbers. We use the usual number names

$$n + 1 := n \cup \{n\},$$

which results in the suggestive notation

$$n + 1 := \{0, \ldots, n\}.$$

In particular $0 \in \mathbb{N}$ and $1 = \{0\}$.

Unfortunately, as clear-cut matters are at the beginning, for easier communication we frequently use jargon rather than rigorous mathematical terms. The trouble with jargon is:

- **Same words or symbols for different things, different words or symbols for equal things.**
- **An abundance of terminology.**

We should always be conscious of and conscientious about this issue.

Example 1.

- A function $x \mapsto f(x)$ usually identified with mapping $f : D(f) \subseteq X \to Y$, $x \mapsto f(x)$, sometimes even with the expression $f(x)$ generating the function[1]. Following the early history of calculus, the function$f$ is also still sometimes referred to as the dependent variable, compare "random variable". By another abuse of logic, functions make their appearance as "constants depending on ...", thus completing the confusion between variables, functions and constants.
- A function $f : D(f) \subseteq X \to Y$, X, Y metric spaces, is called bounded if its range $\mathrm{R}(f) = f[X]$ is a bounded set in Y. If in addition X, Y are linear spaces and f is linear, then f is called bounded if it maps bounded sets into bounded sets.
- $\mathbb{R}^+$ is used for the positive reals $\mathbb{R}_{>0}$ or the non-negative reals $\mathbb{R}_{\geq 0}$ or the additive group $(\mathbb{R}, +)$.

[1] To avoid the possible confusion between $f(x)$ as a mathematical expression with the so-called "free" variable x with the function f, we shall, if the need arises to express a function in this fashion, write $x \mapsto f(x)$. As a typical example rather than writing the ambiguous x^2 meaning the function of squaring, we prefer $x \mapsto x^2$. We may also use the notation $f(\,\cdot\,)$ for the function f. Note that the widely used "$y = x^2$" is not a function but, clearly, an equation. The intended function would be $\{(x, y) \in \mathbb{R}^2 | y = x^2\}$. Usually this casual jargon does not cause too much difficulties, but it is always good to know that if confusion arises, *we can be precise.*

- $\mathbb{R}_+$ is used for the positive reals $\mathbb{R}_{>0}$ or the non-negative reals $\mathbb{R}_{\geq 0}$ or the additive group $(\mathbb{R}, +)$.
- $+$ for numbers, vectors, matrices etc. .
- sets and structures: $(M, +)$ group, jargon: the group M, $x \in M$ is a group element.
- Let $f_k : X \to \mathbb{C}$, $k \in \mathbb{N}$, denote continuous linear mappings (mappings into numbers are called functionals). Then $(f_k)_k$ is called *weak-*-convergent* to f_∞ if $f_k(x) \to f_\infty(x)$ for every $x \in X$, which in turn is equivalent to saying $(f_k)_k$ *converges strongly* to f_∞ or $(f_k)_k$ *converges point-wise* to f_∞.

The advantages of jargon: it may serve to recognize historical achievements, may inspire intuition, may ease speaking, facilitate interpretation and may simplify communication with the scientific main stream. So: jargon is okay, but rigorous mathematics should always be available as a fall-back in case of trouble.

An important concept within set theory is the so-called binary relation, which is a set R of ordered pairs, i.e.

$$r \in R \implies r = (a, b) \text{ for some sets } a, b.$$

The now generally agreed concept of an ordered pair goes back to Kuratowski 1921:

$$(a, b) := \{\{a\}, \{a, b\}\}.$$

Cartesian product of set X_0 with set X_1:

$$X_0 \times X_1 := \{(x_0, x_1) \,|\, x_k \in X_k,\ k = 0, 1\}$$

Recovery of the first component of (a, b).

Big intersection, big union:

$$\bigcap M := \{x | x \in A \text{ for all } A \in M\},$$
$$\bigcup M := \{x | x \in A \text{ for some } A \in M\}.$$

With this we have

$$\begin{aligned} \bigcap (a, b) &= \bigcap \{\{a\}, \{a, b\}\} \\ &= \{a\} \cap \{a, b\} \\ &= \{a\} \end{aligned}$$

and so

$$\begin{aligned} a &= \in \left(\bigcap (a, b)\right) \\ &= \bigcup \{a\} = \bigcap \{a\} \\ &= \bigcup \bigcap (a, b) \\ &= \bigcap \bigcap (a, b). \end{aligned}$$

The recovery formula for b can be given as[2]

$$\begin{aligned} b &= \in \left(\left\{x \in \bigcap (a, b) \,|\, \bigcup (a, b) \setminus \bigcap (a, b) \neq 0 \implies x \in \bigcup (a, b) \setminus \bigcap (a, b)\right\}\right) \\ &= \bigcup \left(\left\{x \in \bigcup (a, b) \,|\, \bigcup (a, b) \setminus \bigcap (a, b) \neq 0 \implies x \in \bigcup (a, b) \setminus \bigcap (a, b)\right\}\right) \\ &= \bigcup \left(\{x \in \{a, b\} \,|\, \{a, b\} \setminus \{a\} \neq 0 \implies x \in \{a, b\} \setminus \{a\}\}\right) \\ &= \bigcup \left(\{x \in \{a, b\} \,|\, a \neq b \implies x = b\}\right). \end{aligned}$$

[2]Note that $\in(\{x\}) := x = \bigcup \{x\} = \bigcap \{x\}$. The domain of $\in(\,\cdot\,)$ is $\{\{x\} \,|\, x \in \mathcal{U}\}$.

We choose[3] to denote

$$bRa := (a,b) \in R.$$

$[M]\,R$ pre-set of a set M under a relation R, $R\,[N]$ post-set of a set N under a relation R. R^{-1} denotes the inverse relation to R, i.e.

$$R^{-1} := \{(x,y) \mid (y,x) \in R\},$$

A correspondence is an ordered pair of a relation R and a Cartesian product $A \times B := \{(x,y) \mid x \in A, y \in B\}$ such that $R \subseteq A \times B$: $(R, A \times B)$.

The inverse correspondence to $(R, A \times B)$ is given by

$$\left(R^{-1}, B \times A\right).$$

A function is simply a right-unique relation f. If f is a function we write

$$x \mapsto f(x) := f.$$

The total post-set $f\,[\mathcal{U}]$ is here also called the range $\mathrm{R}(f)$ of f and the total pre-set $[\mathcal{U}]\,f$ is called the domain $\mathrm{D}(f)$ of f (here $\mathcal{U}$ denotes the class of all sets; the "universe"). A function f is called one-to-one if f^{-1} is also a function.

A mapping is a correspondence $(f, A \times B)$ where f is a function. Notation: $\mathrm{D}(f) = [B]\,f$

$$\begin{aligned} f : D(f) \subseteq A &\to B, \\ x \mapsto f(x) &:= \in (f\,[\{x\}]) = \bigcup f\,[\{x\}]. \end{aligned}$$

A mapping $(f, A \times B)$ is onto if the correspondence $(f, A \times B)$ is right-total, i.e. if $f\,[A] = B$. This concept clearly only makes sense for mappings not for functions.

Analogously a mapping or correspondence is called left-total if $A = [B]\,f$.

A mapping $(f, A \times B)$ is called one-to-one if f is one-to-one. In this case $\left(f^{-1}, B \times A\right)$ is a mapping, i.e.

$$\begin{aligned} f^{-1} : f\,[A] \subseteq B &\to [B]\,f \subseteq A, \\ x \mapsto f^{-1}(x) &:= \in ([\{x\}]\,f) = \bigcup [\{x\}]\,f. \end{aligned}$$

Every function f gives rise to an onto and left-total mapping in an obvious way:

$$\begin{aligned} \mathrm{D}(f) &\to \mathrm{R}(f), \\ x &\mapsto f(x). \end{aligned}$$

A relation R gives rise to a function $x \mapsto R\,[\{x\}]$ and a mapping

$$\begin{aligned} F_R : [\mathcal{U}]\,R &\to 2^{R[\mathcal{U}]}, \\ x &\mapsto R\,[\{x\}]. \end{aligned}$$

Note that F_R^{-1} has little to do with R^{-1} or $F_{R^{-1}}$.

To give some more warnings about *jargon trouble*, here some more examples in connection with the concepts *relation* and *correspondence*.

Example 2.

- ► of a one-to-one function f is often said that "f is invertible", although the inverse f^{-1} of of a function f always exists, but usually is not a function. One should always say "invertible function" not just "invertible" to avoid this terminological problem.

---

[3] Note the permutation in the order of appearance of a, b. This is chosen for consistency with the usual perspective on functions. We shall explain later.

- a correspondence $(R, A \times A)$ is also known as a (directed) graph[4], the elements of R are called edges and if $(a, b) \in R$ then a, b are called end vertices. In this interpretation A would be the set of vertices and the standard notation for the graph would be (R, A).
- a correspondence $(R, A \times B)$ written as a triple (R, A, B) is called a **formal context** in the field of formal concept analysis. A objects, B properties, R incidence relation. formal concept[5] $(U, V) \in 2^A \times 2^B$, i.e. $U \subseteq A$, $V \subseteq B$, if
$$V = [U]\, R =: U',\ U = R\,[V] =: V'.$$
$$\begin{aligned} 2^A \cup 2^B &\to 2^A \cup 2^B \\ S &\mapsto S' \end{aligned}$$
is called derivation[6], a term we know in a totally different context from analysis. U extent of the formal concept (U, V) and V its intent.

- for a correspondence $(R, A \times B)$ the relation R is also referred to as the **graph** of the correspondence, unfortunately also called relation graph although the graph *is* the relation.
- correspondence $(R, A \times B)$ and relation R are as a matter of jargon used equ**i**vocally (the same for mapping and function). So self-confusing that later the relation R of the correspondence $(R, A \times B)$ is referred to as the graph of the relation R.
- due to vigorous self-confusion apparently long forgotten "relations" make their ghastly appearance as so-called "multi-valued functions" (or "multi-valued mappings") or "set-valued functions" (or "set-valued mappings"). The latter is a typical jargon term since of course every function/mapping has sets as values (recall for example that numbers are also sets!).
- The element relation we need from set theory $\varepsilon \subseteq \mathcal{U} \times \mathcal{U}$ fits into our terminology. It is $x \in M \iff (M, x) \in\in$, $\in [U] = \{y|\, (S, y) \in\in,\ S \in U\} = \{y|y \in S,\ S \in U\} = \bigcup U$, $\in [\{S\}] = \{y|y \in T,\ T \in \{S\}\} = \{y|y \in S\} = S$.
$$[X] \in= \{S|\, (S, x) \in\in,\ x \in X\} = \{S|x \in S \text{ for some } x \in X\}\,.$$
$$\{(\{x\}, x)\, |x \in \mathcal{U}\} \subseteq\in$$
The sub-relation $\{(\{x\}, x)\, |x \in \mathcal{U}\}$ is obviously right-unique defining a mapping
$$\begin{aligned} \{\{x\}\, |x \in \mathcal{U}\} &\to \mathcal{U}, \\ \{x\} &\mapsto x, \end{aligned}$$
for which we keep the notation $\in$: $\{\{x\}\, |x \in \mathcal{U}\} \to \mathcal{U}$, it is $\in (\{x\}) = x$. $\{\{x\}\, |x \in \mathcal{U}\}$ is the largest pre-set on which $\in$ will be a function, which supports our round-bracket notation. $\in (S)$ reads as "the member of S" with the implicit assumption that S has only 1 member so that
$$S = \{\in (S)\}\,.$$

[4] In an undirected graph the edges are not ordered pairs but just subsets of cardinality two or, if one allows for "edges connecting a point with itself", of cardinality one or two. In the latter case the graph would be (E, A) with some set
$$E \subseteq \{\{a, b\}\, |a, b \in A\}\,.$$
If multiple edges between the same end vertices are permitted edges need to be indexed. In this case
$$E \subseteq I \times \{\{a, b\}\, |a, b \in A\}\,,$$
where I is an index set such as $\mathbb{N}$ or $\mathbb{Z}$ or for example a finite index set $n + 1 := \{0, \ldots, n\}$, $n \in \mathbb{N}$.

[5] 2^A denotes the so-called power set of A
$$2^A = \{U|U \subseteq A\}\,.$$

[6] Also U' is used in classical set theory to denote the complement $\mathcal{U} \setminus U$ of a set in a prescribed "universe" $\mathcal{U}$, a sufficiently large set.

In other words, $x \mapsto \in (x)$ is the inverse of function that results in singletons:

$$x \mapsto \{x\}.$$

Recall our convention

$$bRa := (a,b) \in R.$$

E.g. $R = \leq$

$$\leq = \left\{(y,x) \in \mathbb{R}^2 | x \leq y\right\},$$

then $\leq [\{3\}] = \{y | 3 \leq y\}$. This choice may be slightly confusing, but it is the price to pay for consistency with the case of functions f, where we prefer the **input** to be **to the right** of the **function symbol** but as the first component in the ordered pair. The right component is then the image $f(x)$ (to support the input-output metaphor for functions). Fortunately, we can keep the resulting potential for confusion small by avoiding to mix the notation bRa and $(a,b) \in R$.

We note that for a function f

$$yfx \iff y = f(x) \iff (x,y) \in f$$

so that our choice for general relations seems to be more intuitive, which is the actual reason for this convention.

Since the mapping

$$\begin{aligned} R : \{p : \{0,1\} \to X_0 \cup X_1 | p(k) \in X_k, k = 0,1\} &\to X_0 \times X_1 := \{(a,b) | a \in X_0, b \in X_1\} \\ p = \{(0, p(0)), (1, p(1))\} &\mapsto (p(0), p(1)) \end{aligned}$$

has

$$\begin{aligned} R^{-1} : X_0 \times X_1 &\to \{p : \{0,1\} \to X_0 \cup X_1 | p(k) \in X_k, k = 0,1\} \\ (a,b) &\mapsto \{(0,a), (1,b)\} \end{aligned}$$

as its inverse, we have one-to-one correspondence

$$(R, \{p : \{0,1\} \to X_0 \cup X_1 | p(k) \in X_k, k = 0,1\} \times (X_0 \times X_1))$$

indeed a bijection R, which allows to identify ordered pairs with left-total mappings in

$$\{p : \{0,1\} \to X_0 \cup X_1 | p(k) \in X_k, k = 0,1\} = X_0 \times X_1.$$

This identification also allows for a consistent and convenient generalization to triples, quadruples, quintuples, ... $(n+1)$-tuples for $n \in \mathbb{N}$, …, N-tuples for an arbitrary set N.

Take[7] $e \in M^N := \{f | f : N \to M\}$ then we may associate with e the N-fold Cartesian product

$$\bigtimes_{n \in N} e(n) := \left\{ t \in \left(\bigcup e[N]\right)^N | t(n) \in e(n), \, n \in N \right\}.$$

The elements $t \in \bigtimes_{n \in N} e(n)$ are written as N-tuples

$$t = (t(n))_{n \in N}.$$

In this sense every mapping $f \in S^N$ can be considered as an element of $\bigtimes_{n \in N} S$ and so the identification

$$f = (f(n))_{n \in N}$$

is motivated. This works for arbitrary sets M, N. If, however, $N = n + 1 := \{0, \ldots, n\}$ (following John von Neumann) we get so-called $(n+1)$-tuples, wher one often writes

$$f_k := f(k), \; k = 0, \ldots, n,$$

---

[7]The notation M^N is somewhat suggestive in so far as for sets M and N with finite cardinalities $\#M, \#N \in \mathbb{N}$ we have

$$\#\left(M^N\right) = (\#M)^{\#N}.$$

Moreover, for $m, n \in \mathbb{N}$ we have that the cardinality of $(m+1)^{(n+1)} = \{0, \ldots, m\}^{\{0, \ldots, n\}}$ is just the number $(m+1)^{(n+1)}$ in the sense of usual arithmetic.

and so

$$f = (f_n)_{n \in n+1} = (f_n)_{n=0,\ldots,n}$$

If $N = \mathbb{N}$ we speak of a sequence. Any $f \in S^{\mathbb{N}}$ or $f \in S^{n+1}$, $n \in \mathbb{N}$, which is a bijection[8] is called an **enumeration** of S.

N arbitrary, $S = 2 \equiv \{0,1\}$

$$2^N = \{P | P : N \to \{0,1\}\}$$

$$\begin{aligned} \{P | P : N \to \{0,1\}\} &\to \{U | U \subseteq N\} \\ P &\mapsto \{x \in N | P(x) = 1\}. \end{aligned}$$

Recall that "$x \in U$ *true*" means $(x \in U) = 1$ and $x \notin U$ means $(x \in U) = 0$ we get as the inverse function

$$\begin{aligned} \{U | U \subseteq N\} &\to \{P | P : N \to \{0,1\}\} \\ U &\mapsto \begin{pmatrix} N & \to \{0,1\} \\ x & \mapsto x \in U \end{pmatrix} \end{aligned}$$

and so we have a bijection between 2^N and the so-called power set $\{V | V \subseteq N\}$ of N, which finally justifies the common notation 2^N for the power set. The mapping

$$\begin{aligned} \chi_V : \mathcal{U} &\to \{0,1\} \\ x &\mapsto (x \in V) \end{aligned}$$

is called characteristic function of V. These observations motivate in particular to make little distinction between sets and their characteristic functions.

In general: if $(G, a, b, \ldots)$ is a set with operations or mappings $a, b, \ldots$ acting on some Cartesian product of the set G one usually re-uses the name of the set G for the structure $(G, a, b, \ldots)$, allowing to write $x \in G$ for members of the first component of the structure $(G, a, b, \ldots)$. Consider for example an "additive group" $(G, +)$ then we can write $x \in G$ meaning x is an element of the group, rather than having to say x is a member of the first component of the group $(G, +)$.

[8] A bijection is a left-total, one-to-one and onto mapping.

CHAPTER 1

Hilbert Spaces

1.1. Inner Product Spaces

Definition 3. A normed linear space[1] $((M, +, (\alpha\cdot)_{\alpha\in\mathbb{K}}), |\cdot|_M)$ is called an **inner-product space** if the norm $|\cdot|_M$ is induced by a functional $\langle\cdot|\cdot\rangle_M : M \times M \to \mathbb{K}$, $(x, y) \mapsto \langle x|y\rangle_M$, satisfying

(1) $\langle x|\cdot\rangle_M : M \to \mathbb{K}$ is a linear functional for every fixed $x \in M$,
(2) $\langle x|y\rangle_M = \overline{\langle y|x\rangle_M}$ for every $x, y \in M$,
(3) $\langle x|x\rangle_M \in \mathbb{R}_{\geq 0}$ for every $x \in M$,
(4) $\langle x|x\rangle_M = 0$ implies $x = 0$ for every $x \in M$,

in the sense that

$$|x|_M := \sqrt{\langle x|x\rangle_M} \text{ for all } x \in M. \tag{1.1.1}$$

We speak of an inner product and as a matter of jargon of an inner-product space M (assuming that linear structure and inner product are clear from the context).

Note that since $\langle x|\cdot\rangle_M : M \to \mathbb{K}$ is linear and the property 2. holds, we also have that $\langle\cdot|x\rangle_M : M \to \mathbb{K}$ satisfies

$$\begin{aligned}\langle\cdot|x\rangle_M(\alpha\cdot u+v) &= \langle\alpha\cdot u+v|x\rangle_M\\ &= \overline{\alpha}\,\langle u|x\rangle_M + \langle v|x\rangle_M\end{aligned}$$

for all $\alpha \in \mathbb{K}$, $u, v \in H$. A functional $F : M \times M \to \mathbb{K}$, such that $F(x,\cdot)$ and $\overline{F(\cdot, x)}$ are both linear, is called sesqui-linear. In the special case $\mathbb{K} = \mathbb{R}$ such a mapping F is called bi-linear.

Definition 4. Let $((M, +, (\alpha\cdot)_{\alpha\in\mathbb{K}}), |\cdot|_M)$ be a real or complex inner product space. For any finite families $u := (u_i)_{i\in n+1}, v := (v_i)_{i\in n+1} \in M^{n+1}$, $n+1 := \{0, \dots, n\}$, $n \in \mathbb{N}$, in M the square matrix

$$\mathcal{G}_{\langle\cdot|\cdot\rangle_M}(u|v) := \left(\langle u_i|\,v_j\rangle_M\right)_{i,j\in n+1}$$

[1] A normed linear space is an ordered pair of a linear space $(M, +, (\alpha\cdot)_{\alpha\in\mathbb{K}})$ and a norm $|\cdot|_M$. A norm is a functional in $\mathbb{R}^M$ with the properties

- $|\alpha x|_M = |\alpha|\,|x|_M$
- $|x+y|_M \leq |x|_M + |y|_M$
- $|x|_M = 0 \implies x = 0$

for all $x, y \in M$, $\alpha \in \mathbb{K}$, where $\mathbb{K}$ denotes the underlying field for the linear space.

of the pair-wise evaluation of the inner product $\langle\cdot|\cdot\rangle_M$ on F and G is called the **Gramian**[2] of F, G. We have for fixed $(v_i)_{i\in n+1\setminus\{k\}}$ that

$$\begin{aligned}M &\to \mathbb{C}^{(n+1)\times(n+1)}\\ v_k &\mapsto \mathcal{G}_{\langle\cdot|\cdot\rangle_M}(F|\,(v_i)_{i\in n+1})\end{aligned}$$

is a linear mapping, $k\in n+1$. Moreover,

$$\overline{\mathcal{G}_{\langle\cdot|\cdot\rangle_M}(F|G)} = \mathcal{G}_{\langle\cdot|\cdot\rangle_M}(G|F).$$

LEMMA 5. *Let $((M,+,(\alpha\cdot)_{\alpha\in\mathbb{K}}),|\cdot|_M)$ be a real or complex inner-product space and $\langle\cdot|\cdot\rangle_M$ its inner product. Moreover, let $u := (u_i)_{i=0,\ldots,n}$, $v := (v_i)_{i=0,\ldots,n}\in M^{\{0,\ldots,n\}}$, $n\in\mathbb{N}$, be a finite families in M. Then we have for the corresponding Gramian matrix $\mathcal{G}_{\langle\cdot|\cdot\rangle_M}(u|v)$ that*

$$z^*\mathcal{G}_{\langle\cdot|\cdot\rangle_M}(u|v)\,w = \langle z\cdot u\,|\,w\cdot v\rangle_M \text{ for all } z,w\in\mathbb{R}^{n+1} \text{ or } z,w\in\mathbb{C}^{n+1}\text{, respectively.} \tag{1.1.2}$$

Here $z\cdot u$ abbreviates the linear combination

$$z\cdot u := \sum_{i=0}^{n} z_i\cdot u_i \text{ for all } z=(z_0,\ldots,z_n)\in\mathbb{R}^{n+1} \text{ or all } z=(z_0,\ldots,z_n)\in\mathbb{C}^{n+1}\text{, respectively.}$$

REMARK 6. This result shows in particular that $\mathcal{G}_{\langle\cdot|\cdot\rangle_M}(u|u)$ is a selfadjoint, positive semi-definite real or complex matrix. Moreover, $\mathcal{G}_{\langle\cdot|\cdot\rangle_M}(u|u)$ is a selfadjoint, positive definite matrix if and only if $u=(u_i)_{i=0,\ldots,n}$ is linearly independent, i.e.

$$\bigwedge_{z\in\mathbb{K}^{n+1}} z\cdot u = 0 \Rightarrow z = 0.$$

COROLLARY 7. *Let $((H,+,(\alpha\cdot)_{\alpha\in\mathbb{K}}),|\cdot|_H)$ be a real ($\mathbb{K}=\mathbb{R}$) or complex ($\mathbb{K}=\mathbb{C}$) inner product space with inner product $\langle\cdot|\cdot\rangle_H$. Then the Cauchy-Schwarz inequality holds*

$$\bigwedge_{x,\,y\in H} |\langle x|y\rangle_H| \le |x|_H\,|y|_H\,. \tag{1.1.3}$$

Equivalently, we have

$$\bigwedge_{x,\,y\in H} |Re\,\langle x|y\rangle_H| \le |x|_H\,|y|_H\,. \tag{1.1.4}$$

If $\langle\cdot|\cdot\rangle_H$ is an inner product, then the equality sign holds

- *in (1.1.3) if and only if x and y are linearly dependent with coefficients in $\mathbb{K}$,*
- *in (1.1.4) if and only if x and y are linearly dependent with coefficients in $\mathbb{R}$.*

---

[2]Somewhat suggestively we may consider u,v as column matrices and then $\mathcal{G}_{\langle\cdot|\cdot\rangle_M}(u|v)$ as a formal matrix product of $u^\top$ and v, where multiplication is carried out in the sense of the inner product. Using the $\langle\cdot|\cdot\rangle_M$ as the corresponding product symbol we have

$$\begin{aligned}\mathcal{G}_{\langle\cdot|\cdot\rangle_M}((u_0,\ldots,u_n)\,|\,(v_0,\ldots,v_n)) &= \begin{pmatrix}u_0\\ \vdots\\ u_n\end{pmatrix}^\top \langle\cdot|\cdot\rangle_M \begin{pmatrix}v_0\\ \vdots\\ v_n\end{pmatrix}\\ &= \begin{pmatrix}u_0 & \cdots & u_n\end{pmatrix}\langle\cdot|\cdot\rangle_M \begin{pmatrix}v_0\\ \vdots\\ v_n\end{pmatrix}\\ &= \begin{pmatrix}\langle u_0|v_0\rangle_M & \cdots & \langle u_0|v_n\rangle_M\\ \vdots & \ddots & \vdots\\ \langle u_n|v_0\rangle_M & \cdots & \langle u_n|v_n\rangle_M\end{pmatrix}.\end{aligned}$$

PROOF. The result is a special case of the previous considerations. Indeed, taking $F = (x,\, y)$ then

$$\mathcal{G}_{\langle\cdot|\cdot\rangle_H}(F|F) := \begin{pmatrix} \langle x|x\rangle_H & \langle x|y\rangle_H \\ \langle y|x\rangle_H & \langle y|y\rangle_H \end{pmatrix}$$

is positive semi-definite and so it must have a non-negative determinant

$$det\left(\mathcal{G}_{\langle\cdot|\cdot\rangle_H}(F|F)\right) = \langle x|x\rangle_H \langle y|y\rangle_H - \left|\langle x|y\rangle_H\right|^2 \geq 0.$$

This yields (1.1.3). That (1.1.4) is actually equivalent to (1.1.3) becomes clear if we replace y by $\kappa \cdot y$ with $\kappa \in \mathbb{C}$ such that $|\kappa| = 1$ and $\kappa \langle x|y\rangle_H \in \mathbb{R}$. The linear dependence results are immediate from (1.1.2) applied to $\langle\cdot|\cdot\rangle_H$ and $Re\ \langle\cdot|\cdot\rangle_H$, respectively. $\square$

The Cauchy-Schwarz inequality shows that an inner product in an inner product space M induces indeed a norm, since it implies the triangle inequality for $x \mapsto \sqrt{\langle x|x\rangle_M}$.

LEMMA 8. *In an inner-product space $(M,\, +,\, (\alpha\cdot)_{\alpha\in\mathbb{K}},\, |\cdot|_M)$ the* ***parallelogram equality*** *holds:*

$$|x-y|_M^2 + |x+y|_M^2 = 2\left(|x|_M^2 + |y|_M^2\right) \tag{1.1.5}$$

for all $x,\, y \in M$.

DEFINITION 9. An inner-product space $(M,\, +,\, (\alpha\cdot)_{\alpha\in\mathbb{K}},\, |\cdot|_M)$, where $(M,\, +,\, (\alpha\cdot)_{\alpha\in\mathbb{K}},\, |\cdot|_M)$ is a normed linear space, is called an **inner-product space** or a **pre-Hilbert space**. A **Hilbert space** is a complete pre-Hilbert space. If $\mathbb{K} = \mathbb{R}$ then we speak more specifically of a **real inner-product space**, a **real pre-Hilbert space** or a **real Hilbert space**, respectively. If $\mathbb{K} = \mathbb{C}$ then we speak of a **complex inner-product space,** a **complex pre-Hilbert space** or a **complex Hilbert space**, respectively.

Note that every complex inner-product space $(M,\, +,\, (\alpha\cdot)_{\alpha\in\mathbb{C}},\, |\cdot|_M)$ with inner product $\langle\,\cdot\,|\,\cdot\,\rangle_M$ is also a real inner product space by limiting the scalar multiplication to elements in $\mathbb{R}$, i.e. $(M,\, +,\, (\alpha\cdot)_{\alpha\in\mathbb{R}},\, |\cdot|_M)$ with inner product $\mathfrak{Re}\,\langle\,\cdot\,|\,\cdot\,\rangle_M$.

Conversely every real inner product space gives rise to a complex inner product space by a process known as complexification. For a real inner product space X we consider the inner product space $X \oplus X$ (the direct sum), which is the set $X \times X$ with component-wise linear structure and the inner product

$$\langle (x,y)\,|\,(u,v)\rangle_{X\oplus X} := \langle x|u\rangle_X + \langle y|v\rangle_X\,.$$

Via

$$X \to X \oplus \{0\} \subseteq X \oplus X$$
$$x \mapsto (x,0) \equiv \begin{pmatrix} x \\ 0 \end{pmatrix} \equiv x \oplus 0$$

X is isometrically embedded into $X \oplus X$. By extending the linear structure of $X \oplus X$ by letting

$$(\alpha + \mathrm{i}\beta)\cdot \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},\ \alpha,\beta \in \mathbb{R},$$

we introduce in $X \oplus X$ a complex linear structure. The mapping

$$X \oplus X \to (X \oplus X) \oplus (X \oplus X)$$
$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \left(\begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} -y \\ x \end{pmatrix} \right) \equiv \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

is a bijection, indeed, chosing the induced component-wise linear structure and inner product, an isometry. The resulting inner product space

$$\left\{ \begin{pmatrix} x & -y \\ y & x \end{pmatrix} |x, y \in H \right\}$$

is called the complexification $(X \oplus X)_{\mathbb{C}}$ of X. One usually introduces the notation

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix} = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} + \begin{pmatrix} 0 & -y \\ y & 0 \end{pmatrix} =: x + \mathrm{i}y$$

and speaks of x as the real part and y as the imaginary part of $x + \mathrm{i}y$. The inner product of $(X \oplus X)_{\mathbb{C}}$ can be guessed from the binomial formula

$$\langle x + \mathrm{i}y | u + \mathrm{i}v \rangle_{(X \oplus X)_{\mathbb{C}}} = \langle x|u\rangle_{H \oplus H} + \langle y|v\rangle_{H \oplus H} + \mathrm{i}\left(\langle x|v\rangle_{H \oplus H} - \langle y|u\rangle_{H \oplus H} \right)$$

and all needed properties can be shown (exercise!).

Note that this has nothing to do with the issue of x and y being "real" in some sense. Indeed, the complex numbers are a complex Hilbert space with inner product

$$(x, y) \mapsto \overline{x}y.$$

They may, however, also be considered as a – then two-dimensional – real Hilbert space $\mathbb{C}_{\mathbb{R}}$ with inner product

$$(x, y) \mapsto \mathfrak{Re}\left(\overline{x}y\right).$$

Its complexification[3]

$$(\mathbb{C}_{\mathbb{R}} \oplus \mathbb{C}_{\mathbb{R}})_{\mathbb{C}}$$

is then four-dimensional, with complex numbers as real and imaginary part.

The parallelogram equality yields the following remarkable consequence.

PROPOSITION 10. *Let* $\left((M, +, (\alpha\cdot)_{\alpha \in \mathbb{K}}), |\cdot|_M\right)$ *be a pre-Hilbert space or a Hilbert space and* K *a convex subset of* H, *i.e.*

$$\bigwedge_{x,\, y \in K} x + [0,1] \cdot (y - x) := \{x + t \cdot (y - x) \,|t \in [0,1]\} \in K,$$

then for any $x \in M$ *and any sequence* $f := (f_n)_{n \in \mathbb{N}}$ *in* K *such that*

$$|x - f_n|_M \to |x - K|_M := \inf\{|x - y|_M \,|y \in K\}, \tag{1.1.6}$$

---

[3]The space $(\mathbb{C}_{\mathbb{R}} \oplus \mathbb{C}_{\mathbb{R}})_{\mathbb{C}} = \left\{ \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \right\}$ is via

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix} \mapsto \begin{pmatrix} x & -\overline{y} \\ y & \overline{x} \end{pmatrix}$$

unitarily equivalent to the (real) space of quaternions.

Taking a complex inner product space X then $X \oplus X$ equipped with the quaternionic multiplication

$$\begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

would lead to a "quaternification" of X. If X has an additive involution $x \mapsto \overline{x}$, which is compatibel with conjugation in the sense that

$$\overline{\alpha x} = \overline{\alpha}\,\overline{x},$$

then $X \oplus X$ can carry a quaternionic structure by identifying $X \oplus X$ with

$$\left\{ \begin{pmatrix} x & -\overline{y} \\ y & \overline{x} \end{pmatrix} |x, y \in H \right\}.$$

the sequence f is also a Cauchy sequence.

REMARK 11. That the distance $|x-K|_M$ between x and K can be approximated by a sequence f of elements in K is clear by the definition of an infimum.

PROOF. Let f be such an extremal sequence in K, i.e. we have (1.1.6). Then by the parallelogram equation, the convexity of K and the extremal property of $|x-K|_M$ we have

$$\begin{aligned} |f_n-x|_M^2+|f_m-x|_M^2 &= 2\left(\left|\frac{1}{2}\cdot(f_n+f_m)-x\right|_M^2+\left|\frac{1}{2}\cdot(f_n-f_m)\right|_M^2\right), \\ &\geq 2\left(|x-K|_M^2+\left|\frac{1}{2}\cdot(f_n-f_m)\right|_M^2\right), \end{aligned}$$

or

$$\frac{1}{2}\,|f_n-f_m|_M^2 \leq \left(|f_n-x|_M^2-|x-K|_M^2\right)+\left(|f_m-x|_M^2-|x-K|_M^2\right) \longrightarrow 0\; as\; n,m\to\infty.$$

□

For Hilbert spaces we now get

COROLLARY 12. *Let $((M,+,(\alpha\cdot)_{\alpha\in\mathbb{K}}),|\cdot|_M)$ be a Hilbert space and K a convex subset of M. Then there is an element $f_{min}\in\bar{K}$ such that*

$$|x-f_{min}|_M=|x-K|_M\,. \tag{1.1.7}$$

Such an f_{min} is uniquely determined.

PROOF. An extremal sequence f is a Cauchy sequence and has therefore - by completeness and definition of closure - a limit $f_{min}:=\lim f\in\bar{K}$. Continuity yields (1.1.7) from (1.1.6) by taking limits. Let now g_{min} be another element with $|x-g_{min}|_M=|x-K|_M$, then applying the parallelogram equality as in the proof of the previous proposition we get

$$\begin{aligned} |f_{min}-x|_M^2+|g_{min}-x|_M^2 &= 2\left(\left|\frac{1}{2}\cdot(f_{min}+g_{min})-x\right|_M^2+\left|\frac{1}{2}\cdot(f_{min}-g_{min})\right|_M^2\right), \\ &\geq 2\left(|x-K|_M^2+\left|\frac{1}{2}\cdot(f_{min}-g_{min})\right|_M^2\right), \end{aligned}$$

and so

$$\begin{aligned} 2\;|x-K|_M^2 &= |f_{min}-x|_M^2+|g_{min}-x|_M^2\,, \\ &\geq \quad 2\;|x-K|_M^2+\frac{1}{2}\,|f_{min}-g_{min}|_M^2\,, \end{aligned}$$

which in turn implies the desired uniqueness. □

PROPOSITION 13. *The completion[4] of a pre-Hilbert space is a Hilbert space.*

PROOF. The properties of the inner product carries over to the completion. □

DEFINITION 14. Let $H:=((M,+,(\alpha\cdot)_{\alpha\in\mathbb{K}}),|\cdot|_M)$ be a real or complex inner-product space, then two elements $x,\,y\in M$ are called **orthogonal** if

$$\langle x|y\rangle_H=0.$$

In this case we write $x\perp y$. If $x\perp y$ for all $y\in N\subseteq M$ we write $x\perp N$. Moreover, we define the so-called ortho-complement as

$$N^{\perp}:=\{x\in M\,|x\perp N\}\,.$$

[4]The completion of a metric space M is the smallest complete space containing M. This completion can always be constructed, see Appendix A.

REMARK 15. In the real case
$$\langle x|y\rangle_H = 0.$$
is equivalent to
$$|x+y|_M^2 = |x|_M^2 + |y|_M^2. \tag{1.1.8}$$
If H is a complex inner-product space, then for two elements $x,\, y \in M$ to be orthogonal is equivalent to
$$|x+y|_M^2 = |x|_M^2 + |y|_M^2 = |\mathrm{i}\, x + y|_M^2. \tag{1.1.9}$$
Note that equation (1.1.8) can be interpreted as Pythagoras' theorem with equations (1.1.9) as the complex variant.

EXAMPLE 16. Let either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Denote the characteristic function of a set N by χ_N, i.e.
$$\chi_N(t) \coloneqq \begin{cases} 1 \text{ for } t \in N \\ 0 \text{ for } t \notin N \end{cases}.$$
Then consider the characteristic functions $\chi_I \in \mathbb{K}^{\mathbb{R}},$ for I any bounded interval in $\mathbb{R}$, and the generated linear space
$$S(\mathbb{R},\mathbb{K}) \coloneqq Lin_{\mathbb{K}} \left\{\chi_I \,|\, I \subset \mathbb{R} \text{ bounded interval}\right\}$$
which, since $S(\mathbb{R},\mathbb{K}) \subset \mathbb{K}^{\mathbb{R}}$, inherits its linear structure from $\mathbb{K}$ by letting
$$(\alpha \cdot x + y)(t) \coloneqq \alpha\, x(t) + y(t) \text{ for all } t \in \mathbb{R},\ x, y \in \mathbb{K}^{\mathbb{R}},\ \alpha \in \mathbb{K}.$$
With
$$\langle x|y\rangle_{L^2(\mathbb{R},\mathbb{K})} \coloneqq \int_{t\in\mathbb{R}} x(t)^* y(t)\, dt \text{ for } x,\, y \in S(\mathbb{R},\mathbb{K}),$$
we have a inner product on the space of step-functions $S(\mathbb{R},\mathbb{K})$, if we identify step functions s_1, s_2 if their difference $s_1 - s_2$ satisfies
$$\int_{\mathbb{R}} |s_1(t) - s_2(t)|^2\, dt = 0.$$
The completion of the resulting pre-Hilbert space X is a Hilbert space denoted by $L^2(\mathbb{R},\mathbb{K})$. The Lebesgue integration theory shows that the elements of $L^2(\mathbb{R},\mathbb{K})$ can be identified with equivalence classes of (real- or complex-valued) **measurable, square integrable functions** with respect to **almost everywhere** equality as equivalence relation. We abbreviate
$$L^2(\mathbb{R}) \coloneqq L^2(\mathbb{R},\mathbb{C}).$$

EXAMPLE 17. Let either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Then consider the characteristic functions $\chi_I \in \mathbb{K}^{\mathbb{R}^{n+1}}$ for $I = I_0 \times \cdots \times I_n$ with I_k bounded intervals in $\mathbb{R}$, $k = 0, \ldots n$, and the generated linear space
$$S(\mathbb{R}^{n+1},\mathbb{K}) \coloneqq Lin_{\mathbb{K}} \left\{\chi_I \,|\, I = I_0 \times \cdots \times I_n \ \subset \mathbb{R}^{n+1},\ I_k \text{ bounded intervals},\ k = 0, \ldots, n\right\}$$
which, since $S(\mathbb{R}^n,\mathbb{K}) \subset \mathbb{K}^{\mathbb{R}^{n+1}}$, inherits its linear structure from $\mathbb{K}$ by letting
$$(\alpha \cdot x + y)(t) \coloneqq \alpha\, x(t) + y(t) \text{ for all } t \in \mathbb{R}^{n+1},\ x, y \in \mathbb{K}^{\mathbb{R}^{n+1}},\ \alpha \in \mathbb{K}.$$
With
$$\langle x|y\rangle_{L^2(\mathbb{R}^{n+1},\mathbb{K})} \coloneqq \int_{t\in\mathbb{R}^{n+1}} x(t)^* y(t)\, dt \text{ for } x,\, y \in S(\mathbb{R}^{n+1},\mathbb{K}),$$
we have an inner product on the space of step-functions $S(\mathbb{R}^{n+1},\mathbb{K})$, if we identify step functions s_1, s_2 if their difference $s_1 - s_2$ satisfies
$$\int_{\mathbb{R}^{n+1}} |s_1(x) - s_2(x)|^2\, dx = 0.$$
The completion of the resulting pre-Hilbert space X is a Hilbert space denoted by $L^2(\mathbb{R}^{n+1},\mathbb{K})$. Again we note that the Lebesgue integration theory shows that elements of $L^2(\mathbb{R}^{n+1},\mathbb{K})$ can be

identified with equivalence classes of (real- or complex-valued) **measurable, square integrable functions** with respect to **almost everywhere** equality as equivalence relation. We use the abbreviation

$$L^2(\mathbb{R}^{n+1}) \coloneqq L^2(\mathbb{R}^{n+1}, \mathbb{C}).$$

Even without measure theory we can get a more managable access to $L^2(\mathbb{R}^{n+1})$. Let $\mathring{C}_\infty(\mathbb{R}^{n+1})$ be the set of infinitely often differentiable function φ defined on $\mathbb{R}^{n+1}$ with $\varphi = 0$ outside of a bounded closed subset. The smallest such set outside of which $\varphi = 0$ is called the **support** of φ and denoted by supp φ. One says the elements of $\mathring{C}_\infty(\mathbb{R}^{n+1})$ have **compact support**. The set $\mathring{C}_\infty(\mathbb{R}^{n+1})$ is a linear space if interpreted as a linear space of functions. By identifying $\varphi \in \mathring{C}_\infty(\mathbb{R}^{n+1})$ with the equivalence class of Cauchy sequences approximating φ with respect to $|\cdot|_{L^2(\mathbb{R}^{n+1})}$ we obtain an embedding of $\mathring{C}_\infty(\mathbb{R}^{n+1})$ in $L^2(\mathbb{R}^{n+1})$.

EXAMPLE 18. Let M be a set and either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Then consider the characteristic functions $\chi_{\{t\}}$, $t \in M$, and

$$X \coloneqq Lin_{\mathbb{K}} \left\{ \chi_{\{t\}} \,|\, t \in M, \right\}$$

with the induced linear structure of $\mathbb{K}^M$. With

$$\langle x|y\rangle_{\ell^2} \coloneqq \sum_{t\in M} \overline{x(t)}\, y(t) \text{ for } x,\, y \in X$$

we have an inner product on X. The completion of the resulting pre-Hilbert space X is a Hilbert space denoted by $\ell^2(M, \mathbb{K})$. The complex Hilbert space $\ell^2(M, \mathbb{C})$ will be denoted by $\ell^2(M)$.

PROPOSITION 19. *Every linear subspace $V \subseteq W$ of a pre-Hilbert space W is a pre-Hilbert space (with $|\cdot|_V = (|\cdot|_W)\big|_V$). If V is closed and W is a Hilbert space then V is also a Hilbert space. Let A be an arbitrary subset of W then $A^\perp$ is a closed linear subspace. If W is a Hilbert space then $A^\perp$ is also a Hilbert space. In any case, we have*

$$A^\perp = (Lin_{\mathbb{K}}\, A)^\perp = \overline{Lin_{\mathbb{K}}\, A}^\perp. \tag{1.1.10}$$

Leaving the proof of this proposition as an exercise we conclude this section and turn our attention to Hilbert spaces.

1.2. Hilbert Spaces

1.2.1. Fundamental Results. A central result of Hilbert space theory is the so-called projection theorem. The following is a first variant.

THEOREM 20. *(Projection Theorem 1) Let H be a real or complex Hilbert space and C a closed subspace of H. Then for any $x \in H$ there is a unique $y \in C$ (called the **orthogonal projection** of x on C) such that*

$$|x - y|_H = |x - C|_H. \tag{1.2.1}$$

This element $y \in C$ is characterized by

$$x - y \perp C.$$

PROOF. The unique existence of a $y \coloneqq y_{min} \in C$ of minimal distance is already known. We need to show the characterization in terms of orthogonality. It also suffices to discuss the complex case. From the extremal properties of $|x - C|_H$ we have

$$\bigwedge_{z\in C} |x - y|_H^2 = |x - C|_H^2 \leq |x - y - z|_H^2. \tag{1.2.2}$$

From $|x - y|_H^2 \leq |x - y - z|_H^2$ we get

$$0 \leq |z|_H^2 - 2\, \mathfrak{Re} \langle x - y|z\rangle_H .$$

Let now $z = \varepsilon \cdot (\omega \cdot u)$ where $\varepsilon \in \mathbb{R}$, $\omega \in \mathbb{C}$, $|\omega| = 1$, $u \in C$, then we see that the quadratic polynomial function in the real variable ε

$$\varepsilon \mapsto \varepsilon^2 |u|_H^2 - 2\,\varepsilon\, \mathfrak{Re}\, \langle x - y \,|\, \omega \cdot u \rangle_H$$

can only be non-negative if

$$\mathfrak{Re}\, \langle x - y \,|\, \omega \cdot u \rangle_H = 0$$

Choosing $\omega = \imath$ and $\omega = 1$ shows that

$$\langle x - y \,|\, u \rangle_H = 0,$$

where $u \in C$ was arbitrary. Thus, we have

$$x - y \perp C. \tag{1.2.3}$$

Conversely, let $y \in C$ be such that (1.2.3) holds. By reverting the above reasoning we obtain again (1.2.2) from which the claim (1.2.1) is obvious. $\square$

Definition 21. A set $o \subseteq H$ is called an **orthonormal set**, if

$$\bigwedge_{u,v \in o} \langle u \,|\, v \rangle_H = \begin{cases} 1 \text{ for } u = v \\ 0 \text{ for } u \neq v \end{cases}.$$

If the cardinality $\#o$ is finite, countable (i.e. $\#o = \#\mathbb{N}$) or non-countable, then o is called a finite, countable or non-countable orthonormal set, respectively. An orthonormal set is called **complete** if the linear span $Lin_{\mathbb{K}}\, o := \{u \in H \,| u \text{ (finite) linear combination of } o\}$ is dense in H, i.e.

$$H = \overline{Lin_{\mathbb{K}}\, o}$$

and $\mathbb{K}$ is the field of scalars of H. In the latter case o is also known as a **Hilbert space basis**.

Example 22. Let M be a set and either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. The Hilbert space $\ell^2(M, \mathbb{K}) := \overline{\left(Lin_{\mathbb{K}} \left\{ \chi_{\{t\}} \,|\, t \in M, \right\},\, f \mapsto \sqrt{\sum_{t \in M} |f(t)|^2} \right)}$ defined as the completion of the pre-Hilbert space $\left(Lin_{\mathbb{K}} \left\{ \chi_{\{t\}} \,|\, t \in M, \right\},\, f \mapsto \sqrt{\sum_{t \in M} |f(t)|^2} \right)$ has $\left\{ \chi_{\{t\}} \middle|\, t \in M \right\}$ as complete orthonormal set.

Remark 23. If a finite family $(u_i)_{i=1,\ldots,n}$, $n \in \mathbb{N}$, of linearly independent elements of a real or complex Hilbert space H with scalar field $\mathbb{K}$ is given, then the projection theorem yields an orthonormal family $(v_i)_{i=1,\ldots,n}$ in the following way: Letting $v_1 := \frac{1}{|u_1|_H} \cdot u_1$, pick *the* normalized element $v_{k+1} \in Lin_{\mathbb{K}} \{u_i | i = 1, \ldots, k+1\}$ orthogonal to the linear subspace $Lin_{\mathbb{K}} \{u_i | i = 1, \ldots, k\}$ satisfying $\langle u_{k+1} | v_{k+1} \rangle_H > 0$. That v_{k+1} is indeed uniquely determined can be seen in the following way. Let $w_{k+1} \in Lin_{\mathbb{K}} \{u_i | i = 1, \ldots, k+1\} \cap Lin_{\mathbb{K}} \{u_i | i = 1, \ldots, k\}^{\perp}$ be normalized then $v_{k+1} = \alpha \cdot w_{k+1}$ for some $\alpha \in \mathbb{C}$ with $|\alpha| = 1$. We have $\langle u_{k+1} | v_{k+1} \rangle_H = \alpha \, \langle u_{k+1} | w_{k+1} \rangle_H$ and making this positive indeed determines α uniquely. By this iterative construction we get

$$Lin_{\mathbb{K}} \{u_i | i = 1, \ldots, k\} = Lin_{\mathbb{K}} \{v_i | i = 1, \ldots, k\} \text{ for all } k \in \mathbb{N}$$

and $\{v_i | i = 1, \ldots, n\}$ an orthonormal set. The procedure, known as **E.Schmidt orthonormalization**, is recursive and extends in the obvious way to countable families of linearly independent elements of H.

Theorem 24. *Let H be a real or complex Hilbert space. Then there is an orthonormal set o such that*

$$H = \overline{Lin_{\mathbb{K}}\, o}$$

holds.

Definition 25. A Hilbert space H is called **finite-dimensional** if it has a finite complete orthonormal set. A Hilbert space H is called **separable** if it has a countable complete orthonormal set.

REMARK 26. The cardinality $\#o$ of a complete orthonormal set o in a Hilbert space H is an invariant for H.

As a consequence of the last theorem we have that apparently any element f of a Hilbert space H can be approximated by a (finite) linear combination of a complete orthonormal set o to any degree of accuracy.

For any fixed number of elements of o used for such an approximation there is indeed a best approximation. The usefulness of orthonormal sets lies to a good deal in the simplicity of finding this best approximation.

THEOREM 27. *Let H be a real or complex Hilbert space with scalar field $\mathbb{K}$ and o a complete orthonormal set in H. Then we have for any finite subset $o_{fin} \subseteq o$ that*

$$\left| x - \sum_{u \in o_{fin}} \langle u|x\rangle \cdot u \right|_H = \left| x - \overline{Lin_{\mathbb{K}}\, o_{fin}} \right|_H .$$

PROOF. Since o_{fin} is finite we have $Lin_{\mathbb{K}}\, o_{fin} = \overline{Lin_{\mathbb{K}}\, o_{fin}}$. That a best approximation in $Lin_{\mathbb{K}}(o_{fin})$ exists follows therefore from the projection theorem. Let $z = \sum_{u \in o_{fin}} \alpha_u \cdot u$ be this best approximation, then we must have

$$x - \sum_{u \in o_{fin}} \alpha_u \cdot u \perp o_{fin}.$$

or for all $v \in o_{fin}$

$$\left\langle v \left| x - \sum_{u \in o_{fin}} \alpha_u \cdot u \right. \right\rangle_H = \langle v|x\rangle_H - \alpha_v = 0.$$

$\square$

LEMMA 28. *Let H be a Hilbert space and o a complete orthonormal set. For any $x \in H$ there is an at most countable orthonormal set $o_{(x)} \subseteq o$ such that*

$$x \in \overline{Lin_{\mathbb{K}}\, o_{(x)}}.$$

PROOF. Since

$$x \in \overline{Lin_{\mathbb{K}}\, o}$$

for a complete orthonormal set o, we have that x can be approximated by (finite) linear combinations of o. Let $o_{1/(n+1)} \subseteq o$, $n \in \mathbb{N}$, be a finite orthonormal set in the Hilbert space H such that

$$\left| x - Lin_{\mathbb{K}}\, o_{1/(n+1)} \right|_H < \frac{1}{n+1} .$$

Clearly, $B_n := \bigcup_{k=0}^{n} o_{1/(k+1)} \subseteq o$ and so B_n is (as a subset of an orthonormal set) an orthonormal set. We certainly have also

$$|x - Lin_{\mathbb{K}}\, B_n|_H < \frac{1}{n+1}.$$

Since any of the B_n is finite, either $o = B_n$ for some $n \in \mathbb{N}$ or the orthonormal system $o_x := \bigcup_{k=0}^{\infty} o_{1/(k+1)} \subseteq o$ is countable – as a union of finite sets – and we have

$$|x - Lin_{\mathbb{K}}\, o_x|_H = 0$$

and so as desired

$$x \in \overline{Lin_{\mathbb{K}}\, o_x}.$$

$\square$

LEMMA 29. *(Bessel's inequality and Parseval's equality) Let H be a Hilbert space. For any finite orthonormal set o_{fin} we have*

$$\left|\sum_{u\in o_{fin}} \langle u|x\rangle_H \cdot u\right|_H^2 = \sum_{u\in o_{fin}} \left|\langle u|x\rangle_H\right|^2 \leq |x|_H^2 \tag{1.2.4}$$

for all $x \in H$. Let $o_{\mathbb{N}}$ be a countable orthonormal set and $(u_i)_{i\in\mathbb{N}}$ an enumeration of $o_{\mathbb{N}}$ then

$$\left|\sum_{i=0}^{\infty} \langle u_i|x\rangle_H \cdot u_i\right|_H^2 = \sum_{i=0}^{\infty} \left|\langle u_i|x\rangle_H\right|^2 \leq |x|_H^2. \tag{1.2.5}$$

If o is a complete orthonormal set, then for any $x \in H$ there is a finite or countable orthonormal system o_x (enumerated as $(v_i)_{i\in\mathbb{N}}$) such that

$$\left|\sum_{i} \langle v_i|x\rangle_H \cdot v_i\right|_H^2 = \sum_{i} \left|\langle v_i|x\rangle_H\right|^2 = |x|_H^2.$$

PROOF. We have by orthonormality of o_{fin}that

$$\begin{aligned} 0 &\leq \left|x - \textstyle\sum_{u\in o_{fin}} \langle u|x\rangle_H \cdot u\right|_H^2, \\ &= |x|_H^2 - 2\,Re\left\langle \textstyle\sum_{u\in o_{fin}} \langle u|x\rangle_H \cdot u \middle| x\right\rangle_H + \left|\textstyle\sum_{u\in o_{fin}} \langle u|x\rangle_H \cdot u\right|_H^2, \\ &= |x|_H^2 - \textstyle\sum_{u\in o_{fin}} \left|\langle u|x\rangle_H\right|^2, \end{aligned} \tag{1.2.6}$$

and

$$\left|\sum_{u\in o_{fin}} \langle u|x\rangle_H \cdot u\right|_H^2 = \sum_{u\in o_{fin}} \left|\langle u|x\rangle_H\right|^2.$$

This proves (1.2.4). From (1.2.4) follows

$$\left|\sum_{i=n}^{m} \langle u_i|x\rangle_H \cdot u_i\right|_H^2 = \sum_{i=n}^{m} \left|\langle u_i|x\rangle_H\right|^2 \leq |x|_H^2 \tag{1.2.7}$$

for all $m, n \in \mathbb{N}, m \geq n$. Thus, we have from the convergence of the numerical series $\sum_{i=0}^{\infty} \left|\langle u_i|x\rangle_H\right|^2$ that $\left(\sum_{i=0}^{n} \langle u_i|x\rangle_H \cdot u_i\right)_{n\in\mathbb{N}}$ is a Cauchy sequence. by completeness the existence of

$$\lim_{n\to\infty} \sum_{i=0}^{n} \langle u_i|x\rangle_H \cdot u_i =: \sum_{i=0}^{\infty} \langle u_i|x\rangle_H \cdot u_i$$

follows. From (1.2.7) now follows (1.2.5) with $n = 0$ and letting $m \to \infty$. Finally, let o be a complete orthonormal set and o_x and B_n, $n \in \mathbb{N}$, as constructed in the proof of the previous lemma. Let $(v_i)_i$ be an enumeration of o_x such that

$$\bigwedge_{n\in\mathbb{N}} \bigvee_{N\in\mathbb{N}} B_n = Lin_{\mathbb{K}}\{v_i \,|\, i = 1, \ldots, N\},$$

($\mathbb{K}$ the underlying scalar field,) then according to Theorem 27 and by Lemma 28 we have

$$\left|x - \sum_{u\in B_n} \langle u|x\rangle_H \cdot u\right|_H = \left|x - \overline{Lin_{\mathbb{K}}\, B_n}\right|_H \to 0 \text{ as } n \to \infty.$$

Letting $n \to \infty$ we get

$$\left|x - \sum_{i} \langle v_i|x\rangle_H \cdot v_i\right|_H = \left|x - \overline{Lin_{\mathbb{K}}\, o_x}\right|_H = 0,$$

where the sum is finite if H is finite dimensional or countable if H is infinite dimensional. □

REMARK 30. The result is apparently independent of the enumeration since the limiting element $x = \sum_i \langle v_i|x\rangle_H \cdot v_i$ is always the same. Moreover, since x is orthogonal to all $u \in o \setminus o_x$, we may write

$$x = \sum_{u \in o} \langle u|x\rangle_H \cdot u,$$

realizing that at most countably many of the coefficient $\langle u|x\rangle_H$, $u \in o$, are nonzero.

THEOREM 31. *(Fischer-Riesz theorem 1) Let H be an infinite dimensional, real or complex Hilbert space with scalar field $\mathbb{K}$. Then an orthonormal set o is complete if and only if*

$$\bigwedge_{x \in H} \sum_{u \in o} |\langle u|x\rangle_H|^2 = |x|_H^2. \tag{1.2.8}$$

Moreover, for any sequence $(\alpha_i)_{i\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ with $\sum_{i=0}^{\infty} |\alpha_i|^2 < \infty$ we have

$$x := \sum_{i=0}^{\infty} \alpha_i \cdot u_i \in H$$

and

$$\alpha_i = \langle u_i|x\rangle_H , \; i = 0, 1, 2, \ldots,$$

for any enumeration $(u_i)_{i\in\mathbb{N}}$ of a countable orthonormal set $o_{\mathbb{N}}$. In any case, we have

$$\left| x - \sum_{u \in o} \langle u|x\rangle_H \cdot u \right|_H^2 = |x|_H^2 - \sum_{u \in o} |\langle u|x\rangle_H|^2 . \tag{1.2.9}$$

PROOF. Equality (1.2.9) follows from (1.2.6) with $o_{fin} = B_n$ and letting $n \to \infty$. If o is a complete orthonormal set then by the previous lemma we have (1.2.8). Conversely, if (1.2.8) holds, then with (1.2.9) we obviously have

$$x \in \overline{Lin_{\mathbb{K}}\, o}.$$

Since $x \in H$ was arbitrary we have completeness, i.e.

$$H = \overline{Lin_{\mathbb{K}}\, o}.$$

Finally, let $(\alpha_i)_{i\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ with $\sum_{i=0}^{\infty} |\alpha_i|^2 < \infty$ and $(v_i)_{i\in\mathbb{N}}$ an enumeration of a countable orthonormal set $o_{\mathbb{N}}$. Observe that by orthonormality

$$\left| \sum_{i=n}^{m} \alpha_i \cdot v_i \right|_H^2 = \sum_{i=n}^{m} |\alpha_i|^2 ,$$

we have the existence of

$$x := \sum_{i=0}^{\infty} \alpha_i \cdot u_i \in H.$$

By continuity we have

$$\left\langle u_j \middle| \sum_{i=0}^{n} \alpha_i \cdot u_i \right\rangle \to \alpha_j \text{ as } n \to \infty.$$

and

$$\left\langle u_j \middle| \sum_{i=0}^{n} \alpha_i \cdot u_i \right\rangle \to \langle u_j|x\rangle_H \text{ as } n \to \infty.$$

By uniqueness of limit we obtain indeed

$$\alpha_i = \langle u_i|x\rangle_H , \; i \in \mathbb{N}.$$

□

A one-to-one and onto, norm-preserving, linear mapping between two Hilbert spaces is called a **unitary** mapping. In this case the two Hilbert spaces are called **unitarily equivalent**.With this notion we can reformulate our last result conveniently. Indeed,$\{\chi_{\{t\}} \mid t \in M, \}$ is a complete orthonormal set in $\ell^2(M, \mathbb{K})$. The cardinality of this orthonormal set coincides with $\# M$.

Now, Theorem 31 can be reformulated in the following way.

Corollary 32. *(Fischer-Riesz theorem 2) Let H be a real or complex Hilbert space with scalar field $\mathbb{K}$. Then an orthonormal set o is complete if and only if the mapping*

$$\begin{aligned} H &\longrightarrow \ell^2(o, \mathbb{K}) \\ x &\longmapsto (\langle u|x\rangle_H)_{u\in o} \end{aligned}$$

is unitary.

Remark 33. If $\# o = n \in \mathbb{N}$ then $\ell^2(o, \mathbb{K})$ can be replaced by $\ell^2(n, \mathbb{K}) \equiv \mathbb{K}^n$, $n := \{0, \dots, n-1\}$, by introducing an enumeration. Likewise, if $\# o = \# \mathbb{N}$ then $\ell^2(o, \mathbb{K})$ can be replaced by $\ell^2(\mathbb{N}, \mathbb{K})$ (or $\ell^2(\mathbb{Z}, \mathbb{K})$ or $\ell^2(\mathbb{Z}^n, \mathbb{K})$, $n \in \mathbb{N}$, depending on preferences and/or convenience).

Example 34. Let either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. The Hilbert space $L^2(\mathbb{R}, \mathbb{K})$ has by construction the set S of step functions as a dense set. The following set h of step functions (the so-called **Haar basis**[5]) is a complete orthonormal set:

$$h := \{\, \phi_{n,m} \,|\, n, m \in \mathbb{Z}\}$$

with

$$\phi_{n,m}(t) := 2^{n/2}\Phi(2^n t - m)$$

where

$$\Phi := 2\,\chi_{[0,1/2[} - \chi_{[0,1[}.$$

Indeed, we find

$$\begin{aligned} \langle \phi_{n,m} \,|\, \phi_{n,m}\rangle_{L^2(\mathbb{R},\mathbb{K})} &= \textstyle\int_{t\in\mathbb{R}} 2^n \; |\Phi(2^n t - m)|^2 \; dt, \\ &= \textstyle\int_{t\in\mathbb{R}} |\Phi(t)|^2 \; dt, \\ &= \textstyle\int_{t\in\mathbb{R}} \left|\chi_{[0,1[}(t)\right|^2 \; dt, \\ &= 1, \end{aligned}$$

and by a similar simplification we only need to show that $\langle \phi_{0,0} \,|\, \phi_{n,m}\rangle_{L^2(\mathbb{R},\mathbb{K})} = 0$ for $n \in \mathbb{N}$ and $m \in \mathbb{Z}$ not both equal to zero. First we find

$$\langle \phi_{0,0} \,|\, \phi_{n,m}\rangle_{L^2(\mathbb{R},\mathbb{K})} = \int_{t\in\mathbb{R}} 2^{n/2}\Phi(t)\,\Phi(2^n t - m)\,dt.$$

Since $\Phi \equiv 0$ outside of $[0,1[$, the integrand is equal to zero if $t \notin [0,1[$ or $2^n t - m \notin [0,1[$. So assume $t \in [0,1[$ and $2^n t - m \in [0,1[$, i.e.

$$\max(0, 2^{-n}m) \le t < \min(1,\, 2^{-n}(m+1)). \tag{1.2.10}$$

For $-m \in \mathbb{N}$ this describes the empty set. For $m \in \mathbb{N}$ the inequality (1.2.10) simplifies to

$$2^{-n}m \le t < \min(1,\, 2^{-n}(m+1)).$$

If $2^{-n}(m+1) \ge 1$ then $2^{-n}m$ must be less than 1 and so

$$m < 2^n \le m+1. \tag{1.2.11}$$

This implies (excluding the case $m = n = 0$) that in this case $m = 2^n - 1$ and $n \in \mathbb{N}$, i.e.

$$\frac{1}{2} \le 1 - 2^{-n} \le t < 1 \text{ for } n \in \mathbb{N} \text{ and } m = 2^n - 1. \tag{1.2.12}$$

[5]The Haar basis is the earliest known example of what became much later known as "wavelets".

In this case we have (since $\Phi(t) = -1$ on $[1/2,\, 1[$)

$$(1.2.13)\qquad \begin{aligned} \langle \phi_{0,0} \,|\, \phi_{n,m}\rangle_{L^2(\mathbb{R},\mathbb{K})} &= -\int_{t\in[1-2^{-n},\,1[} 2^{n/2}\Phi(2^n t - 2^n + 1)\,dt,\\ &= -\int_{t\in[0,\,1[} 2^{-n/2}\Phi(t)\,dt,\\ &= 0. \end{aligned}$$

If $2^{-n}(m+1) < 1$ then (1.2.10) turns into

$$2^{-n}m \leq t < 2^{-n}(m+1).$$

Since $[2^{-n}m,\, 2^{-n}(m+1)[\; \subseteq \big[0,\, 2^{-1}\big[$ or $[2^{-n}m,\, 2^{-n}(m+1)[\; \subseteq \big[2^{-1},\, 1\big[$ we have

$$\Phi(t) = \pm 1 \text{ on } \big[2^{-n}m,\, 2^{-n}(m+1)\big[$$

and we get similar to (1.2.13)

$$(1.2.14)\qquad \begin{aligned} \langle \phi_{0,0} \,|\, \phi_{n,m}\rangle_{L^2(\mathbb{R},\mathbb{K})} &= \pm \int_{t\in[2^{-n}m,\,2^{-n}(m+1)[} 2^{n/2}\,\Phi(2^n t - m)\,dt,\\ &= \pm \int_{t\in[0,\,1[} 2^{-n/2}\Phi(t)\,dt,\\ &= 0. \end{aligned}$$

Thus, orthonormality of h is shown. To see completeness we first note that sums of characteristic functions of the form $\chi_{[2^{-n}m,\,2^{-n}(m+1)[}$ are sufficient to approximate a characteristic function of an arbitrary interval I. This is due to the density of $\{2^{-n}m \,|\, n \in \mathbb{N} \wedge m \in \mathbb{Z}\}$ in $\mathbb{R}$. By rescaling and translation it becomes clear that it suffices to approximate just $\chi_{[0,\,1[}$. The latter, however, can be seen explicitly by noting that

$$\sum_{n\in\mathbb{N}} 2^{-(n+1)/2}\phi_{-n,0}$$

actually represents $\chi_{[0,\,1[}$. According to (1.2.9), for this it is enough to see that

$$\sum_{n\in\mathbb{N}} 2^{-n-1} = 1.$$

The Fischer-Riesz variant 32 (together with remark 33) now yields the unitarity of the mapping

$$\begin{aligned} I_{L^2(\mathbb{R},\mathbb{K})} :\; L^2(\mathbb{R},\mathbb{K}) &\longrightarrow \ell^2(\mathbb{Z}^2,\mathbb{K}),\\ x &\longmapsto \Big(\langle \phi_{n,m} | x\rangle_{L^2(\mathbb{R},\mathbb{K})}\Big)_{n,m\in\mathbb{Z}}. \end{aligned}$$

In particular, we found

$$I_{L^2(\mathbb{R},\mathbb{K})}(\chi_{[0,1[}) = (\gamma_{n,m})_{n,m\in\mathbb{Z}}$$

with

$$\gamma_{n,m} := \begin{cases} 2^{-(n+1)/2} & \text{for } n \in \mathbb{N} \text{ and } m = 0\\ 0 & \text{otherwise} \end{cases}.$$

The rather remarkable consequence of Corollary 32 stating the unitary equivalence of any Hilbert space with an ℓ^2-type space seems to say that everything is known about the structure of Hilbert spaces. This is indeed the case, but as a matter of convenience it may be quite inadvisable to replace a particular Hilbert space H by its unitarily equivalent ℓ^2-space, since e.g. certain elements of H may have rather complicated representations in the ℓ^2- space. The characteristic function $\chi_{[0,1[}$ of the previous example and its representation $(\gamma_{n,m})_{n,m\in\mathbb{Z}}$ may serve as a first indication. Depending on circumstances it may also be of interest to choose a particular orthonormal set rather than just any. We will become more aware of how little is said by the structure result of the Fischer-Riesz theorem as we investigate linear mappings between Hilbert spaces.

EXAMPLE 35. The functions $\chi_{]-1/2,+1/2[}\, exp(2\pi\mathrm{i}\, n\cdot)$, $n\in\mathbb{Z}$, form a complete orthonormal set in the closed subspace

$$L^2\left(]-1/2,+1/2[\right) := \overline{\left\{ f\in S(\mathbb{R},\mathbb{C}) \,\middle|\, \bigvee_{C\subseteq]-1/2,+1/2[\text{ closed}} f=0 \text{ on } \mathbb{R}\setminus C \right\}}.$$

Orthonormality can be seen by direct calculation:

$$\begin{aligned}\langle \chi_{]-1/2,+1/2[}\, exp(2\pi\mathrm{i}\, m\cdot) | \chi_{]-1/2,+1/2[}\, exp(2\pi\mathrm{i}\, n\cdot)\rangle_{L^2(]-1/2,+1/2[)} &= \\ &= \textstyle\int_{\mathbb{R}} \chi_{]-1/2,+1/2[}(t)\, exp(-2\pi\mathrm{i}\, m\, t)\, \chi_{]-1/2,+1/2[}(t)\, exp(2\pi\mathrm{i}\, n\, t)\, dt, \\ &= \textstyle\int_{-1/2}^{1/2} exp(2\pi\mathrm{i}\,(n-m)\,t)\,dt, \\ &= \begin{cases} \frac{exp(\pi\mathrm{i}\,(n-m))-exp(-\pi\mathrm{i}\,(n-m))}{2\pi\mathrm{i}\,(n-m)} = 0 & \text{for } m\neq n \\ 1 & \text{for } m=n \end{cases}.\end{aligned}$$

The completeness of the orthonormal set can be shown. In particular, we have

$$f = \sum_{k\in\mathbb{Z}} \left\langle \chi_{]-1/2,+1/2[}\, exp(2\pi\mathrm{i}\, kn\cdot) \middle| \, f\right\rangle_{L^2(]-1/2,+1/2[)} \chi_{]-1/2,+1/2[}\; exp(2\pi\mathrm{i}\, k\cdot),$$

as the Fourier expansion of $f\in L^2(]-1/2,+1/2[)$. The coefficients

$$\begin{aligned}\left\langle \chi_{]-1/2,+1/2[}\, exp(2\pi\mathrm{i}\, kn\cdot) \middle| \, f\right\rangle_{L^2(]-1/2,+1/2[)} &= \\ &= \langle exp(2\pi\mathrm{i}\, k\cdot) |\, f\rangle_{L^2(]-1/2,+1/2[)}\,, \\ &= \langle exp(2\pi\mathrm{i}\, k\cdot) |\, f\rangle_{L^2(\mathbb{R})}\,,\end{aligned}$$

are the so-called Fourier coefficients of f.

EXAMPLE 36. Let us consider the well-known Gauss distribution function $\gamma\in L^2(\mathbb{R})$ given by $\gamma(x) = exp(-x^2/2)$, $x\in\mathbb{R}$. Letting formally $\mathcal{D} := \frac{1}{\sqrt{2}}\,(m-\partial)$, we shall see that $\{\Gamma_k \,|\, k\in\mathbb{N}\}$, $\Gamma_k := \frac{1}{|\mathcal{D}^k\gamma|_{L^2(\mathbb{R})}}\mathcal{D}^k\gamma$, is an orthonormal set. We first recognize that it is a set of (normalized) linearly independent elements in $L^2(\mathbb{R})$. Infact, $\mathcal{D}^k\gamma$ is of the form $Q_k\,\gamma$, where Q_k is a polynomial[6] of degree k with leading coefficient $2^{k/2}$. This is true for $k=0$ and by the product rule we see

$$\mathcal{D}^{k+1}\gamma = \mathcal{D}\,(Q_k\gamma) = \frac{1}{\sqrt{2}}\,(Q_k'\,\gamma + 2\,m\,Q_k\,\gamma)$$

from which the claim follows by induction for all $k\in\mathbb{N}$. Letting formally $\mathcal{D}^* := \frac{1}{\sqrt{2}}\,(\partial+m)$ we find

$$\mathcal{D}^*\gamma = 0 \tag{1.2.15}$$

and

$$\mathcal{D}^*\mathcal{D} - \mathcal{D}\,\mathcal{D}^* = 1. \tag{1.2.16}$$

Moreover, we calculate by repeated integration by parts that

$$\left\langle \mathcal{D}^k\gamma \,|\, \mathcal{D}^j\gamma\right\rangle_{L^2(\mathbb{R})} = \left\langle \gamma \,|\, \mathcal{D}^{*k}\mathcal{D}^j\gamma\right\rangle_{L^2(\mathbb{R})} \tag{1.2.17}$$

for k, $j\in\mathbb{N}$. Let now $k\in\mathbb{N}$ be given, then we want to show that

$$\mathcal{D}^{*k}\mathcal{D}^j\gamma = 0 \text{ for all } j<k,\ j\in\mathbb{N}.$$

[6]The polynomial Q_k is up to normalization the $k-$th Hermite polynomial. Frequently it is assumed that the leading coefficient should be $+1$ in which case the $k-$th Hermite polynomial would be $\frac{1}{2^{k/2}}Q_k$, $k\in\mathbb{N}$.

That the result holds for $j = 0$ is obvious from (1.2.15). Let now $j < k$ and assume the result already holds for smaller j. We find

$$\begin{aligned}
\mathcal{D}^{*k}\mathcal{D}^{j}\gamma &= \mathcal{D}^{*(k-1)}\mathcal{D}\mathcal{D}^{*}\mathcal{D}^{(j-1)}\gamma+ \\
&\quad +\mathcal{D}^{*(k-1)}\mathcal{D}^{(j-1)}\gamma, \\
&= \mathcal{D}^{*(k-1)}\mathcal{D}\mathcal{D}^{*}\mathcal{D}^{(j-1)}\gamma, \\
&= \mathcal{D}^{*(k-2)}\mathcal{D}^{*}\mathcal{D}\mathcal{D}^{*}\mathcal{D}^{(j-1)}\gamma \\
&= \mathcal{D}^{*(k-2)}\mathcal{D}\mathcal{D}^{*2}\mathcal{D}^{(j-1)}\gamma+ \\
&\quad +\mathcal{D}^{*(k-1)}\mathcal{D}^{(j-1)}\gamma, \\
&= \mathcal{D}^{*(k-2)}\mathcal{D}\mathcal{D}^{*2}\mathcal{D}^{(j-1)}\gamma.
\end{aligned}$$

By iterating this step we get

$$\mathcal{D}^{*k}\mathcal{D}^{j}\gamma = \mathcal{D}\mathcal{D}^{*k}\mathcal{D}^{(j-1)}\gamma$$

and since $\mathcal{D}^{*k}\mathcal{D}^{(j-1)}\gamma = 0$ by induction assumption the claimed result follows. With (1.2.17) the orthogonality of the set $\left\{ \mathcal{D}^{k}\gamma \middle| \, k \in \mathbb{N} \right\}$ follows. An additional normalization yields that $\{ \Gamma_k |\, k \in \mathbb{N} \}$ is an orthonormal set. This orthonormal set is also complete.

Theorem 37. *(Projection Theorem 2) Let H be a real or complex Hilbert space and C a closed subspace of H. Then we have*

$$H = C \oplus C^{\perp} \tag{1.2.18}$$

in the sense of unitary equivalence.

Proof. Clearly, C and $C^{\perp}$ are Hilbert spaces and so is (by definition) $C \oplus C^{\perp}$ we only need to establish the unitary equivalence with H. Let $x \in H$ be given. According to Theorem 20 there is precisely one $z \in C$ such that $|x - z|_H = |x - C|_H$. Then in particular

$$\begin{aligned}
P_C : H &\longrightarrow H \\
x &\longmapsto \in \{u \in C \,|\, |x-u|_H = |x - C|_H\}
\end{aligned} \tag{1.2.19}$$

is a well-defined mapping. Since also from Theorem 20 we know that $x - P_C(x) \in C^{\perp}$ and $P_C(x) \in C$, we have a resulting mapping

$$\begin{aligned}
U_C : H &\longrightarrow C \oplus C^{\perp} \\
x &\longmapsto (P_C(x),\, x - P_C(x))
\end{aligned}.$$

To see that U_C is a linear mapping it suffices to show that P_C is linear. Consider now $v := \alpha \cdot P_C(x) + P_C(y) \in C$, for $\alpha \in \mathbb{K}$, $x, y \in H$. We shall show that v is best approximation of $\alpha \cdot x + y$, then by uniqueness $P_C(\alpha \cdot x + y) = v = \alpha \cdot P_C(x) + P_C(y)$ follows. According to Theorem 20 for this we only need to see that

$$\alpha \cdot x + y - v \in C^{\perp}.$$

This, however, is clear since

$$x - P_C(x) \in C^{\perp} \text{ and } y - P_C(y) \in C^{\perp}$$

and $C^{\perp}$ is a linear space. Thus we have that P_C and so U_C is linear. The linear mapping U_C is also norm-preserving:

$$|U_C(x)|_{C \oplus C^{\perp}} = \sqrt{|P_C(x)|_H^2 + |x - P_C(x)|_H^2} = |x|_H \text{ for all } x \in H.$$

The latter equality follows by noting that in particular $(x - P_C(x)) \perp P_C(x)$. That U_C is one-to-one follows directly from norm-preservation (and linearity). To see that U_C is also onto let

$(u,v) \in C \oplus C^{\perp}$. Define $x := u + v$ then $x - u = v \in C^{\perp}$ and by uniqueness we have $u = P_C(x)$ and $v = x - P_C(x)$, i.e. $U_C(x) = (u,v)$ - as needed. □

REMARK 38. The linear mapping P_C introduced in (1.2.19) is called the **orthogonal projector** associated with C. Writing 1 for the identity mapping, we see that

$P_C\,[H] = C,$

$(1 - P_C)\,[H] = C^{\perp},$

and by uniqueness of the best approximation

$P_C \circ P_C = P_C\,.$

The unitary equivalence expressed in (1.2.18) can be used to identify both sides in which case we may interpret (1.2.18) as what is known as an **orthogonal decomposition** of H. Indeed, from the proof we have that every $x \in H$ can be written in a unique way as $x = u + v$ with $u \in C$ and $v \in C^{\perp}$.

In general an orthogonal projector is defined in terms of the properties noted in remark 38.

DEFINITION 39. Let H be a Hilbert space. Then a linear mapping $P : H \to H$ satisfying

$$P \circ P = P$$

is called **projector** onto $P\,[H]$. A projector is called an **orthogonal projector** if in addition

$$P\,[H] \perp (1 - P)\,[H]\,.$$

Together with P_C constructed above also $(1 - P_C)$ is an orthogonal projector (namely onto $C^{\perp}$). Indeed,

$$(1 - P_C) \circ (1 - P_C) = 1 \circ 1 - P_C \circ P_C = 1 - P_C = P_{C^{\perp}}\,.$$

We note the following important consequence of the projection Theorem 37.

LEMMA 40. *Let H be a Hilbert space with scalar field $\mathbb{K}$ and $M \subseteq H$ a subset. Then*

$$\overline{Lin_{\mathbb{K}} M} = M^{\perp\perp}.$$

Moreover, if $N \subseteq M$ then

$$M^{\perp} \oplus (M^{\perp\perp} \cap N^{\perp}) = N^{\perp}.$$

PROOF. According to Theorem 37 we have

$$H = \overline{Lin_{\mathbb{K}} M} \oplus \overline{Lin_{\mathbb{K}} M}^{\perp}$$

and

$$H = M^{\perp} \oplus M^{\perp\perp}. \tag{1.2.20}$$

We note that with (1.2.3) we have

$$H = \overline{Lin_{\mathbb{K}} M} \oplus M^{\perp}. \tag{1.2.21}$$

A careful comparison of these two decompositions shows the desired equality. □

1.2.2. Constructions of Hilbert Spaces. We have already seen that the process of completion is of central interest in constructing Hilbert spaces, since mostly we have only simple description of a semi- or pre-Hilbert space as a starting point. Another way to obtain Hilbert spaces is to consider closed subspaces of given ones, e.g. $M^{\perp}$ for M a subset. More interesting are processes which create new Hilbert spaces out of other Hilbert spaces.

For sake of simplicity we shall henceforth assume that the underlying scalar field $\mathbb{K} = \mathbb{C}$, since the real case is usually easy to obtain by the same argument allowing only real scalars instead of complex ones. In other words, we shall assume all our Hilbert spaces to be complex Hilbert spaces.

There are three main constructions which we shall encounter more closely: first we re-consider the direct sum $\oplus$ more closely, then the construction of Hilbert spaces based on dualisation.

DEFINITION 41. Let $\mathcal{HS}$ denote the class of all real or complex Hilbert spaces with scalar field $\mathbb{K}$ defined by $\mathcal{HS} := \{H|\, H$ Hilbert space with scalar field $\mathbb{K}\}$ and let $H.$, i.e. $x \mapsto H_x$, be a mapping in $\mathcal{HS}^M$, M a set. The Cartesian product $\underset{x\in M}{\times} H_x$ of $H.$ becomes a linear space by defining the component-wise linear structure

$$(\alpha \cdot v + w)(t) = \alpha \cdot v(t) + w(t) \in H_t \text{ for all } t \in M,\ \alpha \in \mathbb{K},\ v, w \in \underset{x\in M}{\times} H_x .$$

Let W be the subspace generated by all $w \in \underset{x\in M}{\times} H_x$ such that $w(t) = 0 \in H_t$ for all $t \in M$ but one. Then W equipped with the inner product

$$\langle v|w\rangle_{\oplus} := \sum_{t\in M} \langle v(t)\,|\,w(t)\rangle_{H_t} \text{ for all } v, w \in Lin_{\mathbb{K}} W \tag{1.2.22}$$

is a pre-Hilbert space. Its completion is called the **direct sum** of $H.$, denoted by

$$\bigoplus_{x\in M} H_x \,.$$

If $M = \{0, \dots, n\}$, $n \in \mathbb{N}$, we also write

$$H_0 \oplus \cdots \oplus H_n$$

for the direct sum of $(H_0, \dots, H_n)$. The elements $(x_i)_{i=0,\dots,n} \equiv (x_0, \dots, x_n)$ of such a **finite direct sum** will occasionally also be denoted by $x_0 \oplus \cdots \oplus x_n$ or as a column matrix

$$\begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix}.$$

The linearity of the structure is easily checked. Note also that the sum in (1.2.23) is indeed a finite sum, since except for finitely many terms the inner products are zero. If M is a finite set then the pre-Hilbert space constructed here is already complete. The simplicity of the construction of such a direct sum is also reflected in the simplicity to obtain an orthonormal set in $\bigoplus_{x\in M} H_x$ from orthonormal sets in the spaces $H_t,\ t \in M$. Indeed, if o_t is a (complete) orthonormal set in Hilbert space H_t, $t \in M$, then the set of all $\omega \in \underset{x\in M}{\times} o_x$ with $\omega(t) = 0$ for all but one $t \in M$ is a complete orthonormal set in $\bigoplus_{x\in M} H_x$.

$\mathbb{C}^n$ considered as the Hilbert space $\ell^2(n, \mathbb{C})$ is a direct sum of n copies of $\mathbb{C}$ (as Hilbert space $\ell^2(1, \mathbb{C})$):

$$\mathbb{C}^n = \bigoplus_{t=0,\dots,n-1} \mathbb{C}.$$

More generally, we see that by construction e.g.

$$\ell^2(M) = \bigoplus_{t\in M} \mathbb{C},\ M \text{ an arbitrary set.}$$

Let $(H_0, \ldots, H_n)$ be a finite family of Hilbert spaces with common scalar field $\mathbb{K}$. Then $H_0 \times \cdots \times H_n$ with the component-wise linear structure and

equipped with the inner product

$$\langle v|w\rangle_{\oplus} := \sum_{t=0}^{n} \langle v_t \,|\, w_t\rangle_{H_t} \text{ for all } v = (v_0, \ldots, v_n), w = (w_0, \ldots, w_n) \in H_0 \times \cdots \times H_n \tag{1.2.23}$$

is a Hilbert space.

It remains to show that $H_0 \times \cdots \times H_n$ as the described pre-Hilbert space is already complete. By definition of the inner product we have

$$|(v_0, \ldots, v_n)|_{\oplus} = \sqrt{|v_0|^2_{H_0} + \cdots |v_n|^2_{H_n}} \text{ for all } v_j \in H_j,\ j = 0, \ldots, n.$$

So let $(w_j)_j$ be a Cauchy sequence in $H_0 \times \cdots \times H_n$, i.e. $w_j = (w_{j0}, \ldots, w_{jn}) \in H_0 \times \cdots \times H_n$, then $(w_{jk})_j$ is a Cauchy sequence in H_k, $k = 0, \ldots, n$. By completeness of H_k we have $w_{\infty k} := \lim_{j\to\infty} w_{jk} \in H_k$, $k = 0, \ldots, n$, and $w_j \to (w_{\infty 0}, \ldots, w_{\infty n})$ as $j \to \infty$.

Next we discuss another way to construct a new Hilbert space denoted by H' as the so-called dual of a given Hilbert space H. For this construction, let first

$$H' := \left\{ f \in \mathbb{C}^H \,\middle|\, f \text{ linear} \wedge f \text{ Lipschitz continuous} \right\}.$$

Since any mapping into the numbers is called a **functional**, the elements of H' are referred to as **continuous, linear functionals**. On H' we do *not* use the linear structure induced by $\mathbb{C}$, we rather define the following modified linear structure:

$$(\alpha f + g)(t) := \overline{\alpha} f(t) + g(t) \text{ for all } \alpha \in \mathbb{C},\ f, g \in H',\ t \in H. \tag{1.2.24}$$

Thus, H' becomes a linear space.

THEOREM 42. *(Riesz representation theorem) Let H be a complex Hilbert space. For any $f \in H'$ there is a unique $w \in H$ such that*

$$f(x) = \langle w|x\rangle_H \text{ for all } x \in H.$$

PROOF. If $f \equiv 0$ then $w \in H$ must also be the zero element of H. Otherwise, we choose

$$w = \overline{f(x_0)} \cdot x_0$$

for a normalized $x_0 \in ([\{0\}]\, f)^{\perp}$. Here $[\{0\}]\, f := \{x \in H \,|\, f\,(x) = 0\}$ denotes the null space or kernel of f. Uniqueness is clear since a difference of two such representers would have to be orthogonal to H. □

DEFINITION 43. Let H be a complex Hilbert space. The mapping

$$\begin{aligned} R_H : H' &\longrightarrow H \\ f &\longmapsto \in \left(\left\{w \in H \,\middle|\, \textstyle\bigwedge_{x\in H} f(x) = \langle w|x\rangle_H \right\}\right) \end{aligned}$$

is called **Riesz operator** or **Riesz mapping**.

The Riesz mapping enjoys the following properties:

LEMMA 44. *Let H be a complex Hilbert space. The Riesz mapping $R_H : H' \to H$ is a linear mapping and norm-preserving in the sense that*[7]

$$|R_H f|_H = |f|_{Lip} \text{ for all } f \in H'. \tag{1.2.25}$$

[7]Here $|\cdot|_{\mathrm{Lip}}$ denotes the best, i.e. the smallest Lipschitz constant.

PROOF. The linearity follows straightforwardly from the definition of R_H an the properties of the inner product

$$\begin{aligned} \langle R_H(\alpha f + g) \,|\, y\rangle_H &= (\alpha f + g)(y), \\ &= \overline{\alpha} f(y) + g(y), \\ &= \overline{\alpha} \langle R_H f \,|\, y\rangle_H + \langle R_H g \,|\, y\rangle_H , \\ &= \langle \alpha \cdot R_H f + R_H g \,|\, y\rangle_H , \end{aligned}$$

for all f, $g \in H'$, $y \in H$ and $\alpha \in \mathbb{C}$.

Moreover, by the Cauchy-Schwarz inequality we have for all $y \in H$

$$|f(y)| = |\langle R_H f \,|\, y\rangle_H| \le |R_H f|_H \, |y|_H \tag{1.2.26}$$

and so with the linearity of f

$$|f|_{Lip} \le |R_H f|_H \,.$$

Since for $y = R_H f$ equality occurs in (1.2.26), we see that $|R_H f|_H$ is also the best possible constant, i.e. (1.2.25) holds. □

From (1.2.25) we also learn that $|\cdot|_{Lip}$ actually is a pre-Hilbert space norm, since the left-hand side is clearly a pre-Hilbert space norm. Thus, we found that H' equipped with this inner product is a pre-Hilbert space.

LEMMA 45. *Let H be a complex Hilbert space. The Riesz mapping $R_H : H' \to H$ is onto.*

PROOF. Let $y \in H$ be arbitrary. The linear functional $f_y : H \to \mathbb{C}$, $x \mapsto \langle y \,|\, x\rangle_H$, is clearly linear and again by the Cauchy-Schwarz inequality also Lipschitz continuous. Indeed, as in the proof of the previous lemma we have

$$|\langle y \,|\, \cdot\rangle_H|_{Lip} = |y|_H \,.$$

Thus, we have $f_y \in H'$. Moreover, by definition

$$f_y(x) = \langle y \,|\, x\rangle_H = \langle R_H f_y \,|\, x\rangle_H \text{ for all } x \in H$$

and therefore

$$R_H f_y = y. \tag{1.2.27}$$

□

As a consequence we have

THEOREM 46. *Let H be a complex Hilbert space. Then H' is also a Hilbert space (the so-called* ***dual space*** *to H) and the Riesz mapping $R_H : H' \to H$ is unitary.*

PROOF. Given our previous results we only need to show completeness of H'. So, let $(f_n)_n$ be a Cauchy sequence in H'. Then, since R_H is norm-preserving, we also have that $(R_H f_n)_n$ is a Cauchy sequence in H. By the completeness of H we have

$$y := \lim_{n\to\infty} R_H f_n \in H.$$

According to (1.2.27) we $R_H f_y = y$ for $f_y : H \to \mathbb{C}$, $x \mapsto \langle y \,|\, x\rangle_H$. Thus, we find (again by norm-preservation) that from $R_H f_n \to y$ as $n \to \infty$ we obtain

$$f_n \to f_y \text{ as } n \to \infty.$$

Since R_H is norm-preserving it is automatically one-to-one and by the above indeed a unitary mapping with respect to the inner product

$$(f, g) \mapsto \langle R_H f | \, R_H g\rangle_H \,.$$

□

In the following we shall always make use of the implied possibility to identify H with $(H')'$. This implies in particular that $x \in H$ may always be interpreted as a continuous linear functional on H', so that

$$x\,(f) = \overline{f\,(x)}$$

for all $x \in H$ and $f \in H'$. Note that

$$\begin{aligned} x\,(\alpha \cdot f) &= \alpha\, x\,(f) \\ &= \overline{(\alpha \cdot f)\,(x)} \\ &= \overline{\overline{\alpha} f\,(x)} \\ &= \alpha\, \overline{f\,(x)} \\ &= \overline{f\,(\overline{\alpha} \cdot x)} \\ &= (\overline{\alpha} \cdot x)\,(f) \end{aligned}$$

for all $\alpha \in \mathbb{K}$.
In some instances it may also be useful to even identify H with H'. We shall come to this later.

A final construction principle for new Hilbert spaces we shall explore is the so-called tensor product of Hilbert spaces. It can be rooted on the concept of multi-linear forms on Hilbert spaces. Let $(H_i)_{i=0,\ldots,n-1}$ be a family of n (complex) Hilbert spaces. A mapping $\mu :\, H_0 \times \cdots \times H_{n-1} \longrightarrow \mathbb{C}$ is called continuous $n-$linear form on $H_0 \times \cdots \times H_{n-1}$ if $y \mapsto \mu(\cdots, x_{i-1}, y, x_{i+1}, \ldots) \in H_i'$ for $x_k \in H_k$ for $k = 0, \ldots, n-1$, $k \neq i$. The implied linear structure for such $n-$linear forms f, g is given by

$$\bigwedge_{\alpha \in \mathbb{C},\, x \in H_0 \times \cdots \times H_{n-1}} (\alpha\, f + g)(x) := \overline{\alpha}\, f(x) + g(x).$$

The set of special continuous $n-$linear forms $x_0 \otimes \cdots \otimes x_{n-1}$ on $H_0 \times \cdots \times H_{n-1}$ defined by

$$x_0 \otimes \cdots \otimes x_{n-1}(u_0, \ldots, u_{n-1}) := \langle x_0 | u_0 \rangle_{H_0} \cdots \langle x_{n-1} | u_{n-1} \rangle_{H_{n-1}}$$

generate the linear space

$$W_\otimes\left((H_i)_{i=0,\ldots,n-1}\right) := Lin_{\mathbb{C}}\, \{x_0 \otimes \cdots \otimes x_{n-1} |\, x_i \in H_i,\, i = 0, \ldots, n-1\}.$$

Defining

$$\langle x_0 \otimes \cdots \otimes x_{n-1} \,|\, u_0 \otimes \cdots \otimes u_{n-1} \rangle_{H_0 \otimes \cdots \otimes H_{n-1}} := \langle x_0 | u_0 \rangle_{H_0} \cdots \langle x_{n-1} | u_{n-1} \rangle_{H_{n-1}} \tag{1.2.28}$$

for all $x_i, u_i \in H_i$, $i = 0, \ldots, n-1$, we obtain a candidate for an inner product for $W_\otimes$ by sesqui-linear extension. Since the values may depend on the representation, this process may a-priori merely create a sesqui-linear relation, i.e. a relation, which is linear w.r.t. the second component and conjugate linear w.r.t. the first component. From the properties we shall demonstrate it follows, however, that we have actually defined a sesqui-linear functional on $W_\otimes\left((H_i)_{i=0,\ldots,n-1}\right)$. Since sesqui-linearity is assumed by construction, we only need to confirm symmetry, non-negativity and definiteness. Consider arbitrary linear combinations $\sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i}$ and $\sum_j \beta_j\, u_{0,j} \otimes \cdots \otimes u_{n-1,j}$ then using sesqui-linearity we obtain

$$\begin{aligned} &\left\langle \textstyle\sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \,\middle|\, \textstyle\sum_j \beta_j\, u_{0,j} \otimes \cdots \otimes u_{n-1,j} \right\rangle_{H_0 \otimes \cdots \otimes H_{n-1}} = \\ &\qquad = \textstyle\sum_i \sum_j \alpha_i^* \beta_j\, \langle x_{0,i} \otimes \cdots \otimes x_{n-1,i} \,|\, u_{0,j} \otimes \cdots \otimes u_{n-1,j} \rangle_{H_0 \otimes \cdots \otimes H_{n-1}}. \end{aligned}$$

Using (1.2.28) this yields

$$\begin{aligned} &\left\langle \textstyle\sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \,\middle|\, \textstyle\sum_j \beta_j\, u_{0,j} \otimes \cdots \otimes u_{n-1,j} \right\rangle_{H_0 \otimes \cdots \otimes H_{n-1}} = \\ &\qquad = \textstyle\sum_i \sum_j \overline{\alpha_i} \beta_j\, \langle x_{0,i} \,|\, u_{0,j} \rangle_{H_0} \cdots \langle x_{n-1,i} \,|\, u_{n-1,j} \rangle_{H_{n-1}} \end{aligned} \tag{1.2.29}$$

From the latter symmetry follows. Indeed,

$$\begin{aligned}\overline{\left\langle \sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \,\middle|\, \sum_j \beta_j\, u_{0,j} \otimes \cdots \otimes u_{n-1,j} \right\rangle_{H_0 \otimes \cdots \otimes H_{n-1}}} &= \\ = \sum_i \sum_j \alpha_i\, \overline{\beta_j}\, \langle u_{0,j} \,|\, x_{0,i} \rangle_{H_0} \cdots \langle u_{n-1.j} \,|\, x_{n-1,i} \rangle_{H_{n-1}}\,, \\ = \left\langle \sum_j \beta_j\, u_{0,j} \otimes \cdots \otimes u_{n-1,j} \,| \sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \right\rangle_{H_0 \otimes \cdots \otimes H_{n-1}}.\end{aligned}$$

From (1.2.29) we also get

$$\begin{aligned}\left\langle \sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \,\middle|\, \sum_j \alpha_j\, x_{0,j} \otimes \cdots \otimes x_{n-1,j} \right\rangle_{H_0 \otimes \cdots \otimes H_{n-1}} &= \\ = \sum_i \sum_j \overline{\alpha_i} \alpha_j\, \langle x_{0,i} \,|\, x_{0,j} \rangle_{H_0} \cdots \langle x_{n-1,i} \,|\, x_{n-1,j} \rangle_{H_{n-1}}\,.\end{aligned} \tag{1.2.30}$$

Let $(A^{(k)}_{ij})_{i,j}$ be a non-negative root of the (non-negative and selfadjoint) Gramian $\left(\langle x_{k,i} \,|\, x_{k,j} \rangle_{H_k} \right)_{i,j}$, $k = 0, \ldots, n-1$, then we obtain from (1.2.30)

$$\begin{aligned}\left\langle \sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \,\middle|\, \sum_j \alpha_j\, x_{0,j} \otimes \cdots \otimes x_{n-1,j} \right\rangle_{H_0 \otimes \cdots \otimes H_{n-1}} &= \\ = \sum_i \sum_j \sum_{s_0} \cdots \sum_{s_n} \overline{\alpha_i} \alpha_j\, A^{(0)}_{is_0} A^{(0)}_{s_0 j} \cdots A^{(n-1)}_{is_{n-1}} A^{(n-1)}_{s_{n-1} j}\,, \\ = \sum_{s_0} \cdots \sum_{s_{n-1}} \left(\sum_i \alpha_i\, A^{(0)}_{s_0 i} \cdots A^{(n-1)}_{s_{n-1} i} \right)^* \left(\sum_j \alpha_j\, A^{(0)}_{s_0 j} \cdots A^{(n-1)}_{s_{n-1} j} \right).\end{aligned}$$

The last term is, however, the square of the norm of w in $\ell_2(\mathbb{N}^n)$, where

$$w_{s_0 \cdots s_{n-1}} := \sum_i \alpha_i\, A^{(0)}_{s_0 i} \cdots A^{(n-1)}_{s_{n-1} i}$$

for $s = (s_0, \ldots, s_{n-1}) \in \mathbb{N}^n$, and so

$$\left\langle \sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \,\middle|\, \sum_j \alpha_j\, x_{0,j} \otimes \cdots \otimes x_{n-1,j} \right\rangle_{H_0 \otimes \cdots \otimes H_{n-1}} \geq 0.$$

Next, we want to show definiteness of the semi-inner product $\langle \cdot \,|\, \cdot \rangle_{H_0 \otimes \cdots \otimes H_{n-1}}$. So let

$$\left| \sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \right|_{H_0 \otimes \cdots \otimes H_{n-1}} = 0,$$

then from the Cauchy Schwarz inequality follows that

$$\left\langle \sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \,|\, u_0 \otimes \cdots \otimes u_{n-1} \right\rangle_{H_0 \otimes \cdots \otimes H_{n-1}} = 0,$$

for all $u_i \in H_i,\ i = 0, \ldots, n-1$. This, however, is by definition

$$0 = \sum_i \overline{\alpha_i}\, \langle x_{0,i} | u_0 \rangle_{H_0} \cdots \langle x_{n-1,i} | u_{n-1} \rangle_{H_{n-1}} = \left(\sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \right) (u_0, \ldots, u_{n-1})$$

for all $u_i \in H_i,\ i = 0, \ldots, n-1$. In other words the $n-$linear form

$$\sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} = 0.$$

From the Cauchy-Schwarz inequality we may finally conclude independence of the chosen representation. Indeed, if $\sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} = \sum_j \beta_i\, y_{0,i} \otimes \cdots \otimes y_{n-1,i}$ then

$$\left\langle \sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \mid u_0 \otimes \cdots \otimes u_{n-1} \right\rangle_{H_0 \otimes \cdots \otimes H_{n-1}} + \\ - \left\langle \sum_j \beta_i\, y_{0,i} \otimes \cdots \otimes y_{n-1,i} \mid u_0 \otimes \cdots \otimes u_{n-1} \right\rangle_{H_0 \otimes \cdots \otimes H_{n-1}} = \\ = \left(\sum_i \alpha_i\, x_{0,i} \otimes \cdots \otimes x_{n-1,i} \right) (u_0, \cdots, u_{n-1}) - \left(\sum_j \beta_i\, y_{0,i} \otimes \cdots \otimes y_{n-1,i} \right) (u_0, \cdots, u_{n-1}), \\ = 0.$$

With this we can now define the tensor product of Hilbert spaces.

DEFINITION 47. Let $(H_i)_{i=0,\dots,n-1}$ be a family of n (complex) Hilbert spaces. The completion of the pre-Hilbert space $W_\otimes((H_i)_{i=0,\dots,n-1}) := Lin_{\mathbb{C}}\{x_0 \otimes \cdots \otimes x_{n-1} | \, x_i \in H_i,\ i = 0, \dots, n-1\}$ with respect to the norm $|\cdot|_{H_0 \otimes \cdots \otimes H_{n-1}}$ is called the **tensor product** of $(H_i)_{i=0,\dots,n-1}$ and denoted by $\bigotimes_{i=0,\dots,n-1} H_i$ or by $H_0 \otimes \cdots \otimes H_{n-1}$. Let V_i be a subspace of H_i, $i = 0, \dots, n-1$. The pre-Hilbert space $W_\otimes((V_i)_{i=0,\dots,n-1})$ is called the **algebraic tensor product** of $(V_i)_{i=0,\dots,n-1}$ and denoted by $\overset{a}{\bigotimes}_{i=0,\dots,n-1} V_i$ or $V_0 \overset{a}{\otimes} \cdots \overset{a}{\otimes} V_{n-1}$.

The completion of a normed linear or pre-Hilbert space of real- or complex-valued functions is often referred to as a function Banach space or a function Hilbert space, respectively.

EXAMPLE 48. $\ell_2(\mathbb{Z}^n) = \bigotimes_{k=0,\dots,n-1} \ell_2(\mathbb{Z})$

Recalling that the elements of $V_i := \ell_2(\mathbb{Z})$ are complex number sequences, we realize that with $M_i = \mathbb{Z}$ we encounter an instance of the last proposition. The correspondence

$$\left(z_{i_0} \cdots z_{i_{n-1}}\right)_{(i_0,\dots,i_{n-1}) \in \mathbb{Z}^n} \mapsto (z_{i_0})_{i_0 \in \mathbb{Z}} \otimes \cdots \otimes (z_{i_{n-1}})_{i_{n-1} \in \mathbb{Z}}$$

extends to a linear isometry I. Since the elements generating $\ell_2(\mathbb{Z}^n)$ (according to its definition) have product form

$$\begin{aligned} \chi_{\{(i_0,\dots,i_{n-1})\}}(j_0, \dots, j_{n-1}) &= \chi_{\{i_0\}}(j_0) \cdots \chi_{\{i_{n-1}\}}(j_{n-1}) \\ &= \langle \chi_{\{i_0\}} | \chi_{\{j_0\}} \rangle_{\ell^2(\mathbb{Z})} \cdots \langle \chi_{\{i_{n-1}\}} | \chi_{\{j_{n-1}\}} \rangle_{\ell^2(\mathbb{Z})} \\ &= \left(\chi_{\{i_0\}} \otimes \cdots \otimes \chi_{\{i_{n-1}\}}\right) \left(\chi_{\{j_0\}}, \dots, \chi_{\{j_{n-1}\}}\right) \end{aligned}$$

for $(i_0, \dots, i_{n-1})$, $(j_0, \dots, j_{n-1}) \in \mathbb{Z}^n$, we see that I extends indeed to a unitary mapping $\bar{I}$ from $\ell_2(\mathbb{Z}^n)$ onto $\bigotimes_{k=1,\dots,n} \ell_2(\mathbb{Z})$.

EXAMPLE 49. $L_2(\mathbb{R}^n) = \bigotimes_{k=0,\dots,n-1} L_2(\mathbb{R})$

This result follows by the same arguments as the previous example. For this we first need to notice that it is sufficient to use characteristic functions $\chi_{I_0 \times \cdots \times I_{n-1}}$ with $I_k = [a_k, b_k[$, $a_k < b_k$, $k = 0, \dots, n-1$, in order to generate $L_2(\mathbb{R}^n)$, since

$$\left|\chi_{I_0 \times \cdots \times I_{n-1}} - \chi_{J_0 \times \cdots \times J_{n-1}}\right|_{L_2(\mathbb{R}^n)} = 0$$

for $J_k =]a_k, b_k[$, $J_k =]a_k, b_k]$ or $J_k = [a_k, b_k]$, $k = 0, \dots, n-1$. The linear span of these restricted characteristic functions is however a function space. Moreover, we have

$$\chi_{I_0 \times \cdots \times I_{n-1}}(t_0, \dots, t_{n-1}) = \chi_{I_0}(t_0) \cdots \chi_{I_{n-1}}(t_{n-1}) \text{ for } (t_0, \dots, t_{n-1}) \in \mathbb{R}^n,$$

from which the claim follows in the same way as for the previous example by observing that

$$\begin{aligned}\left\langle \chi_{I_0\times\cdots\times I_{n-1}} | \chi_{J_0\times\cdots\times J_{n-1}} \right\rangle_{L^2(\mathbb{R}^n)} &= \langle \chi_{I_0} | \chi_{J_0} \rangle_{L^2(\mathbb{R})} \cdots \left\langle \chi_{I_{n-1}} | \chi_{J_{n-1}} \right\rangle_{L^2(\mathbb{R})} \\ &= \left(\chi_{I_0} \otimes \cdots \otimes \chi_{I_{n-1}}\right)\left(\chi_{J_0}, \ldots, \chi_{J_{n-1}}\right).\end{aligned}$$

EXAMPLE 50. Similar to the previous examples it can be seen that in general the completion $L_2(\mathbb{R}, H)$ of $H-$valued step functions is just $L_2(\mathbb{R}) \otimes H$ for every Hilbert space H.

We also have the following useful density result:

LEMMA 51. *Let $(H_i)_{i=0,\ldots,n-1}$ be a family of n (complex) Hilbert spaces and $(S_i)_{i=0,\ldots,n-1}$ a corresponding family of respective subsets. If $V_i := Lin_{\mathbb{C}} S_i$ is dense in H_i for $i = 0, \ldots, n-1$, then*

$$V_0 \overset{a}{\otimes} \cdots \overset{a}{\otimes} V_{n-1} = Lin_{\mathbb{C}}\{x_0 \otimes \cdots \otimes x_{n-1} \,|\, x_i \in S_i \text{ for } i = 0, \ldots, n-1\}$$

is dense in $H_0 \otimes \cdots \otimes H_{n-1}$.

PROOF. It suffices to approximate **decomposable elements**, i.e. elements of the form $x_0 \otimes \cdots \otimes x_{n-1} \in H_0 \otimes \cdots \otimes H_{n-1}$. Let now $v_i := (v_{i,j})_j$ be a sequence approximating $x_i \in H_i$, $i = 0, \ldots, n-1$. Then we can estimate

$$\begin{aligned}&|x_0 \otimes \cdots \otimes x_{n-1} - v_{0,j} \otimes \cdots \otimes v_{n-1,j}|_{H_0\otimes\cdots\otimes H_{n-1}} \leq \\ &\qquad \leq \textstyle\sum_{k=0}^{n-1} |v_{0.,j} \otimes \cdots \otimes v_{0.k-1j} \otimes (x_k - v_{k,,j}) \otimes x_{k+1} \otimes \cdots \otimes x_{n-1}|_{H_0\otimes\cdots\otimes H_{n-1}}, \\ &\qquad \leq \textstyle\sum_{k=0}^{n-1} (|v_{0,j}| \cdots |v_{0,k-1}|\, |x_k - v_{k,,j}|_{H_k}\, |x_{k+1}|_{H_{k+1}} \cdots |x_{n-1}|_{H_{n-1}}).\end{aligned}$$

Since the last term is a finite sum composed of norms with bounded factors and one factor tending to zero, the right-hand side goes to zero. Thus we have

$$v_{0,j} \otimes \cdots \otimes v_{n-1,j} \to x_0 \otimes \cdots \otimes x_{n-1} \text{ as } j \to \infty.$$

□

Based on this lemma we can now address the issue of generating orthonormal sets in tensor products from the knowledge of orthonormal sets in the "factors".

PROPOSITION 52. *Let $(H_i)_{i=0,\ldots,n-1}$ be a family of n (complex) Hilbert spaces and $(o_i)_{i=0,\ldots,n-1}$ a corresponding family of respective orthonormal sets. Then*

$$[o_0] \otimes \cdots \otimes [o_{n-1}] := \left\{x_0 \otimes \cdots \otimes x_{n-1} \,\middle|\, x_i \in o_i \text{ for } i = 0, \ldots, n-1\right\}$$

is an orthonormal set in $H_0 \otimes \cdots \otimes H_{n-1}$. Moreover, if o_i is complete in H_i for $i = 0, \ldots, n-1$, then $[o_0] \otimes \cdots \otimes [o_{n-1}]$ is a complete orthonormal set in $H_0 \otimes \cdots \otimes H_{n-1}$.

PROOF. Let $x_0 \otimes \cdots \otimes x_{n-1}$, $y_0 \otimes \cdots \otimes y_{n-1} \in o_0 \otimes \cdots \otimes o_{n-1}$ then

$$\begin{aligned}\left\langle x_0 \otimes \cdots \otimes x_{n-1} | y_0 \otimes \cdots \otimes y_{n-1} \right\rangle &= \langle x_0 | y_0 \rangle_{H_0} \cdots \langle x_{n-1} | y_{n-1} \rangle_{H_{n-1}} \\ &= \begin{cases} 1 \text{ if } x_0 = y_0, \cdots, x_{n-1} = y_{n-1} \\ 0 \qquad \text{otherwise} \end{cases},\end{aligned}$$

which shows that indeed $[o_0] \otimes \cdots \otimes [o_{n-1}]$ is an orthonormal set in $H_0 \otimes \cdots \otimes H_{n-1}$. Recalling that for an orthonormal set o_k to be complete means that $V_k := Lin_{\mathbb{C}}\, o_k$ is dense in H_k, $k = 0, \ldots n-1$, the rest of the proposition follows from the previous lemma. □

Since closed linear operators between Hilbert spaces are also just closed linear subspaces of Hilbert spaces, it would be natural to define the tensor product of linear operators as tensor products of these subspaces. As for the direct sum there is, however, a more useful (algebraic) concept of a tensor product for linear operators $A_k \subseteq H_{0k} \oplus H_{1k}$, $k = 0, \dots, n-1$, defined as the linear relation given by

$$(A_0 \overset{\dot a}{\otimes} \cdots \overset{\dot a}{\otimes} A_{n-1})(x_0 \otimes \cdots \otimes x_{n-1}) := A_0 x_0 \otimes \cdots \otimes A_{n-1} x_{n-1} \tag{1.2.31}$$

for all $x_k \in D(A_k)$, $k = 0, \dots, n-1$. The overset ’$\cdot$’ is again used as a reminder of an ’image-wise’ algebraic tensor product[8]. That (1.2.31) defines indeed a linear operator

$$(A_0 \overset{\dot a}{\otimes} \cdots \overset{\dot a}{\otimes} A_{n-1}) : D(A_0) \overset{a}{\otimes} \cdots \overset{a}{\otimes} D(A_{n-1}) \subseteq H_{00} \otimes \cdots \otimes H_{0(n-1)} \to H_{10} \otimes \cdots \otimes H_{1(n-1)}$$

and not just a linear relation is not obvious and needs a careful consideration. We must assure that different representations of an element do not lead to different ’images’. To show right-uniqueness let $(0, w) \in A_0 \overset{\dot a}{\otimes} \cdots \overset{\dot a}{\otimes} A_{n-1}$, i.e. $w = \sum_i \alpha_i A_0 x_{0,i} \otimes \cdots \otimes A_{n-1} x_{n-1,i}$ and $\sum_i \alpha_i x_{0,i} \otimes \cdots \otimes x_{n-1,i} = 0$. The latter is the same as

$$0 = \sum_i \alpha_i \left\langle x_{0,i} | v_0 \right\rangle_{H_{00}} \cdots \left\langle x_{n,i} | v_n \right\rangle_{H_{0(n-1)}}$$

for all $v_k \in H_{1k}$, $k = 0, \dots, n-1$. If $(x_{k,i})_i$ is linearly independent for $k = 0, \dots, n-1$, then all coefficients α_i must be zero. In this case $w = 0$. Let now $(y_{k,j})_j$ be a maximal linearly independent subfamily of $(x_{k,i})_i$ then

$$x_{k,i} = \sum_j a_{k,ij} \, y_{k,j}$$

for suitable coefficients $a_{k,ij}$. Then

$$\sum_i \alpha_i \, x_{0,i} \otimes \cdots \otimes x_{n-1,i} = \sum_i \sum_{j_0} \cdots \sum_{j_{n-1}} \alpha_i \, a_{0,ij_0} \cdots a_{n-1,ij_n} \, y_{0,j_0} \otimes \cdots \otimes y_{n-1,j_{n-1}} = 0$$

and due to linear independence

$$\sum_i \alpha_i \, a_{0,ij_0} \cdots a_{n-1,ij_{n-1}} = 0$$

for all multi-indices $(j_0, \dots, j_{n-1})$ appearing in the sum. Trivially this yields

$$\sum_i \sum_{j_0} \cdots \sum_{j_{n-1}} \alpha_i \, a_{0,ij_0} \cdots a_{n-1,ij_{n-1}} \, A_0 y_{0,j_0} \otimes \cdots \otimes A_{n-1} y_{n-1,j_{n-1}} = 0$$

and by the linearity of the operators $(A_k)_{k=0,\dots,n-1}$

$$\begin{aligned} 0 &= \sum_i \alpha_i \, A_0 \sum_{j_0} a_{0,ij_0} \, y_{0,j_0} \otimes \cdots \otimes A_{n-1} \sum_{j_{n-1}} a_{n-1,ij_{n-1}} y_{n-1,j_{n-1}} \\ &= \sum_i \alpha_i \, A_0 x_{0,i} \otimes \cdots \otimes A_{n-1} x_{n-1,i} = w. \end{aligned}$$

Thus right-uniqueness of $A_0 \overset{\dot a}{\otimes} \cdots \overset{\dot a}{\otimes} A_{n-1}$ is shown.

If $A_0 \overset{\dot a}{\otimes} \cdots \overset{\dot a}{\otimes} A_{n-1}$ is a *closable*, linear operator then we define

$$A_0 \dot{\otimes} \cdots \dot{\otimes} A_{n-1} = \overline{A_0 \overset{\dot a}{\otimes} \cdots \overset{\dot a}{\otimes} A_{n-1}} \, .$$

As for the direct sums of operators we have

$$A \overset{\dot a}{\otimes} B \overset{\dot a}{\otimes} C = (A \overset{\dot a}{\otimes} B) \overset{\dot a}{\otimes} C = A \overset{\dot a}{\otimes} (B \overset{\dot a}{\otimes} C)$$

[8]This is the natural idea for an image-wise algebraic tensor product of linear operators, since a construction analogous to the direct sum case, i.e. $x \mapsto Ax \otimes Bx$, would not yield a *linear* operator.

and in the closable case

$$A \dot{\otimes} B \dot{\otimes} C = (A \dot{\otimes} B) \dot{\otimes} C = A \dot{\otimes} (B \dot{\otimes} C),$$

so that again we may focus our attention on the case of two factors with the general case being included by 'bracketing'. We shall investigate tensor products of operators more closely later.

CHAPTER 2

Linear Operators

2.1. Linear Operators and Relations

We already mentioned the construction of Hilbert spaces as closed subspaces of Hilbert spaces.

DEFINITION 53. Let $H_0 \oplus H_1$ be a direct sum of two Hilbert spaces, i.e. $H_0 \oplus H_1$ is the Hilbert space given by the set $H_0 \times H_1$ equipped with the component-wise linear structure and the norm

$$(x, y) \mapsto \sqrt{|x|_{H_0}^2 + |y|_{H_1}^2}.$$

A subset A of $H_0 \oplus H_1$ defines a relation between H_0 and H_1 or a correspondence $(A, H_0 \times H_1)$. If additionally A is also right-unique we obtain a function $A \subsetneqq H_0 \times H_1$ or a mapping $A : D(A) \subseteq H_0 \to H_1$. If A is a linear subspace then A is also called a **linear relation** or **linear correspondence**, respectively. If A is a linear subspace and a right-unique relation, then A is called a **linear mapping** or **linear operator** from H_0 to H_1. For linear mappings A it is common to use the simplified **multiplicative notation**

$$Ax := A(x) \text{ for } x \in D(A).$$

The linear operator as defined here is of course nothing but a linear mapping in the usual sense.

LEMMA 54. *A linear relation $A \subseteq H_0 \oplus H_1$, H_j, $i = 0, 1$, Hilbert spaces, is a linear operator in the sense of definition 53 if and only if $A : D(A) \subseteq H_0 \to H_1$ is a linear mapping in the usual sense.*

PROOF. The only thing to check is that linearity of a mapping is indeed the same as linearity of the subspace A. This becomes clear by observing that

$$A(\alpha \cdot x + y) = \alpha \cdot Ax + Ay$$

is the same as saying

$$\alpha \cdot (x, Ax) + (y, Ay) = (\alpha \cdot x + y, \alpha \cdot Ax + Ay) \in A.$$

It should be noted that $\alpha\cdot$ as an operator in a Hilbert space H is also a linear operator in this sense. With little risk of confusion we will therefore also in many instances simplify notation by letting

$$\alpha\, x := \alpha \cdot x \text{ for all } \alpha \in \mathbb{K},\ x \in H,$$

where $\mathbb{K}$ denotes the scalar field of H. □

REMARK 55. The restriction of the norm in $H_0 \oplus H_1$ to A is also known as the **graph norm** $|\cdot|_A$:

$$\bigwedge_{(x,y)\in A} \sqrt{|x|_{H_0}^2 + |y|_{H_1}^2} = |(x, y)|_A.$$

In particular in the case where A is a mapping the domain of A becomes an inner product space with norm

$$\bigwedge_{(x,y)\in A} |x|_{D(A)} := \sqrt{|x|_{H_0}^2 + |A(x)|_{H_1}^2} = |(x, y)|_A.$$

The norm $|\cdot|_{D(A)}$ is then also referred to as the graph norm associated with the mapping A. If A is closed then A is a Hilbert space (as a closed subspace of $H_0 \oplus H_1$). With the graph norm the domain $D(A)$ of a closed linear mapping A is also a Hilbert space, which by definition is unitarily equivalent to A. To establish a Hilbert space (as the completion of $D(A)$ or) as the domain $D(A)$ of some (closable[1] or) closed linear mapping A, is one of the standard procedures for constructing Hilbert spaces. Any dense subset of $D(A)$ is known as a **core** of A. In this sense, if A is closable then $D(A)$ is naturally a core for $\bar{A}$.

EXAMPLE 56. To illustrate let us consider

$$\begin{array}{rcl} \operatorname{grad}\big|_{\mathring{C}_\infty(\Omega)} : \mathring{C}_\infty(\Omega) \subseteq L^2(\Omega) & \longrightarrow & \bigoplus_{k=1,\dots,n} L^2(\Omega) \\ \varphi & \longmapsto & \operatorname{grad}\varphi = (\partial_i\varphi)_{i=1,\dots,n} \end{array}.$$

Here $\Omega \subseteq \mathbb{R}^n$ is an open set, $\mathring{C}_\infty(\Omega)$ is the set of infinitely often differentiable function φ defined on $\mathbb{R}^n$ with compact support in Ω, i.e. $\varphi = 0$ outside of a bounded closed subset of Ω. (note that this means in particular $\varphi = 0$ in a neighborhood of the boundary $\partial\Omega$) and $L^2(\Omega)$ is defined as a subspace of $L^2(\mathbb{R}^n)$ generated by the closure of $\mathring{C}_\infty(\Omega) \subseteq \mathring{C}_\infty(\mathbb{R}^n)$, (for the imbedding of $\mathring{C}_\infty(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$ see example 37). ∂_i denotes the partial derivative operation with respect to the $i-$th variable. In the above notation we have for the vector analytical operation '**gradient**' established as a particular Hilbert space operator

$$\operatorname{grad}\big|_{\mathring{C}_\infty(\Omega)} \subseteq L^2(\Omega) \oplus \bigoplus_{k=1,\dots,n} L^2(\Omega).$$

Denoting

$$\mathring{\operatorname{grad}} := \overline{\operatorname{grad}\big|_{\mathring{C}_\infty(\Omega)}}$$

the question arises: Is $\mathring{\operatorname{grad}}$ a linear operator? This is the same as asking: Is the linear operator $\operatorname{grad}\big|_{\mathring{C}_\infty(\Omega)}$ closable? To answer this question let $(x, y^{(i)}) \in \mathring{\operatorname{grad}}$ then there is a sequence $((x_k^{(i)}, y_k^{(i)}))_k$ in $\operatorname{grad}\big|_{\mathring{C}_\infty(\Omega)}$ with

$$(x_k^{(i)}, y_k^{(i)})_k \to (x, y^{(i)}) \text{ as } k \to \infty.$$

By linearity we have $(0, w) := (0, y^{(1)} - y^{(2)}) \in \mathring{\operatorname{grad}}$ and

$$(u_k, \operatorname{grad} u_k) := (x_k^{(1)} - x_k^{(2)}, y_k^{(1)} - y_k^{(2)})_k \to (0, w) \text{ as } k \to \infty.$$

If we can show that $w = (w^{(1)}, \dots, w^{(n)}) \in \bigoplus_{k=1,\dots,n} L^2(\Omega)$ vanishes, then $\mathring{\operatorname{grad}}$ is a closed linear operator. By integration by parts we have (since $u_k \in \mathring{C}_\infty(\Omega)$)

(2.1.1)
$$\bigwedge_{\Phi^{(i)} \in \mathring{C}_\infty(\Omega),\, i=1,\dots,n} \left\langle (\Phi^{(1)}, \dots \Phi^{(n)}) | \operatorname{grad} u_k \right\rangle_{\bigoplus_{k=1,\dots,n} L^2(\Omega)} + \left\langle \operatorname{div}(\Phi^{(1)}, \dots \Phi^{(n)}) |\, u_k \right\rangle_{L^2(\Omega)} = 0$$

with $\operatorname{div}(\Phi^{(1)}, \dots \Phi^{(n)}) := \sum_{i=1,\dots,n} \partial_i \Phi^{(i)}$, the vector analytical divergence. Letting $k \to \infty$ in (2.1.1) we get

$$\bigwedge_{\Phi^{(i)} \in \mathring{C}_\infty(\Omega),\, i=1,\dots,n} \left\langle (\Phi^{(1)}, \dots \Phi^{(n)}) | (w^{(1)}, \dots, w^{(n)}) \right\rangle_{\bigoplus_{k=1,\dots,n} L^2(\Omega)} = \sum_{i=1}^{n} \left\langle \Phi^{(i)} | w^{(i)} \right\rangle_{L^2(\Omega)} = 0.$$

Since $\mathring{C}_\infty(\Omega)$ is dense in $L^2(\Omega)$, we get that $\sum_{i=1}^n \left\langle w^{(i)} | w^{(i)} \right\rangle_{L^2(\Omega)} = 0$ and so $w = 0$.

Thus, the domain $D(\mathring{\operatorname{grad}})$ of the closed, linear operator $\mathring{\operatorname{grad}}$ equipped with the graph norm $|\cdot|_{\mathring{\operatorname{grad}}}$ is a Hilbert space. This Hilbert space (known as the Sobolev space $\mathring{H}_1(\Omega)$ is used to generalize differentiation and the boundary condition of 'vanishing at the boundary' (the so-called homogeneous Dirichlet boundary condition) to arbitrary open sets Ω.

[1] Recall that a mapping A is called a closable mapping if $\overline{A}$ is still a mapping, i.e. right-unique.

In the terminology of definition 53 linear mappings $A:\, D(A) \subseteq H_0 \to H_1$ are just particular pre-Hilbert subspaces of $H_0 \oplus H_1$. By the projection theorem we have

$$H_0 \oplus H_1 = A \oplus A^{\perp}$$

for any closed linear relation A. A closed linear relation closely related to $A^{\perp}$ is of particular interest.

Definition 57. Let A be a relation in $H_0 \oplus H_1$. Then

$$A^* := -\left(A^{\perp}\right)^{-1}$$

will be called the **adjoint relation**, $-A^* = \left(A^{\perp}\right)^{-1}$ is called the **skew-adjoint relation**. If A^* is a linear mapping, it is called the **adjoint operator** and $-A^*$ the **skew-adjoint operator**.

Remark 58. Thus, we have $(u,v) \in A^*$ if and only if

$$\bigwedge_{(x,y)\in A} \langle x|v\rangle_{H_0} = \langle y|u\rangle_{H_1} .$$

The reason for defining this peculiar looking combination of operations lies in the property that this way

$$\langle x|v\rangle_{H_0} = \langle y|u\rangle_{H_1} \tag{2.1.2}$$

for all $(x,y) \in A$ and $(u,v) \in A^*$. In this form the ortho-complement is closer to the construction of adjoint operators (see later). Indeed, if A is right-unique we have that a $u \in H_1$ is in $[H_0]\, A^*$ if and only if there is a $v \in H_0$ such that

$$\bigwedge_{x\in D(A)} \langle x|v\rangle_{H_0} = \langle A\,(x)\,|u\rangle_{H_1} .$$

Moreover, if A and A^* are right-unique we have

$$\bigwedge_{x\in D(A),\, u\in D(A^*)} \langle x|A^*u\rangle_{H_0} = \langle A\,(x)\,|u\rangle_{H_1}$$

and a $u \in H_1$ is in $D\left(A^*\right) = [H_0]\, A^*$ if and only if there is a $v \in H_0$ such that

$$\bigwedge_{x\in D(A)} \langle x|v\rangle_{H_0} = \langle A\,(x)\,|u\rangle_{H_1} .$$

This $v \in H_0$ is of course nothing but A^*u.

Lemma 59. *Let A be a relation in $H_0 \oplus H_1$. Then*[2]

$$\begin{aligned} \left(A^{-1}\right)^{\perp} &= \left(A^{\perp}\right)^{-1}, \\ \left(-A^{\perp}\right)^{-1} &= -\left(A^{\perp}\right)^{-1}, \\ (-A)^{\perp} &= -A^{\perp}, \\ -A^{-1} &= -\left[(-A)^{-1}\right]. \end{aligned} \tag{2.1.3}$$

In the latter case we have, if

$$-[A] = A$$

then

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

[2]Note the difference between $-A = \{(x,y)\,|\,(x,-y) \in A\}$ and $-[A] = \{(x,y)\,|\,(-x.-y) \in A\}$. The latter notation follows our general convention according to which for a function f we have

$$f\,[W] = \{f\,(x)\,|\,x \in W \cap D\,(f)\} .$$

Furthermore,

$$A^* = -\left(A^{-1}\right)^{\perp} = -\left(A^{\perp}\right)^{-1} = \left(-A^{-1}\right)^{\perp} = \left(-A^{\perp}\right)^{-1} = \left((-A)^{-1}\right)^{\perp} = \left((-A)^{\perp}\right)^{-1}.$$

PROOF. The equations follow by easy but tedious calculations. We shall only give an exemplary proof of the third equality, the others being similar and left as excercises for the inclined reader. So let $(x,y) \in (-A)^{\perp}$, i.e.

$$\bigwedge_{(u,v)\in 1\oplus(-1)\,A} \langle (x,y)|(u,v)\rangle_{H_0\oplus H_1} = 0$$

or

$$\bigwedge_{(u,-v)\in A} \langle (x,y)|(u,v)\rangle_{H_0\oplus H_1} = \langle x|u\rangle_{H_0} + \langle \pm y| \pm v\rangle_{H_1} = \langle (x,-y)|(u,-v)\rangle_{H_0\oplus H_1} = 0.$$

Thus, we have

$$\bigwedge_{(u,v)\in A} \langle (x,-y)|(u,v)\rangle_{H_0\oplus H_1} = 0$$

or

$$(x,-y) \in A^{\perp}$$

and so

$$(x,y) \in -A^{\perp}.$$

Since the implications can be reversed, we have shown the last equality. □

As a consequence we obtain the follow result directly from Theorem 40.

THEOREM 60. *Let $H_0 \oplus H_1$ be a direct sum of two (complex) Hilbert spaces and $A \subseteq H_0 \oplus H_1$ a relation. Then*

$$\overline{Lin_{\mathbb{C}}\, A} = A^{**}.$$

PROOF. According to (2.1.3) and by definition of the adjoint relation we have

$$\begin{aligned} A^{**} &= -\left(\left(-\left(A^{\perp}\right)^{-1}\right)^{\perp}\right)^{-1}, \\ &= -\left(-\left(\left(A^{\perp\perp}\right)^{-1}\right)^{-1}\right), \\ &= \overline{Lin_{\mathbb{C}}\, A}. \end{aligned}$$

Here we have used Theorem 40 together with the facts that double negation cancels, double inversion cancels. □

EXAMPLE 61. Noting that

$$T \mapsto \frac{1}{2}\left(T + T^{\top}\right)$$
$$T \mapsto \frac{1}{2}\left(T - T^{\top}\right)$$

are projectors in $\left(L^2(\Omega)\right)^{3\times 3} = \bigoplus_{s=1,2,3}\left(\bigoplus_{k=1,2,3} L^2(\Omega)\right)$ denoted by sym and skew, we define

$$L^2_{\mathrm{sym}}(\Omega) := \mathrm{sym}\left[\left(L^2(\Omega)\right)^{3\times 3}\right].$$

We have

$$\mathrm{sym}\,\mathrm{skew} = \mathrm{skew}\,\mathrm{sym} = 0,\ \mathrm{sym} = 1 - \mathrm{skew}\,.$$

Consider now $\mathring{\mathrm{Grad}}$ defined as the closure of

$$\begin{aligned} \mathrm{Grad}\big|_{\mathring{C}_\infty(\Omega)} : \left(\mathring{C}_\infty(\Omega)\right)^3 \subseteq \left(L^2(\Omega)\right)^3 &\to L^2_{\mathrm{sym}}(\Omega) \\ v &\mapsto \mathrm{Grad}v := \frac{1}{2}\left(d\otimes v + (d\otimes v)^{\top}\right) \end{aligned}$$

where $d\otimes v$ denotes the Jacobian of $v:\Omega\subseteq\mathbb{R}^3\to\mathbb{C}^3$. That $\operatorname{Grad}|_{\mathring{C}_\infty(\Omega)}$ is closable follows similar as for the usual gradient via Gauss' theorem

$$\begin{aligned}\langle \operatorname{Grad} v|\,T\rangle_{L^2(\Omega)}+\langle v|\,\operatorname{div} T\rangle_{L^2(\Omega)} &= \int_\Omega \operatorname{trace}\left((\operatorname{Grad} v\,(x))^*\,T\,(x)\right)\,dx+\int_\Omega v\,(x)^*\operatorname{div} T\,(x)\,dx\\ &= \frac{1}{2}\sum_{k,s}\int_\Omega \partial_k v^{(s)}\,(x)^*\,T_s^k\,(x)\,dx+\\ &+\frac{1}{2}\sum_{k,s}\int_\Omega \partial_s v^{(k)}\,(x)^*\,T_s^k\,(x)\,dx+\\ &+\sum_{k,s}\int_\Omega v^{(s)}\,(x)^*\,\partial_k\,T_s^k\,(x)\,dx \qquad (2.1.4)\\ &= \sum_{k,s}\int_\Omega \partial_s v^{(k)}\,(x)\,T_s^k\,(x)\,dx+\sum_{k,s}\int_\Omega v^{(s)}\,(x)^*\,\partial_k\,T_s^k\,(x)\,dx\\ &= \sum_{k,s}\int_\Omega \partial_k\left(v_s\,(x)^*\,T_{ks}\,(x)\right)\,dx\\ &= 0,\end{aligned}$$

where $T_{ks}=T_{sk}\in\mathring{C}_\infty\left(\mathbb{R}^3\right)$, $s,k\in\{1,2,3\}$.

Similarly, we can define $\mathring{\operatorname{Div}}$ as the closure of

$$\begin{aligned}\operatorname{Div}|_{\mathring{C}_\infty(\Omega)}:\operatorname{sym}\left[\left(\mathring{C}_\infty\,(\Omega)\right)^{3\times 3}\right]\subseteq L^2_{\mathrm{sym}}\,(\Omega)\to\left(L^2\,(\Omega)\right)^3&\\ T\mapsto \operatorname{Div} T\ :=\left(\sum_{k=1}^{3}\partial_k T_{sk}\right)_{s=1,2,3}&.\end{aligned}$$

With this we can now define a generalized Grad and Div by letting – in accordance with the integration by parts formula (2.1.4)–

$$\begin{aligned}\operatorname{Div}&:=-\left(\operatorname{Grad}|_{\mathring{C}_\infty(\Omega)}\right)^*,\\ \operatorname{Grad}&:=-\left(\operatorname{Div}|_{\mathring{C}_\infty(\Omega)}\right)^*.\end{aligned}$$

We read off

$$\begin{aligned}\mathring{\operatorname{Div}}&\subseteq\operatorname{Div}\,,\\ \mathring{\operatorname{Grad}}&\subseteq\operatorname{Grad}\,.\end{aligned}$$

Consequently, we have by definition

$$\langle\operatorname{Grad} v|T\rangle_{L^2(\Omega)}+\langle v|\operatorname{Div} T\rangle_{L^2(\Omega)}=0$$

for all $v\in D\left(\mathring{\operatorname{Grad}}\right)$, $T\in D\,(\operatorname{Div})$ or for all $v\in D\,(\operatorname{Grad})$, $T\in D\left(\mathring{\operatorname{Div}}\right)$ in generalization of the above integration by parts formula (2.1.4).

From the projection theorem follows:

THEOREM 62. *(Projection Theorem 3) Let $H_0\oplus H_1$ be a direct sum of two (complex) Hilbert spaces and $A\subseteq H_0\oplus H_1$ a closed, linear relation. Then we have the orthogonal decompositions*

$$H_0=[\{0\}]\,A\oplus\overline{A^*\,[H_1]}\ \textit{and}\ H_1=[\{0\}]\,A^*\oplus\overline{A\,[H_0]}.$$

PROOF. Let

$$y\perp A\,[H_0]\,.$$

Then

$$\bigwedge_{(u,v)\in A} 0 = \langle y|v\rangle_{H_1} = \langle (0,y)|(u,v)\rangle_{H_0\oplus H_1}$$

or

$$(0,y) \in A^{\perp}.$$

The latter is equivalent

$$(0,y) \in (A^*)^{-1},$$

i.e.

$$y \in (A^*)^{-1}[\{0\}] = [\{0\}]\, A^*.$$

Since the implications can be reversed, the second decomposition is shown to be valid. The first decomposition result follows from the second by replacing A by A^* and using Theorem 60. $\square$

REMARK 63. The name 'Projection Theorem 3' is given for the importance of this result in the applications we have in mind. With $A = P_C$ as the orthogonal projector defined in (1.2.19) and $H = H_0 = H_1$ we recover from Theorem 62 the earlier Theorem 37. It can be shown that $P_C = P_C^*$ and $N(P_C) \equiv P_C^{-1}[\{0\}] = [\{0\}]\, P_C = R(1 - P_C) \equiv (1 - P_C)[H]$.

EXAMPLE 64. This variant of the projection theorem contains the core of the linear solution theory. Indeed, if A is a densely defined, closed linear mapping *with closed range* then

$$Au = f$$

is solvable for all $f \perp [\{0\}]\, A^* = N(A^*)$. If we look for solutions in $\overline{A^*[H_1]} = ([\{0\}]\, A)^{\perp}$ then the solution is unique.

REMARK 65. We can find the best approximation of f in the closed subspace $\overline{A[H_0]}$ by finding $u \in D(A)$ with

$$Au - f \perp A[H_0],$$

i.e.

$$\langle Au|Av\rangle_{H_1} - \langle f|Av\rangle_{H_1} = 0 \tag{2.1.5}$$

for all $v \in D(A)$. Note that for $f \in D(A^*)$ this equivalent to solving the associated so-called Euler equation

$$A^*Au = A^*f. \tag{2.1.6}$$

If we happen to know that $\widehat{A} := A \cap \left(\overline{A^*[H_1]} \oplus \overline{A[H_0]}\right) \subseteq \overline{A^*[H_1]} \oplus \overline{A[H_0]}$ has a bounded inverse, then we see from

$$\langle u|u\rangle_{H_1} = \left\langle \widehat{A}^{-1}Au|\widehat{A}^{-1}Au\right\rangle_{H_1} \leq \left\|\widehat{A}^{-1}\right\|^2 \langle Au|Au\rangle_{H_1}$$

that

$$|Au|_{H_1} \leq \sqrt{|u|_{H_0}^2 + |Au|_{H_1}^2} \leq \sqrt{1 + \left\|\widehat{A}^{-1}\right\|^2}\, |Au|_{H_1}$$

for all $u \in D(A)$. Thus,

$$(u,v) \mapsto \langle Au|Av\rangle_{H_1}$$

is an inner product for $D(A) \cap \overline{A^*[H_1]}$ and finding the best approximation as a solution for (2.1.5) is achieved by simply applying the Riesz representation theorem to the continuous linear functional

$$v \mapsto \langle f|Av\rangle_{H_1}.$$

This is a common idea of finding a best approximation or solving equation (2.1.6) via (2.1.5) known as the variational method.

2.2. Spectrum and Resolvent Set

In many connotations the old linear algebra question arises: Given a linear operator A, for which $\lambda \in \mathbb{C}$ can one always solve the equation

$$(\lambda - A)u = f \tag{2.2.1}$$

for u with given data f? The not so deep first answer is of course: If all such data f are in the range of $(\lambda - A)$! The triviality of this answer may irritate enough to bring us to the deeper question: When is a problem reasonably solvable? A little more subtle pondering would probably lead us to add to the first response, that we would be quite happy to find solutions satisfying the equation not exactly, but to any prescribed degree of accuracy. For this, however, we would want to have some control over errors so that an approximate solution is not too far away from other solutions solving the equation more accurately. Such control would certainly also require to have not more than one solution, since otherwise we have a manifold of solutions with any degree of deviation. This brings us to the three celebrated requirements for being "reasonably solvable" due to Hadamard:

- uniqueness of solution,
- existence of solution (at least for a dense set of given data),
- locally uniformly continuous dependence of solution on the given data.

Transcribing these requirements to the initial question of classifying $\lambda \in \mathbb{C}$, we are led to define the resolvent set $\varrho(A)$. The set $\varrho(A)$ contains all such $\lambda \in \mathbb{C}$ for which equation (2.2.1) can always reasonably be (re)solved. In order to define the resolvent set we need some more concepts, which we introduce here in sufficient generality to make them useful for later discussions.

Definition 66. Let $B \subseteq H_0 \oplus H_1$ be a linear operator, H_0, H_1 (complex) Hilbert spaces. Then B is called a **densely defined** linear operator if $D(B)$ is dense in H_0. B is called a **continuous linear operator** (or **bounded linear operator**), if

$$\bigwedge_{x \in D(A)} |Bx|_{H_1} \leq C\,|x|_{H_0} \tag{2.2.2}$$

for some constant $C \in \mathbb{R}_{>0}$. The best constant in (2.2.2) is referred to as the **operator norm** of B and is denoted by $|B|_{H_0 \to H_1}$ or simply by $||B||$. A linear operator B which is not continuous is called **unbounded, linear operator**.

Remark 67. We comment here once again that for linear operators continuity in the topological sense actually implies Lipschitz continuity (due to the uniform boundedness principle), but for sake of keeping matters elementary we will not pursue this further.

We note that for a linear operator $B : D(B) \subseteq H_0 \to H_1$ we have $B(u - v) = Bu - Bv$ and therefore

$$||B|| = |B|_{Lip}.$$

In particular, we realize that for continuous linear operators the Lipschitz semi-norm is indeed a norm, viz. the operator norm. The set $L(H_0, H_1)$ of bounded linear operators defined on H_0 equipped with the linear structure

$$(\alpha \cdot A + B)(x) = \alpha \cdot Ax + Bx \text{ for all } \alpha \in \mathbb{C},\ A,\ B \in L(H_0, H_1),\ x \in H_0,$$

becomes a normed linear space with $|\cdot|_{L(H_0,H_1)} := |\cdot|_{H_0 \to H_1}$ as norm.

Since the composition $A \circ B$, usually also written in the multiplicative form AB, of an element in $B \in L(H_0, H_1)$ with $A \in L(H_1, H_2)$ is in $L(H_0, H_2)$ (H_2 another Hilbert space), it is worth noting that composition is a continuous operation between $L(H_1, H_2) \times L(H_0, H_1)$ and $L(H_0, H_2)$ in the sense that

$$|ABx|_{H_2} \leq |A|_{L(H_1,H_2)}\,|Bx|_{H_1} \leq |A|_{L(H_1,H_2)}\,|B|_{L(H_0,H_1)}\,|x|_{H_0} \text{ for all } x \in H_0.$$

In other words, the bi-linear mapping $\circ : L(H_1, H_2) \times L(H_0, H_1) \longrightarrow L(H_0, H_2)$ is continuous. Here bi-linearity means

$$(\alpha A + C) \circ (\beta B + D) = \alpha\beta\, A \circ B + \alpha\, A \circ D + \beta\, C \circ B + C \circ D$$

and we have indeed

$$|A \circ B|_{L(H_0,H_2)} \le |A|_{L(H_1,H_2)}\, |B|_{L(H_0,H_1)} \tag{2.2.3}$$

for all $\alpha \in \mathbb{C}$, $A,\, C \in L(H_1, H_2)$ and $B,\, D \in L(H_0, H_1)$. Moreover, we have for all such normed linear spaces $L(H_0, H_1)$:

PROPOSITION 68. *The normed linear space $L(H_0, H_1)$ of continuous linear operators defined on H_0 and with range in H_1 is a Banach space.*

PROOF. We have to show completeness. So, let $(A_k)_k$ be a Cauchy sequence in $L(H_0, H_1)$. Then

$$|A_n x - A_m x|_{H_1} = |(A_n - A_m)x|_{H_1} \le |A_n - A_m|_{H_0 \to H_1}\, |x|_{H_0}.$$

Thus, $(A_n x)_n$ is a Cauchy sequence in H_1 for every $x \in H_0$. Therefore, $\lim_{n\to\infty} A_n x \in H_1$ exists and we define

$$A_\infty(x) := \lim_{n\to\infty} A_n x \text{ for all } x \in H_0.$$

That A_∞ is a linear operator defined on H_0 follows from the continuity of the linear operations:

$$\begin{aligned} A_\infty(\alpha \cdot x + y) &:= \lim_{n\to\infty} A_n(\alpha \cdot x + y) = \lim_{n\to\infty}(\alpha \cdot A_n x + A_n y) = \\ &= \alpha \cdot \lim_{n\to\infty} A_n x + \lim_{n\to\infty} A_n y = \alpha \cdot A_\infty(x) + A_\infty(y). \end{aligned}$$

It remains to show that A_∞ is also bounded. But this follows since Cauchy sequences are bounded and so for all $x \in H_0$ we have

$$|A_n x|_{H_1} \le \sup\{|A_n|_{H_0 \to H_1} \mid\ n \in \mathbb{N}\}|x|_{H_0}.$$

Letting $n \to \infty$ this yields

$$|A_\infty x|_{H_1} \le \sup\{|A_n|_{H_0 \to H_1} \mid\ n \in \mathbb{N}\}|x|_{H_0}.$$

□

One can show an even stronger result.

PROPOSITION 69. *Let $(A_n)_n$ be a sequence in the Banach space $L(H_0, H_1)$ such that*

$$\lim_{n\to\infty} A_n x \textit{ exists for all } x \in H_0. \tag{2.2.4}$$

Then

$$\begin{aligned} A_\infty : H_0 &\longrightarrow H_1 \\ x &\longmapsto \lim_{n\to\infty} A_n x \end{aligned}$$

defines a linear operator $A_\infty \in L(H_0, H_1)$.

REMARK 70. In contrast to the convergence in the Banach space $L(H_0, H_1)$, which is also called **uniform convergence** or **convergence in operator norm**, the convergence concept in (2.2.4) is called **strong convergence** or **point-wise convergence**. In this terminology the previous proposition says that $L(H_0, H_1)$ is also complete with respect to strong (or point-wise) convergence. The proof also shows that this is really a Banach space result, since the structural features of a Hilbert space have not been used at all.

Getting back to the issue of 'reasonable solvability' we now introduce the resolvent set.

DEFINITION 71. Let $A \subseteq H \oplus H$ be a linear operator, H a (complex) Hilbert space. Then

$$\varrho(\lambda) := \left\{\lambda \in \mathbb{C} \,\middle|\, (\lambda - A)^{-1} \text{ is a densely defined, continuous linear operator}\right\}$$

is called the **resolvent set** of A.

As far as controllability of solutions is concerned, we note that

$$\left|(\lambda - A)^{-1}u - (\lambda - A)^{-1}v\right|_H \leq \left|(\lambda - A)^{-1}\right|_{Lip} |u - v|_H$$

holds for all u, $v \in (\lambda - A)H$ and we have

$$\left\|(\lambda - A)^{-1}\right\| = \left|(\lambda - A)^{-1}\right|_{Lip}.$$

Clearly all three of Hadamard's requirements are met for $\lambda \in \varrho(A)$. Writing the resolvent set in the redundant way as

$$\varrho(\lambda) = \left\{ \lambda \in \mathbb{C} \,\middle|\, \overbrace{(\lambda - A)^{-1} \text{ is a linear operator}}^{\alpha(\lambda):=} \wedge \wedge \overbrace{(\lambda - A)^{-1} \text{ is a densely defined linear operator}}^{\beta(\lambda):=} \wedge \wedge \overbrace{(\lambda - A)^{-1} \text{ is a densely defined, continuous linear operator}}^{\gamma(\lambda):=} \right\}.$$

adds no information to the description of $\varrho(A)$, but it may make one of the common classification of failure to be 'reasonably solvable' more transparent.

Definition 72. Let $A \subseteq H \oplus H$ be a linear operator, H a (complex) Hilbert space. Then the complement of the resolvent set $\varrho(A)$ is called the **spectrum** $\sigma(A)$ of A, i.e.

$$\sigma(A) := \mathbb{C} \setminus \varrho(A).$$

The parts of the spectrum are

- the **point spectrum**

$$\begin{aligned} P\sigma(A) &:= \{\lambda \in \mathbb{C} \,|\, \neg\alpha(\lambda)\}, \\ &= \{\lambda \in \mathbb{C} \,|\, (\lambda - A) \text{ is not one} - \text{to} - \text{one}\}, \end{aligned}$$

- the **residual spectrum**

$$\begin{aligned} R\sigma(A) &:= \{\lambda \in \mathbb{C} \,|\, \alpha(\lambda) \wedge \neg\beta(\lambda)\}, \\ &= \left\{\lambda \in \mathbb{C} \,\middle|\, (\lambda - A)^{-1} \text{ is a not densely defined, linear operator}\right\}, \end{aligned}$$

- the **continuous spectrum**

$$\begin{aligned} C\sigma(A) &:= \{\lambda \in \mathbb{C} \,|\, \alpha(\lambda) \wedge \beta(\lambda) \wedge \neg\gamma(\lambda)\}, \\ &= \left\{\lambda \in \mathbb{C} \,\middle|\, (\lambda - A)^{-1} \text{ is an unbounded, densely defined linear operator}\right\}. \end{aligned}$$

There are other spectral parts of occasional interest:

- the **approximate point spectrum**

$$\pi(A) := \left\{ \lambda \in \mathbb{C} \,\middle|\, \bigvee_{(x_n)_n} |x_n|_H = 1 \wedge (\lambda - A)x_n \to 0 \text{ as } n \to \infty \right\},$$

- the **compression spectrum**

$$\sigma_{comp}(A) := \left\{ \lambda \in \mathbb{C} \,\middle|\, \overline{(\lambda - A)H} \neq H \right\},$$

- the **discrete spectrum**

$$\sigma_d(A) := \{\lambda \in P\sigma(A) \,|\, \lambda \text{ isolated point of } \sigma(A) \wedge N(\lambda - A) \text{ finite} - \text{dimensional}\},$$

► the **essential spectrum**

$$\sigma_e(A) := R\sigma\,(A) \cup C\sigma\,(A) \cup \\ \cup \{\lambda \in P\sigma\,(A)\,|\,\lambda \text{ accumulation point of } \sigma(A) \,\vee\, N(\lambda - A) \text{ infinite} - \text{dimensional}\}\,.$$

The following lemma gives a few relations between some parts of the spectrum.

LEMMA 73. *Let $A \subseteq H \oplus H$ be a linear operator, H a (complex) Hilbert space. Then*

$$\sigma(A) = P\sigma(A) \cup R\sigma(A) \cup C\sigma(A). \tag{2.2.5}$$

The spectral parts $P\sigma(A)$, $R\sigma(A)$, $C\sigma(A)$ are disjoint. Moreover, we have

$$P\sigma(A) \cup C\sigma(A) \subseteq \pi(A) \subseteq \sigma(A). \tag{2.2.6}$$

$$R\sigma(A) \subseteq \sigma_{comp}(A) \subseteq R\sigma(A) \cup P\sigma(A), \tag{2.2.7}$$

$$\sigma_d(A) \subseteq P\sigma(A), \tag{2.2.8}$$

$$\sigma_e(A) = \sigma(A) \setminus \sigma_d(A). \tag{2.2.9}$$

PROOF. The equality (2.2.5) following by elementary logic

$$\neg\alpha(\lambda) \vee \neg\beta(\lambda) \vee \neg\gamma(\lambda) \Leftrightarrow \neg\alpha(\lambda) \vee (\alpha(\lambda) \wedge \neg\beta(\lambda)) \vee (\alpha(\lambda) \wedge \beta(\lambda) \wedge \neg\gamma(\lambda)).$$

The disjointness of the union is obvious. Since the condition for being in $\pi(A)$ contradicts continuity of $(\lambda - A)^{-1}$ we have immediately

$$C\sigma(A) \subseteq \pi(A) \subseteq \sigma(A).$$

If $\lambda \in P\sigma(A)$ then there is an element $x \in D(A)$, $|x|_H = 1$, with $(\lambda - A)x = 0$. The constant sequence $(x)_n$ satisfies the condition defining $\pi(A)$ and so also

$$P\sigma(A) \subseteq \pi(A).$$

Clearly, $R\sigma(A) \subseteq \sigma_{comp}(A)$. Since $\lambda \in C\sigma(A)$ contradicts being in $\sigma_{comp}(A)$ and since $\sigma_{comp}(A) \subseteq \sigma(A)$ we have

$$\sigma_{comp}(A) \subseteq \sigma(A) \setminus C\sigma(A) = R\sigma(A) \cup P\sigma(A).$$

That $\sigma_e(A) = \sigma(A) \setminus \sigma_d(A)$ and $\sigma_d(A) \subseteq P\sigma(A)$ is obvious. □

The resolvent set has the remarkable property of being always an open set, which makes the spectrum a closed set. Before we are able to prove this, we need to know more about the **resolvent** which is the linear operator $(\lambda - A)^{-1}$ in case $\lambda \in \varrho(A)$.

PROPOSITION 74. *Let $A \subseteq H \oplus H$ be a linear relation, H a (complex) Hilbert space. Then we have for all $\lambda \in \mathbb{C}$ that*[3]

$$\overline{(\lambda - A)^{-1}} = (\lambda - \bar{A})^{-1}. \tag{2.2.10}$$

[3]For relations $A, B \subseteq H_0 \oplus H_1$ the relation

$$\alpha A + B$$

is defined – in analogy to mappings – right-wise, i.e.

$$\alpha A + B := \left\{ x \oplus y \mid \bigvee_{x \oplus a \in A,\, x \oplus b \in B} y = \alpha a + b \right\},$$

where α is in the underlying field of the Hilbert spaces H_k, $k = 0, 1$. This way such relations obtain a natural additive and scalar multiplicative structure. Note. however, that for a linear structure the additive group structure is usually missing. Indeed, in general

$$A - A \neq H_0 \oplus \{0\}\,,$$

where $H_0 \oplus \{0\}$ is the zero mapping. If

$$A - A \subseteq H_0 \oplus \{0\}\,,$$

then we call A right-unique. Even in this case, in general we may have $A - A \subset H_0 \oplus \{0\}$. If A is right-unique and left-total, then

$$A - A = H_0 \oplus \{0\}\,.$$

If A is a linear operator, then we have for all $\lambda \in \varrho(A)$ that $(\lambda - \bar{A})^{-1} : H \to H$ is a continuous, linear operator with

$$\begin{aligned}\left\|(\lambda - \bar{A})^{-1}\right\| &= \left|(\lambda - \bar{A})^{-1}\right|_{H\to H} = \left|(\lambda - \bar{A})^{-1}\right|_{Lip}\\ &= \left|(\lambda - A)^{-1}\right|_{H\to H} = \left|(\lambda - A)^{-1}\right|_{Lip} = \left\|(\lambda - A)^{-1}\right\|.\end{aligned}$$

If, moreover, A is closable then we have

$$\varrho(A) = \varrho(\bar{A}).$$

Proof. To see (2.2.10) we take $(u, w) \in \overline{(\lambda - A)^{-1}}$ and a sequence $((u_k, w_k))_k$ in $(\lambda - A)^{-1}$ converging to (u, w). Since

$$(x, y) \in (\lambda - A)^{-1} \Leftrightarrow (y, x) \in (\lambda - A) \Leftrightarrow (y, \lambda \cdot y - x) \in A, \tag{2.2.11}$$

we have

$$(w_k,\ -u_k + \lambda \cdot w_k) \to (w,\ -u + \lambda \cdot w) \text{ as } k \to \infty$$

and therefore

$$(w,\ -u + \lambda \cdot w) \in \bar{A}.$$

Using (2.2.11) once again with $\bar{A}$ replacing A we get

$$(u, w) \in (\lambda - \bar{A})^{-1}.$$

Since the reasoning can be reversed, we obtain (2.2.10). If A is a linear operator, then its resolvent $(\lambda - A)^{-1}$ is Lipschitz continuous and densely defined for all $\lambda \in \varrho(A)$. Therefore, its closure $\overline{(\lambda - A)^{-1}}$ exists, is Lipschitz continuous with the same Lipschitz semi-norm and defined on $\overline{(\lambda - A)[H]} = H$. With (2.2.10) we obtain that also $(\lambda - \bar{A})^{-1}$ has these properties. That the Lipschitz semi-norm is equal to the operator norm has already been noted. If $\bar{A}$ is a linear operator then $(\lambda - \bar{A})^{-1}$ is its resolvent. This shows that

$$\varrho(A) \subseteq \varrho(\bar{A}).$$

The reverse inclusion also holds. Let $\lambda \in \varrho(\bar{A})$, then with (2.2.10) we see that $(\lambda - A)^{-1}$ is a continuous linear operator as a restriction of a continuous linear operator. Since $(\lambda - A)^{-1}$ is dense in $(\lambda - \bar{A})^{-1}$ we also get $D((\lambda - A)^{-1})$ dense in $D((\lambda - \bar{A})^{-1}) = H$. In other words $\lambda \in \varrho(A)$. □

In case the linear operator A is closed, the resolvents are all defined on the same domain and thus become comparable. Therefore we will here and later mostly assume that we are in this situation.

Theorem 75. *Let $A \subseteq H \oplus H$ be a closed, linear operator, H a (complex) Hilbert space. Then we have for all λ, $\mu \in \varrho(A)$ that the so-called* ***resolvent equation*** *holds*

$$(\lambda - A)^{-1} - (\mu - A)^{-1} = -(\lambda - \mu)\,(\lambda - A)^{-1}(\mu - A)^{-1}. \tag{2.2.12}$$

The mapping

$$\begin{aligned} R(\cdot,\, A) :\ \varrho(A) \subseteq \mathbb{C} &\longrightarrow L(H, H)\\ \lambda \qquad &\longmapsto (\lambda - A)^{-1}\end{aligned}$$

is analytic on the open set $\varrho(A)$ of $\mathbb{C}$ in the sense that it is complex differentiable. We have for the complex derivative $\partial_{\mathbb{C}}$ of $R(\cdot, A)$

$$\partial_{\mathbb{C}}^n R(\lambda, A) = (-1)^n\, n!\, R(\lambda, A)^{n+1} = (-1)^n\, n!\, (\lambda - A)^{-n-1} \text{ for all } \lambda \in \varrho(A),\ n \in \mathbb{N}. \tag{2.2.13}$$

Moreover, we have that the power series representation

$$R(\lambda, A) = R(\mu, A) \sum_{n=0}^{\infty} (\mu - \lambda)^n\, R(\mu, A)^n \tag{2.2.14}$$

holds in $L(H, H)$ for all $\lambda \in \varrho(A)$ in the open disc $B_{\mathbb{C}}(\mu,\, 1/\left\|R(\mu, A)\right\|)$.

Proof. The resolvent equality follows by a simple calculation:

$$\begin{aligned}(\lambda - A)^{-1} - (\mu - A)^{-1} &= (\lambda - A)^{-1}(1 - (\lambda - A)(\mu - A)^{-1}) \\ &= (\lambda - A)^{-1}((\mu - A) - (\lambda - A))\,(\mu - A)^{-1} \\ &= (\mu - \lambda)\,(\lambda - A)^{-1}(\mu - A)^{-1}.\end{aligned} \tag{2.2.15}$$

Note that the calculation in (2.2.15) are valid since $D(A) = (\lambda - A)^{-1}[H]$. Next we show that $\varrho(A)$ is open. We observe that the geometric series (a special instance of a so-called **Neumann series**, see corollary below)

$$N(\lambda, \mu, A) := \sum_{k=0}^{\infty}(\mu - \lambda)^k\,(\mu - A)^{-(k+1)}$$

converges in $L(H_0, H_1)$ if $\lambda \in B_{\mathbb{C}}\left(\mu,\, \left\|(\mu - A)^{-1}\right\|^{-1}\right)$, $\mu \in \varrho(A)$. □

Corollary 76. *Let $Q \in L(H, H)$ with $||Q|| < 1$ then*

$$(1 - Q)^{-1} = \sum_{k=0}^{\infty} Q^k \text{ in } L(H, H).$$

Proof. The result follows by observing that

$$S \mapsto SQ + 1$$

is a contraction mapping in $L(H, H)$ and that the (by the contraction mapping theorem) unique fixed point $\widehat{S}$ must satisfy

$$\widehat{S}(1 - Q) = 1$$

and so that $\widehat{S} = (1 - Q)^{-1}$, since it is not hard to see that Q and $\widehat{S}$ are commuting. Any starting point yields an iterated sequence converging to this fixed point. Starting with 0 yields the partial sums of $\sum_{k=0}^{\infty} Q^k$ showing convergence and representation at once. □

2.3. Special Classes of Linear Operators

2.3.1. Densely Defined, Closable and Closed Linear Operators and their Adjoints.

We have already introduced densely defined, linear operators, here we would like to investigate a bit deeper, what the consequences are of this property. As it turns out this is precisely that property which warrants the existence of the adjoint operator. In order to get a deeper understanding of the situation, let us start slightly more general than actually needed.

Lemma 77. *Let $A \subseteq H_0 \oplus H_1$ be a linear relation between the complex Hilbert spaces H_0, H_1. The adjoint relation A^* is a linear operator if and only if $A^{-1}[H_1] = [H_1]\,A$ is dense in H_0.*

Proof. As an ortho-complement A^* is certainly an - always closed - linear relation. We only need to show the right-uniqueness for this linear relation. So let $(0, w) \in A^*$. If we can show that w must be zero if and only if $A^{-1}[H_1]$ is dense in H_0, then we are done. We find the following chain of equivalences

$$\begin{aligned}(0, w) \in A^* = \left(-A^{\perp}\right)^{-1} &\Leftrightarrow (w, 0) \in A^{\perp}, \\ &\Leftrightarrow \textstyle\bigwedge_{(x,y)\in A} \langle x|w\rangle_{H_0} + \langle y|0\rangle_{H_1} = 0, \\ &\Leftrightarrow \textstyle\bigwedge_{x\in[H_0]A = A^{-1}[H_1]} \langle x|w\rangle_{H_0} = 0.\end{aligned}$$

□

As an immediate corollary we get.

Corollary 78. *Let $A \subseteq H_0 \oplus H_1$ be a linear operator between the complex Hilbert spaces H_0, H_1. The adjoint relation A^* is a linear operator if and only if A is densely defined.*

Proof. The conclusion is immediate from the previous lemma, if one notices that $D(A) = A^{-1}[H_1] = [H_1]\,A$. $\square$

Using the mapping notation for linear operators we get

Corollary 79. *Let $A : D(A) \subseteq H_0 \to H_1$ be a densely defined, linear operator between the complex Hilbert spaces H_0, H_1 and $A^* : D(A^*) \subseteq H_1 \to H_0$ its adjoint. Then*

$$D(A^*) = \left\{ v \in H_1 \,\middle|\, \bigvee_{w \in H_0} \bigwedge_{x \in D(A)} \langle Ax|v\rangle_{H_1} = \langle x|w\rangle_{H_0} \right\}$$

and

$$A^*v =\in \left\{ w \in H_0 \,\middle|\, \bigwedge_{x \in D(A)} \langle Ax|v\rangle_{H_1} = \langle x|w\rangle_{H_0} \right\}.$$

Moreover, we have

$$\bigwedge_{x \in D(A)} \bigwedge_{y \in D(A^*)} \langle Ax|y\rangle_{H_1} = \langle x|A^*y\rangle_{H_0}. \tag{2.3.1}$$

Proof. We find similar to the argument in the proof of Lemma 77 by direct calculation that

$$\begin{aligned} (v,w) \in A^* = \left(1 \oplus (-1)\,A^{\perp}\right)^{-1} &\Leftrightarrow (w,-v) \in A^{\perp}, \\ &\Leftrightarrow \textstyle\bigwedge_{(x,y)\in A} \langle x|w\rangle_{H_0} - \langle y|v\rangle_{H_1} = 0, \\ &\Leftrightarrow \textstyle\bigwedge_{x \in D(A)} \langle x|w\rangle_{H_0} = \langle Ax|v\rangle_{H_1}. \end{aligned} \tag{2.3.2}$$

The last statement yields the desired characterization of $D(A)$ as well as the description of A^*v. Inserting that $w = A^*v$ by definition we also get (2.3.1). $\square$

Corollary 80. *Let $A : D(A) \subseteq H_0 \to H_1$ be a densely defined, closed linear operator between the complex Hilbert spaces H_0, H_1 and $A^* : D(A^*) \subseteq H_1 \to H_0$ its adjoint. Then, we obtain a continuous linear mapping $A^\diamond : H_1' \to H_A'$, the dual mapping, given by*

$$A^\diamond w := \left(x \mapsto w\,(Ax) = \langle R_{H_1} w|Ax\rangle_{H_1}\right)$$

for $w \in H_1'$. By identifying $H_1 = H_1'$ we obtain

$$A^\diamond v = \left(x \mapsto \langle v|Ax\rangle_{H_1}\right) : H_1 \to H_A'.$$

Then by also identifying $H_0 = H_0'$ we get

$$D(A^*) = [H_0]\,A^\diamond$$

and we obtain

$$A^*v = A^\diamond v.$$

Moreover, we have

$$\bigwedge_{x \in D(A)} \bigwedge_{y \in H_1} \langle Ax|y\rangle_{H_1} = \langle x|A^\diamond y\rangle_{H_0} \tag{2.3.3}$$

and

$$\bigwedge_{x \in D(A)} \bigwedge_{y \in D(A^*)} \langle Ax|y\rangle_{H_1} = \langle x|A^*y\rangle_{H_0}. \tag{2.3.4}$$

PROOF. By definition, for every $v \in H_1$ we have

$$\left(A^{\diamond}v\right)(x) = \langle v|Ax\rangle_{H_1}$$

for all $v \in H_1' = H_1$ and all $x \in D(A)$. This defines the continuous linear operator $A^{\diamond}$ uniquely and indeed by the given formula. Moreover, H_A is densely and continuously embedded into H_0 and so is H_0' in H_A'. Thus, we have $H_A \subseteq H_0 = H_0' \subseteq H_A'$ as a sequence of dense inclusions. For $v \in [H_0]\,A^{\diamond}$ we have now that

$$A^{\diamond}v \in H_0' = H_0$$

and so

$$\langle A^{\diamond}v|x\rangle_{H_0} = \left(A^{\diamond}v\right)(x) = \langle v|Ax\rangle_{H_1},$$

from which we can read off the stated results. □

The last lemma can also be applied to A^* instead of A.

LEMMA 81. *Let $A \subseteq H_0 \oplus H_1$ be a relation between the complex Hilbert spaces H_0, H_1. Then the linear relation $Lin_{\mathbb{C}}\,A$ is a closable linear operator if and only if $(A^*)^{-1}[H_0] = \left(A^{-1}\right)^*[H_0] = [H_0]\,A^*$ is dense in H_1.*

PROOF. First we realize that applying Lemma 77 to A^* we find that A^{**} is a linear operator if and only $(A^*)^{-1}[H_0]$ is dense in H_1. The linear relation A^{**} is however nothing but the linear relation $\overline{Lin_{\mathbb{C}}\,A}$ which is a linear operator if and only if $Lin_{\mathbb{C}}\,A$ is a closable linear operator. □

PROPOSITION 82. *Let $A \subseteq H_0 \oplus H_1$ be a densely defined, linear operator between the complex Hilbert spaces H_0, H_1. The adjoint operator A^* is a densely defined, linear operator if and only if A is a closable linear operator.*

PROOF. The result is just a special case of the previous lemma. Note that $D(A^*) = A^{*-1}[H_0]$. □

A characterization of closedness of an operator is given in our next proposition, which we recall from earlier considerations about constructing Hilbert spaces.

PROPOSITION 83. *Let $A \subseteq H_0 \times H_1$ be a linear operator between the complex Hilbert spaces H_0, H_1. The operator A is a closed, linear operator if and only if $D(A)$ is a Hilbert space with respect to the graph norm $|\cdot|_{D(A)}$.*

In practical cases it is frequently quite obvious that not only the original operator A but also its adjoint A^* is densely defined, so that closability of A follows according to Proposition 82 as a by-product. A concept relevant in this context is explained in our next definition.

DEFINITION 84. Relations $A \subseteq H_0 \times H_1$, $B \subseteq H_1 \times H_0$ between the complex Hilbert spaces H_0, H_1 are called formally adjoint if

$$A \subseteq B^*. \tag{2.3.5}$$

If

$$A \subseteq -B^* \tag{2.3.6}$$

we say that A, B are formally skew-adjoint. In the special case of $A = B$ we speak of formally selfadjoint relation, if $A = -B$ we say A is formally skew-selfadjoint. If even $A = A^*$ or $A = -A^*$ then we say that A is selfadjoint or that A is skew-selfadjoint, respectively.

REMARK 85. Note that the terms "formally adjoint" and "formally skew-adjoint" are indeed symmetric in A and B. Since B^* is always a closed, linear subspace, inclusion (2.3.5) is equivalent to

$$\overline{Lin_{\mathbb{C}}\,A} \subseteq B^*.$$

The latter is in turn equivalent to

$$\overline{Lin_{\mathbb{C}}\,B} \subseteq A^*.$$

Indeed,

$$\overline{Lin_{\mathbb{C}}\, A} \subseteq B^* \Rightarrow B \subseteq \overline{Lin_{\mathbb{C}}\, B} = B^{**} \subseteq \overline{Lin_{\mathbb{C}}\, A}^* = A^*.$$

Similarly in the skew-adjoint case. If A, B are closed, linear and formally adjoint then for a closed linear V satisfying $A \subseteq V \subseteq B^*$ we also have $B \subseteq V^* \subseteq A^*$. Thus, V, B and V^*, A are formally adjoint. Conversely, if $V \subseteq B^*$ and $V^* \subseteq A^*$, we obtain$A \subseteq V$ and so

$$A \subseteq V \subseteq B^*.$$

Note that A and B are formally adjoint if and only if A and $-B$ formally skew-adjoint. Therefore it suffices to focus on one of the cases, the other being a trivial consequence.

We note an interesting relation between adjoints and closability.

Proposition 86. *Let $A \subseteq H_0 \times H_1$, $B \subseteq H_1 \times H_0$ be two linear operators between the complex Hilbert spaces H_0, H_1. Then we have that A and B are formally adjoint if and only if*

$$\bigwedge_{x \in D(A)} \bigwedge_{v \in D(B)} \langle Ax|v\rangle_{H_1} = \langle x|Bv\rangle_{H_0}.$$

If, moreover, both are densely defined then both linear operators are closable.

Proof. For (v, w) to be in A^* means according to earlier calculations, see (2.3.2),

$$\bigwedge_{x \in D(A)} \langle x|w\rangle_{H_0} = \langle Ax|v\rangle_{H_1}. \tag{2.3.7}$$

Since A and B are formally adjoint, we must have $B \subseteq A^*$. Thus, (2.3.7) implies that for all $(v, Bv) \in B \subseteq A^*$we have

$$\bigwedge_{x \in D(A)} \langle x|Bv\rangle_{H_0} = \langle Ax|v\rangle_{H_1}. \tag{2.3.8}$$

Conversely, comparing (2.3.8) the characterization (2.3.7) we see that (2.3.8) implies that $(v, Bv) \in A^*$. Since $v \in D(B)$ was arbitrary, we get indeed

$$B \subseteq A^*.$$

If now in addition A and B are densely defined, then A^* and B^* are closed linear operators, in particular right-unique. Consequently, $\overline{A}$ and $\overline{B}$ must also be right-unique as subsets of B^* and A^*, respectively. $\square$

Example 87. Let us consider once again the vector-analytical operation 'gradient' as a densely defined, linear operator

$$\begin{aligned} \operatorname{grad}\big|_{\mathring{C}_\infty(\Omega)} : \mathring{C}_\infty(\Omega) \subseteq L^2(\Omega) &\longrightarrow \bigoplus_{k=1,\dots,n} L^2(\Omega) \quad , \\ \varphi &\longmapsto \operatorname{grad}\varphi = (\partial_i \varphi)_{i=0,\dots,n-1} \end{aligned}$$

where $\Omega \subseteq \mathbb{R}^n$ is an open set, $\mathring{C}_\infty(\Omega)$ is the set of infinitely often differentiable function φ defined on $\mathbb{R}^n$ with $\varphi = 0$ outside of a bounded closed subset of Ω. By integration by parts we have that for $u \in \mathring{C}_\infty(\Omega)$

$$\bigwedge_{\Phi_i \in \mathring{C}_\infty(\Omega),\, i=0,\dots,n-1} \langle (\Phi_1, \dots \Phi_n)|\operatorname{grad} u\rangle_{\bigoplus_{k=1,\dots,n} L^2(\Omega)} + \langle \operatorname{div}(\Phi_1, \dots \Phi_n)|\, u\rangle_{L^2(\Omega)} = 0$$

where $\operatorname{div}(\Phi_1, \dots \Phi_n) := \sum_{i=0,\dots,n-1} \partial_i \Phi_i$, denotes the vector analytical 'divergence'. In our current terminology this means that the operator

$$\begin{aligned} \operatorname{div}\big|_{\bigoplus_{k=1,\dots,n} \mathring{C}_\infty(\Omega)} : \bigoplus_{k=1,\dots,n} \mathring{C}_\infty(\Omega) \subseteq \bigoplus_{k=1,\dots,n} L^2(\Omega) &\longrightarrow L^2(\Omega) \\ \phi &\longmapsto \operatorname{div}\phi = \sum_{k=1,\dots,n} \partial_i \phi_i \end{aligned}$$

and the operator $\operatorname{grad}\big|_{\mathring{C}_\infty(\Omega)}$ are formally skew-adjoint linear operators. Moreover, since we found that $\mathring{C}_\infty(\Omega)$ is dense in $L^2(\Omega)$, we also have that both operators are closable linear operators. This confirms our earlier finding that

$$\mathring{\operatorname{grad}} := \overline{\operatorname{grad}\big|_{\mathring{C}_\infty(\Omega)}}$$

is a closed linear operator and asserts in addition that also

$$\mathring{\operatorname{div}} := \overline{\operatorname{div}\big|_{\bigoplus_{k=1,\ldots,n}\mathring{C}_\infty(\Omega)}}$$

is also a closed linear operator. Moreover, we have that their adjoints define indeed – as anticipated earlier – closed linear operators

$$\operatorname{grad} := (-\mathring{\operatorname{div}})^*,\ \operatorname{div} := (-\mathring{\operatorname{grad}})^*.$$

By construction we also have

$$\mathring{\operatorname{grad}} \subseteq \operatorname{grad},\ \mathring{\operatorname{div}} \subseteq \operatorname{div}. \tag{2.3.9}$$

Since all these operators are closed, their domains are complex Hilbert spaces with respect to the graph norm denoted by

$$\begin{aligned} H(\mathring{\operatorname{grad}},\Omega) &:= D(\mathring{\operatorname{grad}}),\\ H(\mathring{\operatorname{div}},\Omega) &:= D(\mathring{\operatorname{div}}),\\ H(\operatorname{grad},\Omega) &:= D(\operatorname{grad}),\\ H(\operatorname{div},\Omega) &:= D(\operatorname{div}). \end{aligned}$$

As a consequence of (2.3.9) we have the inclusions

$$H(\mathring{\operatorname{grad}},\Omega) \subseteq H(\operatorname{grad},\Omega),\ H(\mathring{\operatorname{div}},\Omega) \subseteq H(\operatorname{div},\Omega).$$

Applying the projection theorem to this closed subspace configurations, we first get

$$H(\operatorname{grad},\Omega) = H(\mathring{\operatorname{grad}},\Omega) \oplus H(\mathring{\operatorname{grad}},\Omega)^\perp,\ H(\operatorname{div},\Omega) = H(\mathring{\operatorname{div}},\Omega) \oplus H(\mathring{\operatorname{div}},\Omega)^\perp.$$

To conclude this example let us determine the ortho-complements $H(\mathring{\operatorname{grad}},\Omega)^\perp$ and $H(\mathring{\operatorname{div}},\Omega)^\perp$. We first find

$$\Phi \in H(\mathring{\operatorname{div}},\Omega)^\perp \Leftrightarrow \bigwedge_{\Psi\in H(\mathring{\operatorname{div}})} \langle\Phi|\Psi\rangle_{\bigoplus_{k=1,\ldots,n} L^2(\Omega)} + \langle \operatorname{div}\Phi|\operatorname{div}\Psi\rangle_{L^2(\Omega)} = 0\,.$$

According to the above characterization of the domain of the adjoint, we get

$$\operatorname{div}\Phi \in D\left((\mathring{\operatorname{div}})^*\right) = D(\operatorname{grad})$$

and

$$-\operatorname{grad}\operatorname{div}\Phi = -\Phi$$

or

$$\Phi \in N(-\operatorname{grad}\operatorname{div}\ + 1).$$

Since the reasoning can be reversed, we obtain

$$N(-\operatorname{grad}\operatorname{div}\ + 1) = H(\mathring{\operatorname{div}},\Omega)^\perp.$$

Similarly,

$$N(-\operatorname{div}\operatorname{grad}\ + 1) = H(\mathring{\operatorname{grad}},\Omega)^\perp.$$

Thus, we have the orthogonal decompositions

$$H(\operatorname{grad},\Omega) = H(\mathring{\operatorname{grad}},\Omega) \oplus N(-\operatorname{div}\operatorname{grad}\ + 1),\ H(\operatorname{div},\Omega) = H(\mathring{\operatorname{div}},\Omega) \oplus N(-\operatorname{grad}\operatorname{div}\ + 1).$$

These decompositions can be employed to give a first access to the discussion of boundary value problems. Let $f \in H(\mathrm{grad})$ be given then the orthogonal projection u of f onto $N(-\operatorname{div}\operatorname{grad} + 1)$ yields a solution of the equation[4]

$$-\operatorname{div}\operatorname{grad} u + u = 0 \tag{2.3.10}$$

satisfying additionally the 'boundary condition'

$$u - f \in H(\mathring{\mathrm{grad}}, \Omega). \tag{2.3.11}$$

That (2.3.11) is indeed a generalized (inhomogeneous Dirichlet type) boundary condition (generalizing '$u = f$ on $\partial\Omega$') needs some clarification, which we shall, however, postpone. Similarly accepting a condition of the form

$$U - F \in H(\mathring{\mathrm{div}}, \Omega) \tag{2.3.12}$$

as a generalized inhomogeneous Neumann type boundary condition (generalizing '$n \cdot U = n \cdot F$ on $\partial\Omega$', where n denotes the (exterior) normal assuming it exists), we can interpret the second orthogonal decomposition as solving the problem

$$-\operatorname{grad}\operatorname{div} U + U = 0 \tag{2.3.13}$$

with U subject to the boundary condition (2.3.12). Since $U \in H(\mathrm{div}, \Omega)$ by construction, we deduce from (2.3.13) that also $\operatorname{grad}\operatorname{div} U \in H(\mathrm{div}, \Omega)$ and so

$$-\operatorname{div}\operatorname{grad}\operatorname{div} U + \operatorname{div} U = 0.$$

In other words, we have found a solution $\varphi := \operatorname{div} U$ satisfying again equation (2.3.10) and the boundary condition

$$\operatorname{grad}\varphi - F \in H(\mathring{\mathrm{div}}, \Omega).$$

Thus we have solved the inhomogeneous Dirichlet and Neumann type boundary value problems for equation (2.3.10). Following our solution philosophy we should also check for uniqueness and continuous dependence on the data in both cases. Uniqueness in the Dirichlet case is clear: Since the difference u of two solutions satisfies the homogeneous Dirichlet type boundary condition

$$u \in H(\mathring{\mathrm{grad}}, \Omega)$$

and the equation (2.3.10), i.e.

$$u \in N(-\operatorname{div}\operatorname{grad} + 1) := [\{0\}]\,(-\operatorname{div}\operatorname{grad} + 1) = H(\mathring{\mathrm{grad}}, \Omega)^{\perp},$$

it follows that $u = 0$. In the Neumann case we have that the difference φ of two solution satisfies

$$\operatorname{grad}\varphi \in H(\mathring{\mathrm{div}}, \Omega)$$

and (by applying grad to the equation)

$$-\operatorname{grad}\operatorname{div}\operatorname{grad}\varphi + \operatorname{grad}\varphi = 0,$$

i.e.

$$\operatorname{grad}\varphi \in N(-\operatorname{grad}\operatorname{div} + 1) = H(\mathring{\mathrm{div}}, \Omega)^{\perp}.$$

Therefore, we first find $\operatorname{grad}\varphi = 0$ and then, since

$$\varphi = \operatorname{div}\operatorname{grad}\varphi,$$

that $\varphi = 0$. The continuous dependence of the solution u in the Dirichlet case follows from the projection theorem

$$|u|_{H(\mathrm{grad},\Omega)} \leq \sqrt{|u|^2_{H(\mathrm{grad})} + |f - u|^2_{H(\mathrm{grad})}} = |f|_{H(\mathrm{grad})}.$$

---

[4]Solutions for vanishing boundary data of

$$-\operatorname{div}\mathring{\mathrm{grad}}\, u + u = g$$

or

$$-\mathring{\mathrm{div}}\operatorname{grad} u + u = g$$

have already been discussed earlier

In the Neumann case we argue that

$$-\overset{\circ}{\operatorname{div}}\,(\operatorname{grad}\varphi - F) + \varphi = \operatorname{div} F,$$

and so after multiplying by φ in the sense of the $L^2(\Omega)$ we get

$$-\left\langle \varphi|\overset{\circ}{\operatorname{div}}\,(\operatorname{grad}\varphi - F)\right\rangle_{L^2(\Omega)} + \langle\varphi|\varphi\rangle_{L^2(\Omega)} = \langle\varphi|\operatorname{div} F\rangle_{L^2(\Omega)}.$$

Since $\varphi \in D(\overset{\circ}{\operatorname{div}}^*) = D(\operatorname{grad})$ this yields

$$\langle\operatorname{grad}\varphi|\operatorname{grad}\varphi - F\rangle_{\bigoplus_{k=1,\dots,n} L^2(\Omega)} + \langle\varphi|\varphi\rangle_{L^2(\Omega)} = \langle\varphi|\operatorname{div} F\rangle_{L^2(\Omega)}$$

or

$$\langle\operatorname{grad}\varphi|\operatorname{grad}\varphi\rangle_{\bigoplus_{k=1,\dots,n} L^2(\Omega)} + \langle\varphi|\varphi\rangle_{L^2(\Omega)} = \langle\varphi|\operatorname{div} F\rangle_{L^2(\Omega)} + \langle\operatorname{grad}\varphi|\,F\rangle_{\bigoplus_{k=1,\dots,n} L^2(\Omega)}.$$

Finally, noting that $(\operatorname{div} F, F) \in L^2(\Omega) \oplus \bigoplus_{k=1,\dots,n} L^2(\Omega)$ and using the Cauchy-Schwarz inequality in $L^2(\Omega) \oplus \bigoplus_{k=1,\dots,n} L^2(\Omega)$, we get

$$|\varphi|^2_{H(\operatorname{grad},\Omega)} \le |\varphi|_{H(\operatorname{grad},\Omega)}\,|F|_{H(\operatorname{div},\Omega)}$$

or

$$|\varphi|_{H(\operatorname{grad},\Omega)} \le |F|_{H(\operatorname{div},\Omega)}.$$

2.3.2. Continuous, Linear Operators. Various instances of continuous, linear operators we have also already encountered. To illuminate the special qualities of this operator class we consider them now in slightly more detail in particular in conjunction with the concepts studied in the previous subsection. By continuous extension clearly a densely defined, continuous linear operator extends to a continuous linear operator defined on the whole space (with the same operator norm, i.e. Lipschitz (semi-) norm. We have already used these ideas in the discussion of the resolvent operator. The adjoint of a densely defined, continuous, linear operator is also a continuous, linear operator defined on the whole space. Let us collect these fundamental findings in our next proposition. Since continuous linear mappings, also map bounded sets to bounded sets they are also known as bounded, linear mappings, although, such a mapping is only bounded in our earlier sense, if it is the constant linear mapping 0!

PROPOSITION 88. *Let $A \subseteq H_0 \times H_1$, be a continuous, linear operator between complex Hilbert spaces H_0, H_1. Then A is always closable. If the closure $\bar{A}$ is defined on H_0 then A must be densely defined. If A is densely defined, then*

$$\bar{A} \in L(H_0, H_1),\ A^* \in L(H_1, H_0).$$

PROOF. Since A is continuous, it is also Lipschitz continuous and therefore closable. If A is densely defined, then $\bar{A}$ is defined on all of H_0 and is Lipschitz continuous with the same Lipschitz (semi-) norm, i.e. the operator norm. If $\bar{A}$ is defined on H_0 then for $(x, y) \in \bar{A}$ there is a sequence $((x_n, y_n))_n$ in A with

$$x_n \to x \text{ as } n \to \infty \text{ and } y_n \to y \text{ as } n \to \infty.$$

Since $x \in H_0$ can be arbitrary, we thus have always a sequence in $D(A)$ approximating an arbitray $x \in H_0$. In other words, A is densely defined. To investigate A^* we use the convenient characterization of Proposition 79. For fixed $y \in H_1$ consider the linear functional $\langle y|\bar{A}\cdot\rangle_{H_1}$ given by

$$\begin{aligned} H_0 &\to \mathbb{C} \\ x &\mapsto \langle y|\bar{A}x\rangle_{H_1} \end{aligned}.$$

This functional is also continuous (by the Cauchy-Schwarz inequality)

$$\left|\langle y\,|\bar{A}x\,\rangle_{H_1}\right| \le |y|_{H_1}\,|\bar{A}x|_{H_1} \le |y|_{H_1}\,||\bar{A}||\,|x|_{H_0} = |y|_{H_1}\,||A||\,|x|_{H_0}$$

for all $x \in H_0$, $y \in H_1$. Thus, we have $\left\langle y|\bar{A}\cdot\right\rangle_{H_1} \in H_0'$ and applying the corresponding Riesz mapping R_{H_0} we also obtain

$$\bigwedge_{x\in H_0} \left\langle y\,|\bar{A}x\,\right\rangle_{H_1} = \left\langle R_{H_0}\left\langle y|\bar{A}\cdot\right\rangle_{H_1} \middle| x\right\rangle_{H_0}.$$

In other words,

$$y \in D(\bar{A}^*) = D(A^*) \text{ and } A^*y = R_{H_0}\left\langle y\,|\bar{A}\cdot\,\right\rangle_{H_1}.$$

Since $y \in H_1$ was arbitrary, we see that $D(A^*) = H_1$. Moreover,

$$|A^*y|_{H_0} = \left|R_{H_0}\left\langle y|\bar{A}\cdot\right\rangle_{H_1}\right|_{H_0} = \left|\left\langle y|\bar{A}\cdot\right\rangle_{H_1}\right|_{H_0'} \leq |y|_{H_1}\,||A||$$

for all $y \in H_1$. This estimate shows that $A^* \in L(H_1, H_0)$ and $\|A^*\| \leq \|A\|$. Exchanging the role of $\bar{A}$ and A^*, we also find $\|\bar{A}\| = \|A^{**}\| \leq \|A^*\|$. □

We note the following useful results about the interaction between continuous, linear operators with possibly discontinuous, linear operators with regards to taking adjoints.

LEMMA 89. *Let $B \in L(H_0, H_1)$ and let $A \subseteq H_0 \oplus H_1$ be a closed, densely defined, linear operator between complex Hilbert spaces H_0 and H_1. Then*

$$(A + B)^* = A^* + B^*.$$

PROOF. Since for all $x \in D(A^*) \subseteq D(B^*) = H_1$ and all $y \in D(A) \subseteq D(B) = H_0$ we have

$$\langle (A+B)y|x\rangle_{H_1} = \langle Ay\,|x\,\rangle_{H_1} + \langle By\,|x\,\rangle_{H_1} = \langle y\,|(A^* + B^*)x\,\rangle_{H_0},$$

it follows that

$$A^* + B^* \subseteq (A+B)^*.$$

For $x \in D\left((A+B)^*\right)$ we have

$$\bigwedge_{y\in D(A)} \langle (A+B)y|x\rangle_{H_{23}} = \langle y\,|(A+B)^*x\,\rangle_{H_0}.$$

Since B is continuous,

$$\bigwedge_{y\in D(BA)} \langle (A+B)y|x\rangle_{H_1} = \langle Ay\,|x\,\rangle_{H_1} + \langle y\,|B^*x\,\rangle_{H_1} = \langle y\,|(A+B)^*x\,\rangle_{H_0}$$

and we see that $x \in D(A^*)$ and $A^*x = (A+B)^*x - B^*x$. □

LEMMA 90. *Let $B \in L(H_1, H_2)$, $C \in L(H_0, H_0)$ be operators between complex Hilbert spaces H_1 and H_2 and H_0 and H_0, respectively, and let $A \subseteq H_0 \oplus H_1$ be a closed, densely defined, linear operator between complex Hilbert spaces H_0 and H_1. Then*

$$(B\,A)^* = A^*B^*,\ (AC)^* = \overline{C^*A^*}.$$

PROOF. Since for all $x \in D(A^*B^*)$ and all $y \in D(BA)$ we have

$$\langle BAy|x\rangle_{H_2} = \langle Ay\,|B^*x\,\rangle_{H_1} = \langle y\,|A^*B^*x\,\rangle_{H_0},$$

it follows that

$$A^*B^* \subseteq (B\,A)^*.$$

For $x \in D\left((B\,A)^*\right)$ we have

$$\bigwedge_{y\in D(BA)} \langle BAy|x\rangle_{H_2} = \langle y\,|(B\,A)^*x\,\rangle_{H_0}.$$

Since B is continuous,

$$\bigwedge_{y\in D(BA)} \langle Ay|B^*x\rangle_{H_2} = \langle y\,|(B\,A)^*x\,\rangle_{H_0}$$

and we see that $B^*x \in D(A^*)$ and $A^*B^*x = (B\,A)^*x$. Applying our findings we also obtain $(C^*A^*)^* = A^{**}\,C^{**} = A\,C$. Taking adjoints this yields

$$(A\,C)^* = (C^*A^*)^{**} = \overline{C^*A^*}.$$

$\square$

A rather special case of boundedness which is of particular interest here is the case of a unitary operator $U : H_0 \to H_1$.

Proposition 91. *Let $U \in L(H_0, H_1)$ be a unitary operator between complex Hilbert spaces H_0, H_1. If the spaces H_0, H_1 are not equal to $\{0\}$ then*

$$||U|| = 1.$$

Moreover,

$$U^{-1} = U^*$$

and if $H := H_0 = H_1$ we have

$$\sigma(U) \subseteq \partial B_{\mathbb{C}}(0,1).$$

Proof. That

$$||U|| = 1$$

is obvious from the isometric property of U :

$$|Ux|_{H_1} = |x|_{H_0} \text{ for all } x \in H_0.$$

From this we also find by polarization

$$\langle Ux|Uy\rangle_{H_1} = \langle x|y\rangle_{H_0} \text{ for all } x,\, y \in H_0.$$

We read off

$$U^*U = 1 \text{ on } H_0$$

and applying U to this equality and noting that U is assumed to be onto we also get

$$U\,U^* = 1 \text{ on } H_1.$$

In other words,

$$U^{-1} = U^*.$$

In regards to the spectral properties we have from the spectral theorem (variant 3) using $H := H_0 = H_1$ and

$$|(U-\lambda)\,x|_H \geq ||Ux|_H - |\lambda|\,|x|_H| = |1-|\lambda||\,|x|_H$$
$$|(U^*-\lambda)\,x|_H \geq ||U^*x|_H - |\lambda|\,|x|_H| = |1-|\lambda||\,|x|_H$$

for all $x \in H$, that

$$B_{\mathbb{C}}(0,1) \cup (\mathbb{C} \setminus B_{\mathbb{C}}(0,1)) \subseteq \varrho\,(U) \cap \varrho\,(U^*)$$

and so indeed

$$\sigma(U) \subseteq \partial B_{\mathbb{C}}(0,1),\ \sigma(U^*) \subseteq \partial B_{\mathbb{C}}(0,1). \tag{2.3.14}$$

$\square$

Example 92. We consider the so-called Fourier transform $\widetilde{\mathcal{F}}$ as a particular example for a unitary mapping. Let $\varphi \in \mathring{C}_\infty(\mathbb{R})$ then there is some $N \in \mathbb{N}$ such that $\varphi = 0$ outside of the interval

$]-2^{N-1},+2^{N-1}[$. Then the rescaled function $\varphi_N \coloneqq \sigma_{2^N}\,\varphi$ is in $L^2(]-1/2,+1/2[)$, $\sigma_c\,\varphi \coloneqq \sqrt{|c|}\,\varphi(c\cdot)$, $c\in\mathbb{R}\setminus\{0\}$[5].The Fourier coefficients of φ_N are

$$\begin{aligned}\widehat{\varphi}_N(k) &\coloneqq \langle exp(2\pi\mathrm{i}\,k\cdot)|\varphi_N\rangle_{L^2(\mathbb{R})}\\ &= \int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,kx)\,\varphi_N(x)\,dx\\ &= 2^{-N/2}\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,k\,2^{-N}y)\,\varphi(y)\,dy\\ &\eqqcolon 2^{-N/2}\,\widehat{\varphi}(k\,2^{-N}).\end{aligned}$$

According to what we said about the Fourier series $\left(\widehat{\varphi}(k\,2^{-N})\right)_{k\in\mathbb{Z}}$ conserves all the information to reconstruct φ. Since fractions of the form $\frac{k}{2^N}$ are dense in $\mathbb{R}$, we are led to consider the so-called Fourier transform $\varphi\mapsto\widetilde{\mathcal{F}}\varphi$ for arbitrary $\varphi\in\mathring{C}_\infty(\mathbb{R})$, where $\left(\widetilde{\mathcal{F}}\varphi\right)(x)\coloneqq\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,\varphi(y)\,dy$. We also see that $\widetilde{\mathcal{F}}\varphi\in L^2(\mathbb{R})$, since φ has compact support, and

$$\begin{aligned}\int_{\mathbb{R}}\left|\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,\varphi(y)\,dy\right|^2 dx &= \int_{\mathbb{R}}\tfrac{1}{1+x^2}\left|\int_{\mathbb{R}}(D-\mathrm{i})\,exp(-2\pi\mathrm{i}\,x\,y)\,\varphi(y)\,dy\right|^2 dx,\\ &= \int_{\mathbb{R}}\tfrac{1}{1+x^2}\left|\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,(D+\mathrm{i})\varphi(y)\,dy\right|^2 dx,\\ &\le \int_{\mathbb{R}}\tfrac{1}{1+x^2}\left|\int_{\mathbb{R}}|(D+\mathrm{i})\varphi(y)|\,dy\right|^2 dx,\\ &\le \int_{\mathbb{R}}\tfrac{1}{1+x^2}\int_{\mathbb{R}}\tfrac{1}{1+y^2}\,dy\int_{\mathbb{R}}(1+y^2)\,|(D+\mathrm{i})\varphi(y)|^2\,dy\,dx\\ &= \pi^2\int_{\mathbb{R}}(1+y^2)\,|(D+\mathrm{i})\varphi(y)|^2\,dy = |(m-\mathrm{i})(D-\mathrm{i})\varphi|^2_{L^2(\mathbb{R})}\,,\end{aligned} \tag{2.3.15}$$

where m symbolizes multiplication by the argument, i.e. $(m\varphi)(x)\coloneqq x\varphi(x)$ for $x\in\mathbb{R}$, and D abbreviates $\frac{1}{2\pi\mathrm{i}}\partial$. This suggests to establish the Fourier transform $\widetilde{\mathcal{F}}$ as the closure of the mapping

$$\begin{array}{ccc}\mathring{C}_\infty(\mathbb{R})\subseteq L^2(\mathbb{R}) & \to & L^2(\mathbb{R})\\ \varphi & \mapsto & \widetilde{\mathcal{F}}\varphi\end{array}.$$

That this is well-defined will follow by showing that the Fourier transform is an isometry. The Fischer-Riesz theorem shows that for all sufficiently large $N\in\mathbb{N}$

$$\sum_{k=-\infty}^{\infty} 2^{-N}\left|\widetilde{\mathcal{F}}\varphi(k\,2^{-N})\right|^2 = \int_{\mathbb{R}}|\varphi(y)|^2\,dy.$$

We define

$$W_N(x)\coloneqq\begin{cases}\vdots\\ \left|\widetilde{\mathcal{F}}\varphi(k\,2^{-N})\right|^2 & \text{for } x\in[k\,2^{-N},\,(k+1)\,2^{-N}[,\ k\in\mathbb{Z}\\ \vdots\end{cases}$$

and apparently obtain a non-negative step function W_N with countably many steps approximating $\left|\widetilde{\mathcal{F}}\varphi(\cdot)\right|^2$ point-wise as $N\to\infty$ and we have

$$\int_{\mathbb{R}} W_N(x)\,dx = \int_{\mathbb{R}}|\varphi(y)|^2\,dy$$

for all sufficiently large $N\in\mathbb{N}$. Now by Fatou's lemma $\left|\widetilde{\mathcal{F}}\varphi(\cdot)\right|^2$ is integrable and

$$\int_{\mathbb{R}}\left|\widetilde{\mathcal{F}}\varphi(x)\right|^2 dx \le \int_{\mathbb{R}}|\varphi(y)|^2\,dy.$$

[5]Note that the rescaling operation σ_c is defined in such a way to make it a unitary mapping in $L^2(\mathbb{R})$. Indeed, $\int_{\mathbb{R}}|(\sigma_c\varphi)(x)|^2dx = \int_{\mathbb{R}}|\varphi(c\,x)|^2|c|\,dx = \int_{\mathbb{R}}|\varphi(x)|^2\,dx$

Moreover, similar to (2.3.15) we have

$$\begin{aligned}
\left|\widetilde{\mathcal{F}}\varphi(x)\right| &= \left|\int_{\mathbb{R}} exp(-2\pi \mathrm{i}\, x\, y)\, \varphi(y)\, dy\right|, \\
&= \tfrac{1}{\sqrt{1+x^2}} \left|\int_{\mathbb{R}} (D-\mathrm{i})\, exp(-2\pi \mathrm{i}\, x\, y)\, \varphi(y)\, dy\right|, \\
&= \tfrac{1}{\sqrt{1+x^2}} \left|\int_{\mathbb{R}} exp(-2\pi \mathrm{i}\, x\, y)\, (D-\mathrm{i})\varphi(y)\, dy\right|, \\
&\leq \tfrac{1}{\sqrt{1+x^2}} \int_{\mathbb{R}} |(D-\mathrm{i})\varphi(y)|\, dy, \\
&\leq \pi^{1/2}\, \tfrac{1}{\sqrt{1+x^2}} \sqrt{\int_{\mathbb{R}} (1+y^2)\, |(D-\mathrm{i})\varphi(y)|^2\, dy} = \\
&\qquad = \pi^{1/2}\, \tfrac{1}{\sqrt{1+x^2}}\, |(m-\mathrm{i})(D-\mathrm{i})\varphi|_{L^2(\mathbb{R})}\,,
\end{aligned}$$

and so also

$$\begin{aligned}
\left|\partial\widetilde{\mathcal{F}}\varphi(x)\right| &= \left|\partial \int_{\mathbb{R}} exp(-2\pi \mathrm{i}\, x\, y)\, \varphi(y)\, dy\right|, \\
&= 2\pi \left|\int_{\mathbb{R}} exp(-2\pi \mathrm{i}\, x\, y)\, y\, \varphi(y)\, dy\right|, \\
&\leq 2\pi^{3/2}\, \tfrac{1}{\sqrt{1+x^2}}\, |(m-\mathrm{i})(D-\mathrm{i})(m\varphi)|_{L^2(\mathbb{R})}\,, \\
&\leq 2\pi^{3/2}\, \tfrac{1}{\sqrt{1+x^2}} \left(\left|(m-\mathrm{i})^2(D-\mathrm{i})\varphi\right|_{L^2(\mathbb{R})} + |(m-\mathrm{i})\varphi|_{L^2(\mathbb{R})} + |(m-\mathrm{i})(D-\mathrm{i})\varphi|_{L^2(\mathbb{R})}\right).
\end{aligned}$$

With this and observing that

$$|\left(\widetilde{\mathcal{F}}\varphi\right)(x) - \left(\widetilde{\mathcal{F}}\varphi\right)(y)| \leq \sup\left\{|\partial\widetilde{\mathcal{F}}\varphi(t)|\,\middle|\, t \in [k\, 2^{-N},\, (k+1)\, 2^{-N}[\right\}\, 2^{-N}$$

for all $x,\, y \in [k\, 2^{-N},\, (k+1)\, 2^{-N}[$, we get

$$|\left(\widetilde{\mathcal{F}}\varphi\right)(y)| \leq |\left(\widetilde{\mathcal{F}}\varphi\right)(x)| + C\, \frac{1}{\sqrt{1+\frac{k^2}{2^{2N}}}} 2^{-N}$$

for all $x,\, y \in [k\, 2^{-N},\, (k+1)\, 2^{-N}[$. This proves that

$$W_N(x) \leq 2\, |\left(\widetilde{\mathcal{F}}\varphi\right)(x)|^2 + U_N(x)$$

with

$$U_N(x) := 2C^2 \begin{cases} \vdots \\ \frac{1}{2^{2N}+k^2} & \text{for } x \in [k\, 2^{-N},\, (k+1)\, 2^{-N}[,\ k \in \mathbb{Z} \\ \vdots \end{cases} .$$

Since

$$\int_{\mathbb{R}} U_N(x)\, dx = 2C^2\, 2^{-N} \sum_{k\in\mathbb{Z}} \frac{1}{2^{2N}+k^2} \leq 2C^2\, 2^{-N} \sum_{k\in\mathbb{Z}} \frac{1}{k^2},$$

we have (by so-called dominated convergence) that even

$$\int_{\mathbb{R}} \left|\left(\widetilde{\mathcal{F}}\varphi\right)(x)\right|^2 dx = \int_{\mathbb{R}} |\varphi(y)|^2\, dy.$$

Since $\varphi \in \mathring{C}_\infty(\mathbb{R})$ was arbitrary, this proves that $\widetilde{\mathcal{F}}\big|_{\mathring{C}_\infty(\mathbb{R})}$ is an isometry and extends by continuity to its closure $\widetilde{\mathcal{F}}$[6]. Let us calculate its adjoint $\widetilde{\mathcal{F}}^*$. The adjoint operator $\widetilde{\mathcal{F}}^*$ must satisfy

$$\left\langle \widetilde{\mathcal{F}}\varphi \,|\, \psi \right\rangle_{L^2(\mathbb{R})} = \left\langle \varphi \,|\, \widetilde{\mathcal{F}}^*\psi \right\rangle_{L^2(\mathbb{R})} \tag{2.3.16}$$

[6] It should be noted that this way $\widetilde{\mathcal{F}}\varphi$ is defined even if $\varphi \notin L^1(\mathbb{R})$ and so the integral defining the Fourier transform originally does *not* make sense.

where it suffices to consider $\varphi,\,\psi \in \mathring{C}_\infty(\mathbb{R})$. We calculate using Fubini's theorem

$$\begin{aligned}\left\langle \widetilde{\mathcal{F}}\varphi\,|\,\psi\right\rangle_{L^2(\mathbb{R})} &= \textstyle\int_{\mathbb{R}}\int_{\mathbb{R}} exp(2\pi\mathrm{i}\,x\,y)\,\varphi(y)^* dy\,\psi(x)\,dx,\\ &= \textstyle\int_{\mathbb{R}}\int_{\mathbb{R}} exp(2\pi\mathrm{i}\,x\,y)\,\varphi(y)^*\psi(x)\,dx\,dy,\\ &= \textstyle\int_{\mathbb{R}}\varphi(y)^*\int_{\mathbb{R}} exp(2\pi\mathrm{i}\,x\,y)\,\psi(x)\,dx\,dy,\end{aligned}$$

which shows that

$$(\widetilde{\mathcal{F}}^*\psi)(y) = \int_{\mathbb{R}} exp(2\pi\mathrm{i}\,x\,y)\,\psi(x)\,dx$$

for all $\psi \in \mathring{C}_\infty(\mathbb{R})$. We observe that $\widetilde{\mathcal{F}}^*$ is just the Fourier transform rescaled by (-1). Indeed, we find (by simple substitution)

$$\widetilde{\mathcal{F}}^* = \sigma_{-1}\widetilde{\mathcal{F}} = \widetilde{\mathcal{F}}\,\sigma_{-1} \tag{2.3.17}$$

(first in $\mathring{C}_\infty(\mathbb{R})$ then by continuity in $L^2(\mathbb{R})$). Recalling that scaling is an isometry (indeed unitary) we see that $\widetilde{\mathcal{F}}^*$ is isometric as a composition of to isometries. Moreover, we have applying (2.3.16) with $\psi = \widetilde{\mathcal{F}}\varphi$

$$\langle\varphi\,|\,\varphi\rangle_{L^2(\mathbb{R})} = \left\langle\widetilde{\mathcal{F}}\varphi\,|\,\widetilde{\mathcal{F}}\varphi\right\rangle_{L^2(\mathbb{R})} = \left\langle\varphi\,|\,\widetilde{\mathcal{F}}^*\widetilde{\mathcal{F}}\varphi\right\rangle_{L^2(\mathbb{R})}$$

for all $\varphi \in L^2(\mathbb{R})$. By polarization we find

$$\langle\varphi\,|\,\psi\rangle_{L^2(\mathbb{R})} = \left\langle\varphi\,|\,\widetilde{\mathcal{F}}^*\widetilde{\mathcal{F}}\psi\right\rangle_{L^2(\mathbb{R})}$$

for all $\varphi,\,\psi \in L^2(\mathbb{R})$ and so

$$\widetilde{\mathcal{F}}^*\widetilde{\mathcal{F}} = 1_{L^2(\mathbb{R})}.$$

With (2.3.17) we see also

$$\widetilde{\mathcal{F}}\widetilde{\mathcal{F}}^* = \widetilde{\mathcal{F}}\,\sigma_{-1}\,\widetilde{\mathcal{F}} = \widetilde{\mathcal{F}}^*\widetilde{\mathcal{F}} = 1_{L^2(\mathbb{R})}.$$

This finally confirms that $\widetilde{\mathcal{F}}$ is also onto and so unitary. The adjoint Fourier transform $\widetilde{\mathcal{F}}^*$ is therefore its inverse $\widetilde{\mathcal{F}}^{-1}$. According to the above result we must have $\sigma(\widetilde{\mathcal{F}}) \subseteq \partial B_{\mathbb{C}}(0,1)$. Indeed, the spectrum will turn out to be

$$\sigma(\widetilde{\mathcal{F}}) = P\sigma(\widetilde{\mathcal{F}}) = \{+1, +\mathrm{i}\,-1, -\mathrm{i}\}. \tag{2.3.18}$$

To find the spectrum we shall explicitely construct the eigensolutions of $\widetilde{\mathcal{F}}$.

Recall the Gauss distribution function $\gamma \in L^2(\mathbb{R})$ given by $\gamma(x) = exp(-\pi x^2)$, $x \in \mathbb{R}$. We shall see that it is an eigensolution of $\widetilde{\mathcal{F}}$ associated with the eigenvalue $+1$. The function

$$\widehat{\gamma}(x) := \int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,\gamma(y)\,dy$$

is well-defined and in $L^2(\mathbb{R})$ by the estimate (2.3.15) applied to γ. We need to show that $\gamma = \widehat{\gamma}$. We calculate $\partial\gamma(x) = -2\pi\,x\,\gamma\,(x)$ and with an integration by parts

$$\begin{aligned}\partial\widehat{\gamma}(x) &= -2\pi\mathrm{i}\textstyle\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,y\,\gamma(y)\,dy,\\ &= \imath\textstyle\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,\partial\gamma(y)\,dy,\\ &= -2\pi x\textstyle\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,\gamma(y)\,dy,\\ &= -2\pi x\,\widehat{\gamma}(x).\end{aligned}$$

Thus, γ, $\widehat{\gamma}$ satisfy the same homogeneous ordinary differential equation. Moreover, we find $\gamma(0) = 1$ and

$$\widehat{\gamma}(0) = \int_{\mathbb{R}}\gamma(y)\,dy = 1. \tag{2.3.19}$$

By the classical uniqueness of solutions for initial value problems for linear ordinary differential equations we therefore must have $\gamma = \widehat{\gamma}$. (Equation (2.3.19) can be shown by applying the residuum calculus). Letting[7] $\mathcal{D} := -\mathrm{i}\sqrt{\pi}\,(D + \mathrm{i}\,m)$ with $D := \frac{1}{2\pi\mathrm{i}}\partial$ we find

$$\widetilde{\mathcal{F}}\mathcal{D}\gamma = (-\mathrm{i})\,\mathcal{D}\widetilde{\mathcal{F}}\gamma = (-\mathrm{i})\,\mathcal{D}\gamma. \tag{2.3.20}$$

Indeed,

$$\begin{aligned} \int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,\mathcal{D}\varphi(y)\,dy &= -\mathrm{i}\sqrt{\pi}\,\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y) D\varphi(y)\,dy + \\ &\quad +\sqrt{\pi}\,\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,y\,\varphi(y)\,dy, \\ &= -\mathrm{i}\sqrt{\pi}\,x\,\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,\varphi(y)\,dy + \\ &\quad -\mathrm{i}\sqrt{\pi}\,D\,\int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,\varphi(y)\,dy, \\ &= -\mathrm{i}\,\mathcal{D}\,\widehat{\varphi}(x) \end{aligned}$$

for all $\varphi \in C_1(\mathbb{R})$ with sufficient decay at infinity and thus certainly for the function γ. From (2.3.20) we now see that $-\mathrm{i}$ belongs to the point spectrum of $\mathcal{F}$ and $\mathcal{D}\gamma$ is an eigensolution associated with this eigensolution. By induction we see

$$\widetilde{\mathcal{F}}\,\mathcal{D}^k\gamma = (-\mathrm{i})^k\,\mathcal{D}^k\gamma$$

for all $k \in \mathbb{N}$. Thus, $\{(-\mathrm{i})^k \,|\, k = 0,1,2,3\} \subseteq P\sigma(\mathcal{F})$. We shall see shortly that $\{\Gamma_k \,|\, k \in \mathbb{N}\}$, $\Gamma_k := \frac{1}{|\mathcal{D}^k\gamma|_{L^2(\mathbb{R})}}\mathcal{D}^k\gamma$, is a complete orthonormal set. Assuming this for now we would indeed confirm that (2.3.18) holds. In this case, it can be seen that the operator

$$\phi \mapsto \sum_{k\in\mathbb{N}} \frac{1}{(-\mathrm{i})^k - \lambda}\,\langle \Gamma_k \,|\, \phi\rangle_{L^2(\mathbb{R})}\,\Gamma_k$$

is a well-defined bounded operator in $L^2(\mathbb{R})$ for every $\lambda \notin \{1, -\mathrm{i} - 1, \mathrm{i}\}$ and is indeed the resolvent $(\widetilde{\mathcal{F}} - \lambda)^{-1}$. Recall that $\{\,\Gamma_k |\, k \in \mathbb{N}\}$ is already an orthonormal set. We need to show completeness. So, let $f \perp \Gamma_k$ for all $k \in \mathbb{N}$, then noting that $\Gamma_k = Q_k\,\gamma$, where Q_k is a polynomial of degree k, we see that

$$f \perp m^k\,\gamma \text{ for all } k \in \mathbb{N}. \tag{2.3.21}$$

We also note $\gamma\, f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and $exp(2\pi\mathrm{i}\,x\,\cdot)\,\gamma \in L^2(\mathbb{R})$ and so

$$F(x) := (\widetilde{\mathcal{F}}\gamma f)(x) = \int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,\gamma(y)\,f(y)\,dy = \langle exp(2\pi\mathrm{i}\,x\,\cdot)\,\gamma \,|\, f\rangle_{L^2(\mathbb{R})}\,. \tag{2.3.22}$$

From the integral representation we see that F is well-defined even for $x \in \mathbb{C}$, since $exp(2\pi\mathrm{i}\,x\,\cdot)\gamma \in L^2(\mathbb{R})$ for all $x \in \mathbb{C}$. Moreover, we realize that F is analytic. We calculate the Taylor coefficients by observing that

$$(\partial^k F)(x) = (-2\pi\mathrm{i})^k \int_{\mathbb{R}} exp(-2\pi\mathrm{i}\,x\,y)\,y^k\,\gamma(y)\,f(y)\,dy = (-2\pi\mathrm{i})^k\,\left\langle exp(2\pi\mathrm{i}\,x\,\cdot)\,m^k\,\gamma \,|\, f\right\rangle_{L^2(\mathbb{R})}\,.$$

Consequently, by (2.3.21)

$$\frac{1}{k!}(\partial^k F)(0) = (-2\pi\mathrm{i})^k\,\left\langle m^k\,\gamma \,|\, f\right\rangle_{L^2(\mathbb{R})} = 0$$

---

[7]Due to the usage of different manuscripts the operator $\mathcal{D}$ defined earlier is different from the one defined here. The connection is via unitary equivalence (see later):

$$\sigma_{\sqrt{2\pi}}\,\frac{1}{\sqrt{2}}\,(m - \partial)\,\sigma^*_{\sqrt{2\pi}} = -\mathrm{i}\sqrt{\pi}\,(D + \mathrm{i}\,m)\,.$$

The unitary re-scaling operation σ_α with $\alpha \in \mathbb{R}\setminus\{0\}$ is given by

$$(\sigma_\alpha\varphi)\,(x) = \sqrt{|\alpha|}\,\varphi\,(\alpha x)$$

for $x \in \mathbb{R}$ and $\varphi \in \mathring{C}_\infty\,(\mathbb{R})$. Note that

$$\sigma^*_\alpha = \sigma_{\alpha^{-1}}\,.$$

for all $k \in \mathbb{N}$. For an analytic function this means that $F = 0$. Applying the inverse Fourier transform to (2.3.22) we now get $\gamma f = 0$ and so $f = 0$. This finally shows the completeness of the orthonormal set $\{\Gamma_k \mid k \in \mathbb{N}\}$ and thus completes the proof of (2.3.18). Moreover, the Fourier transform can be represented by

$$\widetilde{\mathcal{F}}\varphi = \sum_{k \in \mathbb{N}} (-\mathrm{i})^k \ \langle \Gamma_k \mid f \rangle_{L^2(\mathbb{R})} \ \Gamma_k.$$

It is sometimes more appropriate to use a re-scaled version $\mathcal{F}$ of the Fourier transform $\widetilde{\mathcal{F}}$ given by

$$\mathcal{F} \coloneqq \mathcal{L}_0 \coloneqq \sigma^*_{\sqrt{2\pi}} \, \widetilde{\mathcal{F}} \, \sigma_{\sqrt{2\pi}},$$

a particular instance of the so-called Fourier-Laplace transform, which is a family of unitary transforms $\mathcal{L}_\nu \coloneqq \mathcal{F} \ \exp(-\nu m)$, $\nu \in \mathbb{R}$, where

$$(\exp(-\nu m)\, \varphi)\,(x) \coloneqq \exp(-\nu x)\, \varphi\,(x)$$

for $x \in \mathbb{R}$ and $\varphi \in \mathring{C}_\infty(\mathbb{R})$.

To make this unitary we introduce the exponential weight function $t \mapsto \exp(-\nu\, t)$, $\nu \in \mathbb{R}$, and consider the weighted L^2-space $H_{\nu,0}$ generated by completion of $\mathring{C}_\infty(\mathbb{R})$ with respect to the inner product $\langle \,\cdot \mid \cdot\, \rangle_{\nu,0}$

$$(\varphi, \psi) \mapsto \int_{\mathbb{R}} \varphi\,(t)^* \, \psi\,(t) \ \exp(-2\nu t)\, dt.$$

The associated norm will be denoted by $|\cdot|_{\nu,0}$. The multiplication operator

$$\begin{gathered} \mathring{C}_\infty(\mathbb{R}) \subseteq H_{\nu,0} \to \mathring{C}_\infty(\mathbb{R}) \subseteq H_{0,0} = L^2(\mathbb{R}) \\ \varphi \mapsto \exp(-\nu m)\, \varphi \end{gathered}$$

clearly has a unitary extension, which we shall denote by $\exp(-\nu m)$, where the m serves as a reminder for 'multiplication'.

Its inverse will be denoted by $\exp(\nu m)$.

2.3.3. Hermitean, Skew-Hermitean, Symmetric, Skew-Symmetric, Selfadjoint and Skew-Selfadjoint Operators. Already from linear algebra we recall that this operator class is of particular interest. In the unbounded case matters become, however, slightly more intricate.

Definition 93. Let $A \subseteq H \oplus H$ be linear operator in complex Hilbert space H. We call A a **formally selfadjoint operator** or **Hermitean** if A is contained in the adjoint relation

$$A \subseteq A^*.$$

We call A a **formally skew-selfadjoint operator** or **skew-Hermitean** if A is contained in the negative adjoint relation

$$A \subseteq -A^* = \left(A^{\perp}\right)^{-1}.$$

If, moreover, A^* is an operator, i.e. if A is densely defined, then we call A **symmetric**. If even

$$A = A^*,$$

then we call A **selfadjoint**. A symmetric operator A is called **essentially selfadjoint** if $\bar{A}$ is selfadjoint, i.e. if

$$\bar{A} = A^*.$$

Frequently densely defined operators are of interest for which

$$A \subseteq -A^*$$

holds. Such operators are called anti-symmetric or **skew-symmetric**. If

$$A = -A^*$$

then A is called **skew-selfadjoint**. A skew-symmetric operator A is called **essentially skew-selfadjoint** if $\bar{A}$ is skew-selfadjoint, i.e. if

$$\bar{A} = -A^*.$$

REMARK 94. Note that A is formally selfadjoint if and only if A is formally adjoint to itself. Since for a (skew-)symmetric operator A we see that A^* must also be a densely defined, linear operator, we also realize that every (skew-)symmetric operator is closable.

LEMMA 95. *Let $A \subseteq H \oplus H$ be linear operator in complex Hilbert space H. Then A is Hermitean if and only if*

$$\bigwedge_{x,y\in D(A)} \langle Ax|y\rangle_H = \langle x|Ay\rangle_H . \tag{2.3.23}$$

Condition (2.3.23) implies[8]

$$w(A) := \left\{ \lambda \in \mathbb{C} | \bigvee_{x\in S_H(0,1)=\dot{B}_H(0,1)} \lambda = \langle x|Ax\rangle_H \right\} \subseteq \mathbb{R}. \tag{2.3.24}$$

The operator A is symmetric if and only if A is Hermitean and densely defined. The operator A is selfadjoint if and only if A is symmetric and $D(A^*) \subseteq D(A)$.

PROOF. For A to satisfy $A \subseteq A^*$ means

$$\begin{aligned} \textstyle\bigwedge_{x\in D(A)}(x, Ax) \in A^* &\Leftrightarrow \textstyle\bigwedge_{x\in D(A)}(Ax, -x) \in A^{\perp}, \\ &\Leftrightarrow \textstyle\bigwedge_{x\in D(A)} \bigwedge_{y\in D(A)} \langle y|Ax\rangle_H + \langle Ay| -x\rangle_H = 0, \\ &\Leftrightarrow \textstyle\bigwedge_{x,y\in D(A)} \langle Ax|y\rangle_H = \langle x|Ay\rangle_H . \end{aligned}$$

Condition (2.3.23) implies as a special case

$$\langle Ax|x\rangle_H = \langle x|Ax\rangle_H = \overline{\langle Ax|x\rangle_H} \text{ for all } x \in D(A).$$

This proves that (2.3.24) holds.

The characterization of A being symmetric is by definition. In the symmetric case we also read off that

$$D(A) \subseteq D(A^*). \tag{2.3.25}$$

For a selfadjoint operator A we have even $D(A) = D(A^*)$ which implies of course $D(A^*) \subseteq D(A)$. Conversely, from $D(A^*) \subseteq D(A)$ and (2.3.25) we get $D(A^*) = D(A)$. Since $A \subseteq A^*$, but $D(A^*) = D(A)$, we must have

$$A = A^*.$$

□

For the spectral allocation we find with $\mathbb{C}^{\pm} := [\mathbb{R}] \pm \mathrm{i}\, [\mathbb{R}_{>0}]$:

PROPOSITION 96. *Let $A \subseteq H \oplus H$ be a closed, symmetric operator in complex Hilbert space H. Then*

$$P\sigma(A) \cup C\sigma(A) \subseteq \mathbb{R}.$$

There are three possibilities:

(1) $\mathbb{C} \setminus \mathbb{R} \subseteq R\sigma(A)$,
(2) $\mathbb{C}^+ \subseteq R\sigma(A)$ and $\mathbb{C}^- \subseteq \varrho(A)$, or
$\mathbb{C}^- \subseteq R\sigma(A)$ and $\mathbb{C}^+ \subseteq \varrho(A)$,
(3) $\mathbb{C} \setminus \mathbb{R} \subseteq \varrho(A)$.

In these cases we have correspondingly

---

[8]The set $w(A)$ is called the numerical range of an operator $A \subseteq H \oplus H$.

(1) $\sigma(A) = \mathbb{C}$,
(2) $\sigma(A) = \overline{\mathbb{C}^+}$, or
$\sigma(A) = \overline{\mathbb{C}^-}$,
(3) $\sigma(A) \subseteq \mathbb{R}$.

Case 3. characterizes the selfadjointness of A.

Proof. We find with (2.3.24) that $\mathbb{C}\setminus\mathbb{R} \subseteq \varrho(A)\cup R\sigma(A)$. Since there are two open, connected components $\mathbb{C}^{\pm}$ in $\mathbb{C}\setminus\mathbb{R}$, the three cases are indeed the only possibilities. The allocation of the spectrum follows with the closedness of $\sigma(A)$ in $\mathbb{C}$. If A is selfadjoint, then the projection theorem yields

$$H = N(\overline{\lambda} - A) \oplus \overline{(\lambda - A)\,[H]}.$$

Since for $\lambda \in \mathbb{C}\setminus\mathbb{R}$ we have $N(\overline{\lambda} - A) = \{0\}$, it follows that $(\lambda - A)^{-1}$ is always densely defined for such λ. Therefore, we must be in case 3. Conversely, let case 3. hold, then we have

$$(\lambda - A) \subseteq (\lambda - A^*).$$

We get for $\lambda \in \mathbb{C}\setminus\mathbb{R}$

$$(\lambda - A)^{-1} \subseteq (\lambda - A^*)^{-1}.$$

Since, however, $(\lambda - A)^{-1}$ is already defined on all of H and $\left(\left(\overline{\lambda} - A\right)^{-1}\right)^* = (\lambda - A^*)^{-1}$ is also a mapping, this yields

$$(\lambda - A)^{-1} = (\lambda - A^*)^{-1}. \tag{2.3.26}$$

For every closed linear operator $B \subseteq H \oplus H$ with non-empty resolvent set, we have for $\lambda \in \varrho(B)$ that

$$D(B) = (\lambda - B)^{-1}\,[H]\,.$$

Consequently, (2.3.26) implies

$$D(A) = D(A^*).$$

□

The fact that the numerical range of a symmetric operator is real allows for a specific ordering of such operators.

Definition 97. Let A, $B \subseteq H \oplus H$ be operators in complex Hilbert space H. We define[9]

$$A \le B :\Leftrightarrow \bigwedge_{x\in D(A)\cap D(B)} \mathfrak{Re}\,\langle x|Ax\rangle_H \le \mathfrak{Re}\,\langle x|Bx\rangle_H\,.$$

If $A \ge 0$ then we call A **non-negative**[10]. If $A \ge 0$ and $Ax = 0 \Rightarrow x = 0$ for all $x \in H$ then A is called **positive (definite)**[11]. If $A \ge \varepsilon$ for some $\varepsilon \in \mathbb{R}_{>0}$, then A is called **strictly positive (definite)**[12].

For Hermitean operators we may omit taking the real part since the numerical range is already real[13].

[9]This is only an abbreviation since in general it is **not** a semi-order relation due to domain issues.

[10]Non-negativity is characterized by $\mathfrak{Re}\,w(A) \ge 0$.

[11]Note that for the purposes of these concepts we consider the Hilbert space H as a real Hilbert space. Frequently non-negativity is only defined for symmetric operators. This is largely unnecessary since for complex Hilbert spaces and continuous linear operators non-negativity – in the sense of non-negative numerical range – actually implies symmetry.

[12]This is characterized by $\mathfrak{Re}\,w(A) \ge \varepsilon > 0$.

[13]As noted already, one frequently finds positive definiteness only defined for Hermitean operators – a custom, which we do not follow here.

EXAMPLE 98.

(1) Let $A \subseteq H_0 \oplus H_1$ be a closed, densely defined linear operator, then
$$A^*A$$
is a non-negative selfadjoint operator.

(2) For A as in the previous example we have that
$$\begin{pmatrix} 0 & -A^* \\ A & 0 \end{pmatrix}$$
is skew-selfadjoint in $H_0 \oplus H_1$.
   (a) $A = \operatorname{grad}$ or $A = \mathring{\operatorname{grad}}$ (acoustics),
   (b) $A = \operatorname{Grad}$ or $A = \mathring{\operatorname{Grad}}$ (elasticity),
   (c) or $A = \operatorname{curl}$ $A = \mathring{\operatorname{curl}}$ (Maxwell's equations).

(3) Let $P : H \to H$ be an orthogonal projector. Letting $C := P[H]$ we have the canonical embedding of C in H given by
$$\begin{aligned} \iota_C : C &\to H, \\ x &\mapsto x. \end{aligned}$$
Then
$$P = \iota_C \iota_C^*.$$
Indeed, for $x \in C$ we have
$$\begin{aligned} \langle x|y\rangle_H = \langle \iota_C x|y\rangle_H &= \langle x|\iota_C^* y\rangle_C \\ &= \langle \iota_C x|\iota_C \iota_C^* y\rangle_H \\ &= \langle x|\iota_C \iota_C^* y\rangle_H \ . \end{aligned}$$
$Py =: z \in C$ is characterized by
$$\langle x|z\rangle_C = \langle \iota_C x|\iota_C z\rangle_H = \langle \iota_C x|y\rangle_H$$
for all $x \in C$. We read off $\iota_C^* y = z$ and so $\iota_C \iota_C^* y = \iota_C z = z = Py$, i.e.
$$P = \iota_C \iota_C^*.$$
In particular P is selfadjoint and
$$\begin{aligned} \iota_C^* : H &\to C, \\ x &\mapsto Px. \end{aligned}$$

2.3.4. Normal Operators. Unitary operators in $L(H, H)$ and selfadjoint operators in complex Hilbert space H are together part of a larger class of operators the so-called normal operators.

DEFINITION 99. Let $A \subseteq H \oplus H$ be a closed, densely defined, linear operator in the complex Hilbert space H. If
$$A^*A = A\,A^* \tag{2.3.27}$$
then A is called normal.

Note that (2.3.27) is an equality between selfadjoint operators. We first collect some basic properties of normal operators.

THEOREM 100. *Let $A \subseteq H \oplus H$ be a normal operator in complex Hilbert space H. Then A and A^* are commuting and A^*, λA and $(\lambda - A)$, $\lambda \in \mathbb{C}$, are also normal. A is maximal in the sense that there is no proper extension B of A with B normal. Moreover, we have*
$$D(A) = D(A^*),\ N(A) = N(A^*) \text{ and } \overline{A[H]} = \overline{A^*[H]}.$$

Proof. That A^* is normal follows directly from the symmetry of condition (2.3.27). From

$$(\lambda - A)^* = (\overline{\lambda} - A^*)$$

we get

$$\begin{aligned}(\lambda - A)(\lambda - A)^* &= (\lambda - A)\,(\overline{\lambda} - A^*),\\ &= |\lambda|^2 - \lambda\, A^* - \overline{\lambda}\, A + A\, A^*,\\ &= |\lambda|^2 - \lambda\, A^* - \overline{\lambda}\, A + A^* A\,,\\ &= (\overline{\lambda} - A^*)\,(\lambda - A),\\ &= (\lambda - A)^*\,(\lambda - A).\end{aligned}$$

This is the normality of $(\lambda - A)$, $\lambda \in \mathbb{C}$. The normality of $\lambda\, A$ follows also by a straight-forward calculation

$$\lambda\, A\,(\lambda\, A)^* = |\lambda|^2\, A\, A^* = |\lambda|^2\, A^* A = (\lambda\, A)^*(\lambda\, A).$$

For A we also see that

$$\langle Ax \,|\, Ax\rangle_H = \langle x \,|\, A^* Ax\rangle_H = \langle x \,|\, A\, A^* x\rangle_H = \langle A^* x \,|\, A^* x\rangle_H \tag{2.3.28}$$

for all $x \in D(A^* A) = D(A\, A^*)$. We have that $D(A^* A) = D(A\, A^*)$ is dense in $D(A)$ and $D(A^*)$ and we see with (2.3.28) (and the closedness of A and A^*) that $D(A) = D(A^*)$. Let now $A \subseteq B$ with B normal then $B^* \subseteq A^*$ and so $D(A) \subseteq D(B) = D(B^*) \subseteq D(A^*) = D(A)$. This proves that we must have $D(A) = D(B)$ and consequently

$$A = B.$$

The validity of (2.3.28) also implies that $N(A) = N(A^*)$ and consequently

$$N(A)^\perp = \overline{A\,[H]} = \overline{A^*\,[H]} = N(A^*)^\perp.$$

Finally, if A has empty resolvent set then A and A^* are commuting by definition. If $\lambda \in \varrho(A) \neq \emptyset$ then we have from (2.3.27) that

$$(\lambda - A)\, A^* x = A^*(\lambda - A)x$$

for all $x \in D(A^* A) = D(A\, A^*)$. Thus, we have with $y = (\lambda - A)\, x \in D(A^*) = D(A)$

$$(\lambda - A)^{-1}\, A^* y = A^*(\lambda - A)^{-1} y.$$

Since any $y \in D(A^*)$ can be represented in this form by $x := (\lambda - A)^{-1} y \in D(A^* A)$ (note $Ax = \lambda\,(\lambda - A)^{-1} y - y \in D(A^*) = D(A)$!), we have

$$(\lambda - A)^{-1}\, A^* \subseteq A^*(\lambda - A)^{-1},$$

i.e. A and A^* commute in the sense of the earlier definition. □

For the spectrum of a normal operator we find

Proposition 101. *Let $A \subseteq H \oplus H$ be a normal operator in complex Hilbert space H. Then*

$$P\sigma(A)^* = P\sigma(A^*)$$

and even

$$N(\lambda - A) = N(\overline{\lambda} - A^*).$$

Moreover, we have

$$R\sigma(A) = \emptyset,\ C\sigma(A)^* = C\sigma(A^*).$$

Proof. By the projection theorem we have

$$H = N(\lambda - A) \oplus \overline{(\overline{\lambda} - A^*)\,[H]} = N(\overline{\lambda} - A^*) \oplus \overline{(\lambda - A)\,[H]}.$$

Since $(\lambda - A)$ is normal by the previous theorem, we get $N(\lambda - A) = N(\overline{\lambda} - A^*)$, thus in particular $P\sigma(A)^* = P\sigma(A^*)$. Since now

$$H = N(\lambda - A) \oplus \overline{(\lambda - A)\,[H]},$$

we see that $R\sigma(A) = \emptyset$. Using that generally for normal operators $\sigma(A)^* = \sigma(A^*)$, the last equality for the continuous spectra follows. $\square$

Definition 102. Let $A \subseteq H \oplus H$ be a normal operator in complex Hilbert space H. Then

$$\begin{aligned} \mathfrak{Re}\ A &:= \tfrac{1}{2}\overline{(A + A^*)}, \\ \mathfrak{Im}\ A &:= \tfrac{1}{2\,\mathrm{i}}\overline{(A - A^*)}, \end{aligned}$$

are called **real part** and **imaginary part** of A.

Real and imaginary part of a normal operator are apparently closed symmetric linear operators. Indeed, they turn out to be selfadjoint.

Proposition 103. *Let $A \subseteq H \oplus H$ be a normal operator in complex Hilbert space H. Then $\mathfrak{Re}\ A$ and $\mathfrak{Im}\ A$ are commuting, selfadjoint operators and*

$$A = \mathfrak{Re}\ A + \mathrm{i}\ \mathfrak{Im}\ A. \tag{2.3.29}$$

In particular,

$$D(A) = D(\mathfrak{Re}\ A) \cap D(\mathfrak{Im}\ A).$$

Moreover, we have

$$|(\mathfrak{Re}\ A + \mathrm{i}\ \mathfrak{Im}\ A)x|_H^2 = |(\mathfrak{Re}\ A)x|_H^2 + |(\mathfrak{Im}\ A)x|_H^2 \tag{2.3.30}$$

for all $x \in D(A)$, and

$$|A| = \sqrt{(\mathfrak{Re}\ A)^2 + (\mathfrak{Im}\ A)^2} = |A^*|\,.$$

Example 104. As a unitary operator the Fourier transform $\mathcal{F}$ is also normal with real part $\mathcal{F}_{\cos} := Re\,\mathcal{F} = \frac{1}{2}(\mathcal{F} + \mathcal{F}^*)$ as the so-called Fourier cosine transform and with imaginary part $\mathcal{F}_{\sin} := -\,\mathfrak{Im}\ \mathcal{F} = -\frac{1}{2\mathrm{i}}(\mathcal{F} - \mathcal{F}^*)$ as the so-called Fourier sine transform. Being real and imaginary part of a normal operator, the Fourier cosine and the Fourier sine transform must be selfadjoint and their spectrum must therefore be real. According to our findings we already know that

$$Re\,\sigma(\mathcal{F}) = \{+1, -1\} \subseteq \sigma(\mathcal{F}_{\cos}),\ \mathfrak{Im}\ \sigma(\mathcal{F}) = \{+1, -1\} \subseteq \sigma(\mathcal{F}_{\sin}).$$

The relation between $\mathcal{F}_{\cos}$, $\mathcal{F}_{\sin}$ to $\mathcal{F}$ yields further insight. We have (with σ_{-1} as the reflection at the origin)

$$\begin{aligned} \mathcal{F}_{\cos} &= \tfrac{1}{2}(\mathcal{F} + \mathcal{F}^*), \\ &= \tfrac{1}{2}(1 + \sigma_{-1})\,\mathcal{F}, \\ &= \mathcal{F}\,\tfrac{1}{2}(1 + \sigma_{-1}), \\ &= \tfrac{1}{2}(1 + \sigma_{-1})\,\mathcal{F}^*, \\ &= \mathcal{F}^*\,\tfrac{1}{2}(1 + \sigma_{-1})\,; \\ \mathcal{F}_{\sin} &= -\tfrac{1}{2\mathrm{i}}(\mathcal{F} - \mathcal{F}^*), \\ &= \tfrac{1}{2}(1 - \sigma_{-1})\,\mathrm{i}\mathcal{F}, \\ &= \mathrm{i}\mathcal{F}\,\tfrac{1}{2}(1 - \sigma_{-1}), \end{aligned} \tag{2.3.31}$$

It is not hard to see that $\frac{1}{2}(1 + \sigma_{-1})$, $\frac{1}{2}(1 - \sigma_{-1}) = 1 - \frac{1}{2}(1 + \sigma_{-1})$, are the orthogonal projections onto the (almost everywhere) even and odd functions in $L^2(\mathbb{R})$. Therefore we also have $0 \in \sigma(\mathcal{F}_{\cos})$ and $0 \in \sigma(\mathcal{F}_{\sin})$. Moreover, we calculate with (2.3.31)

$$\mathcal{F}_{\sin}\mathcal{F}_{\cos} = \mathcal{F}_{\cos}\mathcal{F}_{\sin} = 0,\ \mathcal{F}_{\sin}\,\mathcal{F}_{\sin} = \frac{1}{2}(1 - \sigma_{-1}),\ \mathcal{F}_{\cos}\mathcal{F}_{\cos} = \frac{1}{2}(1 + \sigma_{-1}). \tag{2.3.32}$$

Thus, we see that $\mathcal{F}_{\cos}$, $\mathcal{F}_{\sin}$ are unitary on the subspaces $\frac{1}{2}(1 \pm \sigma_{-1})\, L^2(\mathbb{R})$, respectively. Together with the selfadjointness of $\mathcal{F}_{\cos}$, $\mathcal{F}_{\sin}$ we have that

$$\sigma(\mathcal{F}_{\cos}) = \sigma(\mathcal{F}_{\sin}) = P\sigma(\mathcal{F}_{\cos}) = P\sigma(\mathcal{F}_{\sin}) = \{0, +1, -1\}.$$

We can identify $\frac{1}{2}(1 \pm \sigma_{-1})\, L^2(\mathbb{R})$ with $L^2(\mathbb{R}_{>0})$ via the unitary correspondence $E_\pm : L^2(\mathbb{R}_{>0}) \to \frac{1}{2}(1 \pm \sigma_{-1})\, L^2(\mathbb{R})$ with

$$(E_\pm f)(x) \coloneqq \frac{1}{\sqrt{2}} \begin{cases} f(x) & \text{for } x \in \mathbb{R}_{>0} \\ \pm f(-x) & \text{otherwise} \end{cases},\ x \in \mathbb{R}.$$

$\sqrt{2}E_+ f$ will be refered to as the even extension of f and $\sqrt{2}E_- f$ as the odd extension of f. Indeed, we find

$$\begin{aligned} |E_\pm\varphi|^2_{L^2(\mathbb{R})} &= \int_{\mathbb{R}} |(E_\pm\varphi)(x)|^2\, dx, \\ &= \tfrac{1}{2}\int_{\mathbb{R}_{>0}} |\varphi(x)|^2\, dx + \tfrac{1}{2}\int_{\mathbb{R}_{<0}} |\varphi(-x)|^2\, dx, \\ &= \int_{\mathbb{R}_{>0}} |\varphi(x)|^2\, dx = |\varphi|^2_{L^2(\mathbb{R}_{>0})}, \end{aligned}$$

for all $\varphi \in L^2(\mathbb{R}_{>0})$ and so we obtain the unitary mappings

$$E_+^*\, \mathcal{F}_{\cos}\, E_+ : L^2(\mathbb{R}_{>0}) \to L^2(\mathbb{R}_{>0})\,,\ E_-^*\, \mathcal{F}_{\sin} E_- : L^2(\mathbb{R}_{>0}) \to L^2(\mathbb{R}_{>0}).$$

Here is $E_\pm^* = E_\pm^{-1}$ the inverse of the unitary mapping $E_\pm$ characterized by

$$E_\pm^* : \frac{1}{2}(1 \pm \sigma_{-1})\left[L^2(\mathbb{R})\right] \to L^2(\mathbb{R}_{>0}),\ \phi \mapsto \sqrt{2}\,\phi\big|_{\mathbb{R}_{>0}}.$$

For 'nice' functions, e.g. for elements of $\mathring{C}_\infty(\mathbb{R}_{>0})$ we get the following integral representations

$$(E_+^*\, \mathcal{F}_{\cos}\, E_+\phi)(x) = \sqrt{\frac{2}{\pi}} \int_{\mathbb{R}_{>0}} \cos(xy)\, \phi(y)\, dy,\ (E_-^*\, \mathcal{F}_{\sin} E_-\phi)(x) = \sqrt{\frac{2}{\pi}} \int_{\mathbb{R}_{>0}} \sin(xy)\, \phi(y)\, dy$$

for all $\phi \in \mathring{C}_\infty(\mathbb{R}_{>0})$. The transforms

$$E_+^*\, \mathcal{F}_{\cos}\, E_+ : L^2(\mathbb{R}_{>0}) \to L^2(\mathbb{R}_{>0})\,,\ E_-^*\, \mathcal{F}_{\sin} E_- : L^2(\mathbb{R}_{>0}) \to L^2(\mathbb{R}_{>0})$$

are also referred to as Fourier cosine transform and Fourier sine transform, respectively. We shall use $\widetilde{\mathcal{F}}_{\cos}$ and $\widetilde{\mathcal{F}}_{\sin}$ to denote these variants on generalized square integrable functions on $\mathbb{R}_{>0}$.

With $\mathcal{F}$ we also have that $exp(\mathrm{i}\,\pi/4)\,\mathcal{F}$ is unitary. Considering the real and imaginary part of $exp(\mathrm{i}\,\pi/4)\,\mathcal{F}$ leads to an interesting situation. We find

$$\mathfrak{Re}\ (exp(\mathrm{i}\,\pi/4)\,\mathcal{F}) = \frac{1}{\sqrt{2}}(\mathcal{F}_{\cos} + \mathcal{F}_{\sin}),\ \mathfrak{Im}\ (exp(\mathrm{i}\,\pi/4)\,\mathcal{F}) = \frac{1}{\sqrt{2}}(\mathcal{F}_{\cos} - \mathcal{F}_{\sin}),$$

and therefore with $H_\pm \coloneqq \mathcal{F}_{\cos} \pm \mathcal{F}_{\sin}$

$$H_+ = \mathfrak{Re}\left(\sqrt{2}\, exp(\mathrm{i}\,\pi/4)\,\mathcal{F}\right)$$
$$H_- = \mathfrak{Im}\left(\sqrt{2}\, exp(\mathrm{i}\,\pi/4)\,\mathcal{F}\right)$$

and

$$\begin{aligned} |H_\pm\varphi|^2_{L^2(\mathbb{R})} &= |(\mathcal{F}_{\cos} \pm \mathcal{F}_{\sin})\varphi|^2_{L^2(\mathbb{R})}\,, \\ &= |\mathcal{F}_{\cos}\varphi|^2_{L^2(\mathbb{R})} \pm 2\, Re\, \langle \mathcal{F}_{\cos}\varphi\, |\mathcal{F}_{\sin}\varphi\rangle_{L^2(\mathbb{R})} + |\mathcal{F}_{\sin}\varphi|^2_{L^2(\mathbb{R})}\,, \\ &= |\mathcal{F}_{\cos}\varphi|^2_{L^2(\mathbb{R})} + |\mathcal{F}_{\sin}\varphi|^2_{L^2(\mathbb{R})}\,, \\ &= |\mathcal{F}\varphi|^2_{L^2(\mathbb{R})} = |\varphi|^2_{L^2(\mathbb{R})} \end{aligned}$$

for $\varphi \in L^2(\mathbb{R})$. The transforms $H_\pm$ are known as Hartley transforms. They are both unitary, selfadjoint and real. In particular,

$$\sigma(H_\pm) = P\sigma(H_\pm) = \{+1, -1\}$$

and

$$\begin{aligned} H_+H_- &= H_-H_+ \\ &= \mathcal{F}_{\cos}\mathcal{F}_{\cos} - \mathcal{F}_{\sin}\mathcal{F}_{\sin} \\ &= \frac{1}{2}(1+\sigma_{-1}) - \frac{1}{2}(1-\sigma_{-1}) \\ &= \sigma_{-1}, \\ H_+H_+ &= H_-H_- \\ &= \mathcal{F}_{\cos}\mathcal{F}_{\cos} + \mathcal{F}_{\sin}\mathcal{F}_{\sin} \\ &= \frac{1}{2}(1+\sigma_{-1}) + \frac{1}{2}(1-\sigma_{-1}) \\ &= 1. \end{aligned}$$

On $\mathring{C}_\infty(\mathbb{R})$ we have

$$(H_\pm\varphi)(\omega) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} (\cos(\omega t) \pm \sin(\omega t))\ \varphi(t)\ dt$$

and

$$(H_-\varphi)(\omega) = (H_+\varphi)(-\omega)$$

for $\omega \in \mathbb{R}$ or

$$H_- = \sigma_{-1}H_+.$$

2.3.5. Positive-Definite Operators.

2.3.5.1. *Hadamard's requirements – revisited.*
We recall Hadamard's celebrated requirements for well-posedness. The task to establish well-posedness based on a relation P under consideration consists in finding metric spaces X, Y, such that $P \subseteq X \times Y$ and that in view of finding x for given y satisfying

$$(x, y) \in \overline{P}$$

is a "reasonable" problem in the sense that

(1) $\overline{P}$ left-unique (**uniqueness**): P^{-1} is a closable mapping,
(2) $\overline{P}$ is right-total (**existence**): P^{-1} is densely defined,
(3) $\overline{P}^{-1}$ is continuous (**continuous dependence on the data**): P^{-1} is locally uniformly continuous in Y, in the sense that for every point z in Y there is a ball $B(z,r)$, $r \in]0,\infty[$, such that $P^{-1}|_{B(z,r)} = P^{-1} \cap (B(z,r) \times X)$ is uniformly continuous.

In short: $\overline{P}^{-1} : Y \to X$ is continuous[14].

2.3.5.2. *Strict Positive-Definiteness as a Key to Well-Posedness.*
It is somewhat unfortunate, but well in the spirit of our underlying theme of confusion by jargon, that "$A : D(A) \subseteq H \to H$ positive" is used synonymously to "$A : D(A) \subseteq H \to H$ non-negative" (or to "$A : D(A) \subseteq H \to H$ non-negative-definite"). For sake of clarity we shall reserve the statement "$A : D(A) \subseteq H \to H$ non-negative" rather than any of the alternatives for the case

$$\mathfrak{Re}\langle x|Ax\rangle \geq 0 \text{ for } x \in D(A).$$

[14] If $\overline{P}$ is itself a mapping then Hadamard's requirements can be rephrased as saying: $\overline{P}$ is injective and onto with a continuous inverse.

$A : D(A) \subseteq H \to H$ positive definite[15]:

$$\mathfrak{Re}\,\langle x|Ax\rangle > 0 \text{ for } x \in D(A) \setminus \{0\}\,.$$

$A : D(A) \subseteq H \to H$ strictly positive definite[16] (also called strictly accretive):

$$\mathfrak{Re}\,\langle x|Ax\rangle \geq c_0\,\langle x|x\rangle$$

for some $c_0 \in \,]0,\infty[$ and all $x \in \mathrm{D}(A)$ or

$$\inf\left\{\mathfrak{Re}\,\langle x|Ax\rangle_H \,|x \in \mathrm{D}(A) \cap B_H(0,1)\right\} =: c_0 > 0.$$

It is remarkable that the latter constraint is of relevance in just about all the basic well-posedness results of partial differential equations.

One idea fits all!

Indeed, let $A : D(A) \subseteq H \to H$ be a closed and densely defined linear operator satisfying

$$\begin{aligned} \mathfrak{Re}\,\langle x|Ax\rangle \geq c_0\,\langle x|x\rangle\,,\ \mathfrak{Re}\,\langle y|A^*y\rangle \geq c_0\,\langle y|y\rangle \text{ for some } c_0 \in \,]0,\infty[\\ \text{and all } x \in \mathrm{D}(A)\,,\ y \in D(A^*)\,. \end{aligned} \tag{2.3.33}$$

Then, the problem of finding for any given $f \in H$ a solution $x \in D(A)$ such that

$$Ax = f$$

is well-posed, indeed this is a rather particular case of Hadamard's requirements.

Theorem 105. *Let $A : D(A) \subseteq H \to H$ be a closed and densely defined linear operator satisfying (2.3.33). Then for every given $f \in H$ there is a unique solution $x \in D(A)$ such that*

$$Ax = f.$$

Moreover, the solution x depends continuously on the data f via the estimates:

$$|x| \leq \frac{1}{c_0}\,|f|_H\,. \tag{2.3.34}$$

Proof. From (2.3.33) we deduce that

$$|x|_H^2 \leq \frac{1}{c_0}\,\mathfrak{Re}\,\langle x|Ax\rangle \leq |x|_H\,|Ax|_H$$

and so

$$|x|_H \leq \frac{1}{c_0}\,|Ax|_H$$

for all $x \in D(A)$. This shows that A^{-1} is a linear mapping and that

$$\left|A^{-1}f\right|_H \leq \frac{1}{c_0}\,|f|_H$$

for all $f \in A[H]$. This proves already the estimate (2.3.34). Moreover, by this very estimate any Cauchy sequence $(f_k)_k$ in $A[H]$ results in a Cauchy sequence of solutions $(x_k)_k$ in H. Due to the closedness of A we have that $x_\infty := \lim_{k\to\infty} x_k \in D(A)$ and $Ax_\infty = \lim_{k\to\infty} Ax_k$. Consequently

---

[15]The added "definite" is needed due to the common confusion about "non-negative" and "positive". From a more logical point of view "positive" would be sufficient to properly denote this case.

We mention in passing that a similar confusion occurs with "contractive" and "non-expansive". A contractive linear mapping acting in a – say – Hilbert space H would (compare Banach's contraction mapping theorem) be a linear mapping $U : H \to H$ with operator norm

$$\|U\| := \sup\left\{|Ux|_H \mid x \in B_H(0,1)\right\} < 1.$$

In contrast, one calls a linear mapping $U : H \to H$ non-expansive if $\|U\| \leq 1$. The expectation that a so-called one-parameter contraction semi-group $V = (V(t))_{t\in[0,\infty[}$, which is by the way strictly speaking a monoid homomorphism (a monoid is a semi-group with identity element), between the monoid $([0,\infty[\,,+)$ and the monoid $(L(H,H)\,,\circ)$, would be a family of contractive mappings is foiled by common practice, which says that it is just a family of non-expansive mappings. Indeed, since by definition of a one-parameter semi-group $V(0)$ is the identity and so $\|V(0)\| = 1$, there cannot be a one-parameter semi-group of contractions.

[16]Note that we do not assume – as is frequently done – that A be selfadjoint here.

$A[H]$ is closed. It remains to see that $A[H]$ is dense in H and so equal to H. This, however, follows from the orthogonal decomposition

$$H = \overline{A[H]} \oplus [\{0\}]\, A^*$$

since A^*shares in particular the property of being injective with A, i.e. $[\{0\}]\, A^* = \{0\}$. □

Remark 106. Obviously, there are many possible generalizations, which would get us closer to the general Hadamard's requirements. We note in particular that the proof of Theorem 105 would work likewise if (2.3.33) is replaced by requiring

$$\begin{aligned} &|\langle x|Ax\rangle_H| \geq c_0 \langle x|x\rangle_H\,,\ |\langle y|A^*y\rangle_H| \geq c_0 \langle y|y\rangle_H \text{ for some } c_0 \in\,]0,\infty[\\ &\qquad\qquad\text{and all } x \in \mathrm{D}(A)\,,\ y \in D(A^*)\,. \end{aligned} \tag{2.3.35}$$

This option is occasionally employed in the context of sesqui-linear form methods. For sake of simplicity we shall, however, not pursue this line of thought.

A typical application of Theorem 105 is given by the solution theory of elliptic partial differential equations.

We first record a particular case of Theorem 105

Corollary 107. *Let $A : X \to X$ be a strictly positive definite, continuous, linear operator in the Hilbert space X. Then for every $f \in X$ there is a unique $u \in X$ such that*

$$Au = f.$$

Indeed, solutions depend continuously on the data in the sense that we have a continuous linear operator $A^{-1} : X \to X$ with

$$u = A^{-1}f.$$

Note that since continuous linear operators and continuous sesqui-linear forms are equivalent, the last corollary is the so-called Lax-Milgram theorem. Indeed,

$$(u, v) \mapsto \langle u|Av\rangle_X$$

is a continuous sesqui-linear form on X and conversely if $\beta\langle\,\cdot\,|\,\cdot\,\rangle$ is a continuous sesqui-linear form on X then

$$\overline{\beta\langle\,\cdot\,|v\rangle} \in X^*$$

and utilizing the Riesz map $R_X : X^* \to X$ we get with the Riesz representation theorem

$$\left\langle R_X \beta\overline{\langle\,\cdot\,|v\rangle}|u\right\rangle_X = \overline{\beta\langle u|v\rangle}$$

and so

$$\beta\langle u|v\rangle = \left\langle u|R_X\beta\overline{\langle\,\cdot\,|v\rangle}\right\rangle_X.$$

Note that

$$Av \coloneqq R_X\beta\overline{\langle\,\cdot\,|v\rangle},\ v \in X, \tag{2.3.36}$$

defines indeed a bounded linear operator on X. Strict positive definiteness for the corresponding operator A results in the so-called coercivity[17] of the sesqui-linear form β:

$$\mathfrak{Re}\,\beta\langle u|u\rangle \geq c\,\langle u|u\rangle_X \tag{2.3.37}$$

for some $c \in\,]0,\infty[$ and all $u \in X$. Thus, as an equivalent formulation of the previous corollary we get the following.

[17]Following Remark 106 the strict positive definiteness can be weakened slightly to requiring merely

$$\bigwedge_{u\in X} |\beta\langle u|u\rangle| \geq c\,\langle u|u\rangle_X$$

yielding the same well-posedness out-come. This option is frequently utilized in applications.

Corollary 108. *(Lax-Milgram theorem) Let $\beta\langle\,\cdot\,|\,\cdot\,\rangle$ be a continuous, coercive sesqui-linear form on a Hilbert space X. Then for every $f\in X^*$ there is a unique $u\in X$ such that*

$$\beta\langle u|v\rangle=f(v)$$

for all $v\in X$.

In other words, β generates via (2.3.36) a mapping $A:X\to X$ such that for every given $g\in X$ there is a unique $u\in X$ such that

$$\beta\langle u|v\rangle=R_X^*g(v)=\langle Au|v\rangle_X=\langle g|v\rangle_X$$

for all $v\in X$ and so we have a solution theory for the equation

$$Au=g.$$

Keeping in mind the latter approach has been utilized extensively just for elliptic type problems it may be interesting to note that its generalization in the form of Theorem 105 is – as we shall see – perfectly sufficient to to solve elliptic, parabolic and hyperbolic systems in a single approach.

2.3.5.3. *A note on the Sesqui-Linear Forms Method.* Consider a continuous linear bijection $C:H_1\to H_0$ and a continuous linear operator $A:H_0\to H_0$ with

$$\mathfrak{Re}\,\langle x|Ax\rangle_0\geq c_0\,\langle x|x\rangle_0$$

for some $c_0\in\,]0,\infty[$ and all $x\in H_0$. With

$$C^\diamond:H_0=H_0'\to H_{-1}$$

where

$$\bigwedge_{x\in D(C)=H_1}\langle y|Cx\rangle_0=:(C^\diamond y)(x)$$

or

$$C^\diamond y:=\langle y|C\,\cdot\,\rangle_0\,,$$

we consider

$$C^\diamond AC:H_1\to H_{-1}.$$

Consider now the uniquely solvable equation

$$C^\diamond ACw=f\in H_1'=:H_{-1}.$$

Then this is equivalent to

$$\begin{aligned}\beta\langle v|w\rangle&=\langle ACv|Cw\rangle_0\\&=(C^\diamond ACv)(w)\,.\end{aligned}$$

As an application of this observation:

Consider the continuous and dense embedding

$$\begin{aligned}\iota:H_1&\to H_0\\x&\mapsto x\end{aligned}$$

we obtain

$$\begin{aligned}\iota^{-1}:H_1\subseteq H_0&\to H_1\\x&\mapsto x\end{aligned}$$

and

$$\left(\iota^{-1}\right)^*:D\left(\left(\iota^{-1}\right)^*\right)\subseteq H_1\to H_0$$

given by

$$y\mapsto\in\left(\left\{f|\left\langle\iota^{-1}x|y\right\rangle_{H_1}=\langle x|f\rangle_{H_0}\text{ for all }x\in H_1\right\}\right).$$

We let $\widetilde{C} \coloneqq \left|\iota^{-1}\right| : H_1 \subseteq H_0 \to H_0$ and choose for C the mapping

$$\begin{aligned} C : H_1 &\to H_0, \\ x &\mapsto \widetilde{C}x. \end{aligned}$$

Then the above consideration applies. Moreover, we have

$$\widetilde{C}A\widetilde{C} = C^{\diamond}AC \cap (H_0 \oplus H_0).$$

Since $\widetilde{C}A\widetilde{C}$ is closed, invertible (inherited from $C^{\diamond}AC$) and onto we have by the closed graph theorem

$$0 \in \varrho\left(\widetilde{C}A\widetilde{C}\right).$$

Moreover,

$$\left(\widetilde{C}A\widetilde{C}\right)^* = \widetilde{C}A^*\widetilde{C}.$$

Indeed,

$$\begin{aligned} \left\langle \widetilde{C}A\widetilde{C}u|x \right\rangle_0 &= \langle u|f\rangle_0 \\ &= \left\langle u|\left(C^{\diamond}A^*C\right)\left(C^{\diamond}A^*C\right)^{-1}f \right\rangle_0 \\ &= \left\langle u|\left(\widetilde{C}A^*\widetilde{C}\right)C^{-1}\left(A^*\right)^{-1}\left(C^{\diamond}\right)^{-1}f \right\rangle_0 \\ &= \left\langle \left(\widetilde{C}A\widetilde{C}\right)u|C^{-1}\left(A^*\right)^{-1}\left(C^{\diamond}\right)^{-1}f \right\rangle_0 \end{aligned}$$

for all $u \in D\left(\widetilde{C}A\widetilde{C}\right)$, which yields

$$x = C^{-1}\left(A^*\right)^{-1}\left(C^{\diamond}\right)^{-1}f$$

and

$$f = \left(\widetilde{C}A^*\widetilde{C}\right)x.$$

In practice, $\widetilde{C}$ comes up via a polar decomposition (see Appendix A)

$$Z = U\widetilde{C}$$

of an operator $Z : D(Z) \subseteq H_0 \to X$ ($U : \overline{Z^*[X]} \to \overline{Z[H_0]}$ unitary) and replacing $A : X \to X$ by $U^*AU : \overline{Z^*[X]} \to \overline{Z^*[X]}$ we have

$$\widetilde{C}U^*AU\widetilde{C} = Z^*AZ$$

resulting in a solution theory for the equation

$$Z^*AZu = f.$$

Example 109. The paradigmatic application here is $\widetilde{C} = \sqrt{-\operatorname{div}\mathring{\operatorname{grad}} + 1}$, $Z = \begin{pmatrix} \mathring{\operatorname{grad}} \\ 1 \end{pmatrix}$: $D\left(\mathring{\operatorname{grad}}\right) \subseteq L^2(\Omega) \eqqcolon H_0 \to \left(L^2(\Omega)\right)^n \oplus L^2(\Omega)$ and

$$\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} : \left(L^2(\Omega)\right)^n \oplus L^2(\Omega) \to \left(L^2(\Omega)\right)^n \oplus L^2(\Omega) \eqqcolon X$$

strictly positive definite[18], which leads to the solution theory for

$$\begin{pmatrix} -\operatorname{div} & 1 \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \operatorname{grad} \\ 1 \end{pmatrix} u = -\operatorname{div} A \operatorname{grad} u + u = f \in H_0$$

[18] Actually we merely need to require that $A : \mathring{\operatorname{grad}} \to \mathring{\operatorname{grad}}$ is a bi-Lipschitz-continuous bijection (recalling that $\mathring{\operatorname{grad}}$ is a closed subspace of $\left(L^2(\Omega)\right)^n \oplus L^2(\Omega)$).

with boundary condition

$$u \in D\left(\mathring{\operatorname{grad}}\right) =: H_1.$$

2.3.6. Congruent, Similar and Equivalent. For comparing relations and operators the following three concepts are useful.

DEFINITION 110. Consider a linear relation $A \subseteq H_0 \oplus H_1$ and a linear relation $B \subseteq X_0 \oplus X_1$, H_k, X_k, $k \in \{0,1\}$, Hilbert spaces. Then A, B are called equivalent if

$$A = \begin{pmatrix} U_0 & 0 \\ 0 & U_1 \end{pmatrix} B$$

for some continuous bijections $U_k : X_k \to H_k$, $k \in \{0,1\}$.

If $H_0 = H_1$, $X_0 = X_1$, we can refine this comparison. If in this case $U_0 = U_1$ the relations A, B are called similar, if however $U_0^{-1} = U_1^*$ then they are called congruent. If A, B are similar and congruent, we call them unitarily similar or unitarily congruent[19].

We mostly use this terminology for mappings A, B. Clearly, in this case A, B are equivalent[20] if

$$A = U_1 B U_0^{-1}$$

for some continuous linear bijections $U_k : X_k \to H_k$, $k \in \{0,1\}$.

Two mappings $A : D(A) \subseteq H_0 \to H_0$, $B : D(B) \subseteq X_0 \to X_0$ are similar[21] if

$$A = U_0 B U_0^{-1}$$

for some continuous linear bijection $U_0 : X_0 \to H_0$.

Two mappings $A : D(A) \subseteq H_0 \to H_0$, $B : D(B) \subseteq X_0 \to X_0$ are congruent[22] if

$$A = U_1 B U_1^*$$

for some continuous linear bijection $U_1 : X_0 \to H_0$.

Two mappings $A : D(A) \subseteq H_0 \to H_0$, $B : D(B) \subseteq X_0 \to X_0$ are unitarily similar or unitarily congruent if

$$A = U_1 B U_1^*$$

for some unitary mapping $U_1 : X_0 \to H_0$.

[19] What is called "congruent" in geometry usually means "unitarily congruent".

[20] The importance of this concept is that equivalence of relations preserves for example well-posedness.

[21] Similarity preserves spectra.

[22] Congruence preserves symmetry and skew-symmetry.

Literature on Functional Analysis and Some Applications

[Akhiezer-Glazman 1993] N.I. Akhiezer and I.M. Glazman. *Theory of linear operators in Hilbert space. Transl. from the Russian and with a preface by Merlynd Nestell (Two volumes bound as one). Repr. of the 1961 and 1963 transl.* New York, NY: Dover Publications. xiv, 147, iv, 1993.

[Kato 1995] T. Kato. *Perturbation theory for linear operators.* Grundlehren der mathematischen Wissenschaften. Springer, 1995.

[Pi 1989] R. Picard. *Hilbert space approach to some classical transforms.* John Wiley, New York, 1989.

[Pi-McGhee 2011] R. Picard and D. McGhee. *Partial differential equations. A unified Hilbert space approach.* de Gruyter Expositions in Mathematics 55. Berlin: de Gruyter. xviii, 2011.

[1] M. Reed and B. Simon. *Methods of modern mathematical physics I-IV.* Academic Press, 1972-79.

[2] F. Riesz, B. Szokefalvi-Nagy, and L. F. Boron. *Functional analysis / Frigyes Riesz and Bela Sz-Nagy ; translated from the 2nd French edition by Leo F. Boron.* Blackie, London :, 1956.

[3] K. Schmüdgen. *Unbounded self-adjoint operators on Hilbert space.* Dordrecht: Springer, 2012.

[Tanabe 1979] H. Tanabe. *Equations of evolution. Translated from Japanese by N. Mugibayashi and H. Haneda.* Monographs and Studies in Mathematics. 6. London - San Francisco - Melbourne: Pitman. XII, 1979.

[4] A. E. Taylor and D. C. Lay. *Introduction to functional analysis. 2nd ed.* New York - Chichester - Brisbane: John Wiley & Sons. XI, 467 p. $ 23.95 , 1980.

[5] J. Weidmann. *Linear Operators in Hilbert Spaces.* Springer Verlag, New York, 1980.

[Wloka 1982] J. Wloka. *Partielle Differentialgleichungen.* B.G. Teubner Stuttgart, 1982.

[Yosida 1974] K. Yosida. *Functional analysis.* Springer-Verlag; Berlin et al., 1974.

Part 2

On the Theory of Evolutionary Equations

CHAPTER 1

On the General Structure of Evolutionary Problems

1.1. Introduction

It may come as a surprise that the simple ideas of the last section may also be the key to solving time-varying problems.

We start by observing that on closer inspection of initial boundary value problems of mathematical physics, in particular those describing wave propagation phenomena one is inclined to describe their general form as

$$\begin{aligned}\partial_0 V + AU &= f \text{ on } \mathbb{R}_{>0},\\ V\left(0+\right) &= \Phi,\end{aligned}$$

where A is commonly skew-selfadjoint in a suitable Hilbert space setting. We shall indeed prefer to consider this problem on the whole real time-line and to by-pass the full construction of associated Sobolev lattices, see [**Pi-McGhee 2011**], we shall assume – without loss of generality – that $\Phi = 0$. This turns our problem into

$$\partial_0 V + AU = f \text{ on } \mathbb{R}. \tag{1.1.1}$$

Following this lead, the abstract evolutionary problem

$$\partial_0 V + AU = f \text{ on } \mathbb{R}.$$

is now completed by an additional rule frequently referred to as a "material law", which for simplicity we assume to be time-translation invariant and more precisely of the form

$$V = M\left(\partial_0^{-1}\right)U, \tag{1.1.2}$$

where $z \mapsto M\left(z\right)$ is bounded-operator-valued and analytic in an open ball $B_{\mathbb{C}}\left(r, r\right)$ with some positive radius r centered at r.

1.2. The Time Derivative

It is well-known that $\frac{1}{\mathrm{i}}\,\partial_0$ can be established as a selfadjoint operator in the space $L^2\left(\mathbb{R}\right)$ of equivalence classes of square-integrable complex-valued functions on $\mathbb{R}$. The space $\mathring{C}_\infty\left(\mathbb{R}\right)$ of smooth complex-valued functions with compact support is densely embedded in the domain. Indeed, this case is occasionally used as a simple example for an explicit spectral representation, which here is provided by the Fourier transform $\mathcal{F}$ given as the unitary extension of

$$\begin{aligned}\mathring{C}_\infty\left(\mathbb{R}\right) \subseteq L^2\left(\mathbb{R}\right) &\to L^2\left(\mathbb{R}\right)\\ \varphi &\mapsto \widehat{\varphi}\end{aligned}$$

with

$$\widehat{\varphi}\left(x\right) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \exp\left(-\mathrm{i}\,x\,t\right)\,\varphi\left(t\right)\,dt,\; x \in \mathbb{R}.$$

As a spectral representation the Fourier transform makes $\frac{1}{\mathrm{i}}\,\partial_0$ unitarily congruent to the multiplication by the argument operator m given by $\left(m\varphi\right)\left(x\right) = x\varphi\left(x\right)$ for $x \in \mathbb{R}$ and $\varphi \in \mathring{C}_\infty\left(\mathbb{R}\right)$:

$$\boxed{\frac{1}{\mathrm{i}}\,\partial_0 = \mathcal{F}^* m\,\mathcal{F}.}$$

Recall our earlier introduction of an exponential weight function $t \mapsto \exp\left(-\nu\, t\right)$, $\nu \in \mathbb{R}$, and the weighted L^2-space $H_{\nu,0}$ generated by completion of $\mathring{C}_\infty\left(\mathbb{R}\right)$ with respect to the inner product $\langle\,\cdot\,|\,\cdot\,\rangle_{\nu,0}$

$$(\varphi,\psi) \mapsto \int_{\mathbb{R}} \varphi\left(t\right)^* \psi\left(t\right)\,\exp\left(-2\nu t\right) dt.$$

The associated norm was denoted by $|\cdot|_{\nu,0}$. The multiplication operator

$$\begin{gathered}\mathring{C}_\infty\left(\mathbb{R}\right) \subseteq H_{\nu,0} \to \mathring{C}_\infty\left(\mathbb{R}\right) \subseteq H_{0,0} = L^2\left(\mathbb{R}\right)\\ \varphi \mapsto \exp\left(-\nu m\right)\varphi\end{gathered}$$

with

$$\left(\exp\left(-\nu m\right)\varphi\right)\left(x\right) = \exp\left(-\nu x\right)\varphi\left(x\right),\ x \in \mathbb{R},$$

clearly has a unitary extension, which we denote by $\exp\left(-\nu m\right)$, where the m serves as a reminder for '**m**ultiplication'. Its inverse will be denoted by $\exp\left(\nu m\right)$.

Thus, the operator

$$\frac{1}{\mathrm{i}}\,\partial_\nu := \exp\left(\nu m\right)\frac{1}{\mathrm{i}}\,\partial_0 \exp\left(-\nu m\right)$$

defines a unitarily congruent operator $\frac{1}{\mathrm{i}}\,\partial_\nu$, which is now selfadjoint in $H_{\nu,0}$.

We shall use the notation $\partial_{0,\nu}$ (or simply ∂_0 if the parameter ν is clear from the context) for the normal operator $\partial_\nu + \nu$, which is justified since indeed

$$\left(\partial_\nu + \nu\right)\varphi = \partial_0\varphi$$

for $\varphi \in \mathring{C}_\infty\left(\mathbb{R}\right)$.

Obviously we have that the spectrum of ∂_ν is purely imaginary. In fact, the spectrum $\sigma\left(\partial_\nu\right)$ is also purely continuous spectrum:

$$\boxed{\sigma\left(\partial_\nu\right) = \sigma_c\left(\partial_\nu\right) = \mathrm{i}\left[\mathbb{R}\right].}$$

Thus, in particular for $\nu \in \mathbb{R}\setminus\{0\}$ we have the bounded invertibility of $\partial_0 = \partial_\nu + \nu$.

With $\mathcal{L}_\nu := \mathcal{F}\,\exp\left(-\nu m\right)$

$$\boxed{\frac{1}{\mathrm{i}}\,\partial_\nu = \mathcal{L}_\nu^* m\,\mathcal{L}_\nu.}$$

1.3. Evolutionary Dynamics and Material Laws

We shall now consider the initially stated evolutionary problem in precise terms. For this we need to extend the operators ∂_0, A to the tensor product spaces $H_{\nu,0} \otimes H$ by interpreting A as $1_{H_{\nu,0}} \otimes A$ with $1_{H_{\nu,0}} : H_{\nu,0} \to H_{\nu,0}$ as the identity operator in $H_{\nu,0}$ and the time derivative ∂_0 as $\partial_0 \otimes 1_H$, where $1_H : H \to H$ denotes the identity operator in H.

In this sense, our aim is to be able to find $U \in H_{\nu,0} \otimes H$ such that for a given $f \in H_{\nu,0} \otimes H$ we have

$$(\partial_0 M(\partial_0^{-1}) + A)\, U = f. \tag{1.3.1}$$

The operator $M(\partial_0^{-1})$ will be referred to as the material operator.

Here $(M(z))_{z\in B_{\mathbb{C}}(r,r)}$ is a uniformly bounded, holomorphic family of linear operators in H with $r \geq \frac{1}{2\nu}$. It is

$$\boxed{M(\partial_0^{-1}) \coloneqq \mathcal{L}_\nu^* \, M\left(\frac{1}{\mathrm{i}\, m+\nu}\right) \mathcal{L}_\nu.}$$

To warrant a solution theory we require an additional constraint on such causal materials: There should be a constant $c \in \mathbb{R}_{>0}$ such that

$$\boxed{\begin{aligned} &\mathfrak{Re}\left(\left\langle U \middle| \left(z^{-1}M(z) + A\right) U\right\rangle_H\right) \geq c\,\langle U|\,U\rangle_H\,, \\ &c > 0.\end{aligned}} \tag{posdef}$$

for all $z \in B_{\mathbb{C}}(r,r)$ and $U \in D(A)$. If A is skew-selfadjoint this reduces to

$$\mathfrak{Re}\left(\left\langle U \middle| z^{-1}M(z)\, U\right\rangle_H\right) \geq c\,\langle U|\,U\rangle_H$$

for all $z \in B_{\mathbb{C}}(r,r)$ and $U \in D(A)$.

1.4. Solution Theory

THEOREM 111. *Let $(M(z))_{z\in B_{\mathbb{C}}(r,r)}$ be a holomorphic family of uniformly bounded linear operators on H, $\nu \geq \frac{1}{2r}$, satisfying our definiteness condition (posdef) and A skew-selfadjoint in H, then we have for every $f \in H_{\nu,0}\otimes H$ a unique solution $U \in H_{\nu,0}\otimes H$ of the problem*

$$(\partial_0 M(\partial_0^{-1}) + A)\, U = f.$$

Moreover, the solution depends continuously on the data in $H_{\nu,0}\otimes H$ and is causal in the sense that

$$\chi_{]-\infty,a]}(m_0)\left(\partial_0 M(\partial_0^{-1}) + A\right)^{-1}\left(1 - \chi_{]-\infty,a]}(m_0)\right) = 0 \tag{1.4.1}$$

for one $a \in \mathbb{R}$ (and so for all $a \in \mathbb{R}$).

To bye-pass the details of this result (see Appendix B for more details) let us be more specific by restricting our attention to the simple case

$$M(\partial_0^{-1}) = M_0 + \partial_0^{-1} M_1.$$

The needed positivity requirement is satisfied if M_0 is selfadjoint and with ι_0 and ι_1 the canonical embeddings of $M_0[H]$ and $[\{0\}]\,M_0$, respectively,

$$\nu\, \iota_0^* M_0 \iota_0 + \iota_0^*\, \mathfrak{Re}\, M_1 \iota_0 \geq c_1 \tag{1.4.2}$$

and

$$\mathfrak{Re}\begin{pmatrix} \partial_0 \iota_0^* M_0 \iota_0 + \iota_0^* M_1 \iota_0 & \iota_0^* M_1 \iota_1 \\ \iota_1^* M_1 \iota_0 & \iota_1^* M_1 \iota_1 \end{pmatrix} = \begin{pmatrix} \nu \iota_0^* M_0 \iota_0 + \iota_0^* M_1 \iota_0 & \iota_0^*\, \mathfrak{Re}\, M_1 \iota_1 \\ \iota_1^*\, \mathfrak{Re}\, M_1 \iota_0 & \iota_1^*\, \mathfrak{Re}\, M_1 \iota_1 \end{pmatrix} \geq c_0,$$

for some positive constants $c_1 \geq c_0$, which in turn is true when

$$\begin{pmatrix} \nu\iota_0^* M_0 \iota_0 + \iota_0^* M_1 \iota_0 & 0 \\ 0 & \iota_1^* \mathfrak{Re}\, M_1 \iota_1 - \iota_1^* \mathfrak{Re}\, M_1 \iota_0 \left(\nu\iota_0^* M_0 \iota_0 + \iota_0^* M_1 \iota_0\right)^{-1} \iota_0^* \mathfrak{Re}\, M_1 \iota_1 \end{pmatrix} \geq c_0.$$

The latter is the case if

$$\iota_1^* \mathfrak{Re}\, M_1 \iota_1 \geq c_0 + \frac{1}{c_1} \left\| \iota_0^* \mathfrak{Re}\, M_1 \iota_1 \right\|^2 . \tag{1.4.3}$$

A typical situation where this occurs is if $\iota_0^* M_0 \iota_0 \geq \varepsilon_0 > 0$, in which case

$$\nu\, \iota_0^* M_0 \iota_0 + \iota_0^* \mathfrak{Re}\, M_1 \iota_0 \geq \nu\varepsilon_0 - \left\| \iota_0^* \mathfrak{Re}\, M_1 \iota_1^* \right\|$$

and so

$$\iota_1^* \mathfrak{Re}\, M_1 \iota_1 \geq c_0 + \frac{1}{\nu\varepsilon_0 - \left\| \iota_0^* \mathfrak{Re}\, M_1 \iota_1 \right\|} \left\| \iota_0^* \mathfrak{Re}\, M_1 \iota_1 \right\|^2 ,$$

which is valid for some positive c_0 for all sufficiently large ν. Thus, in our simplified situation we may focus on the case

$$\iota_0^* M_0 \iota_0,\ \iota_1^* \mathfrak{Re}\, M_1 \iota_1 \text{ strictly positive definite.} \tag{1.4.4}$$

Literature on Evolutionary Equations

[1] R. Picard. *Hilbert space approach to some classical transforms.* John Wiley, New York, 1989.

[2] R. Picard. A structural observation for linear material laws in classical mathematical physics. *Math. Methods Appl. Sci.*, 32(14):1768–1803, 2009.

[3] R. Picard. Mother Operators and their Descendants. *Journal of Mathematical Analysis and Applications*, 403(1):54–62, 2013.

[4] R. Picard and D. McGhee. *Partial differential equations. A unified Hilbert space approach.* de Gruyter Expositions in Mathematics 55. Berlin: de Gruyter. xviii, 2011.

[5] R. Picard. A class of evolutionary problems with an application to acoustic waves with impedance type boundary conditions. In *Spectral Theory, Mathematical System Theory, Evolution Equations, differential and difference Equations*, volume 221 of *Operator Theory: Advances and Applications*, pages 533–548. Springer Basel, 2012.

[6] R. Picard, S. Trostorff, and M. Waurick. On a comprehensive Class of Linear Control Problems. Technical report, TU Dresden, 2012. arXiv:1208.3140.

[7] R. Picard, S. Trostorff, and M. Waurick. On Evolutionary Equations with Material Laws Containing Fractional Integrals. Technical report, TU Dresden, 2013. arXiv:1304.7620, submitted.

[8] R. Picard, S. Trostorff, M. Waurick, and M. Wehowski. On Non-autonomous Evolutionary Problems. *J. Evol. Equ.*, 13(4):751–776, 2013.

CHAPTER 2

Some Applications

2.1. Visco-Elastic Media

2.1.1. Visco-Elastic Solids. The system of visco-elasticity is formally given as

$$\begin{aligned} \partial_0 \varrho + \operatorname{div} \varrho\, \partial_0 u &= 0 \\ \operatorname{Div} T + f &= \varrho\, \partial_0^2 u \end{aligned}$$

where u denotes the displacement field, T the stress tensor, ϱ mass density. Here $\operatorname{Div} T = \left(\sum_{k=1}^{3} \partial_k T_{jk}\right)_{j=1,2,3}$.
A somewhat simplified linearization leads to considering

$$\begin{aligned} \partial_0 \varrho + \varrho_0 \operatorname{div} \partial_0 u &= 0 \\ \operatorname{div} T + f &= \varrho_0\, \partial_0^2 u, \end{aligned}$$

where now $\varrho_0 : L^2(\Omega) \to L^2(\Omega)$ is assumed to be bounded, selfadjoint and strictly positive definite.

In order to see that also this system leads to one of the above abstract form we need to implement some re-formulation. With $v := \partial_0 u$ we first derive from the definition

$$\mathcal{E} := \operatorname{Grad} u,$$

where $\operatorname{Grad} u := \frac{1}{2}\left(\partial \otimes u + (\partial \otimes u)^\top\right)$ denotes the symmetric part of the Jacobi matrix $\partial \otimes u$, another first order dynamic equation

$$\partial_0 \mathcal{E} = \operatorname{Grad} v.$$

Since with the matrix trace operation $A \mapsto \operatorname{trace} A$

$$\begin{aligned} \operatorname{div} u &= \operatorname{trace} \partial \otimes u \\ &= \operatorname{trace} \operatorname{Grad} u \\ &= \operatorname{trace} \mathcal{E}, \end{aligned}$$

we can formally summarize the system in the form

$$\begin{aligned} \partial_0 \begin{pmatrix} 0 \\ \varrho_0\, v \\ \mathcal{E} \end{pmatrix} + \begin{pmatrix} \varrho_0^{-1} & 0 & \operatorname{trace} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \varrho \\ v \\ \mathcal{E} \end{pmatrix} + \\ + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\operatorname{Div} \\ 0 & -\operatorname{Grad} & 0 \end{pmatrix} \begin{pmatrix} \varrho \\ v \\ T \end{pmatrix} = \begin{pmatrix} 0 \\ f \\ 0 \end{pmatrix}. \end{aligned}$$

The system is completed by linear material relations of various forms.

The Kelvin-Voigt Model

This class of materials is characterized by a material relation of the form

$$T = C\mathcal{E} + D\partial_0\mathcal{E}, \tag{2.1.1}$$

where the elasticity tensor C and the viscosity tensor D are assumed to be modeled as bounded, selfadjoint, strictly positive definite mappings in a Hilbert space H_{sym} of $L^2(\Omega)$-valued, selfadjoint 3×3-matrices, with the inner product induced by the Frobenius norm

$$(\Phi, \Psi) \mapsto \int_\Omega \operatorname{trace}\left(\Phi(x)^* \Psi(x)\right) \, dx\,.$$

Noting that the first equation of our system is trivial, once the other equations have been solved, we may indeed – assuming for example Dirichlet boundary conditions – reduce the system to

$$\begin{aligned}\begin{pmatrix} \varrho_0\partial_0 & -\operatorname{Div} \\ -\overset{\circ}{\operatorname{Grad}} & \left(\partial_0^{-1}C + D\right)^{-1} \end{pmatrix}\begin{pmatrix} v \\ T \end{pmatrix} &= \begin{pmatrix} \varrho_0\partial_0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v \\ T \end{pmatrix} + \\ &+ \begin{pmatrix} 0 & 0 \\ 0 & \left(\partial_0^{-1}C + D\right)^{-1} \end{pmatrix}\begin{pmatrix} v \\ T \end{pmatrix} + \\ &+ \begin{pmatrix} 0 & -\operatorname{Div} \\ -\overset{\circ}{\operatorname{Grad}} & 0 \end{pmatrix}\begin{pmatrix} v \\ T \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}.\end{aligned}$$

This leads to a reduced material law operator

$$\begin{aligned}M\left(\partial_0^{-1}\right) &= \begin{pmatrix} \varrho_0 & 0 \\ 0 & 0 \end{pmatrix} + \\ &+ \partial_0^{-1}\begin{pmatrix} 0 & 0 \\ 0 & \left(\partial_0^{-1}C + D\right)^{-1} \end{pmatrix}.\end{aligned}$$

This is the so-called Kelvin-Voigt model of visco-elasticity. The case $C = 0$ leads to a system for a purely viscous behavior (**Newton model**). On the other hand, if C is strictly positive definite, then the limit case $D = 0$ leads to the **standard system for elastic solids**.

The Maxwell Model

An alternative description of visco-elastic solids is due to Maxwell. The material relation here is of the form

$$\partial_0\mathcal{E} = C^{-1}\partial_0 T + D^{-1}T,$$

where we have re-used the named C, D for the coefficients although they of course have different meanings here.

We see that in this case

$$\mathcal{E} = \left(C^{-1} + \partial_0^{-1}D^{-1}\right)T$$

and the material law in this case takes on the form

$$V = \begin{pmatrix} \varrho_0 & 0 \\ 0 & C^{-1} + \partial_0^{-1}D^{-1} \end{pmatrix}\begin{pmatrix} v \\ T \end{pmatrix}.$$

We see that the Maxwell model leads for C strictly positive definite to a regular material law. The limit case of vanishing C^{-1} formally recovers the earlier mentioned Newton model.

The Poynting-Thomson Model (The Linear Standard Model)

The linear standard model or Poynting-Thomson model is based on a generalization of the Maxwell model involving another coefficient operator R and has the form

$$\partial_0\mathcal{E} + R\,\mathcal{E} = C^{-1}\partial_0 T + D^{-1}T.$$

Solving for $\mathcal{E}$ yields

$$\begin{aligned}\mathcal{E} &= (\partial_0 + R)^{-1}\left(C^{-1}\partial_0 + D^{-1}\right)T\\ &= (\partial_0 + R)^{-1}\left(C^{-1}(\partial_0 + R) + D^{-1} - C^{-1}R\right)T\\ &= C^{-1}T + \left(1 + R\,\partial_0^{-1}\right)^{-1}\left(D^{-1} - C^{-1}R\right)\partial_0^{-1}T,\end{aligned}$$

leading to a slightly more complex material law.

Even more complex materials such as the **Burgers model** and suitable **models of "type** (p,q)**"** can be seen to be of the general form . The latter class leads to a material law operator $M\left(\partial_0^{-1}\right)$ given as a rational operator-valued function of ∂_0^{-1}, e. g.

$$M\left(\partial_0^{-1}\right) = Q_0\left(\partial_0^{-1}\right)^{-1} Q_1\left(\partial_0^{-1}\right)$$

with $Q_k(z) = \sum_{s=0}^{N_k} R_{ks} z^s$, where the coefficients R_{ks} are bounded linear operators in H_{sym}, $s = 0, \ldots N_k$, $k = 0, 1$, R_{00} invertible, (or finite products of terms of this form).

Also material laws involving fractional positive powers of ∂_0^{-1} are utilized in applications. Since $z \mapsto z^\alpha$ is analytic in $[\mathbb{R}_{>0}] + \mathrm{i}\,[\mathbb{R}]$ for every $\alpha \in \mathbb{R}$ also such material laws are covered by our abstract approach.

2.1.2. Visco-Elastic Stokes Fluids. Linear visco-elastic fluids are described by a slightly different material relation, a modified Kelvin-Voigt model, of the form

$$T = C\,\mathcal{E} + D\,\partial_0\mathcal{E} - \operatorname{trace}^*\kappa_0\varrho$$

with another coefficient operator $\kappa_0 : L^2(\Omega) \to L^2(\Omega)$ bounded, selfadjoint, strictly positive definite. Thus,

$$\mathcal{E} = (C + D\partial_0)^{-1}T + (C + D\partial_0)^{-1}\operatorname{trace}^*\kappa_0\varrho.$$

For D strictly positive definite we may re-write the resulting system

$$\begin{aligned}\varrho_0^{-1}\,\partial_0\varrho + \operatorname{div}\partial_0 u &= 0\\ \operatorname{Div} T + f &= \varrho\,\partial_0^2 u\\ T &= C\,\mathcal{E} + D\,\partial_0\mathcal{E} - \operatorname{trace}^*\kappa_0\varrho\end{aligned}$$

in the form ($v = \partial_0 u$)

$$\begin{pmatrix} \kappa_0 \varrho_0^{-1} \partial_0 & \kappa_0 \operatorname{trace} \mathring{\operatorname{Grad}} & 0 \\ 0 & \varrho_0 \partial_0 & -\operatorname{Div} \\ \left(\partial_0^{-1} C + D\right)^{-1} \operatorname{trace}^* \kappa_0 & -\mathring{\operatorname{Grad}} & \left(\partial_0^{-1} C + D\right)^{-1} \end{pmatrix} \begin{pmatrix} \varrho \\ v \\ T \end{pmatrix} = $$
$$= \begin{pmatrix} 0 \\ f \\ 0 \end{pmatrix},$$

where we have chosen again the Dirichlet boundary condition to illustrate the procedure.

A simple row operation translates this into a more symmetric form

$$\begin{pmatrix} \kappa_0 \varrho_0^{-1} \partial_0 + \kappa_0 \operatorname{trace} \left(\partial_0^{-1} C + D\right)^{-1} \operatorname{trace}^* \kappa_0 & 0 & \kappa_0 \operatorname{trace} \left(\partial_0^{-1} C + D\right)^{-1} \\ 0 & \varrho_0 \, \partial_0 & -\operatorname{Div} \\ \left(\partial_0^{-1} C + D\right)^{-1} \operatorname{trace}^* \kappa_0 & -\mathring{\operatorname{Grad}} & \left(\partial_0^{-1} C + D\right)^{-1} \end{pmatrix} \begin{pmatrix} \varrho \\ v \\ T \end{pmatrix} = $$
$$= \begin{pmatrix} 0 \\ f \\ 0 \end{pmatrix}.$$

$$\begin{aligned} M\left(\partial_0^{-1}\right) &= \begin{pmatrix} \kappa_0 \varrho_0^{-1} & 0 & 0 \\ 0 & \varrho_0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \\ &+ \partial_0^{-1} \begin{pmatrix} \kappa_0 \operatorname{trace} \left(\partial_0^{-1} C + D\right)^{-1} \operatorname{trace}^* \kappa_0 & 0 & \kappa_0 \operatorname{trace} \left(\partial_0^{-1} C + D\right)^{-1} \\ 0 & 0 & 0 \\ \left(\partial_0^{-1} C + D\right)^{-1} \operatorname{trace}^* \kappa_0 & 0 & \left(\partial_0^{-1} C + D\right)^{-1} \end{pmatrix}, \\ &= \begin{pmatrix} \kappa_0 \varrho_0^{-1} & 0 & 0 \\ 0 & \varrho_0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \partial_0^{-1} \begin{pmatrix} \kappa_0 \operatorname{trace} D^{-1} \operatorname{trace}^* \kappa_0 & 0 & \kappa_0 \operatorname{trace} D^{-1} \\ 0 & 0 & 0 \\ D^{-1} \operatorname{trace}^* \kappa_0 & 0 & D^{-1} \end{pmatrix} + \\ &- \partial_0^{-2} \begin{pmatrix} \kappa_0 \operatorname{trace} \left(\partial_0^{-1} C + D\right)^{-1} C D^{-1} \operatorname{trace}^* \kappa_0 & 0 & \kappa_0 \operatorname{trace} \left(\partial_0^{-1} C + D\right)^{-1} C D^{-1} \\ 0 & 0 & 0 \\ \left(\partial_0^{-1} C + D\right)^{-1} C D^{-1} \operatorname{trace}^* \kappa_0 & 0 & \left(\partial_0^{-1} C + D\right)^{-1} C D^{-1} \end{pmatrix}. \end{aligned}$$

The case of incompressible media, where $\partial_0 \varrho = 0$, is a "singular limit case" in so far as even requirement (posdef) does not hold anymore. The well-known difficulties in solving the Stokes system, which are in general not well-posed in an $L^2(\Omega)$-setting, are an indication of how degenerate a problem may be if it does not satisfy (posdef).

2.2. Thermoelasticity

We consider the following general thermo-elastic system for simplicity in the case of Dirichlet boundary conditions:

$$\partial_0 V + \begin{pmatrix} 0 & \operatorname{Div} & 0 & 0 \\ \mathring{\operatorname{Grad}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \operatorname{div} \\ 0 & 0 & \mathring{\operatorname{grad}} & 0 \end{pmatrix} \begin{pmatrix} v \\ T \\ \vartheta \\ Q \end{pmatrix} = \begin{pmatrix} f \\ 0 \\ g \\ 0 \end{pmatrix}$$

with a material law of the form

$$V = M\left(\partial_0^{-1}\right)\begin{pmatrix} v \\ T \\ \vartheta \\ Q \end{pmatrix}.$$

Here

$$M\left(\partial_0^{-1}\right) =$$

$$= \begin{pmatrix} \varrho_0 & 0 & 0 & 0 \\ 0 & C^{-1} & C^{-1}\Gamma & 0 \\ 0 & \Gamma^* C^{-1} & w + \Gamma^* C^{-1}\Gamma & 0 \\ 0 & 0 & 0 & q_0 + q_2\left(\alpha + \beta\partial_0\right)^{-1} \end{pmatrix}.$$

We see that

$$\begin{pmatrix} \varrho_0 & 0 & 0 & 0 \\ 0 & C^{-1} & C^{-1}\Gamma & 0 \\ 0 & \Gamma^* C^{-1} & w + \Gamma^* C^{-1}\Gamma & 0 \\ 0 & 0 & 0 & q_0 + q_2\left(\alpha + \beta\partial_0\right)^{-1} \end{pmatrix}$$

is by symmetric Gauss elimination equivalent to the block diagonal form

$$\begin{pmatrix} \varrho_0 & 0 & 0 & 0 \\ 0 & C^{-1} & 0 & 0 \\ 0 & 0 & w & 0 \\ 0 & 0 & 0 & q_0 + q_2\left(\alpha + \beta\partial_0\right)^{-1} \end{pmatrix}.$$

E.g. the issue of

$$M_0 = \begin{pmatrix} \varrho_0 & 0 & 0 & 0 \\ 0 & C^{-1} & C^{-1}\Gamma & 0 \\ 0 & \Gamma^* C^{-1} & w + \Gamma^* C^{-1}\Gamma & 0 \\ 0 & 0 & 0 & q_0 \end{pmatrix}$$

being strictly positive definite hinges on the strict positive-definiteness of ϱ_0, C, w, q_0.

For $q_0 = 0$ the above system is known as a type 3 thermo-elastic system. With $\alpha = 0$ we obtain the special case of thermo-elasticity with second sound, i.e. with the Cattaneo modification of the heat transport. The so-called type 2 thermo-elastic system results by letting $q_2 = 0$.

We point out that the well-known Biot system, which describes consolidation of a linearly elastic porous medium, can be reformulated so that up to physical interpretations it has the same form as the thermo-elastic system (with Cattaneo modification). The coupling operator Γ of thermo-elasticity is in the poro-elastic case given as $\Gamma = \operatorname{trace}^*\alpha$, where α is a coupling parameter.

2.3. Piezo-Electro-Magnetism

Here we have a system of the form

$$\partial_0 V + A \begin{pmatrix} v \\ T \\ E \\ H \end{pmatrix} = \begin{pmatrix} f \\ 0 \\ -J \\ 0 \end{pmatrix},$$

where a possible choice of boundary conditions would for example lead to the skew-selfadjoint block operator matrix

$$A = \begin{pmatrix} 0 & \mathrm{Div} & 0 & 0 \\ \overset{\circ}{\mathrm{Grad}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathrm{rot} \\ 0 & 0 & -\overset{\circ}{\mathrm{rot}} & 0 \end{pmatrix}.$$

This system needs to be completed by suitable material relations. A known coupling mechanism is initially described in the form

$$\begin{aligned} T &= C\,\mathcal{E} - dE - q\,H, \\ D &= d^*\mathcal{E} + \varepsilon E + e\,H, \\ B &= q^*\,\mathcal{E} + e^* E + \mu H. \end{aligned}$$

Initial and final spaces of the additional bounded, linear coefficient operators q and e are clear from these equation and for sake of brevity we shall not elaborate on this. As has been already noted in the above, for a proper reformulation we need to solve for $\mathcal{E}$ to obtain suitable material relations. We find

$$\begin{aligned} \mathcal{E} &= C^{-1}T + C^{-1}dE + C^{-1}q\,H, \\ D &= \eta^* C^{-1}T + \left(\varepsilon + d^*C^{-1}d\right)E + d^*C^{-1}q\,H + e\,H, \\ B &= q^*\,C^{-1}T + q^*\,C^{-1}dE + q^*C^{-1}q\,H + e^*\,E + \mu H. \end{aligned}$$

Thus, we obtain the material law

$$V = M\left(\partial_0\right)^{-1} \begin{pmatrix} v \\ T \\ E \\ H \end{pmatrix}$$

with

$$M\left(\partial_0^{-1}\right) = \begin{pmatrix} \varrho_0 & 0 & 0 & 0 \\ 0 & C^{-1} & C^{-1}d & C^{-1}q \\ 0 & d^*C^{-1} & \left(\varepsilon + d^*C^{-1}d\right) & d^*C^{-1}q + e \\ 0 & q^*C^{-1} & q^*C^{-1}d + e^* & \mu + q^*C^{-1}q \end{pmatrix}.$$

By block diagonalizing this via symmetric Gaussian elimination we obtain from

$$\begin{pmatrix} \varrho_0 & 0 & 0 & 0 \\ 0 & C^{-1} & C^{-1}d & C^{-1}q \\ 0 & d^*C^{-1} & \varepsilon + d^*C^{-1}d & d^*C^{-1}q + e \\ 0 & q^*C^{-1} & q^*C^{-1}d + e^* & \mu + q^*C^{-1}q \end{pmatrix}$$

the block diagonal operator matrix

$$\begin{pmatrix} \varrho_0 & 0 & 0 & 0 \\ 0 & C^{-1} & 0 & 0 \\ 0 & 0 & \varepsilon & 0 \\ 0 & 0 & 0 & \mu - e^*\varepsilon^{-1}e \end{pmatrix}.$$

Thus, the given form of material relations only leads to a regular material law in the above sense if in addition to the strict positive-definiteness of the selfadjoint bounded operators ϱ_0, C, ε and μ we require

$$\mu \geq \mu_0 + e^*\varepsilon^{-1}e$$

for some constant $\mu_0 \in \mathbb{R}_{>0}$.

2.4. Thermo-Piezo-Electro-Magnetism

We shall conclude our example collection by coupling also heat transport effects into to our earlier simple version of a piezo-electro-magnetic system. We base our consideration on the material relations suggested by R.D. Mindlin. We are led to the system

$$\partial_0 V + A \begin{pmatrix} v \\ T \\ E \\ H \\ \sqrt{\varrho_0 \Theta}\,\vartheta \\ Q \end{pmatrix} = \begin{pmatrix} f \\ 0 \\ -J \\ 0 \\ \sqrt{\varrho_0 \Theta}\, g \\ 0 \end{pmatrix}.$$

Here A is a skew-selfadjoint operator such as

$$A = \begin{pmatrix} 0 & \operatorname{Div} & 0 & 0 & 0 & 0 \\ \mathring{\operatorname{Grad}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \operatorname{rot} & 0 & 0 \\ 0 & 0 & -\mathring{\operatorname{rot}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\left(\varrho_0\Theta\right)^{-1/2}\ \operatorname{div} \\ 0 & 0 & 0 & 0 & -\mathring{\operatorname{grad}}\left(\varrho_0\Theta\right)^{-1/2} & 0 \end{pmatrix}.$$

The subtle adjustment of taking $\sqrt{\varrho_0\Theta}\vartheta$ in place of ϑ has been implemented to accommodate the particular formulation of heat conduction employed in the literature and to maintain symmetry in the $L^2(\Omega)$-type sense. This, also makes the skew-selfadjointness more evident, then by making adjustments via a change of inner products.

The so-called specific heat capacity Θ can also be allowed to be a bounded, selfadjoint, strictly positive definite mapping in $L^2(\Omega)$ commuting with ϱ_0. The material relations are initially given in the form

$$\begin{aligned} T &= C\,\mathcal{E} - dE - \lambda\vartheta, \\ D &= d^*\mathcal{E} + \varepsilon E + p\,\vartheta, \\ B &= \mu\,H, \\ m^{-1}\Theta^{-1}\sigma &= \lambda^*\mathcal{E} + p^*E + \alpha\,\vartheta. \end{aligned}$$

Applying our earlier reasoning, we solve for $\mathcal{E}$ and obtain

$$\begin{aligned} \mathcal{E} &= C^{-1}T + C^{-1}dE + C^{-1}\lambda\vartheta, \\ D &= d^*C^{-1}T + \left(\varepsilon + d^*C^{-1}d\right)E + \left(p + d^*C^{-1}\lambda\right)\vartheta, \\ B &= \mu\,H, \\ m^{-1}\Theta^{-1}\sigma &= \lambda^*C^{-1}T + \left(p^* + \lambda^*C^{-1}d\right)E + \left(\alpha + \lambda^*C^{-1}\lambda\right)\vartheta. \end{aligned}$$

Thus,

$$V = M\left(\partial_0^{-1}\right)\begin{pmatrix} v \\ T \\ E \\ H \\ \sqrt{\varrho_0\Theta}\,\vartheta \\ Q \end{pmatrix},$$

where $M\left(\partial_0^{-1}\right)$ is of the block form

$$\begin{pmatrix} m_{00} & m_{01} & m_{02} \\ m_{01}^* & m_{11} & m_{12} \\ m_{02}^* & m_{12}^* & m_{22} \end{pmatrix}$$

with block diagonal selfadjoint operator matrices

$$m_{00} = \begin{pmatrix} \varrho_0 & 0 \\ 0 & C^{-1} \end{pmatrix}, \quad m_{11} = \begin{pmatrix} \left(\varepsilon + d^*C^{-1}d\right) & 0 \\ 0 & \mu \end{pmatrix},$$

$$m_{22} = \begin{pmatrix} \left(\varrho_0\Theta\right)^{-1/2}\left(\alpha + \lambda^*C^{-1}\lambda\right)\left(\varrho_0\Theta\right)^{-1/2} & 0 \\ 0 & \kappa^{-1}\partial_0^{-1} \end{pmatrix}.$$

Moreover

$$\begin{aligned} m_{01} &= \begin{pmatrix} 0 & 0 \\ C^{-1}d & 0 \end{pmatrix}, \\ m_{02} &= \begin{pmatrix} 0 & 0 \\ C^{-1}\lambda\left(\varrho_0\Theta\right)^{-1/2} & 0 \end{pmatrix}, \\ m_{12} &= \begin{pmatrix} \left(p + d^*C^{-1}\lambda\right)\left(\varrho_0\Theta\right)^{-1/2} & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

If we assume that ϱ_0, C, ε, μ, α, κ are all bounded, selfadjoint and strictly positive definite in suitable $L^2(\Omega)$-type spaces the problem is covered by our general setting.

Literature for the Application Part

[1] Jacob Bear and Yehuda Bachmat. *Introduction to modelling of transport phenomena in porous media.* Theory and Applications of Transport in Porous Media. 4. Dordrecht etc.: Kluwer Academic Publishers. xxiv, 553 p. , 1990.

[2] B. A. Bednarcyk. A Fully Micro/Macro Theory for Thermo-Electro-Magento-Elasto-Plastic Composite Laminates. Technical report, NASA, 2002. Technical Report 211468.

[3] Albrecht Bertram. *Elasticity and Plasticity of Large Deformations: An Introduction.* Berlin: Springer. xiv, 326 p., 2005.

[4] M. A. Biot. General theory of three-dimensional consolidation. *J. appl. Physics, Lancaster Pa.,*, 12:155–164, 1941.

[5] Carlo Cattaneo. Sulla conduzione del calore. *Atti Semin. Mat. Fis. Univ., Modena*, 3:83–101, 1949.

[6] Song Jiang and Reinhard Racke. *Evolution equations in thermoelasticity.* Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics. 112. Boca Raton, FL: Chapman & Hall/CRC., 308 p., 2000.

[7] Des F. McGhee and R. Picard. A note on anisotropic, inhomogeneous, poro-elastic media. *Mathematical Methods in the Applied Sciences*, 33:313–322, 2010.

[8] R.D. Mindlin. Equations of high frequency vibrations of thermopiezoelectric crystal plates. *Int. J. Solids Structures*, 10:625–637, 1974.

[9] R. Picard. *Hilbert space approach to some classical transforms.* John Wiley, New York, 1989.

[10] R. Picard. Evolution Equations as operator equations in lattices of Hilbert spaces. *Glasnik Matematicki Series III*, 35:111–136, 2000.

[11] R. Picard. On a model of electromagnetic field propagation in ferro-electric media. *J. Math. Anal. Appl.*, 328(I):655–675, 2007.

[12] R. Picard. Linear thermo-elasticity in nonsmooth media. *Mathematical Methods in the Applied Sciences*, 28:2183–2199, 2005.

[13] R. Seeliger. *Elektronentheorie der Metalle*, volume 5 (Physik) of *Encyklyklopaedie der mathematischen Wissenschaften. Mit Einschluss ihrer Anwendungen.*, pages 777–878. Teubner, Leipzig, 1922.

[14] R.E. Showalter. Diffusion in poro-elastic media. *J. Math. Anal. Appl.*, 251(1):310–340, 2000.

APPENDIX A

Metric Spaces

In analysis, topological structure is often provided by a measure of distance, i.e. a metric.

DEFINITION 112. Let M be a non-empty set. A mapping $d: M \times M \longrightarrow \mathbb{R}$ is called a **semi-metric** on M if

(1) $\bigwedge_{x \in M} d(x,x) = 0$,
(2) d is symmetric, i.e. $\bigwedge_{x,\, y \in M} d(x,y) = d(y,x)$,
(3) d satisfies the 'triangle inequality', i.e.
$$\bigwedge_{x,\, y,\, z \in M} d(x,y) \le d(x,z) + d(y,z)\ .$$

REMARK 113. Note that the property
$$\bigwedge_{x,\, y \in M} d(x,y) \ge 0$$
actually follows by setting $x = y$ in 3. Moreover, by the triangle inequality, for any $x,\ y,\ u,\ v \in M$
$$d(x,y) \le d(x,u) + d(u,v) + d(v,y),$$
and
$$d(u,v) \le d(u,x) + d(x,y) + d(y,v).$$
Thus, using symmetry, we obtain the inequality
$$|d(x,y) - d(u,v)| \le d(x,u) + d(y,v) \text{ for all } x,\ y,\ u,\ v \in M. \tag{A.0.1}$$

DEFINITION 114. Let d be a semi-metric on a set M. If in addition d is definite, i.e.
$$\bigwedge_{x,\, y \in M} d(x,y) = 0 \ \Rightarrow\ x = y,$$
then d is called a **metric** on M.

The topology induced by such a (semi-) metric is based on the concept of an **open ball** $B_d(x,r) := \{y \in M |\, d(x,y) < r\}$ with **center** $x \in M$ and **radius** $r \in \mathbb{R}_{>0}$, where $\mathbb{R}_{>0}$ denotes the set of positive real numbers. If the metric is clear from the context we shall simply write $B(x,r)$ for this ball.

A (semi-) metric on a non-empty set M defines a topology (i.e. a set of open sets).

DEFINITION 115. Let M be a non-empty set. The pair $(M,\, d)$ is called a **(semi-) metric space** if d is a **(semi-) metric on** M**.** $(M,\, \mathcal{O})$ is a topological space where the topology is given by
$$\mathcal{O} := \left\{ A \subseteq M \mid \bigwedge_{x \in A} \bigvee_{r \in \mathbb{R}_{>0}} B_d(x,r) \subseteq A \right\} \tag{A.0.2}$$
where d is a (semi-) metric.

The notion of boundedness resides in (semi-) metric spaces.

Definition 116. Let (M, d) be a (semi-) metric space. A subset $A \subseteq M$ is called **bounded** if its **diameter** $\sup \{d(x,y) \,|\, x, y \in A\}$ is finite. A mapping $f : N \longrightarrow M$ defined on a set N is called bounded if its **range** $f[N]$ is bounded.

Continuity of mappings between (semi-) metric spaces can now be given the familiar characterization:

Proposition 117. *Let $f : D(f) \subseteq M \longrightarrow N$ be a mapping and (M, d_M) and (N, d_N) (semi-) metric spaces. Then f is **continuous at** $x \in D(f)$ if and only if*

$$\bigwedge_{\varepsilon\in\mathbb{R}_{>0}} \bigvee_{\delta\in\mathbb{R}_{>0}} \bigwedge_{y\in D(f)} d_M(x,y) < \delta \Longrightarrow d_N(f(x), f(y)) < \varepsilon.$$

*The mapping f is **continuous in** $U \subseteq D(f)$ if and only if*

$$\bigwedge_{x\in U} \bigwedge_{\varepsilon\in\mathbb{R}_{>0}} \bigvee_{\delta\in\mathbb{R}_{>0}} \bigwedge_{y\in D(f)} d_M(x,y) < \delta \Longrightarrow d_N(f(x), f(y)) < \varepsilon.$$

*The mapping f is called **continuous** if f is continuous in $D(f)$.*

A concept of continuity particular to (semi-) metric spaces is the concept of uniform continuity.

Definition 118. Let $f : D(f) \subseteq M \longrightarrow N$ be a mapping and (M, d_M) and (N, d_N) (semi-) metric spaces. Then f is called **uniformly continuous** if

$$\bigwedge_{\varepsilon\in\mathbb{R}_{>0}} \bigvee_{\delta\in\mathbb{R}_{>0}} \bigwedge_{x,\, y\in D(f)} d_M(x,y) < \delta \Longrightarrow d_N(f(x), f(y)) < \varepsilon\,.$$

The function f is called **locally uniformly continuous** if

$$\bigwedge_{z\in D(f)} \bigvee_{\eta\in\mathbb{R}_{>0}} \bigwedge_{\varepsilon\in\mathbb{R}_{>0}} \bigvee_{\delta\in\mathbb{R}_{>0}} \bigwedge_{x,\, y\in B(z,\eta)\cap D(f)} d_M(x,y) < \delta \Longrightarrow d_N(f(x), f(y)) < \varepsilon.$$

The function f is called[1] **M-locally uniformly continuous** or**Cauchy continuous** if

$$\bigwedge_{z\in M} \bigvee_{\eta\in\mathbb{R}_{>0}} \bigwedge_{\varepsilon\in\mathbb{R}_{>0}} \bigvee_{\delta\in\mathbb{R}_{>0}} \bigwedge_{x,\, y\in B(z,\eta)\cap D(f)} d_M(x,y) < \delta \Longrightarrow d_N(f(x), f(y)) < \varepsilon.$$

A particular and easily characterized example of uniform continuity is given by the so-called Lipschitz continuity.

Definition 119. Let $f : D(f) \subseteq M \longrightarrow N$ be a mapping and (M, d_M) and (N, d_N) (semi-) metric spaces. Then f is called **locally Lipschitz continuous** if

$$\bigwedge_{z\in D(f)} \bigvee_{r,\,\delta\in\mathbb{R}_{>0}} \bigwedge_{x,y\in B_{d_M}(z,r)\cap D(f)} d_N(f(x), f(y)) \le \delta\, d_M(x,y)$$

and f is called **M-locally Lipschitz continuous** if

$$\bigwedge_{z\in M} \bigvee_{r,\,\delta\in\mathbb{R}_{>0}} \bigwedge_{x,y\in B_{d_M}(z,r)\cap D(f)} d_N(f(x), f(y)) \le \delta\, d_M(x,y).$$

If there is a constant $\delta \in \mathbb{R}_{>0}$ such that

$$\bigwedge_{x,\, y\in D(f)} d_N(f(x), f(y)) \le \delta\, d_M(x,y).$$

[1]Note the slight change of topology here from the relative topology of $D(f)$ to the topology of M.

then f is called **(globally) Lipschitz continuous**. Any such constant δ is called a **Lipschitz constant**. The **best Lipschitz constant** of a globally Lipschitz continuous f is given by

$$|f|_{Lip} := \inf \left\{ \delta \in \mathbb{R}_{>0} \,\middle|\, \bigwedge_{x,\, y \in D(f)} d_N(f(x), f(y)) \leq \delta\, d_M(x, y) \right\}.$$

Finally, if f satisfies

$$\bigwedge_{x,\, y \in D(f)} d_N(f(x), f(y)) = d_M(x, y),$$

then f is called a **(semi-) isometry**.

Extending mappings by closure is fundamental to many of the elementary structural constructions of functional analysis. A mapping $f : D(f) \subseteq M \longrightarrow N$ can be considered as a particular set of pairs in $M \times N$, i.e. we identify

$$f = \{(x, y) \in M \times N \mid y = f(x)\}.$$

We shall consider $M \times N$ as a metric space with metric

$$\begin{aligned} (M \times N) \times (M \times N) &\to \mathbb{R} \\ ((x_0, y_0), (x_1, y_1)) &\mapsto \sqrt{d_M(x_0, x_1)^2 + d_N(y_0, y_1)^2}. \end{aligned}$$

We shall refer to this metric space as the direct sum space

$$M \oplus N$$

with metric $d_{M \oplus N}$. Then it is clear what the closure $\bar{f}$ of f means. A mapping f is called **closable** if $\overline{f}$ is also a mapping. It is important to note that uniformly continuous mappings between metric spaces are closable. More precisely, we have

Proposition 120. *Let (M, d_M) be a (semi-) metric space and (N, d_N) a metric space and let $f : D(f) \subseteq M \longrightarrow N$ be (locally) uniformly continuous. Then f is closable and its closure $\overline{f}$ is (locally) uniformly continuous.*

Proof. Let $x \in M$ with (x, y), $(x, z) \in \bar{f}$. Let $(x_n)_n$ be a sequence in $D(f)$ converging to x, such that

$$\begin{aligned} f(x_{2n}) &\to y \\ f(x_{2n+1}) &\to z \end{aligned}$$

as $n \to \infty$. Then, there is a mapping $N_x : \mathbb{R}_{>0} \to \mathbb{N}$ such that

$$d_M(x_n, x) < \delta$$

for $n \in \mathbb{N}_{\geq N_x(\delta)}$. By the locally uniform continuity of f we have a mapping $\delta_x : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ such that

$$d_N(f(x_n), f(x_m)) < \varepsilon \tag{A.0.3}$$

for all $n, m \in \mathbb{N}_{\geq N_x(\delta_x(\varepsilon))}$. This shows that $(f(x_n))_n$ is a Cauchy sequence, which can have at most one limit and therefore

$$y = z.$$

This shows that $\overline{f}$ is right-unique and so that f is closable.

To see the locally uniform continuity of $\bar{f}$, let $z \in M$ and $x \in D(\bar{f})$ such that $d_M(x, z) < \delta_x(\varepsilon/2)/2$ and let analogously $u \in D(\bar{f})$ such that $d_M(u, z) < \delta_u(\varepsilon/2)/2$ with a sequence $(u_n)_n$ in $D(f)$ converging to u, such that

$$\begin{aligned} f(x_n) &\to \overline{f}(x) \\ f(u_n) &\to \overline{f}(u) \end{aligned}$$

as $n \to \infty$. We have

$$d_M(v,z) < \delta_z(\varepsilon/2) \wedge d_M(w,z) < \delta_z(\varepsilon/2) \Longrightarrow d_N(f(v),f(w)) < \varepsilon/2$$

for any given $\varepsilon \in \mathbb{R}_{>0}$ and we see that $d_M(x_n,z) < \delta_z(\varepsilon/2)/2 + d_M(x_n,x) < \delta_z(\varepsilon/2)$ and $d_M(u_n,z) < \delta_z(\varepsilon/2)/2 + d_M(u_n,u) < \delta_z(\varepsilon/2)$ and so

$$d_N(f(x_n), f(u_n)) < \varepsilon/3$$

for all sufficiently large $n \in \mathbb{N}$. Consequently, we find

$$d_N(\bar{f}(x),\bar{f}(u)) \leq d_N(\bar{f}(x), f(x_n)) + d_N(f(x_n), f(u_n)) + d_N(f(u_n), \bar{f}(u)) < \varepsilon$$

for all sufficiently large $n \in \mathbb{N}$. Since $\varepsilon \in \mathbb{R}_{>0}$ and $z \in M$ were arbitrary, this shows the locally uniform continuity of $\overline{f}$. We may e.g. choose $\delta := \delta_z(\varepsilon/2)/2$ to have

$$d_N(\bar{f}(x),\bar{f}(u)) < \varepsilon$$

for all $x, u \in D(\overline{f})$ with

$$d_M(x,z),\ d_M(u,z) < \delta\,.$$

If f is uniformly continuous, then δ is independent of the choice of $z \in M$ in the above reasoning. Thus, $\bar{f}$ is uniformly continuous in this case. □

Definition 121. A metric space (N, d_N) is called **complete** if every Cauchy sequence in it converges, i.e. has a limit,

Lemma 122. *Let M be a metric spaces and N be a complete metric space. Let $f\ D(f) \subseteq M \to N$ be a locally uniformly continuous mapping, where $D(f)$ is dense in M. Then $\bar{f}$ is a locally uniformly continuous mapping with $D(\bar{f}) = M$.*

Proof. The result is clear $D(\bar{f}) = \overline{D(f)} = M$. □

Completeness is a fundamental property throughout analysis and so it is comforting to know that there is a canonical construction of a complete metric space from any semi-metric space (M, d). The more usual procedure is to construct the completion of a metric space having first, if necessary, produced a metric space from a semi-metric space by taking equivalence classes of metrically indistinguishable elements, but we choose to show that the construction works in one step directly from a semi-metric space. The first step in this construction is the transition from points in M to Cauchy sequences in M. Consider the set of Cauchy sequences[2] in M, i.e. the set

$$CS(M) := \left\{ f \in M^{\mathbb{N}} \middle| f \text{ Cauchy sequence} \right\}.$$

[2]It is interesting to note that a Cauchy sequence $f = (f_n)_{n\in\mathbb{N}}$ is characterized as a uniformly continuous mapping

$$\begin{aligned} f:\ \mathbb{N} &\to M\,,\\ n &\mapsto f_n \end{aligned}$$

where $\mathbb{N}$ is considered as a metric space with metric

$$(x,y) \mapsto \left|\frac{1}{1+x} - \frac{1}{1+y}\right| = \frac{1}{(1+x)(1+y)}|x-y|\,.$$

This space is not complete, since $(n)_{n\in\mathbb{N}}$ is a Cauchy sequence without limit in $\mathbb{N}$. Its completion in the sense explained in the following is $\mathbb{N} \cup \{\infty\}$, where ∞ is the equivalence class of all unbounded sequences in $\mathbb{N}$. This yields an alternative view on sequences as mappings

$$\begin{aligned} f \circ \left(\frac{1}{\cdot} - 1\right):\ \frac{1}{1+[\mathbb{N}]} \subset \mathbb{R} &\to M\\ \frac{1}{1+n} &\to f_n\,, \end{aligned}$$

where $\frac{1}{1+[\mathbb{N}]}$ is a subspace of $\mathbb{R}$ as the standard metric space with $|\cdot - \cdot|$ as metric. Note that

$$\begin{aligned} \mathbb{N} &\to \frac{1}{1+[\mathbb{N}]}\\ n &\mapsto \frac{1}{1+n} \end{aligned}$$

is a bijective isometry.

and define the injective mapping

$$\Phi\,:\,M\to CS(M),\quad x\longmapsto f_{(x)}:=\in\left\{f\in CS(M)\,\middle|\,\bigwedge_{s\in\mathbb{N}}f_s=x\right\}.$$

Thus we may identify M with the subset $\Phi\left[M\right]=\left\{f_{(x)}\in CS(M)\,|\,x\in M\right\}$ so we may write $M\subseteq CS(M)$. Defining the equivalence relation

$$\bigwedge_{f,g\in CS(M)}f\sim g\;:\Longleftrightarrow\;d(f,g):=\lim_{s\to 0}d(f(s),g(s))\;\to\;0$$

and corresponding equivalence classes

$$[f]:=\{g\in CS(M)\,|\,f\sim g\}\,,$$

we shall see that the following result holds.

Proposition 123. *We have that*

$$\widetilde{d}([f]\,,\,[g]):=\lim_{s\to 0}d(f_s,g_s)\tag{A.0.4}$$

is a well-defined metric on

$$\widetilde{M}:=\{\,[f]|\,f\;\in CS(M)\}\,.$$

Moreover, M is dense in $\widetilde{M}$ in the sense that $\widetilde{\Phi\left[M\right]}\;:=\;\left\{\left[f_{(x)}\right]\in\widetilde{M}\,|\,x\in M\right\}$is dense in the metric space $\left(\widetilde{M},\,\widetilde{d}\right)$, and $\left(\widetilde{M},\,\widetilde{d}\right)$ is complete.

Proof. For $[f]\,,\,[g]\in\widetilde{M}$, by inequality (A.0.1) and the fact that $f,\,g\in CS(M)$, we have

$$|d(f_m,g_m)-d(f_n,g_n)|\le d(f_m,f_n)+d(g_m,g_n)\;\to\;0\quad\text{as}\quad m,\,n\to\infty$$

which shows that $(d(f_n,\,g_n))_{n\in\mathbb{N}}$is a Cauchy sequence in $\mathbb{R}$ so that $\widetilde{d}([f]\,,\,[g]):=\lim_{n\to\infty}d(f_n,g_n)$ exists.

Now, let $u\in[f]$ and $v\in[g]$: we need to show that

$$\widetilde{d}([f]\,,\,[g])=\widetilde{d}([u]\,,\,[v]).\tag{A.0.5}$$

By inequality (A.0.1) we have

$$|d(f_n,g_n)-d(u_n,v_n)|\le d(f_n,u_n)+d(v_n,g_n)$$

from which, since $\lim_{n\to\infty}d(f_n,u_n)=\lim_{n\to\infty}d(g_n,v_n)=0$, the desired equality (A.0.5) follows. That the well-defined mapping $\widetilde{d}$ is a semi-metric follows from the semi-metric properties of d and the standard limit theorems in $\mathbb{R}$. To see that definiteness holds for $\widetilde{d}$ we only need to notice that

$$\widetilde{d}([f]\,,\,[g])=0\;\Leftrightarrow\;\lim_{n\to\infty}d(f_n,g_n)=0\;\Leftrightarrow\;f\sim g\;\Leftrightarrow\;[f]=[g]\,.$$

Finally, to see that the metric space $\left(\widetilde{M},\,\widetilde{d}\right)$ is also complete, we take a Cauchy sequence $F=\left(F_n\right)_n$ in $\widetilde{M}$, i.e. $F_n\in\widetilde{M}$. Let $f^{(n)}\in CS(M)$ be such that $F_n=\left[f^{(n)}\right]_{d_M}$. Again applying (A.0.1) we see

$$\begin{aligned}&\left|d(f_r^{(n)},\,f_u^{(m)})-d(f_s^{(n)},\,f_t^{(m)})\right|\le\\&\qquad\le d(f_r^{(n)},\,f_s^{(n)})+d(f_u^{(m)},\,f_t^{(m)})\end{aligned}\tag{A.0.6}$$

The sequence f is a Cauchy sequence if and only if $f\circ\left(\frac{1}{\cdot}-1\right)$ is uniformly continuous. The sequence f is convergent if and only if $f\circ\left(\frac{1}{\cdot}-1\right)$ has a continuous extension to 0. For the latter note that also

$$\begin{aligned}\mathbb{N}\cup\{\infty\}\;&\to\;\frac{1}{1+[\mathbb{N}]}\cup\{0\}\;,\\ n\;&\mapsto\;\frac{1}{1+n}\;,\;n\in\mathbb{N}\,,\\ \infty\;&\mapsto\;0\end{aligned}$$

is a bijective isometry.

Now, we observe that by the fact that $f^{(n)}$ is a Cauchy sequence, there is a mapping $N_1 : \mathbb{N} \longrightarrow \mathbb{R}_{>0}$ such that we have

$$d(f_r^{(n)},\, f_t^{(n)}) < \frac{1}{n+1} \tag{A.0.7}$$

for all $r, t\, > N_1(n),\ r,\, t,\, n \in \mathbb{N}$. We may assume that N_1 is strictly monotone, since otherwise we may replace N_1 by

$$n \mapsto \sup\left\{N_1\,(r)\,\middle|\, r = 0, \ldots, n\right\} + n\,.$$

From (A.0.6) and (A.0.7) we get

$$\left|d(f_r^{(n)},\, f_u^{(m)}) - d(f_s^{(n)},\, f_t^{(m)}\right| < \frac{1}{n+1} + \frac{1}{m+1} \tag{A.0.8}$$

for $r, s\, > N_1(n)$. $u, t\, > N_1(m)$, $n, m \in \mathbb{N}$.

To show completeness, we need to find a $f^{(\infty)} \in CS(M)$, such that $F \to F_\infty := \left[f^{(\infty)}\right]$. This will be done by a diagonal construction. We obtain that $g^{(n)} := f^{(n)}\,(N_1\,(n) + (\,\cdot\,))$ is also a Cauchy sequence as a subsequence of $f\,(n)$. Moreover, according to (A.0.7) we have $g^{(n)} \sim f^{(n)}$, therefore $F_n = \left[g^{(n)}\right]$, for all $n \in \mathbb{N}$, as well as (according to (A.0.8))

$$\left|d(g_r^{(n)},\, g_r^{(m)}) - d(g_t^{(n)},\, g_t^{(m)})\right| < \frac{1}{n+1} + \frac{1}{m+1} \tag{A.0.9}$$

for all $r,\, t \in \mathbb{N}$. Going in (A.0.9) to the limit with respect to t yields

$$\left|d(g_r^{(n)},\, g_r^{(m)}) - \widetilde{d}(F_n,\, F_m)\right| \leq \frac{1}{n+1} + \frac{1}{m+1} \tag{A.0.10}$$

for all $r \in \mathbb{N}$. We now define the diagonal sequence $f^{(\infty)} := \left(g_n^{(n)}\right)_{n\in\mathbb{N}}$ claiming $f^{(\infty)}$ to be a Cauchy sequence. Indeed, by (A.0.7) and (A.0.10) we have

$$\begin{aligned} d(g_n^{(n)},\, g_m^{(m)}) &\leq d(g_n^{(n)},\, g_n^{(m)}) + d(g_n^{(m)},\, g_m^{(m)}), \\ &\leq \tfrac{1}{n+1} + \tfrac{1}{m+1} + \widetilde{d}(F_n,\, F_m) + \tfrac{1}{m+1}, \end{aligned}$$

which clearly demonstrates the Cauchy sequence property of $f^{(\infty)}$. Moreover, using (A.0.7) and (A.0.10) we find

$$\begin{aligned} d(g_m^{(n)},\, f_m^{(\infty)}) &= d(g_m^{(n)},\, g_m^{(m)}) \\ &\leq d(g_m^{(n)},\, g_n^{(n)}) + d(g_n^{(n)},\, g_m^{(m)}) \\ &< 2\left(\frac{1}{n+1} + \frac{1}{m+1}\right) + \widetilde{d}(F_n,\, F_m) \end{aligned}$$

and so

$$\widetilde{d}\,(F_n, F_\infty) \leq \frac{2}{n+1} + \limsup_{m\to\infty}\, \widetilde{d}(F_n,\, F_m) \to 0$$

as $n \to \infty$, which shows the desired convergence of $(F_n)_{n\in\mathbb{N}}$ to $F_\infty = \left[f^{(\infty)}\right]$. □

This new metric space$\left(\widetilde{M},\, \widetilde{d}\right)$ is called the **completion** of $(M,\, d)$.

By the above identification we have $M \subseteq CS(M)$. The following proposition clarifies the relation between M and $\widetilde{M}$.

PROPOSITION 124. *Let $(M,\, d)$ be a semi-metric space and $\left(\widetilde{M},\, \widetilde{d}\right)$ its completion. Define an equivalence relation $\approx$ on M by*

$$x \,\approx\, y \,:\Longleftrightarrow\, d(x, y) = 0. \tag{A.0.11}$$

Then

(1) the set $M_\approx := \{x_\approx \,|\, x \in M\}$ of equivalence classes $x_\approx := \{y \in M \,|\, x \approx y\}$ becomes a metric space when equipped with the metric

$$d_\approx(x_\approx, y_\approx) := d(x, y) \tag{A.0.12}$$

and

$$\bigwedge_{x,y \in M} d_\approx(x_\approx, y_\approx) = \widetilde{d}\left(\left[f_{(x)}\right],\, \left[f_{(y)}\right]\right); \tag{A.0.13}$$

(2) the imbedding

$$M_\approx \hookrightarrow \widetilde{M}$$

given by $x_\approx \mapsto \left[f_{(x)}\right]$ for $x \in M$ is an isometry and hence is uniformly continuous ;

(3) the imbedding is also dense and so we may consider $M_\approx$ as a dense subset of $\widetilde{M}$.

Proof. That $\approx$ defines an equivalence relation on M follows easily from the properties of a semi-metric. To see that $d_\approx$ is well-defined, we only need to recall (A.0.1):

$$|d(x, y) - d(u, v)| \leq d(x, u) + d(y, v) \text{ for all } x,\, y,\, u,\, v \in M,$$

so that, if $x \approx u$ and $y \approx v$, it follows easily that $d(x, y) = d(u, v)$ and hence

$$d_\approx(x_\approx, y_\approx) = d(x, y) = d(u, v) = d_\approx(u_\approx, v_\approx)$$

which shows that $d_\approx$ is a well-defined mapping into the real numbers. Clearly, the semi-metric properties are inherited from d. Definiteness follows easily from (A.0.12) and (A.0.11). Realizing that since $f,\, g \in M$ are identified with constant sequences, we must have

$$d(f, g) = \widetilde{d}([f]\,,\,[g]) \text{ for all } f,\, g \in M,$$

yields (A.0.13). Statement 2 is just rephrasing (A.0.13). To prove statement 3 we show that for any $[f] \in \widetilde{M}$ we have that the constant sequences $f^{(k)} := (f_k)_{n\in\mathbb{N}}$, $k \in \mathbb{N}$, satisfy

$$\left[f^{(k)}\right] \to [f] \text{ in } \widetilde{M} \text{ as } k \to \infty.$$

For this we need to show that $\limsup_{n\to\infty} d(f_k, f_n)$ goes to zero as $k \to \infty$. This, however, follows from the Cauchy convergence of f. $\square$

Remark 125. Since $(M,\, d)$ is assumed to be merely a semi-metric space the canonical mapping $x \longmapsto [x]$, relating x with its equivalence class $[x]$ is a semi-isometry.

Example 126. Let $(M,\, d)$ be a metric space. Then we define a new metric d_1 given by

$$(x, y) \mapsto \sup\left\{ \frac{|d(x, a) - d(y, a)|}{(1 + d(x, a))(1 + d(y, a))} \,\middle|\, a \in M \right\}.$$

To see that d_1 is indeed a metric we first note that only the triangle inequality is in doubt. The triangle inequality in $\mathbb{R}$ yields

$$\left|\frac{1}{1+x} - \frac{1}{1+y}\right| \leq \left|\frac{1}{1+x} - \frac{1}{1+z}\right| + \left|\frac{1}{1+z} - \frac{1}{1+y}\right| \tag{A.0.14}$$

for $x, y, z \in \mathbb{R}_{\geq 0}$. Replacing in (A.0.14) x, y, z by $d(x, a)$, $d(y, a)$, $d(z, a)$, respectively, we get

$$\frac{|d(x, a) - d(y, a)|}{(1 + d(x, a))(1 + d(y, a))} \leq \frac{|d(x, a) - d(z, a)|}{(1 + d(x, a))(1 + d(z, a))} + \frac{|d(z, a) - d(y, a)|}{(1 + d(z, a))(1 + d(y, a))}$$

for all $x, y, z, a \in M$. Taking the supremum with respect to $a \in M$ (first on the right-hand side then on the left-hand side), we see that d_1 satisfies the triangle inequality. Moreover,

$$\begin{aligned} d_1(x, y) &= \sup\left\{ \tfrac{|d(x,a)-d(y,a)|}{(1+d(x,a))(1+d(y,a))} \,\middle|\, a \in M \right\}, \\ &\leq \sup\left\{ \tfrac{d(x,y)}{(1+d(x,a))(1+d(y,a))} \,\middle|\, a \in M \right\}, \\ &\leq d(x, y) \end{aligned}$$

for all $x, y \in M$. Thus, a Cauchy sequence with respect to d is also a Cauchy sequence with respect to d_1 . Conversely, a Cauchy sequence with respect to d_1, which is *bounded* with respect to d is also a Cauchy sequence with respect to d. Indeed,

$$d_1(x,y) \geq \frac{|d(x,a) - d(y,a)|}{(1 + d(x,a))(1 + d(y,a))}$$

for all $x, y, a \in M$. Choosing $a = y$ we get

$$d_1(x,y) \geq \frac{d(x,y)}{(1 + d(x,y))} \geq \frac{1}{2} d(x,y)$$

for all $x, y \in M$, with $d(x,y) \leq 1$. This proves the above assertion.

There may be, however, Cauchy sequences with respect to d_1, which are *unbounded* with respect to the metric d. In this case, there is no limit in (M, d) to which such a sequence might converge, even if (M, d) is assumed to be complete. If $(M,\, d)$ is complete, then the completion of $(M,\, d_1)$ contains one additional element, the equivalence class of unbounded sequences[3], which we shall denote by ∞.

We note that the process of completion is compatible with (locally) uniform continuity of mappings. Next, we are turning our attention to a surprisingly elementary but incredibly useful result.

THEOREM 127. ***(Contraction mapping theorem)*** *Let $(M,\, d_M)$ be a complete metric space and $F : M \longrightarrow M$ a Lipschitz continuous with best Lipschitz constant $|F|_{Lip} < 1$, i.e. F is a* ***contraction*** *in M. Then F has a unique fixed point $\widehat{x}$, i.e. a unique element $\widehat{x} \in M$ with $F(\widehat{x}) = \widehat{x}$. Moreover, we have for any $x_0 \in M$ that $F^n(x_0) \to \widehat{x}$ as $n \to \infty$ and the following error estimates hold:*

$$\bigwedge_{n \in \mathbb{N}.} d_M(F^n(x_0), \widehat{x}) \leq |F|^n_{Lip}\; d_M(x_0, \widehat{x}), \tag{A.0.15}$$

and

$$\bigwedge_{n \in \mathbb{N}.} d_M(F^n(x_0), \widehat{x}) \leq \frac{|F|^n_{Lip}}{1 - |F|_{Lip}} d_M(F(x_0), x_0). \tag{A.0.16}$$

PROOF. By assumption we have

$$d_M(F(x), F(y)) \leq |F|_{Lip}\, d_M(x,y) \tag{A.0.17}$$

for all $x,\, y \in M$ and by taking $y = \widehat{x}$ to be a fixed point of F we get

$$d_M(F(x), \widehat{x}) \leq |F|_{Lip}\, d_M(x, \widehat{x})$$

from which (A.0.15) follows by induction. If also x is a fixed point, we obtain

$$d_M(x, \widehat{x}) \leq |F|_{Lip}\, d_M(x, \widehat{x})$$

which implies $d_M(x, \widehat{x}) = 0$ and so indeed $x = \widehat{x}$, since $|F|_{Lip} < 1$. Thus, a fixed point is uniquely determined.

The rest of the theorem follows by comparison with the geometric series. We first show that $(F^n(x_0))_{n \in \mathbb{N}}$ is a Cauchy sequence. By induction we find

$$d_M(F^{n+1}(x_0), F^n(x_0)) \leq |F|^n_{Lip}\; d_M(F(x_0), x_0)$$

for all $n \in \mathbb{N}$ and then for $m \in \mathbb{N}$, $m \geq n$,

$$\begin{aligned} d_M(F^n(x_0), F^{m+1}(x_0)) &\leq \textstyle\sum_{k=n}^m d_M(F^k(x_0), F^{k+1}(x_0)), \\ &\leq \textstyle\sum_{k=n}^m |F|^k_{Lip}\; d_M(F(x_0), x_0). \end{aligned} \tag{A.0.18}$$

[3]If $M = \pm\mathbb{R}_{\geq 0}$ and $d := |\cdot - \cdot|$ and then this equivalence class is usually denoted by $\pm\infty$, respectively.

Since $\sum_{k=0}^{\infty} |F|_{Lip}^{k}$ converges as a geometric series, we have as desired that $(F^n(x_0))_{n\in\mathbb{N}}$ is a Cauchy sequence. By assumption $(M,\, d_M)$ is complete and so $\widehat{z} := \lim_{n\to\infty} F^n(x_0)$ exists in M. Letting $m \to \infty$ in (A.0.18) we thus obtain

$$\begin{aligned} d_M(F^n(x_0), \widehat{z}) &\leq \textstyle\sum_{k=n}^{\infty} |F|_{Lip}^{k} \; d_M(F(x_0), x_0), \\ &= |F|_{Lip}^{n} \; \textstyle\sum_{k=0}^{\infty} |F|_{Lip}^{k} \; d_M(F(x_0), x_0), \\ &= \tfrac{|F|_{Lip}^{n}}{1-|F|_{Lip}} d_M(F(x_0), x_0), \end{aligned} \tag{A.0.19}$$

which is (by renaming) error estimate (A.0.16). It remains to show that $\widehat{z}$ is actually a fixed point of F. This follows from the continuity of F

$$F(\widehat{z}) = F(\lim_{n\to\infty} F^n(x_0)) = \lim_{n\to\infty} F^{n+1}(x_0) = \widehat{z}.$$

□

A fundamental concept in analysis that is particularly simple to characterize in metric spaces is compactness.

Definition 128. A subset C of a metric space is called (sequentially) compact if it satisfies the Bolzano-Weierstrass property, i.e. for every sequence $(x_s)_{s\in\mathbb{N}}$ in C there is a convergent subsequence with limit in C.

Finally, we recall that compact sets are also closed and bounded. (we leave the proof of this implication as an exercise).

APPENDIX B

Polar Decomposition

Definition 1. Let A, $B \subseteq H \oplus H$ be Hermitean operators in complex Hilbert space H. We define

$$A \leq B :\Leftrightarrow \bigwedge_{x \in D(A) \cap D(B)} \langle x|Ax\rangle_H \leq \langle x|Bx\rangle_H .$$

If $A \geq 0$ then we call A **non-negative**. If $A \geq 0$ and $Ax = 0 \Rightarrow x = 0$ for all $x \in H$ then A is called **positive (definite)**. If $A \geq \varepsilon$ for some $\varepsilon \in \mathbb{R}_{>0}$, then A is called **strictly positive (definite)**.

For later use we consider higher powers of selfadjoint operators.

Proposition 2. *Let $A : D(A) \subseteq H \to H$ a strictly positive selfadjoint operator in H. Then A^j is selfadjoint for all $j \in \mathbb{N}$.*

Proof. The result is obviously true for $j = 0$ and by assumption for $j = 1$. Consider now the obviously Hermitean operator A^{j+1}. Since by induction hypothesis $D(A^j)$ must be dense in H, we also have $A^{-1}D(A^j)$ dense in $A^{-1}H = D(A)$. Since $D(A)$ is also dense in H, we have that $A^{-1}D(A^j)$ must be dense in H. But $A^{-1}D(A^j) = D(A^{j+1})$, thus A^{j+1} is densely defined and therefore symmetric. Moreover, we have $0 \in \varrho(A^{j+1})$, since $A^{-j-1} = (A^{-1})^{(j+1)}$ features the corresponding resolvent. Thus, A^{j+1} must be selfadjoint. $\square$

Proposition 3. *Let $A \subseteq H \oplus H$ and $B \subseteq H \oplus H$ be two commuting selfadjoint operators in complex Hilbert space H. Then $A \geq 0$ and $B \geq 0$ implies $AB \geq 0$ and $BA \geq 0$.*

Proof. If $C \in L(H, H)$ is selfadjoint, then $C^2 \geq 0$ for

$$\left\langle x \,\middle|\, C^2 x \right\rangle_H = \langle Cx \,|Cx\rangle \geq 0 \text{ for all } x \in H.$$

For $A = 0$ the result is trivial, therefore we assume $A \neq 0$. In a first step, we shall assume $A \in L(H, H)$ and try to express A as a sum of squares. Defining recursively

$$\begin{aligned} A_0 &:= ||A||^{-1} A, \\ A_{n+1} &:= A_n - A_n^2, \end{aligned}$$

we obtain a sequence $(A_n)_{n \in \mathbb{N}}$ in $L(H, H)$. We claim

$$0 \leq A_n \leq 1 \tag{B.20}$$

for all $n \in \mathbb{N}$. The proof of this claim will be by induction.

Clearly, $A_0 \geq 0$. Moreover,

$$\begin{aligned} \langle x|x\rangle_H - \langle x|A_0 x\rangle_H &= \langle x|x\rangle_H - ||A||^{-1} \langle x|Ax\rangle_H , \\ &\geq \langle x|x\rangle_H - ||A||^{-1} ||A|| \, |x|_H^2 = 0 \end{aligned}$$

and so also

$$A_0 \leq 1.$$

Assume now (B.20) holds for some $n \in \mathbb{N}$. Consider

$$\begin{aligned} \left\langle x \,\middle|\, A_n^2 (1 - A_n) x \right\rangle_H &= \langle x|A_n(1 - A_n)A_n x\rangle_H , \\ &= \langle A_n x \,|\, (1 - A_n)A_n x\rangle_H . \end{aligned}$$

The latter term is non-negative by the induction assumption. Therefore

$$A_n^2(1 - A_n) \geq 0$$

and by exchanging the role of A_n and $(1 - A_n)$ we also have

$$A_n(1 - A_n)^2 \geq 0.$$

As the sum of two non-negative selfadjoint operators in $L(H, H)$ must be also non-negative, we have

$$A_{n+1} = A_n(1 - A_n)^2 + A_n^2(1 - A_n) \geq 0.$$

Moreover, since $(1 - A_n)$ and A_n^2 are non-negative, we also have

$$1 - A_{n+1} = 1 - A_n + A_n^2 \geq 0$$

and so (B.20) is shown to be true for all $n \in \mathbb{N}$. A simple induction now yields that

$$\sum_{k=0}^{n} A_k^2 = A_0 - A_{n+1} \leq A_0. \tag{B.21}$$

Indeed, for $n = 0$ we have $A_0^2 = A_0 - A_1$. Moreover,

$$\sum_{k=0}^{n+1} A_k^2 = \sum_{k=0}^{n} A_k^2 + A_{n+1}^2 = A_0 - A_{n+1} + A_{n+1}^2 = A_0 - A_{n+2} \leq A_0$$

so that (B.21) follows for all $n \in \mathbb{N}$. From (B.21) we see that

$$\sum_{k=0}^{n} \langle A_k x \,|A_k x\rangle_H = \sum_{k=0}^{n} \left\langle x \,\middle|A_k^2 x\right\rangle_H = \left\langle x \,\middle|\sum_{k=0}^{n} A_k^2\, x\right\rangle_H \leq \langle x\,|A_0 x\rangle_H$$

for every $n \in \mathbb{N}$. From this we conclude that

$$\sum_{k=0}^{\infty} |A_k x|_H^2 \leq \|A_0\|\ |x|_H^2$$

and so in particular

$$A_n x \to 0 \text{ as } n \to \infty.$$

This finally shows that as desired

$$\sum_{k=0}^{n} A_k^2\, x = A_0 x - A_{n+1} x \to A_0 x \text{ as } n \to \infty$$

or

$$\sum_{k=0}^{\infty} A_k^2\, x = A_0 x \text{ for all } x \in H.$$

From this representation of A_0 in terms of squares the stated result now follows. Since B commutes with A, it also commutes with any polynomial of A and so with each A_n, $n \in \mathbb{N}$. Since $B \geq 0$ we have for $k = 0, \ldots, n$, $n \in \mathbb{N}$,

$$\begin{aligned} \left\langle x \,\middle|\textstyle\sum_{k=0}^{n} A_k^2 B x\right\rangle_H &= \left\langle x \,\middle|B \textstyle\sum_{k=0}^{n} A_k^2 x\right\rangle_H \\ &= \textstyle\sum_{k=0}^{n} \left\langle x \,\middle|B A_k^2 x\right\rangle_H \\ &= \textstyle\sum_{k=0}^{n} \langle x\,|A_k B A_k x\rangle_H \\ &= \textstyle\sum_{k=0}^{n} \langle A_k x\,|B A_k x\rangle_H \\ &\geq \langle A_n x\,|B A_n x\rangle_H \geq 0 \end{aligned} \tag{B.22}$$

for all $x \in D(B)$. By invoking the closedness of B and letting $n \to \infty$ in (B.22) we get

$$\langle x\,|ABx\rangle_H = \langle x\,|BAx\rangle_H = ||A||\ \langle x\,|BA_0 x\rangle_H \geq 0$$

for all $x \in D(B)$. Now let $A \subseteq H \oplus H$ be a possibly unbounded, non-negative, selfadjoint operator and $\varepsilon \in \mathbb{R}_{>0}$. Then $A + \varepsilon \geq \varepsilon$ and so $(A + \varepsilon)^{-1} \in L(H, H)$, $\left\|(A + \varepsilon)^{-1}\right\| \leq \varepsilon^{-1}$ and $(A + \varepsilon)^{-1} \geq 0$.

By assumption B commutes with A, i.e. with resolvents of A, e.g. $(A+\varepsilon)^{-1}$. Applying our previous findings we have

$$\left\langle x \,\middle|\, B(A+\varepsilon)^{-1}x \right\rangle_H = \left\langle x \,\middle|\, (A+\varepsilon)^{-1}Bx \right\rangle_H \geq 0 \text{ for all } x \in D(B).$$

With $y = (A+\varepsilon)^{-1}x$ and observing that $By = (A+\varepsilon)^{-1}Bx \in D(A)$ this implies

$$\langle y \,|(A+\varepsilon)By \rangle_H = \langle y \,|B(A+\varepsilon)y \rangle_H \geq 0 \text{ for all } y \in (A+\varepsilon)^{-1}D(B) \subseteq D(A)\cap D(B).$$

We note that every $y \in D(A)$ with $(A+\varepsilon)y \in D(B)$ can be written as

$$y = (A+\varepsilon)^{-1}(A+\varepsilon)y$$

and so

$$\langle y \,|(A+\varepsilon)By \rangle_H = \langle y \,|B(A+\varepsilon)y \rangle_H \geq 0 \text{ for all } y \in D(A) \text{ with } (A+\varepsilon)y \in D(B).$$

Further specializing

$$\langle y \,|(A+\varepsilon)By \rangle_H = \langle y \,|B(A+\varepsilon)y \rangle_H \geq 0 \text{ for all } y \in D(A)\cap D(B) \text{ with } Ay \in D(B) \text{ and } By \in D(A).$$

Now letting $\varepsilon \to 0+$ we get

$$\langle y \,|ABy \rangle_H = \langle y \,|BAy \rangle_H \geq 0 \text{ for all } y \in D(A)\cap D(B) \text{ with } Ay \in D(B) \text{ and } By \in D(A).$$

Since both cases are analogous, let us focus on AB. In order to show

$$AB \geq 0$$

we need to establish that for every $x \in D(B)$ with $Bx \in D(A)$ we can find an approximation $y \in D(A)\cap D(B)$ with $Ay \in D(B)$ and $By \in D(A)$ to every degree of accuracy. Since B commutes with A we have for $\eta \in \mathbb{R}_{>0}$

$$y_\eta := (\eta A+1)^{-1}x \in D(A)\cap D(B)$$

as well as

$$Ay_\eta := A(\eta A+1)^{-1}x = \eta^{-1}(x-y_\eta) \in D(B)$$

and

$$By_\eta := (\eta A+1)^{-1}Bx \in D(A).$$

Letting $\eta \to 0+$ we find

$$(\eta A+1)^{-1}z - z = -\eta\,(\eta A+1)^{-1}Az \to 0 \text{ as } \eta \to 0+$$

for all $z \in D(A)$ and since $\left\|(\eta A+1)^{-1}\right\| \leq 1$ and $D(A)$ dense in H,

$$(\eta A+1)^{-1}u - u \to 0 \text{ as } \eta \to 0+$$

for all $u \in H$. Therefore,

$$y_\eta - x = (\eta A+1)^{-1}x - x \to 0 \text{ as } \eta \to 0+$$

and

$$By_\eta - Bx = (\eta A+1)^{-1}Bx - Bx \to 0 \text{ as } \eta \to 0+\,.$$

This, however, was our claim. □

Now we shall employ our findings in Proposition 3 to obtain a monotone convergence result for selfadjoint operators.

THEOREM 4. *Let $(A_n)_{n\in\mathbb{N}}$ be a sequence in $L(H,H)$, H a complex Hilbert space, of commuting selfadjoint operators and further suppose that $(A_n)_{n\in\mathbb{N}}$ is non-decreasing, i.e.*

$$\bigwedge_{n\in\mathbb{N}} A_n \leq A_{n+1}$$

and bounded above by a selfadjoint operator $B \in L(H,H)$ commuting with all A_n, $n \in \mathbb{N}$, in the sense that

$$\bigwedge_{n\in\mathbb{N}} A_n \leq B.$$

Then $A_\infty \in L(H,H)$ defined by

$$A_\infty x := \lim_{n\to\infty} A_n x \text{ for all } x \in H,$$

is a selfadjoint operator and commutes with B and all A_n, $n \in \mathbb{N}$. Moreover, we have

$$A_\infty \leq B.$$

Proof. That A_∞ is a well-defined bounded, linear operator is clear from Proposition 69 as soon as it can be shown that $\lim_{n\to\infty} A_n x$ exists for all $x \in H$. Consider the sequence $(C_n)_n := (B - A_n)_n$. Clearly, we have

$$C_n \geq C_{n+1} \geq 0$$

and all C_n commute with each other, $n \in \mathbb{N}$. Therefore, we have for $n > m$ by the previous proposition that

$$C_n(C_m - C_n) \geq 0 \text{ and } (C_m - C_n)C_m \geq 0$$

or

$$C_m^2 \geq C_n\, C_m \geq C_n^2. \tag{B.23}$$

This shows that $\left(\left\langle x \,\middle|\, C_n^2 x \right\rangle_H\right)_n$ is a non-increasing sequence of non-negative numbers and therefore convergent:

$$c(x) := \lim_{n\to\infty} \left\langle x \,\middle|\, C_n^2 x \right\rangle_H .$$

Moreover, with (B.23) we also have

$$\langle x \,|\, C_n\, C_m x \rangle_H \to c(x) \text{ as } m, n \to \infty.$$

We find

$$\begin{aligned}
|C_n x - C_m x|_H^2 &= \langle (C_n - C_m)x \,|\, (C_n - C_m)x \rangle_H , \\
&= \left\langle x \,\middle|\, (C_n - C_m)^2 x \right\rangle_H , \\
&= \left\langle x \,\middle|\, C_n^2 x \right\rangle_H - \langle x \,|\, C_n\, C_m x \rangle_H - \langle x \,|\, C_m\, C_n x \rangle_H + \left\langle x \,\middle|\, C_m^2 x \right\rangle_H , \\
&= \left\langle x \,\middle|\, C_n^2 x \right\rangle_H - 2\, \langle x \,|\, C_n\, C_m x \rangle_H + \left\langle x \,\middle|\, C_m^2 x \right\rangle_H \\
&\to c(x) - 2\, c(x) + c(x) = 0 \text{ as } n, m \to \infty.
\end{aligned}$$

By the completeness of H we have that

$$\lim_{n\to\infty} A_n x = B - \lim_{n\to\infty} C_n x$$

exists for all $x \in H$. Thus, A_∞ is a well-defined, linear operator in $L(H,H)$. It remains to be seen that A_∞ is selfadjoint and bounded above by B. For selfadjointness it suffices to show that A_∞ is Hermitean. But this is obvious, since

$$\begin{aligned}
\langle y \,|\, A_\infty x \rangle_H = \langle y \,|\, \lim_{n\to\infty} A_n x \rangle_H = \lim_{n\to\infty} \langle y \,|\, A_n x \rangle_H = \\
= \lim_{n\to\infty} \langle A_n y \,|\, x \rangle_H = \langle \lim_{n\to\infty} A_n y \,|\, x \rangle_H = \langle A_\infty y \,|\, x \rangle_H
\end{aligned}$$

for all $x, y \in H$. From

$$A_n B x = B A_n x$$

follows by letting $n \to \infty$ that

$$A_\infty B x = B A_\infty x$$

for all $x \in H$. Finally, we have from $A_n \leq B$ for all $n \in \mathbb{N}$ that

$$\langle x \,|\, A_n x \rangle_H \leq \langle x \,|\, B x \rangle_H$$

and again letting $n \to \infty$

$$\langle x \,|\, A_\infty x \rangle_H \leq \langle x \,|\, B x \rangle_H$$

for all $x \in H$, i.e.

$$A_\infty \leq B.$$

□

There is an interesting mechanism which obtains a selfadjoint operator from every densely defined, closable operator.

PROPOSITION 5. *Let $A \subseteq H_0 \oplus H_1$, be a densely defined, closable, linear operator between complex Hilbert spaces H_0, H_1. Then*

$$A^* \bar{A}$$

is selfadjoint. We have

$$\sigma(A^* \bar{A}) \subseteq [0, \infty[$$

and $D(A^*\bar{A})$ is dense in $D(\bar{A})$ with respect to the graph norm of $\bar{A}$.

PROOF. First, we see that $A^*\bar{A}$ is Hermitean

$$\langle x \,|\, A^*\bar{A}x \rangle_H = \langle \bar{A}x \,|\, \bar{A}x \rangle_H = \langle A^*\bar{A}x \,|\, x \rangle_H \text{ for all } x \in D(A^*\bar{A}).$$

Moreover,

$$\begin{aligned} w(A^*\bar{A}) &= \left\{ \langle x \,|\, A^*\bar{A}x \rangle_H \,\middle|\, x \in D(A^*\bar{A}) \wedge |x|_H = 1 \right\}, \\ &= \left\{ \langle \bar{A}x \,|\, \bar{A}x \rangle_H \,\middle|\, x \in D(A^*\bar{A}) \wedge |x|_H = 1 \right\}, \\ &\subseteq [0, \infty[. \end{aligned}$$

Next, we notice that $A^*\bar{A} + 1$ is onto. Since A is closable, the domain $D(\bar{A})$ of its closure $\bar{A}$ is a complex Hilbert space with respect to the graph norm. Solving the equation

$$\left(A^*\bar{A} + 1\right) u = f \tag{B.24}$$

is equivalent to finding $u \in D(\bar{A})$ such that

$$\bigwedge_{x \in D(\bar{A})} \langle \bar{A}u \,|\, \bar{A}x \rangle_H + \langle u \,|\, x \rangle_H = \langle f \,|\, x \rangle_H . \tag{B.25}$$

Indeed, (B.24) implies (B.25) after multiplying (B.24) by x in the sense of the inner product of H. But also conversely, if u satisfies (B.25) then we read off that $\bar{A}u$ must be in $D(A^*)$ and that

$$A^*\bar{A}u = f - u.$$

A solution $u \in D(\bar{A})$ of (B.25), however, is easily found by noting that $\langle f \,|\cdot \rangle_H$ is a bounded, linear functional on $D(\bar{A})$ with respect to the graph norm $|\cdot|_{D(\bar{A})}$ of $\bar{A}$:

$$|\langle f \,|x \rangle_H| \le |f|_H \, |x|_H \le |f|_H \sqrt{|x|_H^2 + |\bar{A}x|_H^2} = |f|_H \; |x|_{D(\bar{A})} \text{ for all } x \in D(\bar{A}). \tag{B.26}$$

An apparently unique solution is provided by

$$u := (A^*\bar{A} + 1)^{-1} f = R_{D(\bar{A})} \left(\langle f \,|\cdot \rangle_H \right),$$

where $R_{D(\bar{A})}$ is the associated Riesz mapping. Moreover, for this solution $u = (A^*\bar{A} + 1)^{-1} f$ we find with (B.26) and (B.25) the estimate

$$\left|(A^*\bar{A} + 1)^{-1} f\right|_{D(\bar{A})} \le |f|_H \text{ for all } f \in H.$$

Thus, we have $-1 \in \varrho(A^*\bar{A})$. We also have that $A^*\bar{A}$ is densely defined. Let $x \in D(A^*\bar{A})^\perp$, where the ortho-complement is taken in $D(\bar{A})$, then

$$\bigwedge_{y \in D(A^*\bar{A})} \langle \bar{A}y \,|\, \bar{A}x \rangle_H + \langle y \,|\, x \rangle_H = \langle A^*\bar{A}y \,|\, x \rangle_H + \langle y \,|\, x \rangle_H = 0.$$

Since we found that $A^*\bar{A} + 1$ is onto, we get

$$\bigwedge_{y \in H} \langle y \,|\, x \rangle_H = 0$$

and so that $x = 0$. Thus, $D(A^*\bar{A})$ is dense in $D(\bar{A})$ and by assumption $D(\bar{A})$ dense in H, therefore $A^*\bar{A}$ densely defined:

$$|y - f|_H \le |y - x|_H + |x - f|_H \le |y - x|_{D(\bar{A})} + |x - f|_H \text{ for all } f \in H, \; x \in D(\bar{A}), \; y \in D(A^*\bar{A}).$$

Thus, we have found that $A^*\bar{A}$ is symmetric, $w(A^*\bar{A}) \subseteq [0,\infty[$ and $-1 \in \varrho(A^*\bar{A})$. Thus, we have that $A^*\bar{A}$ is selfadjoint and

$$\sigma(A^*\bar{A}) \subseteq [0,\infty[.$$

$\square$

EXAMPLE 6. We have according to their definition in example 87 (and using the notations from there) that $-\overset{\circ}{\operatorname{div}}\operatorname{grad}$ and $-\operatorname{div}\overset{\circ}{\operatorname{grad}}$ are of the form $A^*\bar{A}$ and therefore selfadjoint (in $L_2(\Omega)$) with

$$\sigma(-\operatorname{div}\overset{\circ}{\operatorname{grad}}) \subseteq [0,\infty[,\ \sigma(-\overset{\circ}{\operatorname{div}}\operatorname{grad}) \subseteq [0,\infty[.$$

Rephrasing this, we have solved the boundary value problems:

(1) $-\operatorname{div}\operatorname{grad} u - \lambda u = f \in L_2(\Omega)$, $u \in H(\overset{\circ}{\operatorname{grad}})$ (Dirichlet boundary value problem with homogeneous Dirichlet type boundary condition),
(2) $-\operatorname{div}\operatorname{grad} u - \lambda u = f \in L_2(\Omega)$, $\operatorname{grad} u \in H(\overset{\circ}{\operatorname{div}})$ (Neumann boundary value problem with homogeneous Neumann type boundary condition)

for all $\lambda \in \varrho(A)$, in particular for $\lambda \in \mathbb{C} \setminus [0,\infty[$.

The non-negativity of the operator $A^*\bar{A}$ allows for defining a non-negative root. For this result we make use of the convergence result of Theorem 4. First we treat the case for a bounded, non-negative selfadjoint operator.

LEMMA 7. *Let $A \in L(H,H)$ be selfadjoint in complex Hilbert space H and non-negative. Then there exists a non-negative selfadjoint $B \in L(H,H)$ such that*

$$B^2 = A.$$

This operator B commutes with all $C \in L(H,H)$, which commute with A.

PROOF. If $A = 0$ we take $B = 0$. Let now $A \neq 0$. Since A is bounded, we may consider w.l.o.g. $|A| \leq 1$ otherwise replace A by $|A|^{-1} A$. If B is a root of the latter then $|A|^{1/2} B$ is a root of the original A. So let $|A| \leq 1$ and define recursively

$$\begin{aligned} B_0 &:= 0, \\ B_{n+1} &:= B_n + \tfrac{1}{2}(A - B_n^2) \end{aligned} \tag{B.27}$$

for $n \in \mathbb{N}$. By an elementary induction we see that B_n is a real polynomial in A and therefore $B_n \in L(H,H)$ is selfadjoint and commutes with every $C \in L(H,H)$ which commutes with A, $n \in \mathbb{N}$. In particular, all the B_n are commuting with each other, $n \in \mathbb{N}$. In order to apply Theorem 4 we show now that $(B_n)_n$ is a non-decreasing sequence and bounded above by 1. To demonstrate that

$$B_n \leq 1 \tag{B.28}$$

for all $n \in \mathbb{N}$ we procede by induction. Clearly, (B.28) is true for $n = 0$. Now let (B.28) hold for a particular $n \in \mathbb{N}$. The desired induction step follows by observing that

$$1 - B_{n+1} = 1 - B_n - \frac{1}{2}(A - B_n^2) = \frac{1}{2}(1 - B_n)^2 + \frac{1}{2}(1 - A) \geq 0.$$

To demonstrate that $(B_n)_n$ is non-decreasing we first notice that trivially

$$0 = B_0 \leq B_1 = \frac{1}{2}A.$$

Now let

$$B_{n+1} \geq B_n \tag{B.29}$$

for a particular $n \in \mathbb{N}$. Then

$$B_{n+2} - B_{n+1} = B_{n+1} - B_n - \frac{1}{2}(B_{n+1}^2 - B_n^2) = \frac{1}{2}((1 - B_{n+1}) + (1 - B_n))\,(B_{n+1} - B_n) \geq 0$$

as a product of non-negative operators in $L(H,H)$. This confirms the claimed monotonicity property (B.29) for all $n \in \mathbb{N}$. According to Theorem 4 we have B as the strong limit of the sequence $(B_n)_n$. Since

$$\langle x\,|x\rangle_H \geq \langle x\,|B_n x\rangle_H \geq 0$$

for all $x \in H$, we get (by taking the limit)

$$\langle x\,|x\rangle_H \geq \langle x\,|Bx\rangle_H \geq 0,$$

i.e. $1 \geq B \geq 0$. That B commutes with every $C \in L(H,H)$ which commutes with A carries over from the sequence elements to the limit by continuity. That B is the desired square root follows by noting that letting $n \to \infty$ in (B.27) yields

$$B = B + \frac{1}{2}(A - B^2)$$

from which

$$B^2 = A$$

follows. □

We shall now proceed to construct the square root of a strictly positive, unbounded selfadjoint operator.

Lemma 8. *Let $A \subseteq H \oplus H$ be a selfadjoint, strictly positive operator in complex Hilbert space H. Then there is a selfadjoint, strictly positive $B \subseteq H \oplus H$ such that*

$$B^2 = A.$$

Moreover, $D(A)$ is dense in $D(B)$ considered as a complex Hilbert space with respect to the graph norm of B. If for $\varepsilon \in \mathbb{R}_{>0}$ we have that ε^2 is a lower bound of A, then B has ε as lower bound. The operator B commutes with every $C \in L(H,H)$, which commutes with A.

Proof. Since A is strictly positive, there is a constant $\varepsilon \in \mathbb{R}_{>0}$ such that $A \geq \varepsilon^2$, i.e.

$$\langle y\,|Ay\rangle_H \geq \left\langle y\,\middle|\varepsilon^2\, y\right\rangle_H = \varepsilon^2\,\langle y\,|y\rangle_H = \varepsilon^2\,|y|_H^2 \tag{B.30}$$

for all $y \in D(A)$. Since A is selfadjoint, we have therefore

$$\sigma(A) \subseteq \left[\varepsilon^2, \infty\right[.$$

In particular, $A^{-1} \in L(H,H)$. By the previous lemma we have a square root, which in anticipation of a later uniqueness result we denote by $\sqrt{A^{-1}}$, such that $\sqrt{A^{-1}} \geq 0$ and

$$\sqrt{A^{-1}}\sqrt{A^{-1}} = A^{-1}.$$

From $\sqrt{A^{-1}}x = 0$ we get $\sqrt{A^{-1}}\sqrt{A^{-1}}x = A^{-1}x = 0$ and so (by applying A) $x = 0$. Thus, we have a well-defined linear operator $B := \sqrt{A^{-1}}^{-1}$ with domain $\sqrt{A^{-1}}H$. Next we would like to show that

$$B^2 = A.$$

We have for $x \in D(A)$

$$\sqrt{A^{-1}}\sqrt{A^{-1}}Ax = x \in D(B) = \sqrt{A^{-1}}H$$

and

$$Bx = \sqrt{A^{-1}}^{-1}x = \sqrt{A^{-1}}Ax.$$

Thus, $Bx \in D(B)$and $B^2x = Ax$. This shows

$$A \subseteq B^2.$$

Conversely, let $x \in D(B^2)$ and $y := B^2x$ then $A^{-1}y = \sqrt{A^{-1}}^2\,B^2x = \sqrt{A^{-1}}\,Bx = x \in D(A)$ and so $B^2x = y = Ax$, i.e.

$$B^2 \subseteq A.$$

Since $D(A) = D(B^2) \subseteq D(B) \subseteq H$, we see that B is densely defined. We know that B^{-1} is selfadjoint and bounded, in particular

$$B^{-1} = (B^{-1})^* = (B^*)^{-1}.$$

Consequently, we also get

$$B = B^*.$$

By construction $B^{-1} \geq 0$ and so with $x = By$ for arbitrary $y \in D(B)$

$$\langle x \,|B^{-1}x \rangle_H = \langle By\,|y\rangle_H \geq 0.$$

Thus, we found $B \geq 0$. Since $A = B^*B = B^2$, we have $D(A)$ dense in $D(B)$ by Proposition 5. We have by estimate (B.30) that $A - \varepsilon^2 = B^2 - \varepsilon^2 = (B - \varepsilon)(B + \varepsilon) \geq 0$. We want to show that $B - \varepsilon \geq 0$. For this we notice $B + \varepsilon \geq \varepsilon > 0$ and therefore $(B + \varepsilon)^{-1} \in L(H, H)$. Moreover, $\left\|(B + \varepsilon)^{-1}\right\| \leq \varepsilon^{-1}$ and $(B + \varepsilon)^{-1} \geq 0$. Thus, we obtain a bounded, selfadjoint, non-negative operator $(B + \varepsilon)^{-1}$ commuting with $A - \varepsilon^2$. By Proposition 3 we have

$$(B - \varepsilon) \supseteq (B + \varepsilon)^{-1}(A - \varepsilon^2) \geq 0$$

and by the density of $D(A)$ in $D(B)$ we conclude

$$B \geq \varepsilon.$$

Lastly, we show the commutativity property. The square root B^{-1} of A^{-1} commutes with every $C \in L(H, H)$, which commutes with A^{-1}. However,

$$C\,A \subseteq A\,C$$

implies

$$C\,A^{-1} = A^{-1}C,$$

and so

$$C\,B^{-1} = B^{-1}C,$$

which in turn implies as desired

$$C\,B \subseteq BC.$$

□

Finally we want to remove the constraint of *strict* positivity.

THEOREM 9. *Let $A \subseteq H \oplus H$ be a selfadjoint, non-negative operator in complex Hilbert space H. Then there is a unique, selfadjoint, non-negative $B \subseteq H \oplus H$ such that B commutes with every $C \in L(H, H)$, which commutes with A and*

$$B^2 = A.$$

Moreover, $D(A)$ is dense in $D(B)$ considered as a complex Hilbert space with respect to the graph norm of B.

REMARK 10. The uniqueness of B motivates us to speak of *the* square root of A and denote it by $A^{1/2}$ or $\sqrt{A}$.

PROOF. We shall approximate the operator A by strictly positive selfadjoint operators $A+\varepsilon^2 \geq \varepsilon^2$, $\varepsilon \in \mathbb{R}_{>0}$, by letting $\varepsilon \to 0+$. Consider the unique square root $B_\varepsilon := \sqrt{A + \varepsilon^2} \geq \varepsilon$, $\varepsilon \in \mathbb{R}_{>0}$, and the associated mapping

$$\begin{aligned} f_y : \mathbb{R}_{>0} &\to \mathbb{C} \\ \varepsilon &\mapsto \langle y\,|B_\varepsilon y\rangle_H \end{aligned}$$

for $y \in D(A)$ fixed. We first note that all square roots $B_\varepsilon \geq \varepsilon$, $\varepsilon \in \mathbb{R}_{>0}$, are commuting on $D(B)$. Indeed, we find B_ε^{-1} commutes with $A + \varepsilon^2$ and so also with A. Thus, all B_ε^{-1} are commuting, $\varepsilon \in \mathbb{R}_{>0}$. From

$$B_{\varepsilon_1}^{-1}B_{\varepsilon_2}^{-1} = B_{\varepsilon_2}^{-1}B_{\varepsilon_1}^{-1}$$

we get

$$B_{\varepsilon_1}^{-1}\, B_{\varepsilon_2} \subseteq B_{\varepsilon_2}\, B_{\varepsilon_1}^{-1}, \tag{B.31}$$

i.e.

$$B_{\varepsilon_1}^{-1} B_{\varepsilon_2} x = B_{\varepsilon_2}\, B_{\varepsilon_1}^{-1} x \tag{B.32}$$

for all $x \in D(B_{\varepsilon_2})$, ε_1, $\varepsilon_2 \in \mathbb{R}_{>0}$. Since $D(A) = D(A+\varepsilon^2)$ is dense in $D(B_\varepsilon)$ for all $\varepsilon \in \mathbb{R}_{>0}$, we get from (B.32) that also

$$B_{\varepsilon_1}\, B_{\varepsilon_2} y = B_{\varepsilon_2}\, B_{\varepsilon_1} y \tag{B.33}$$

for all $y \in D(A)$, ε_1, $\varepsilon_2 \in \mathbb{R}_{>0}$. With this we now find for f_y

$$\begin{aligned} f_y(\varepsilon_1) - f_y(\varepsilon_2) &= \left\langle y \,\middle|\left(\sqrt{A+\varepsilon_1^2} - \sqrt{A+\varepsilon_2^2}\right) y \right\rangle_H, \\ &= \left\langle y \,\middle|(B_{\varepsilon_1}+B_{\varepsilon_2})^{-1}(B_{\varepsilon_1}+B_{\varepsilon_2})(B_{\varepsilon_1}-B_{\varepsilon_2})\, y \right\rangle_H, \\ &= \left\langle y \,\middle|(B_{\varepsilon_1}+B_{\varepsilon_2})^{-1}\left(B_{\varepsilon_1}^2-B_{\varepsilon_2}^2\right) y \right\rangle_H, \\ &= \left(\varepsilon_1^2-\varepsilon_2^2\right)\left\langle y \,\middle|(B_{\varepsilon_1}+B_{\varepsilon_2})^{-1}\, y \right\rangle_H \end{aligned} \tag{B.34}$$

Since $(B_{\varepsilon_1}+B_{\varepsilon_2}) \geq \varepsilon_1+\varepsilon_2$ we have $\left\|(B_{\varepsilon_1}+B_{\varepsilon_2})^{-1}\right\| \leq \frac{1}{\varepsilon_1+\varepsilon_2}$ and so we have

$$|f_y(\varepsilon_1) - f_y(\varepsilon_2)| \leq |\varepsilon_1-\varepsilon_2|\,|y|_H^2,$$

i.e. the Lipschitz continuity of f_y. Therefore, $f_y(0+)$ exists by continuous extension. Moreover, we see from (B.34) that f_y is non-decreasing. The property (B.31) also yields as in the proof of Theorem 4

$$B_{\varepsilon_1}^{-1}(B_{\varepsilon_2}-B_{\varepsilon_1}) \geq 0,\ (B_{\varepsilon_2}-B_{\varepsilon_1})B_{\varepsilon_2}^{-1} \geq 0$$

for all ε_1, $\varepsilon_2 \in \mathbb{R}_{>0}$with $\varepsilon_2 \geq \varepsilon_1$. Thus, we also get with $y \in D(A)$ and $x := B_{\varepsilon_1} y \in D(B_{\varepsilon_2})$, $z := B_{\varepsilon_2} y \in D(B_{\varepsilon_1})$ (using (B.33)!)

$$\begin{aligned} 0 &\leq \left\langle x \,\middle|B_{\varepsilon_1}^{-1}(B_{\varepsilon_2}-B_{\varepsilon_1})x \right\rangle_H = \langle y\,|(B_{\varepsilon_2}-B_{\varepsilon_1})B_{\varepsilon_1}\, y\rangle_H, \\ 0 &\leq \left\langle z \,\middle|(B_{\varepsilon_2}-B_{\varepsilon_1})B_{\varepsilon_2}^{-1}\, z \right\rangle_H = \langle y\,|B_{\varepsilon_1}(B_{\varepsilon_2}-B_{\varepsilon_1})y\rangle_H, \end{aligned}$$

or

$$\langle y\,|B_{\varepsilon_1}B_{\varepsilon_1}\, y\rangle_H \leq \langle y\,|(B_{\varepsilon_2}B_{\varepsilon_1}\, y\rangle_H \leq \langle y\,|B_{\varepsilon_2}B_{\varepsilon_2}\, y\rangle_H$$

for all $y \in D(A)$. Further, re-fining the reasoning in the proof of Theorem 4 we get

$$\begin{aligned} |B_{\varepsilon_1} y - B_{\varepsilon_2} y|_H^2 &= \langle (B_{\varepsilon_1}-B_{\varepsilon_2})y\,|(B_{\varepsilon_1}-B_{\varepsilon_2})y\rangle_H, \\ &= \left\langle y\,\middle|(B_{\varepsilon_1}-B_{\varepsilon_2})^2 y\right\rangle_H, \\ &= \left\langle y\,\middle|B_{\varepsilon_1}^2 y\right\rangle_H - \langle y\,|B_{\varepsilon_1}\, B_{\varepsilon_2} y\rangle_H - \langle y\,|B_{\varepsilon_2}B_{\varepsilon_1} y\rangle_H + \left\langle y\,\middle|B_{\varepsilon_2}^2 y\right\rangle_H, \\ &= \left\langle y\,\middle|B_{\varepsilon_1}^2 y\right\rangle_H - 2\,\langle y\,|B_{\varepsilon_1}\, B_{\varepsilon_2} y\rangle_H + \left\langle y\,\middle|B_{\varepsilon_2}^2 y\right\rangle_H, \\ &\leq (\varepsilon_2^2-\varepsilon_1^2)\,|y|_H^2 \\ &\leq (\varepsilon_2+\varepsilon_1)\,|y|_H^2\,(\varepsilon_2-\varepsilon_1). \end{aligned}$$

We therefore have with

$$Cy := \lim_{\varepsilon\to 0+} B_\varepsilon y \text{ for all } y \in D(A),$$

a densely defined linear operator C. Moreover, taking the limit $\varepsilon_1 \to 0+$ we obtain with $\varepsilon = \varepsilon_2$

$$|B_\varepsilon y - Cy|_H^2 \leq \varepsilon^2\,|y|_H^2.$$

This shows that $\overline{(C-B_\varepsilon)} \in L(H,H)$, indeed

$$\left\|\overline{(C-B_\varepsilon)}\right\| \leq \varepsilon. \tag{B.35}$$

As in the bounded case it is easily seen that C is non-negative, symmetric and therefore in particular also closable. Denote the closure of C by B_0 then

$$B_0 = B_\varepsilon + \overline{(C - B_\varepsilon)}. \tag{B.36}$$

This shows that $D(B_\varepsilon) = D(B_0)$ is independent of $\varepsilon \in \mathbb{R}_{>0}$ and (for small values of $\varepsilon \in \mathbb{R}_{>0}$) that B_0 can be considered as a small perturbation of B_ε. Therefore, we find for $\varepsilon < 1$ that

$$(B_0 + 1)^{-1} = \left(1 + (B_\varepsilon + 1)^{-1}\,\overline{(C - B_\varepsilon)}\right)^{-1}(B_\varepsilon + 1)^{-1} \in L(H, H).$$

Thus we have $-1 \in \varrho(B_0)$ and since $B_0 \geq 0$ it follows that B_0 is selfadjoint. We want to show that this B_0 is the desired square root. Let $x \in D(A)$ then $x \in D(B_0)$, $B_\varepsilon x \in D(B_0)$ and

$$B_\varepsilon B_\varepsilon x = Ax + \varepsilon^2 x.$$

Since $B_\varepsilon x \to B_0 x$ as $\varepsilon \to 0+$we have with (B.36)

$$B_0 B_\varepsilon x = Ax + \varepsilon^2 x + \overline{(C - B_\varepsilon)} B_\varepsilon x \to Ax \text{ as } \varepsilon \to 0+$$

and so $B_0 x \in D(B_0)$ and

$$B_0\, B_0 x = Ax \text{ for all } x \in D(A).$$

Having shown that $A \subseteq B_0^2$, we now see that

$$A + 1 \subseteq B_0^2 + 1 = (B_0 + \mathrm{i})(B_0 - \mathrm{i})$$

or

$$(A + 1)^{-1} \subseteq (B_0^2 + 1)^{-1} = (B_0 - \mathrm{i})^{-1}(B_0 + \mathrm{i})^{-1}.$$

Since, however, $(A + 1)^{-1} \in L(H, H)$, we must have equality, i.e.

$$(A + 1)^{-1} = (B_0^2 + 1)^{-1}$$

or

$$A + 1 = B_0^2 + 1$$

and so finally

$$A = B_0^2.$$

Thus, we have established existence of a square root of A. That B_0 commutes with any $C \in L(H, H)$ commuting with A carries over from B_ε as $\varepsilon \to 0$. We have that C also commutes with $A + \varepsilon^2$, thus also

$$CB_\varepsilon \subseteq B_\varepsilon C.$$

Letting $\varepsilon \to 0$ we get

$$CB_0 \subseteq B_0 C.$$

Let now $B^2 = A$ with $B \geq 0$ selfadjoint and such that it commutes with every $C \in L(H, H)$ commuting with A. We want to show that $B = B_0$. Since B commutes with itself we first have

$$(B - \lambda)^{-1}\, Ax = (B - \lambda)^{-1}\, B^2 x = B^2 (B - \lambda)x = A\,(B - \lambda)x,$$

for all $x \in D(A)$ and so we see

$$(B - \lambda)^{-1}\, A \subseteq A\,(B - \lambda)^{-1}$$

for $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Since B commutes with A, it also commutes with B_0, i.e.

$$(B - \lambda)^{-1}\, B_0 \subseteq B_0\,(B - \lambda)^{-1}.$$

In particular, we read off that for $z \in D(B_0)$ we have $x := (B - \lambda)^{-1} z \in D(B_0) \cap D(B)$ and

$$B_0\, B\, x = B\, B_0\, x$$

Consider now $y := (B - B_0)x \in D(A)$ for $z \in D(A)$. We calculate

$$\langle y\,|By\rangle_H + \langle y\,|B_0 y\rangle_H = \langle y\,|(B + B_0)(B - B_0)x\rangle_H\,.$$

However, since B and B_0 commute, we find

$$(B + B_0)(B - B_0)x = B^2 x - B_0^2 x = Ax - Ax = 0.$$

From the non-negativity of B and B_0 and using the existence part of this proof we get

$$\langle Dy\,|Dy\rangle_H = \langle y\,|By\rangle_H = \langle y\,|Cy\rangle_H = \langle Ey\,|Ey\rangle_H = 0$$

where $E^2 = C$ and $D^2 = B$. Thus, we find $Dy = Ey = 0$ and so

$$0 = D^2 y = By,\ 0 = E^2 y = B_0 y.$$

Consequently, we get

$$|y|_H^2 = |Bx - B_0 x|_H^2 = \langle x\,|(B - B_0)(B - B_0)x\rangle_H = \langle x\,|(B - B_0)y\rangle_H = 0.$$

We have shown that

$$(B - \lambda)^{-1} Bz = B(B - \lambda)^{-1} z = B_0 (B - \lambda)^{-1} z = (B - \lambda)^{-1} B_0 z$$

for all $z \in D(A)$. Therefore, also

$$Bz = B_0 z \text{ for all } z \in D(A).$$

Since $D(A)$ is dense in $D(B)$ and in $D(B_0)$, we get $B = B_0$. □

Remark 11. As an application consider a compact linear operator $A : H_0 \to H_0$, then $A^*A = |A|^*$ is a non-negative, selfajoint compact operator in H_0. The notation $L_\infty(H_0, H_0)$ of the normed linear space of all such compact operators used earlier is motivated by the following observation. Since $A^*A :\ H_0 \to H_0$ is also compact (as a composition of a bounded operator with a compact one) and since selfadjoint operators have real spectrum, we see that

$$|Ax|^2 = \langle x|A^*Ax\rangle \le \max \sigma\,(A^*A)\,|x|^2$$

with equality holding for elements in the null space of the largest eigenvalue, which must be 0 or a point in the discrete spectrum. Thus,

$$|A|_{L(H_0,H_0)} = \sqrt{\max \sigma\,(A^*A)} = \max\ \sqrt{\sigma\,(A^*A)}.$$

Later we shall find that $\sqrt{\sigma\,(A^*A)} = \sigma\left(\sqrt{A^*A}\right)$ and so the operator norm of corresponds to the sup-norm in $\sigma\left(\sqrt{A^*A}\right)$, which is usually indicated by the index ∞. This characterization of $L_\infty(H_0, H_0)$ leads the way to other normed spaces of operators denoted by $L_p(H_0, H_0)$, $p \in [1, \infty[$, (so-called Schatten classes) and given by $\left\{A \in L_\infty(H_0, H_0)\middle|\ \sum_{\lambda\in\sigma\left(\sqrt{A^*A}\right)} \lambda^p < \infty\right\}$ equipped with the linear structuure of $L(H_0, H_0)$ and the norm

$$A \mapsto |A|_p := \left(\sum_{\lambda\in\sigma\left(\sqrt{A^*A}\right)} \lambda^p\right)^{1/p}.$$

Of particular interest is the space $L_2(H_0, H_0)$, which is a Hilbert space. Elements of $L_2(H_0, H_0)$ are referred to as Hilbert-Schmidt operators. Elements of $L_1(H_0, H_0)$ are called trace class operators.

Being able to define the square root of non-negative, selfadjoint operator gives rise to an interesting representation result known as **polar decomposition**. It turns out that an arbitrary closed, densely defined operator is really quite closely related to a selfadjoint operator.

Theorem 12. *Let $A \subseteq H_0 \oplus H_1$ be a closed, densely defined, linear operator between complex Hilbert spaces H_0 and H_1. Then there is an isometry $U : \overline{A^*H_1} \to H_1$ from the closed subspace $\overline{A^*[H_1]}$ onto $\overline{A[H_0]}$ such that*

$$A = U\,|A|,$$

where $|A| := \sqrt{A^*A}$.

PROOF. We first observe that

$$\langle Ay|Ay\rangle_{H_1} = \langle y\,|A^*Ay\,\rangle_{H_0} = \left\langle y\,\middle|\sqrt{A^*A}\,\sqrt{A^*A}\,y\,\right\rangle_{H_0} = \langle |A|\,y\,||A|\,y\,\rangle_{H_0}$$

for all $y \in D(A^*A)$. Since $D(A^*A)$ is dense in $D(A)$ and in $D(|A|)$, we must have

$$D(A) = D(|A|)$$

and

$$\bigwedge_{y\in D(A)=D(|A|)} \langle Ay|Ay\rangle_{H_1} = \langle |A|\,y\,||A|\,y\,\rangle_{H_0}.$$

In particular, this implies

$$N\left(|A|\right) = N(A) \subseteq D(A) = D(|A|),$$

and consequently (by the projection theorem)

$$\overline{A^*\,[H_1]} = \overline{|A|\,[H_0]}. \tag{B.37}$$

From this we have that $B := |A|\big|_{\overline{A^*H_1}} : D(|A|)\cap\overline{A^*\,[H_1]} \subseteq \overline{A^*\,[H_1]} \to \overline{A^*\,[H_1]}$ is still non-negative and Hermitean. Indeed, we shall see that B is one-to-one and selfadjoint in $\overline{A^*\,[H_1]}$. First, let $Bx = 0$ then $x \in N(|A|)$ and also $x \in N(|A|)^\perp = \overline{A^*\,[H_1]}$. This shows that $x = 0$. Next, we see that B is densely defined. Since $D(|A|)$ is dense in $H_0 = N(|A|) \oplus \overline{A^*\,[H_1]}$, we get from the continuity of orthogonal projector P onto $\overline{A^*\,[H_1]}$ that $P\,[D(|A|)]$ is dense in $P\,[H_0] = \overline{A^*\,[H_1]}$, but since $(1-P)\,[H_0] = N(|A|) \subseteq D(|A|)$ we also have $P\,[D(|A|)] \subseteq D(|A|)\cap\overline{A^*\,[H_1]}$. This confirms that B is densely defined. But B is also closed. Assume $x_n \to x$ in $\overline{A^*\,[H_1]}$ as $n\to\infty$ and $B\,x_n = |A|\,x_n \to y$ in $\overline{A^*\,[H_1]}$ as $n \to \infty$. Since $|A|$ is closed, it follows that x is not only in $\overline{A^*\,[H_1]}$ but we also have $x \in D(|A|)$ and $|A|\,x = Bx = y$. That the closed, symmetric operator B is also selfadjoint follows if we can show that e.g. $B+1$ is onto. Since $|A| \geq 0$ is selfadjoint, we have that $|A|+1$ is onto. Let now $f \in \overline{A^*\,[H_1]}$ be given, then there is $x \in D(|A|)$ such that

$$|A|x + x = f.$$

From (B.37) we see that $x = f - |A|x \in \overline{A^*\,[H_1]} = \overline{|A|\;[H_0]}$, i.e. $x \in D(B)$. This proves that

$$Bx + x = f.$$

We clearly have $B^{-1} : |A|\,[H_0] \subseteq \overline{A^*\,[H_1]} \to D(B) \subseteq \overline{A^*\,[H_1]}$ and so

$$A = A\,B^{-1}\,|A|. \tag{B.38}$$

To conclude the argument we claim that

$$AB^{-1} : |A|\;[H_0] \subseteq \overline{A^*\,[H_1]} \to \overline{A\,[H_0]}$$

is a bounded, linear operator with a unitary closure. By (B.38) we see that the range $\left(A\,B^{-1}|A|\right)[D(A)] = A\,[D(A)]$ is dense in $\overline{A\,[H_0]}$. Since also $|A|\,[H_0]$ dense in $\overline{|A|\,[H_0]} = \overline{A^*\,[H_1]}$, we only need to show that AB^{-1} is norm-preserving. We find with $y = |A|x$, $x \in D(A) = D(|A|)$,

$$\left\langle AB^{-1}y|AB^{-1}y\right\rangle_{H_1} = \langle Ax\,|Ax\,\rangle_{H_1} = \langle |A|\,x\,||A|\,x\,\rangle_{H_0} = \langle y\,|y\,\rangle_{H_0}.$$

Thus, with the isometry $U := \overline{AB^{-1}} \subseteq H_0 \oplus H_1$ we get from (B.38) as claimed

$$A = U\,|A|.$$

□

COROLLARY 13. *Let $A \subseteq H_0 \oplus H_1$ be a closed, densely defined, linear operator between complex Hilbert spaces H_0 and H_1. Then there is an isometry $U : \overline{A^*\,[H_1]} \subseteq H_0 \to \overline{A\,[H_0]} \subseteq H_1$ such that*

$$A = U\,|A| = |A^*|\,U,\; A^* = |A|\,U^* = U^*\,|A^*|$$

on $\overline{A^*\,[H_1]}$ and $\overline{A\,[H_0]}$, respectively.

PROOF. The result follows by applying the polar decomposition also to A^* and taking adjoints in the resulting formulas. The result then follows by Lemma 90. That U^* is the appropriate unitary mapping for the polar representation associated with A^* follows by observing that

$$|A^*| = U\,|A|\,U^*.$$

Indeed, with $G \coloneqq U\,|A|\,U^*$ we get

$$G^2 = U\,|A|\,U^*U\,|A|\,U^* = (U\,|A|)\,(|A|\,U^*) = AA^*.$$

By the uniqueness of the root operator

$$G = |A^*|\,.$$

Then

$$A^* = |A|\,U^* = U^*\,(U\,|A|\,U^*) = U^*\,|A^*|\,.$$

□

APPENDIX C

General Material Laws

C.1. The Paley-Wiener Theorem

We slightly rephrase the related concepts and results from [**Yosida 1974**], p. 162-165.

Definition 14. The Hardy-Lebesgue space $\mathbb{HL}$ is defined as the subspace

$$\mathbb{HL} := \left\{ f \in \mathbb{C}^{\mathbb{R}-\mathrm{i}\,\mathbb{R}_{>0}} \middle| \, f \text{ analytic } \wedge \bigwedge_{\varepsilon\in\mathbb{R}>0} f\left(\,\cdot\, - \mathrm{i}\,\varepsilon\right) \in L^2(\mathbb{R}) \wedge \sup\left\{ \left|f\left(\,\cdot\, - \mathrm{i}\,\varepsilon\right)\right|_0 \mid \varepsilon \in \mathbb{R}_{>0}\right\} < \infty \right\}$$

of $\mathbb{C}^{\mathbb{R}-\mathrm{i}\,\mathbb{R}_{>0}}$ with the usual image-wise linear structure. $\mathbb{HL}$ equipped with

$$f \mapsto \sup\left\{ \left|f\left(\,\cdot\, - \mathrm{i}\,\varepsilon\right)\right|_0 \mid \varepsilon \in \mathbb{R}_{>0}\right\}$$

becomes a Banach space for which we use the same name.

We first formulate a converse version of what is called the Paley-Wiener theorem (note that we consider $L^2(\mathbb{R}_{>0}) := \left\{ f \in L^2(\mathbb{R}) \mid \chi_{\mathbb{R}_{l<0}} f = 0 \right\}$ as a subspace of $L^2(\mathbb{R})$).

Theorem 15. *Let $g \in L^2(\mathbb{R}_{>0})$ be given, then*

$$\begin{aligned} f :\ \mathbb{R} - i\,\mathbb{R}_{>0} &\to \mathbb{C} \\ z &\mapsto \left(\mathcal{L}_{-\,\mathfrak{Im}\,z}\, g\right)(\mathfrak{Re}\,z) \end{aligned}$$

is well-defined and in $\mathbb{HL}$. Moreover, f is a continuous extension of $\mathcal{L}_0 g$ in the sense that

$$f\left(\,\cdot\, - i\,\varepsilon\right) \overset{\varepsilon\to 0+}{\to} \mathcal{L}_0 g$$

in $L^2(\mathbb{R})$. In particular, we have

$$\bigwedge_{\eta\in\mathbb{R}_{>0}} \left|f\left(\,\cdot\, - i\,\eta\right)\right|_0 = \left|\exp\left(-\eta m_0\right) g\right|_0 = |g|_{\eta,0} \leq |g|_0$$

and

$$\sup\left\{ \left|f\left(\,\cdot\, - i\,\varepsilon\right)\right|_0 \mid \varepsilon \in \mathbb{R}_{>0}\right\} = \sup\left\{ \left|\exp\left(-\eta m_0\right) g\right|_0 \mid \eta \in \mathbb{R}_{>0}\right\} = |g|_0\ .$$

Corollary 16. *The Paley-Wiener mapping*

$$\begin{aligned} \mathrm{PW} : L^2(\mathbb{R}_{>0}) &\to \mathbb{HL} \\ g &\mapsto \left(z \mapsto \left(\mathcal{L}_{-\,\mathfrak{Im}\,z}\, g\right)(\mathfrak{Re}\,z)\right) \end{aligned}$$

is an isometry.

This result is complemented by the following characterization of $\mathcal{L}_0\left[L^2(\mathbb{R}_{>0})\right]$.

Theorem 17. *(Paley-Wiener theorem) Let $f \in \mathbb{HL}$, then $f\left(\,\cdot\, - i\,0+\right)$ exists in $L^2(\mathbb{R})$ in the sense of $L^2(\mathbb{R})$-convergence and we have*

$$\mathcal{L}_0^*\left(f\left(\,\cdot\, - i\,0+\right)\right) \in L^2(\mathbb{R}_{>0})\ .$$

Corollary 18. *The Paley-Wiener mapping* PW *is an isometric bijection between the Hilbert space $L^2(\mathbb{R}_{>0})$ and the Banach space $\mathbb{HL}$.*

Since the latter result means that $\mathbb{HL}$ is a Hilbert space if equipped with the inner product

$$(f,h)\mapsto \left\langle \mathrm{PW}^{-1}(f)\;\middle|\;\mathrm{PW}^{-1}(h)\right\rangle_0$$

we also obtain:

Corollary 19. *The Paley-Wiener mapping* PW *is a* ***unitary*** *mapping between the Hilbert spaces* $L^2(\mathbb{R}_{>0})$ *and* $\mathbb{HL}$.

The Paley-Wiener theorem has a straight-forward extension to Hilbert space-valued functions via tensor product extensions. Noting that for any Hilbert space H

$$L^2(\mathbb{R}_{>0},H)=L^2(\mathbb{R}_{>0})\otimes H,$$

we obtain by continuous extension a unitary mapping

$$\begin{aligned}\mathrm{PW}\otimes 1_H:\ L^2(\mathbb{R}_{>0},H)=L^2(\mathbb{R}_{>0})\otimes H&\to \mathbb{HL}\otimes H\\ \phi\otimes w&\mapsto(\mathrm{PW}\phi)\otimes w\,,\end{aligned}$$

where $1_H: H\to H$ denotes the identity operator in H. Thus, we have another corollary.

Corollary 20. *The Paley-Wiener mapping* $\mathrm{PW}\otimes 1_H$ *is a* ***unitary*** *mapping between the Hilbert spaces* $L^2(\mathbb{R}_{>0},H)$ *and* $\mathbb{HL}\otimes H$.

Recall that $L^2(\mathbb{R}_{>0})\otimes H$ can be described as the completion of the algebraic tensor product $\mathring{C}_\infty(\mathbb{R}_{>0})\underset{a}{\otimes}H$, which is the linear space generated by simple H-valued functions of the form $t\mapsto\varphi(t)\,w$ with $\varphi\in\mathring{C}_\infty(\mathbb{R}_{>0})$ and $w\in H$.

C.2. Causality

We first adapt the concept of time support to the Hilbert-space-valued situation. Here we utilize again the concept of an algebraic tensor product.

Definition 21. Let f be a given linear functional on $\mathring{C}_\infty(\mathbb{R})\underset{a}{\otimes}H$, then we say $f=0$ in I if $f(\phi\otimes w)=0$ for all $w\in H$ and $\phi\in\mathring{C}_\infty(\mathbb{R})$ with $\operatorname{supp}\phi\subseteq I$, I open in $\mathbb{R}$. Moreover, we define what we shall refer to as the time support $\operatorname{supp}_0 f$ of f as

$$\operatorname{supp}_0 f:=\mathbb{R}\setminus\bigcup\{I\,|\,I\text{ open in }\mathbb{R}\ \wedge\ f=0\text{ in }I\}.$$

We are now ready to formulate the corresponding concept of causality (somewhat simplified to by-pass more intricate matters associated with Sobolev lattices).

Definition 22. Let $W: D(W)\subseteq\mathbb{C}^{\mathring{C}_\infty(\mathbb{R})\underset{a}{\otimes}H}\to\mathbb{C}^{\mathring{C}_\infty(\mathbb{R})\underset{a}{\otimes}H}$ be a mapping from linear functionals to linear functionals such that

$$\inf\operatorname{supp}_0 f\le\inf\operatorname{supp}_0 W(f)\tag{C.1}$$

for all $f\in D(W)$, then we call W (forward) causal. If

$$\sup\operatorname{supp}_0 f\ge\sup\operatorname{supp}_0 W(f)\tag{C.2}$$

for all $f\in D(W)$, then we call W backward causal.

Here we interpret $\inf\operatorname{supp}_0 f=\infty$ if $\operatorname{supp}_0 f=\emptyset$ and $\inf\operatorname{supp}_0 f=-\infty$ if $\operatorname{supp}_0 f$ is not bounded below, so that (C.1) is only restrictive[1] if we take f with $\operatorname{supp}_0 f$ bounded below. Analogously, we interpret $\sup\operatorname{supp}_0 f=-\infty$ if $\operatorname{supp}_0 f=\emptyset$ and $\sup\operatorname{supp}_0 f=\infty$ if $\operatorname{supp}_0 f$ is not bounded above, so that (C.2) is only restrictive if we take f with $\operatorname{supp}_0 f$ bounded above. It is in this sense that the jargon phrase "as long as f is zero, so is $W(f)$", a variant of which is used in the wording of the main result, is made precise.

[1]Note that $W(0)=0$ for any forward or backward causal mapping W.

Of course, these formal concepts of time support and causality become useful only if associated with a suitable topology and corresponding continuity concepts.

For our purposes we want to generalize the concept of functions of ∂_0 to operator-valued functions of ∂_0 . For this we first need to extend the operators ∂_0 to the tensor product spaces $H_{\nu,0} \otimes H$ by interpreting ∂_0 henceforth as the operator $\partial_0 \otimes 1_H$. It is customary to write again ∂_0 for this extended time-derivative $\partial_0 \otimes 1_H$. Moreover, we need to extend the Fourier-Laplace transform to $H_{\nu,0} \otimes H$. We shall re-utilize the notation $\mathcal{L}_\nu$ and the name Fourier-Laplace transform for the unique unitary extension of

$$\begin{aligned} \mathring{C}_\infty(\mathbb{R}) \underset{a}{\otimes} H \subseteq H_{\nu,0} \otimes H &\to H_{0,0} \otimes H \\ \varphi \otimes w &\mapsto (\mathcal{L}_\nu \varphi) \otimes w \end{aligned}$$

to $H_{\nu,0} \otimes H$. With this extended Fourier-Laplace transform we will be able to describe the class of material laws. Let $(M(z))_{z \in B_{\mathbb{C}}(r,r)}$ be a holomorphic family of uniformly bounded linear operators, then we define, following [**Pi-McGhee 2011**], for $\nu > \frac{1}{2r}$

$$M\left(\partial_0^{-1}\right) := \mathcal{L}_\nu^* M\left(\frac{1}{\mathrm{i}\, m_0 + \nu}\right)\mathcal{L}_\nu .$$

Note that for $r \in \mathbb{R}_{>0}$

$$\begin{aligned} B_{\mathbb{C}}(r,r) &\to \mathrm{i}\,\mathbb{R} + \mathbb{R}_{>1/(2r)} \\ z &\mapsto z^{-1} \end{aligned}$$

is a bijection.

Theorem 23. *Let $(M(z))_{z \in B_{(r,r)}}$ be a holomorphic family of uniformly bounded linear operators on H and $\nu > \frac{1}{2r}$ then $M\left(\partial_0^{-1}\right) : H_{\nu,0} \otimes H \to H_{\nu,0} \otimes H$ is forward causal in the sense that $M\left(\partial_0^{-1}\right)$ restricted to $\mathring{C}_\infty(\mathbb{R}) \overset{a}{\otimes} H$ considered as a subspace of $\mathbb{C}^{\mathring{C}_\infty(\mathbb{R}) \overset{a}{\otimes} H}$ by interpreting $\phi \otimes w$ as the functional*

$$\begin{aligned} \phi \otimes w : \mathring{C}_\infty(\mathbb{R}) \overset{a}{\otimes} H &\to \mathbb{C} \\ \psi \otimes v &\mapsto (\phi \otimes w)(\psi \otimes v) := \int_{\mathbb{R}} \phi(t)^* \, \psi(t) \; dt \; \langle w|v\rangle \end{aligned}$$

for every $\phi, \psi \in \mathring{C}_\infty(\mathbb{R})$, $w, v \in H$, is forward causal according to the above definition.

Proof. For the time-translation τ_h given as the bounded linear extension of

$$\begin{aligned} \tau_h : \mathring{C}_\infty(\mathbb{R}) \underset{a}{\otimes} H \subseteq H_{\nu,0} \otimes H &\to H_{\nu,0} \otimes H \\ \phi \otimes w &\mapsto (\tau_h \phi) \otimes w \end{aligned}$$

with

$$(\tau_h \phi)(t) := \phi(t+h)$$

for all $\phi \in \mathring{C}_\infty(\mathbb{R})$, t, $h \in \mathbb{R}$, we have

$$\tau_h = \exp(h\partial_0)$$

and so the commutator relation

$$\tau_h M\left(\partial_0^{-1}\right) = M\left(\partial_0^{-1}\right) \tau_h .$$

Thus, to test for causality, we may assume without loss of generality that $\phi \in \mathring{C}_\infty(\mathbb{R})$ has support $\operatorname{supp}(\phi)$ with

$$\inf \operatorname{supp}(\phi) = 0$$

and due to translation invariance we only need to show that

$$\inf \operatorname{supp}_0\left(M\left(\partial_0^{-1}\right)\phi \otimes w\right) \geq 0$$

for all $w \in H$. Obviously, $\phi \otimes w \in L^2(\mathbb{R}_{>0}, H)$ and so by the converse of the Paley-Wiener theorem $\mathcal{L}_\nu \phi \otimes w \in \mathbb{HL} \otimes H$. With the assumed holomorphy and uniform boundedness of $(M(z))_{z \in B_{(r,r)}}$ we obtain

$$M\left(\frac{1}{\mathrm{i}\, m_0 + \nu}\right) \mathcal{L}_\nu \phi \otimes w \in \mathbb{HL} \otimes H.$$

By the Paley-Wiener theorem we get

$$\mathcal{L}_0^* M\left(\frac{1}{\mathrm{i}\, m_0 + \nu}\right) \mathcal{L}_\nu \phi \otimes w \in L^2(\mathbb{R}_{>0}, H)$$

or

$$M\left(\partial_0^{-1}\right) \phi \otimes w = \mathcal{L}_\nu^* M\left(\frac{1}{\mathrm{i}\, m_0 + \nu}\right) \mathcal{L}_\nu \phi \otimes w = \exp(\nu m_0) \mathcal{L}_0^* M\left(\frac{1}{\mathrm{i}\, m_0 + \nu}\right) \mathcal{L}_\nu \phi \otimes w \in H_{\nu,0} \otimes H$$

and

$$\inf \operatorname{supp}_0 \left(M\left(\partial_0^{-1}\right) \phi \otimes w\right) \geq 0.$$

$\square$

Remark 24. We note here that the seemingly rather restrictive assumption on the material law operator $M\left(\partial_0^{-1}\right)$ is largely unavoidable in order to maintain causality. Simple examples for a causal operator are the 'delay operator' τ_{-h}, $h \in \mathbb{R}_{>0}$, or positive fractional powers of ∂_0^{-1}, $\alpha \in \mathbb{R}_{\geq 0}$, (fractional integration). Indeed

$$(\partial_0^\alpha)_{\alpha \in \mathbb{R}}$$

is a continuous one-parameter group of (forward) causal operators in $H_{\nu,-\infty}$, $\nu \in \mathbb{R}_{>0}$.[**Schmüdgen 2012**]

Literature for Appendices

[Pi-McGhee 2011] R. Picard and D. F. McGhee. *Partial Differential Equations: A unified Hilbert Space Approach*, volume 55 of *De Gruyter Expositions in Mathematics*. De Gruyter. Berlin, New York. 518 p., 2011.

[Schmüdgen 2012] K. Schmüdgen. *Unbounded self-adjoint operators on Hilbert space.* Dordrecht: Springer, 2012.

[Yosida 1974] K. Yosida. *Functional analysis.* Springer-Verlag; Berlin et al., 1974.