FANCy

an online documentation project on

Functional Analysis and Non-Commutative Geometry

Founded by Markus J. Pflaum

January 26, 2023

Authors

The following people have contributed to the $\mathcal{FANCy}$ Project, in alphabetical order:

Frédéric Pierre-Jean Latrémolière
Paul Mitchener
Markus J. Pflaum
Daniel Spiegel

Copyright

Contents

Introduction

The $\mathcal{FANC}y$ Project is an open source textbook on functional analysis, the mathematics of quantum mechanics, noncommutative geometry and related topics.

Attribution

The main author of this work is Markus J. Pflaum.

Chapter I.2, *General Topology*, is based on the script *Foundations of Point Set Topology* by Frédéric Latrémolière submitted to Libri Mathematicae under the GNU FDL v1.3. The chapter is also part of the CRingProject.

Chapter I.3, *Measure and Integration theory*, is based on the notes *Measure Theory and Integration* by Paul Mitchener submitted to Libri Mathematicae under the GNU FDL v1.3.

M.J. Pflaum is the author of Chapter II.1, *Topological Vector Spaces*.

Chapter II.3 *Hilbert Spaces* and Chapter II.4 *C^*-Algebras* have been written by Markus J. Pflaum and Daniel Spiegel.

Part I.

Fundamentals

I.1. Tools from Analysis

1.1. Some useful inequalities

In this section we collect several inequalities from real analysis which will be of use later in this monograph.

1.1.1 Theorem (Young's inequality) *Let $a, b \geqslant 0$, and assume that $p, q > 1$ satisfy the relation $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$ab \leqslant \frac{1}{p} a^p + \frac{1}{q} b^q \,.$$

Equality holds if and only if $a^p = b^q$.

Proof. Since the second derivative $\exp''$ of the exponential function attains only positive values, the function exp is strictly convex that means satisfies

$$\exp\big(\lambda x + (1-\lambda) y\big) \leqslant \lambda \exp(x) + (1-\lambda)\exp(y)$$

for all $x, y \in \mathbb{R}$ and $\lambda \in [0,1]$ with equality holding true if and only if $x = y$ or $\lambda \in \{1, 0\}$. Putting $x = p \ln a$, $y = q \ln b$, and $\lambda = \frac{1}{p}$ one obtains

$$ab = \exp\big(\lambda x + (1-\lambda) y\big) \leqslant \lambda \exp(x) + (1-\lambda)\exp(y) = \frac{1}{p} a^p + \frac{1}{q} b^q \,.$$

Equality holds if and only if $x = y$ which is equivalent to $a^p = b^q$. □

1.1.2 Theorem (Cauchy–Schwarz inequality for sums) *Let $v, w \in \mathbb{C}^n$. Then*

$$\Big| \sum_{i=1}^n v_i \overline{w_i} \Big|^2 \leqslant \Big(\sum_{i1}^n |v_i|^2 \Big) \Big(\sum_{i=1}^n |w_i|^2 \Big).$$

Equality holds true if and only if v and w are linearly dependant.

Proof. Let us use the *inner product* notation

$$\langle v, w \rangle := \sum_{i=1}^n v_i \overline{w_i} \quad \text{for } v, w \in \mathbb{C}^n.$$

Then the ℓ^2-*norm*

$$\|v\| := \left(\sum_{i=1}^n |v_i|^2 \right)^{1/2} = \langle v, v \rangle^{1/2}$$

is well-defined and non-negative for any $v \in \mathbb{C}^n$. If $\|v\| = 0$ or $\|w\| = 0$, then $v = 0$ or $w = 0$, and the claim is trivial. So we assume $\|v\|, \|w\| > 0$ and compute

$$\begin{aligned} 0 &\leqslant \big\langle \|w\|v - \|v\|w, \|w\|v - \|v\|w \big\rangle = \sum_{i=1}^n \big(\|w\|v_i - \|v\|w_i\big)\big(\|w\|\overline{v_i} - \|v\|\overline{w_i}\big) = \\ &= \sum_{i=1}^n \|w\|^2 v_i\overline{v_i} - \|w\|\|v\|v_i\overline{w_i} - \|w\|\|v\|w_i\overline{v_i} + \|v\|^2 w_i\overline{w_i} = \\ &= 2\|v\|\|w\|\Big(\|v\|\|w\| - \mathfrak{Re}\langle v, w\rangle\Big). \end{aligned} \tag{1.1.1}$$

Now choose $c \in \mathbb{C}$ with $|c| = 1$ such that $c\langle v, w\rangle = |\langle v, w\rangle|$. Replacing v by cv in inequality (1.1.1) and observing that $\|cv\|$ and $\|w\|$ are positive then entails

$$0 \leqslant \|cv\|\|w\| - \mathfrak{Re}\langle cv, w\rangle = \|v\|\|w\| - \mathfrak{Re}(c\langle v, w\rangle) = \|v\|\|w\| - |\langle v, w\rangle|,$$

which is the claimed Cauchy–Schwartz inequality for sums in abbreviated form.

Equality holds true if and only if $\|w\|cv - \|v\|w = 0$. So if $\|v\|\|w\| = |\langle v, w\rangle|$, then v and w are linearly dependant. To show the converse, assume that $av = bw$ for some $a, b \in \mathbb{C}$ with $(a, b) \neq (0, 0)$. Because we consider the nontrivial case where both v and w are nonzero, we can assume without loss of generality that $b = 1$. But then

$$|\langle v, w\rangle| = |\langle v, av\rangle| = |a|\|v\|^2 = \|v\|\,\|w\|\ ,$$

hence equality holds in this case. The proof is finished. □

1.1.3 Besides the ℓ^2-norm on $\mathbb{C}^n$ one has the so-called ℓ^p-norms $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}_{\geqslant 0}$ for $p \geqslant 1$. They are defined by

$$\|v\|_p = \left(\sum_{k=1}^n |v_k|^p\right)^{1/p} \quad \text{for } v \in \mathbb{C}^n\ .$$

The *maximum norm* or ℓ^∞-norm $\|\cdot\|_\infty$ is given by

$$\|v\|_\infty = \sup\big\{|v_k| \;\big|\; k = 1, \ldots, n\big\}\ .$$

The ℓ^p-norms are all norms indeed as we will later see.

1.1.4 Theorem (Hölder's inequality for sums) *Let $p, q \in [1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$\sum_{k=1}^n |v_k w_k| \leqslant \|v\|_p \cdot \|w\|_q \quad \textit{for all } v, w \in \mathbb{C}^n\ .$$

Proof. If $p = 1$ or $q = 1$ the claim is immediate, because then $q = \infty$ or $p = \infty$, respectively, and the two estimates

$$\sum_{k=1}^n |v_k w_k| \leqslant \left(\sum_{k=1}^n |v_k|\right) \cdot \sup\big\{|w_k| \;\big|\; k = 1, \ldots, n\big\}$$

and

$$\sum_{k=1}^{n} |v_k w_k| \leqslant \left(\sum_{k=1}^{n} |w_k| \right) \cdot \sup \left\{ |v_k| \;\middle|\; k = 1, \ldots, n \right\}$$

obviously hold. So we can assume $1 < p, q < \infty$. Moreover we can assume that both v and w are nonzero because otherwise the claim is trivial. Now observe that by Young's inequality

$$\frac{|v_k|}{\|v\|_p} \cdot \frac{|w_k|}{\|w\|_q} = \left(\frac{|v_k|^p}{\|v\|_p^p} \right)^{1/p} \cdot \left(\frac{|w_k|^q}{\|w\|_q^q} \right)^{1/q} \leqslant \frac{1}{p} \frac{|v_k|^p}{\|v\|_p^p} + \frac{1}{q} \frac{|w_k|^q}{\|w\|_q^q} \quad \text{for } k = 1, \ldots, n \; .$$

Summing over all k gives

$$\sum_{k=1}^{n} \frac{|v_k|}{\|v\|_p} \cdot \frac{|w_k|}{\|w\|_q} \leqslant \frac{1}{p} \frac{\|v\|_p^p}{\|v\|_p^p} + \frac{1}{q} \frac{\|w\|_q^q}{\|w\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1 \; .$$

Multiplication of both sides by $\|v\|_p \cdot \|w\|_q$ entails Hölder's inequality. □

1.1.5 Theorem (Minkowski's inequality for sums) *Let $p \in [1, \infty)$. Then*

$$\|v + w\|_p \leqslant \|v\|_q + \|w\|_p \quad \textit{for all } v, w \in \mathbb{C}^n \; .$$

Proof. For $p = 1$ the claim is trivial, likewise for $p = \infty$. So assume $1 < p < \infty$ and put $q := \frac{p}{p-1}$. Then $\frac{1}{p} + \frac{1}{q} = 1$, and we can apply Hölder's inequality to compute

$$\begin{aligned} \|v + w\|_p^p &= \sum_{k=1}^{n} |v_k + w_k|^p \leqslant \sum_{k=1}^{n} |v_k| \, |v_k + w_k|^{p-1} + |v_k| \, |v_k + w_k|^{p-1} \leqslant \\ &\leqslant \|v\|_p \cdot \left(|v_k + w_k|^{(p-1)q} \right)^{1/q} + \|w\|_p \cdot \left(|v_k + w_k|^{(p-1)q} \right)^{1/q} = \\ &= \left(\|v\|_p + \|w\|_p \right) \, \|v + w\|_p^{p/q} \; . \end{aligned}$$

Minkowski's inequality follows. □

I.2. General Topology

2.1. The category of topological spaces

Topologies and continuous maps

2.1.1 Definition Let X be a set. By a topology on X on understands a set $\mathcal{T}$ of subsets of X such that:

(Top0) The sets X and $\emptyset$ are both elements of $\mathcal{T}$.

(Top1) The union of any collection of elements of $\mathcal{T}$ is again in $\mathcal{T}$ that means if $(U_i)_{i \in I}$ is a family of elements $U_i \in \mathcal{T}$, then $\bigcup_{i \in I} U_i \in \mathcal{T}$.

(Top2) The intersection of finitely many elements of $\mathcal{T}$ is again in $\mathcal{T}$ that means for every natural n and $U_1, \ldots, U_n \in \mathcal{T}$ one has $\bigcap_{i=1}^{n} U_i \in \mathcal{T}$.

A pair $(X, \mathcal{T})$ is a called a *topological space* when X is a set and $\mathcal{T}$ a topology on X. Moreover, a subset U of X is called *open* if $U \in \mathcal{T}$ and *closed* if $\complement_X U \in \mathcal{T}$.

2.1.2 Remarks (a) Strictly speaking, Axiom (Top0) can be derived from Axioms (Top1) and (Top2), since the union of an empty family of subsets of X coincides with $\emptyset$, and the intersection of an empty family of subsets of X coincides with X. Nevertheless, it is useful to require it, since in proofs one often shows Axiom (Top1) only for non-empty families of open sets, and Axiom (Top2) only for the case of the intersection of two open subsets. Then it is necessary to verify Axiom (Top0), too, when one wants to prove that a given set of subsets of X is a topology.

(b) When using the notation $\mathcal{T}_X$ for a topology we always mean that $\mathcal{T}_X$ is a topology on the space X.

2.1.3 Examples (a) For every set X the power set $\mathcal{P}(X)$ is a topology on X. It is called the *discrete* or *strongest* topology on X.

(b) The set $\{\emptyset, X\}$ is another topology on a set X called the *indiscrete* or *trivial* or *weakest* topology on X. Unless X is empty or has only one element, the discrete and indiscrete topologies differ.

(c) Let S be a set $\{0, 1\}$. Then the set $\{\emptyset, \{1\}, \{0, 1\}\}$ is a topology on S which does neither coincide with the discrete nor the indiscrete topology. The set S with this topology is called *Sierpiński space*. The closed sets of the Sierpiński space are $\emptyset$, $\{0\}$ and S.

(d) The *standard topology* on the set of real numbers $\mathbb{R}$ consists of all subsets $U \subset \mathbb{R}$ such that for each $x \in U$ there are real numbers a, b satisfying $a < x < b$ and $(a, b) \subset U$. The standard topology on $\mathbb{R}$ will be denoted by $\mathcal{T}_{\mathbb{R}}$.

Let us show that $\mathcal{T}_{\mathbb{R}}$ is a topology on $\mathbb{R}$ indeed. Obviously $\varnothing$ and $\mathbb{R}$ are elements of $\mathcal{T}_{\mathbb{R}}$. Let $U, V \in \mathcal{T}_{\mathbb{R}}$ and $x \in U \cap V$. Then there are $a, b, c, d \in \mathbb{R}$ such that $x \in (a, b) \subset U$ and $x \in (c, d) \subset V$. Put $e := \max\{a, c\}$ and $f := \min\{b, d\}$. Then $x \in (e, f) \subset U \cap V$, which proves $U \cap V \in \mathcal{T}_{\mathbb{R}}$. If $(U_i)_{i \in I}$ is a family of elements $U_i \in \mathcal{T}_{\mathbb{R}}$ and $x \in \bigcup_{i \in I} U_i$, then there exists an $j \in I$ with $x \in U_j$. Choose $a, b \in \mathbb{R}$ such that $x \in (a, b) \subset U_j$. Then $x \in (a, b) \subset \bigcup_{i \in I} U_i$, which proves $\bigcup_{i \in I} U_i \in \mathcal{T}_{\mathbb{R}}$. If not mentioned differently, we always assume the set of real numbers to be equipped with the standard topology. The standard topology coincides with the metric topology induced by the euclidean metric on $\mathbb{R}$, see **??**. One therefore often calls $\mathcal{T}_{\mathbb{R}}$ the *euclidean topology* on $\mathbb{R}$. We will use these terms interchangeably.

(e) The *standard topology* $\mathcal{T}_{\mathbb{Q}}$ on the set of rational numbers $\mathbb{Q}$ is defined analogously. It consists of all subset $U \subset \mathbb{Q}$ such that for each $x \in U$ there exist rational numbers a, b with $a < x < b$ and $(a, b) \subset U$. Like for the reals one proves that $\mathcal{T}_{\mathbb{Q}}$ is a topology on $\mathbb{Q}$. Unless mentioned differently it is always assumed that $\mathbb{Q}$ comes equipped with the standard topology. Like for $\mathbb{R}$, the standard topology on $\mathbb{Q}$ coincides with the *euclidean topology* on $\mathbb{Q}$ which is the one induced by the euclidean metric.

(f) Let X be a set, and let $\mathcal{T}_{\mathrm{cof}}$ denote the set of all subset of X which are either empty or have finite complement in X. Then $\mathcal{T}_{\mathrm{cof}}$ is a topology on X called the *cofinite topology.*

(g) Let X be a set, and let $\mathcal{T}_{\mathrm{coc}}$ denote the set of all subset of X which are either empty or have countable complement in X. Then $\mathcal{T}_{\mathrm{coc}}$ is a topology on X called the *cocountable topology.*

(h) Let X be a (nonempty) set, $(Y, \mathcal{T})$ be a topological space, and $f : X \to Y$ a function. Define

$$f^*\mathcal{T} := f^{-1}\mathcal{T} := \{f^{-1}(U) \in \mathcal{P}(X) \mid U \in \mathcal{T}\} .$$

Then $(X, f^*\mathcal{T})$ is a topological space. One calls $f^*\mathcal{T}$ the *initial topology on X with respect to f* or the *topology on X induced by f*.

Let us verify that $f^*\mathcal{T}$ is a topology on X indeed. By $f^{-1}(Y) = X$ and $f^{-1}(\varnothing) = \varnothing$ the sets X and $\varnothing$ are in $f^*\mathcal{T}$. Now let $(V_i)_{i \in I}$ be a family of elements of $f^*\mathcal{T}$. In other words we have, for each $i \in I$, $V_i = f^{-1}(U_i)$ for some $U_i \in \mathcal{T}$. Then $U := \bigcup_{i \in I} U_i \in \mathcal{T}$ and

$$\bigcup_{i \in I} V_i = \bigcup_{i \in I} f^{-1}(U_i) = f^{-1}\Big(\bigcup_{i \in I} U_i\Big) = f^{-1}(U) \in f^*\mathcal{T} .$$

Finally, let $V_1, \ldots, V_n \in f^{-1}\mathcal{T}$. Then, by definition, there exist $U_1, \ldots, U_n \in \mathcal{T}$ such that $V_i = f^{-1}(U_i)$ for $i = 1, \ldots, n$. Thus $U := \bigcap_{i=1}^n U_i \in \mathcal{T}$ and

$$\bigcap_{i=1}^n V_i = \bigcap_{i=1}^n f^{-1}(U_i) = f^{-1}\Big(\bigcap_{i=1}^n U_i\Big) = f^{-1}(U) \in f^*\mathcal{T} .$$

(i) Let $(X, \mathcal{T})$ be a topological space, Y a (nonempty) set, and $g : X \to Y$ a function. Define $g_*\mathcal{T} \subset \mathcal{P}(Y)$ as the set of all $U \subset Y$ such that $g^{-1}(U) \in \mathcal{T}$. Then $g_*\mathcal{T}$ is a topology on Y. It is called the *final topology on Y with respect to g* or the *topology on Y induced by g*. If $g : X \to Y$

is a *quotient map* that means that g is surjective, then the final topology on Y induced by g is also called the *quotient topology on X induced by g.*

Let us show why $g_*\mathcal{T}$ is a topology on Y. Obviously, $Y, \varnothing \in g_*\mathcal{T}$. Let $(U_i)_{i\in I}$ be a family of elements of $g_*\mathcal{T}$. Then $g^{-1}(U_i) \in \mathcal{T}$ for all $i \in I$ which entails

$$g^{-1}\Big(\bigcup_{i\in I} U_i\Big) = \bigcup_{i\in I} g^{-1}(U_i) \in \mathcal{T},$$

hence $\bigcup\limits_{i\in I} U_i \in g_*\mathcal{T}$. If $U_1, \ldots U_k \in g_*\mathcal{T}$, then

$$g^{-1}(U_1 \cap \ldots \cap U_k) = \bigcap_{i=1}^{k} g^{-1}(U_i) \in \mathcal{T}.$$

So $U_1 \cap \ldots \cap U_k \in g_*\mathcal{T}$ and the claim is proved.

2.1.4 Section 2.2 on fundamental examples collects several more examples of topologies. For now, we will work out a few basic properties of topologies and their structure preserving morphisms, the continuous maps defined below.

2.1.5 Definition Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be two topological spaces and assume that $f : X \to Y$ is a function. One says that f is *continuous* if for all $U \in \mathcal{T}_Y$ the preimage $f^{-1}(U)$ is open in X. The map f is called *open* if $f(V)$ is open in Y for all $V \in \mathcal{T}_X$.

2.1.6 Example Any constant function $c : X \to Y$ between two topological spaces is continuous since the preimage of an open set in Y is either the full set X or empty depending on whether the image of c is contained in the open set or not.

2.1.7 Theorem and Definition (a) *The identity map* id_X *on a topological space* $(X, \mathcal{T}_X)$ *is continuous and open.*

(b) *Let* $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ *and* $(Z, \mathcal{T}_Z)$ *be three topological spaces. Assume that* $f : X \to Y$ *and* $g : Y \to Z$ *are maps. If* f *and* g *are both continuous, so is* $g \circ f$. *If* f *and* g *are both open, then* $g \circ f$ *is open as well.*

(c) *Topological spaces as objects together with continuous maps as morphisms form a category. It is called the* category of topological spaces *and will be denoted by* Top.

Proof. It is obvious by definition that the identity map id_X is continuous and open. Now assume that f and g are continuous and let $U \in \mathcal{T}_Z$. Then $g^{-1}(U) \in \mathcal{T}_Y$ by continuity of g. Hence $f^{-1}(g^{-1}(U)) \in \mathcal{T}_X$ by continuity of f. So $g \circ f$ is continuous. If f and g are open maps, and $V \in \mathcal{T}_X$, then $f(V) \in \mathcal{T}_Y$ and $g \circ f(V) = g(f(V)) \in \mathcal{T}_Z$. Hence the composition of two open maps is open, too. The rest of the claim follows immediately. □

Comparison of topologies

2.1.8 The initial topology $f^*\mathcal{T}_Y$ induced by a function $f : X \to Y$ between topological spaces is a subset of the topology on X if and only if f is continuous. This motivates the following definition.

2.1.9 Definition Let X be a set. Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be two topologies on X. One says that $\mathcal{T}_1$ is *finer* or *stronger* than $\mathcal{T}_2$ and $\mathcal{T}_2$ is *coarser* or *weaker* than $\mathcal{T}_1$ when $\mathcal{T}_2 \subset \mathcal{T}_1$.

2.1.10 Of course, inclusion induces an order relation on topologies on a given set. A remarkable property is that any nonempty subset of the ordered set of topologies on a given set always admits a greatest lower bound.

2.1.11 Theorem *Let X be a set. Let $\mathfrak{S}$ be a nonempty set of topologies on E. Then the set*

$$\mathcal{T}_{\mathfrak{S}} := \bigcap_{\mathcal{T}\in\mathfrak{S}} \mathcal{T} = \left\{ U \in \mathcal{P}(X) \mid U \in \mathcal{T} \text{ for all } \mathcal{T} \in \mathfrak{S} \right\}$$

is a topology on X and it is the greatest lower bound of $\mathfrak{S}$, where the order between topologies is given by inclusion. In other words, $\mathcal{T}_{\mathfrak{S}}$ is the finest topology contained in each topology from $\mathfrak{S}$.

Proof. We first show that $\mathcal{T}_{\mathfrak{S}}$ is a topology. Since each $\mathcal{T} \in \mathfrak{S}$ is a topology on X, we have $\varnothing, X \in \mathcal{T}$ for all $\mathcal{T} \in \mathfrak{S}$. Hence $\varnothing, X \in \mathcal{T}_{\mathfrak{S}}$.

Let $(U_i)_{i\in I}$ be a nonempty family of elements $U_i \in \mathcal{T}_{\mathfrak{S}}$. Let $\mathcal{T} \in \mathfrak{S}$ be arbitrary. By definition of $\mathcal{T}_{\mathfrak{S}}$, we have $U_i \in \mathcal{T}$ for all $i \in I$. Since $\mathcal{T}$ is a topology, $\bigcup_{i\in I} U_i \in \mathcal{T}$. Hence, as $\mathcal{T}$ was arbitrary, $\bigcup_{i\in I} U_i \in \mathcal{T}_{\mathfrak{S}}$.

Now, let $U_1, \ldots, U_n \in \mathcal{T}_{\mathfrak{S}}$. Let $\mathcal{T} \in \mathfrak{S}$ be arbitrary. By definition of $\mathcal{T}_{\mathfrak{S}}$, we have $U_1, \ldots, U_n \in \mathcal{T}$. Therefore, $U_1 \cap \ldots \cap U_n \in \mathcal{T}$ since $\mathcal{T}$ is a topology. Since $\mathcal{T}$ was arbitrary in $\mathfrak{S}$, we conclude that $U_1 \cap \ldots \cap U_n \in \mathcal{T}_{\mathfrak{S}}$ by definition.

So $\mathcal{T}_{\mathfrak{S}}$ is a topology on X. By construction, $\mathcal{T}_{\mathfrak{S}} \subset \mathcal{T}$ for all $\mathcal{T} \in \mathfrak{S}$, so $\mathcal{T}_{\mathfrak{S}}$ is a lower bound for $\mathfrak{S}$. Assume given a new topology $\mathcal{Q}$ on X such that $\mathcal{Q} \subset \mathcal{T}$ for all $\mathcal{T} \in \mathfrak{S}$. Let $U \in \mathcal{Q}$. Then we have $U \in \mathcal{T}$ for all $\mathcal{T} \in \mathfrak{S}$. Hence by definition $U \in \mathcal{T}_{\mathfrak{S}}$. So $\mathcal{Q} \subset \mathcal{T}_{\mathfrak{S}}$ and thus $\mathcal{T}_{\mathfrak{S}}$ is the greatest lower bound of $\mathfrak{S}$. □

2.1.12 Corollary *Let X be a set, $(Y, \mathcal{T})$ be a topological space, and $f : X \to Y$ a map. The coarsest topology on X which makes f continuous is the initial topology $f^*\mathcal{T}$.*

Proof. Let $\mathfrak{S}$ be the set of all topologies on X such that f is continuous. By definition, $f^*\mathcal{T}$ is a lower bound of $\mathfrak{S}$. Moreover, $f^*\mathcal{T} \in \mathfrak{S}$. Hence $f^*\mathcal{T}$ is the coarsest topology making the function $f : X \to Y$ continuous. □

2.1.13 Proposition *Let $(X, \mathcal{T})$ be a topological space, Y a set, and $g : X \to Y$ a map. The finest topology on Y which makes g continuous is the final topology $g_*\mathcal{T}$.*

Proof. Let $\mathcal{S}$ be a topology on Y so that $g : (X, \mathcal{T}) \to (Y, \mathcal{S})$ is continuous. Let $U \in \mathcal{S}$. Then $g^{-1}(U) \in \mathcal{T}$ by continuity of $g : (X, \mathcal{T}) \to (Y, \mathcal{S})$. Hence $U \in g_*\mathcal{T}$ by definition, and $\mathcal{S} \subset \mathcal{T}$. Since $g : (X, \mathcal{T}) \to (Y, g_*\mathcal{T})$ is continuous by definition, the claim follows. □

2.1.14 We can use Theorem 2.1.11 to define other interesting topologies. Note that trivially $\mathcal{P}(X)$ is a topology on a given set X, so given any $\mathcal{S} \subset \mathcal{P}(X)$ there is at least one topology containing $\mathcal{S}$. From this:

2.1.15 Proposition and Definition *Let X be a set, and $\mathcal{S}$ a subset of $\mathcal{P}(X)$. The greatest lower bound of the set*

$$\mathfrak{S} = \{\mathcal{T} \in \mathcal{P}(\mathcal{P}(X)) \mid \mathcal{T} \text{ is a topology on } X \,\&\, \mathcal{S} \subset \mathcal{T}\}$$

is the coarsest topology on X containing $\mathcal{S}$. We call it the topology generated by $\mathcal{S}$ on X *and denote it by $\mathcal{T}_{\mathcal{S}}$. The topology $\mathcal{T}_{\mathcal{S}}$ consists of unions of finite intersections of elements of $\mathcal{S}$ that means*

$$\mathcal{T}_{\mathcal{S}} = \left\{ U \in \mathcal{P}(X) \mid \exists J \,\forall j \in J \,\exists n_j \in \mathbb{N} \,\exists U_{j,1}, \ldots, U_{j,n_j} \in \mathcal{S} : \, U = \bigcup_{j \in J} \bigcap_{k=1}^{n_j} U_{j,k} \right\} .$$

Proof. By definition of $\mathfrak{S}$ and Theorem 2.1.11, $\mathcal{T}_{\mathfrak{S}} = \bigcap_{\mathcal{T} \in \mathfrak{S}} \mathcal{T}$ is a topology on X which contains $\mathcal{S}$. Hence $\mathcal{T}_{\mathfrak{S}}$ is an element of $\mathfrak{S}$ and a subset of any element of $\mathfrak{S}$. The first claim follows. To verify the second, observe that it suffices to show that

$$\mathcal{R} := \left\{ U \in \mathcal{P}(X) \mid \exists J \,\forall j \in J \,\exists n_j \in \mathbb{N} \,\exists U_{j,1}, \ldots, U_{j,n_j} \in \mathcal{S} : \, U = \bigcup_{j \in J} \bigcap_{k=1}^{n_j} U_{j,k} \right\}$$

is a topology. The set $\mathcal{R}$ being a topology namely entails $\mathcal{T}_{\mathcal{S}} \subset \mathcal{R}$ because $\mathcal{S} \subset \mathcal{R}$. The inclusion $\mathcal{R} \subset \mathcal{T}_{\mathcal{S}}$ is clear by definition, since $\mathcal{T}_{\mathcal{S}}$ is a topology containing $\mathcal{S}$. So let us show that $\mathcal{R}$ is a topology. Obviously $\varnothing$ and X are elements of $\mathcal{R}$ because $\bigcup_{i \in \varnothing} U_i = \varnothing$ and $\bigcap_{k=1}^{0} U_k = X$. Now assume that $(U_i)_{i \in I}$ is a family of elements of $\mathcal{R}$. Then there exists for each $i \in I$ a set J_i and for every $j \in J_i$ a natural number $n_{i,j}$ together with elements $U_{i,j,1}, \ldots, U_{i,j,n_{i,j}} \in \mathcal{S}$ such that

$$U_i = \bigcup_{j \in J_i} \bigcap_{k=1}^{n_{i,j}} U_{i,j,k} .$$

Put $J := \bigcup_{i \in I} \{i\} \times J_i$. Then

$$U := \bigcup_{i \in I} U_i = \bigcup_{i \in I} \bigcup_{j \in J_i} \bigcap_{k=1}^{n_{i,j}} U_{i,j,k} = \bigcup_{(i,j) \in J} \bigcap_{k=1}^{n_{i,j}} U_{i,j,k} \in \mathcal{R} .$$

Last assume $U_1, \ldots U_n \in \mathcal{T}$ where $n \in \mathbb{N}$. Then one can find for each $i \in \{1, \ldots, n\}$ a set J_i and for every $j \in J_i$ a natural number $n_{i,j}$ together with elements $U_{i,j,1}, \ldots, U_{i,j,n_{i,j}} \in \mathcal{S}$ such that

$$U_i = \bigcup_{j \in J_i} \bigcap_{k=1}^{n_{i,j}} U_{i,j,k} .$$

Put $J := J_1 \times \ldots \times J_n$. Then

$$U := \bigcap_{i=1}^{n} U_i = \bigcap_{i=1}^{n} \bigcup_{j \in J_i} \bigcap_{k=1}^{n_{i,j}} U_{i,j,k} = \bigcup_{(j_1, \ldots, j_n) \in J} \bigcap_{k_1=1}^{n_{1,j_1}} U_{1,j_1,k_1} \cap \ldots \cap \bigcap_{k_n=1}^{n_{n,j_n}} U_{n,j_n,k_n} \in \mathcal{R} .$$

Hence $\mathcal{R}$ is a topology, indeed, and the proposition is proved. □

2.1.16 Definition Let X be a set, and $\mathcal{T}$ a topology on X. One calls a subset $\mathcal{S} \subset \mathcal{T}$ a *subbase* (or *subbasis*) of the topology if $\mathcal{T}$ coincides with $\mathcal{T}_{\mathcal{S}}$. If in addition $X = \bigcup_{S \in \mathcal{S}} S$, the subbase $\mathcal{S}$ is said to be *adequate*.

Bases of topologies

2.1.17 When inducing a topology from a family $\mathcal{B}$ of subsets of some set X, the fact that $\mathcal{B}$ enjoys the following property greatly simplifies the description of the topology $\mathcal{T}_{\mathcal{B}}$ generated by $\mathcal{B}$.

2.1.18 Definition Let X be a set. A *base* (or *basis*) on X is a subset $\mathcal{B}$ of the powerset $\mathcal{P}(X)$ such that

(Bas1) $X = \bigcup_{B \in \mathcal{B}} B$,

(Bas2) For all $B_1, B_2 \in \mathcal{B}$ and all $x \in B_1 \cap B_2$ there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subset B_1 \cap B_2$.

The main purpose for this definition stems from the following theorem:

2.1.19 Theorem *Let X be some set. Let $\mathcal{B}$ be a base on X. Then the topology generated by $\mathcal{B}$ coincides with the set of unions of elements of $\mathcal{B}$ that means*

$$\mathcal{T}_{\mathcal{B}} = \left\{ \bigcup_{B \in \mathcal{U}} B \in \mathcal{P}(\mathcal{P}(X)) \;\middle|\; \mathcal{U} \subset \mathcal{B} \right\} .$$

Proof. Denote, for this proof, the set $\left\{ \bigcup_{B \in \mathcal{U}} B \;\middle|\; \mathcal{U} \subset \mathcal{B} \right\}$ by $\mathcal{S}$ and let us abbreviate $\mathcal{T}_{\mathcal{B}}$ by $\mathcal{T}$. We wish to prove that $\mathcal{T} = \mathcal{S}$. First, note that $\mathcal{B} \subset \mathcal{S}$ by construction. By definition, $\mathcal{B} \subset \mathcal{T}$. Since $\mathcal{T}$ is a topology, it is closed under arbitrary unions. Hence $\mathcal{S} \subset \mathcal{T}$. To prove the converse, it is sufficient to show that $\mathcal{S}$ is a topology. As it contains $\mathcal{B}$, and $\mathcal{T}$ is the smallest such topology, this will provide us with the inverse inclusion. By definition, $\bigcup_{B \in \varnothing} B = \varnothing$ and thus $\varnothing \in \mathcal{S}$. By assumption, since $\mathcal{B}$ is a base, $X = \bigcup_{B \in \mathcal{B}} B$ so $X \in \mathcal{S}$. As the union of unions of elements in $\mathcal{B}$ is a union of elements in $\mathcal{B}$, $\mathcal{S}$ is closed under abritrary unions. Now, let B_1, B_2 be elements of $\mathcal{B}$. If $B_1 \cap B_2 = \varnothing$ then $B_1 \cap B_2 \in \mathcal{S}$. Assume that B_1 and B_2 are not disjoints. Then by definition of a base, for all $x \in B_1 \cap B_2$ there exists $B_x \in \mathcal{B}$ such that $x \in B_x$ and $B_x \subset B_1 \cap B_2$. So

$$B_1 \cap B_2 = \bigcup_{x \in B_1 \cap B_2} B_x \ ,$$

and therefore, by definition, $B_1 \cap B_2 \in \mathcal{S}$. We conclude that the intersection of two arbitary elements in $\mathcal{S}$ is again in $\mathcal{S}$ by using the distributivity of the union with respect to the intersection. □

2.1.20 Definition We shall say that a base $\mathcal{B}$ on a set X is a *base for a topology* $\mathcal{T}$ on X when the smallest topology containing $\mathcal{B}$ coincides with $\mathcal{T}$, in other words when $\mathcal{T} = \mathcal{T}_{\mathcal{B}}$.

The typical usage of the preceding theorem comes from the following result.

2.1.21 Corollary *Let $\mathcal{B}$ be a base for a topology $\mathcal{T}$ on X. A subset U of X is in $\mathcal{T}$ if and only if for evry $x \in U$ there exists $B \in \mathcal{B}$ such that $x \in B$ and $B \subset U$.*

Proof. We showed that any open set for the topology $\mathcal{T}$ is a union of elements in $\mathcal{B}$. Hence if $x \in U$ for $U \in \mathcal{T}$ then there exists $B \in \mathcal{B}$ such that $x \in B$ and $B \subset U$. Conversely, if U is some subset of X such that for all $x \in U$ there exists $B_x \in \mathcal{B}$ such that $x \in B_x$ and $B_x \subset U$, then $U = \bigcup_{x \in U} B_x$ and thus $U \in \mathcal{T}$. □

The last result in this section is a useful tool for showing continuity of a map.

2.1.22 Proposition *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be two topological spaces, $\mathcal{A}$ a base for the topology $\mathcal{T}_X$ and $\mathcal{B}$ a base for the topology $\mathcal{T}_Y$. Assume further that $f : X \to Y$ is a map. Then the following are equivalent:*

(i) *The map f is continuous.*

(ii) *For every open $V \subset Y$ and all $x \in f^{-1}(V)$ there exists $A \in \mathcal{A}$ such that $x \in U$ and $f(A) \subset V$.*

(iii) *For every $B \in \mathcal{B}$ the preimage $f^{-1}(B)$ is open in X.*

Proof. Obviously, (i) implies (iii).

Assume that (iii) holds and that $V \subset Y$ is open. Let $x \in f^{-1}(V)$ and put $y = f(x)$. Then $y \in V$. Since $\mathcal{B}$ is a base for the topology $\mathcal{T}_Y$ there exists $B \in \mathcal{B}$ such that $x \in B \subset V$. By assumption $f^{-1}(B)$ is open in X and $x \in f^{-1}(B)$. Since $\mathcal{A}$ is a base for the topology $\mathcal{T}_X$, there exists $A \in \mathcal{A}$ such that $x \in A \subset f^{-1}(B)$. Since $f^{-1}(B) \subset f^{-1}(V)$, (ii) follows.

Now assume that (ii) holds true. Let $V \subset Y$ be open, and choose for every $x \in f^{-1}(V)$ a base element $A_x \in \mathcal{A}$ such that $x \in A_x \subset f^{-1}(V)$. Then $f^{-1}(V) = \bigcup_{x \in f^{-1}(V)} A_x$ which is open in X. Hence f is continuous. □

2.2. Examples and categorical constructions of topological spaces

This section provides various examples and constructions of topological spaces which will be used all along in this monograph.

The order topology

2.2.1 Proposition *Let $(X, \leqslant)$ be a totally ordered set, and assume that $\infty, -\infty$ are two symbols not in X. Define $[-\infty, \infty]_X = X \cup \{-\infty, \infty\}$ and extend $\leqslant$ to $[-\infty, \infty]_X$ by requiring $x \leqslant y$ for $x, y \in [-\infty, \infty]_X$ to hold when $x, y \in X$ and $x \leqslant y$, when $x = -\infty$, or when $y = \infty$. Then $[-\infty, \infty]_X$ together with the relation $\leqslant$ becomes a totally ordered set as well, and the embedding $X \hookrightarrow [-\infty, \infty]_X$ is order-preserving.*

Proof. By definition, the relation $\leqslant$ on $[-\infty, \infty]_X$ is reflexive, and any two elements of $[-\infty, \infty]_X$ are comparable. Also by definition, $x \leqslant -\infty$ is equivalent to $x = -\infty$ and $\infty \leqslant y$ equivalent to $y = \infty$. Since the restriction of $\leqslant$ to X is antisymmetric by assumption, $\leqslant$ therefore is an antisymmetric relation on $[-\infty, \infty]_X$. Using the definition of $\leqslant$ again one finally observes that for $x, y, z \in [-\infty, \infty]_X$ the following implications hold true.

$$\begin{aligned}
-\infty \leqslant y \,\&\, y \leqslant z &\implies -\infty \leqslant z \\
x \leqslant -\infty \,\&\, -\infty \leqslant z &\implies x = -\infty \leqslant z \\
x \leqslant y \,\&\, y \leqslant -\infty &\implies x = y = -\infty \\
x \leqslant y \,\&\, y \leqslant \infty &\implies x \leqslant \infty \\
x \leqslant \infty \,\&\, \infty \leqslant z &\implies x \leqslant \infty = z \\
\infty \leqslant y \,\&\, y \leqslant z &\implies \infty = y = z \ .
\end{aligned}$$

Since its restriction to X is already transitive, transitivity of $\leqslant$ now follows and the proposition is proved. □

2.2.2 Remark For the rest of this paragraph we always assume that an ordered set $(X, \leqslant)$ does not contain the symbols $\infty, -\infty$, and that $[-\infty, \infty]_X$ and the extended order relation $\leqslant$ are defined as in the preceding proposition.

2.2.3 Definition For a totally ordered set $(X, \leqslant)$, define *intervals* with boundaries $x, y \in [-\infty, \infty]$ as follows:

$$\begin{aligned}
(x,y) &:= (x,y)_X := \{z \in [-\infty,\infty] \mid x < z < y\} , \\
[x,y) &:= [x,y)_X := \{z \in [-\infty,\infty] \mid x \leqslant z < y\} , \\
(x,y] &:= (x,y]_X := \{z \in [-\infty,\infty] \mid x < z \leqslant y\} , \\
[x,y] &:= [x,y]_X := \{z \in [-\infty,\infty] \mid x \leqslant z \leqslant y\} .
\end{aligned}$$

The intervals $(x, y)_X$ are called *open intervals*, intervals of the form $[x, y]_X$ are called *closed intervals*, and intervals of the form $[x, y)_X$ or $(x, y]_X$ are the *half-open intervals*.

2.2.4 Remarks (a) Note that in case $x = y$ only the closed interval $[x, x]_X$ is non-empty. In case $y < x$ all the intervals $(x, y)_X$, $[x, y)_X$, $(x, y]_X$, and $[x, y]_X$ are empty.

(b) We mostly use the notation (x, y), $[x, y]$, etc. for intervals and denote the X in intervals only when otherwise some ambiguity could appear.

2.2.5 Definition Let $(X, \leqslant)$ be a totally ordered set. Then the topology generated by the set

$$\mathcal{I}_X = \{(x,y) \in \mathcal{P}(X) \mid x, y \in [-\infty,\infty] \;\&\; x \leqslant y\}$$

is called the *order topology* on X. It is usually denoted $\mathcal{T}_{(X,\leqslant)}$.

2.2.6 Proposition *Let $(X, \leqslant)$ be a totally ordered set. Then the set $\mathcal{I}_X$ is a base for the order topology on X. A subbase of the order topology is given by the set $\mathcal{S}_X$ of* rays *(x, ∞) and $(-\infty, y)$, where x, y run through the elements of X.*

Proof. Since X is totally ordered, so is $[-\infty, \infty]$. It is immediate that $(x, y) \cap (x', y') = (w, z)$ if w is the largest of x and x' and z is the smallest of y and y'. Hence $\mathcal{I}_X$ is a base of the order topology.

Since $(x, \infty) \cap (-\infty, y) = (x, y)$ for $x \leqslant y$, the set $\mathcal{S}_X$ is a subbase of the order topology. □

2.2.7 Example The standard topology on $\mathbb{R}$ from Example 2.1.3 (d) is the order topology. Likewise, the standard topology on $\mathbb{Q}$ coincides with the order topology.

2.2.8 Remark If X neither has a minimum nor a maximum, one usually denotes the space $[-\infty, \infty]$ by $\overline{X}$. This notation fits with the understanding that $\overline{}$ denotes the closure operation, because the closure of X in $[-\infty, \infty]$ with respect to the order topology coincides with the full space $[-\infty, \infty]$ under the assumptions made.

Extending the ordered set of real numbers $(\mathbb{R}, \leqslant)$ in that way gives the so-called *extended real number system* $\bar{\mathbb{R}}$.

The subspace topology

2.2.9 Proposition and Definition *Let $(X, \mathcal{T})$ be a topological space. Let $S \subset X$ and $\iota : S \hookrightarrow X$ the canonical embedding. Then initial topology $\iota^*\mathcal{T}$ coincides with*

$$\mathcal{T}_S^X := \{U \cap S \in \mathcal{P}(S) \mid U \in \mathcal{T}\} .$$

One calls $\mathcal{T}_Y^X$ the subspace *or* trace topology *on S. Sometimes one says that $\mathcal{T}_Y^X$ is the* topology induced by $(X, \mathcal{T})$.

Proof. The claim follows immediately from the definition of the initial topology $\iota^*\mathcal{T}$. □

Just as easy is the following observation:

2.2.10 Proposition *Let $(X, \mathcal{T})$ be a topological space, and $S \subset X$ a subset. Let $\mathcal{B}$ be a basis for $\mathcal{T}$. Then the set*

$$\mathcal{B}_S^X := \{B \cap S \in \mathcal{P}(S) \mid B \in \mathcal{B}\}$$

is a basis for the subspace topology on S induced by $(X, \mathcal{T})$.

Proof. Trivial exercise. □

2.2.11 Example The default topologies on $\mathbb{N}$ and Z are the subspace topologies induced by the standard topology on $\mathbb{R}$. Since $\{n\} = (n - \frac{1}{2}, n + \frac{1}{2}) \cap \mathbb{Z}$ for all $n \in \mathbb{Z}$, we see that the natural topologies on $\mathbb{N}$ and $\mathbb{Z}$ are in fact the discrete topologies. The topology on $\mathbb{Q}$ induced by the standard topology on $\mathbb{R}$ coincides with the default topology on $\mathbb{Q}$ (which is, as pointed out above, the same as the order topology).

The quotient topology

The product topology

2.2.12 Definition Let I be some nonempty set. Let us assume given a family $(X_i, \mathcal{T}_i)_{i \in I}$ of topological spaces. Consider the cartesian product $X := \prod_{i \in I} X_i$ and denote for each $j \in I$ by $\pi_j : X \to X_j$, $(x_i)_{i \in I} \mapsto x_j$ the projection on the i-th coordinate. The initial topology on X with respect to the

basic open set of the cartesian product $\prod_{i \in I} E_i$ is a set of the form $\prod_{i \in I} U_i$ where $\{i \in I : U_i \neq E_i\}$ is finite and for all $i \in I$, we have $U_i \in \mathcal{T}_i$.

2.2.13 Definition Let I be some nonempty set. Let us assume given a family $(E_i, \mathcal{T}_i)_{i \in I}$ of topological spaces. The product topology on $\prod_{i \in I} E_i$ is the smallest topology containing all the basic open sets.

2.2.14 Proposition *Let I be some nonempty set. Let us assume given a family $(E_i, \mathcal{T}_i)_{i \in I}$ of topological spaces. The collection of all basic open sets is a basis on the set $\prod_{i \in I} E_i$.*

Proof. Trivial exercise. □

2.2.15 Remark The product topology is not just the basic open sets on the cartesian products: there are many more open sets!

2.2.16 Proposition *Let I be some nonempty set. Let us assume given a family $(E_i, \mathcal{T}_i)_{i\in I}$ of topological spaces. The product topology on $\prod_{i\in I} E_i$ is the initial topology for the the set $\{p_i : i \in I\}$ where $p_i : \prod_{j\in I} E_j \to E_i$ is the canonical surjection for all $i \in I$.*

Proof. Fix $i \in I$. Let $V \in \mathcal{T}_{E_i}$. By definition, $p_i^{-1}(V) = \prod_{j\in I} U_j$ where $U_j = E_j$ for $j \in I\backslash\{i\}$, and $U_i = V$. Hence $p_i^{-1}(V)$ is open in the product topology. As V was an arbitrary open subset of E_i, the map p_i is continuous by definition. Hence, as i was arbitrary in I, the initial topology for $\{p_i : i \in I\}$ is coarser than the product topology.

Conversely, note that the product topology is generated by $\{p_i^{-1}(V) : i \in I, V \in \mathcal{T}_{E_i}\}$, so it is coarser than the initial topology for $\{p_i : i \in I\}$. This concludes this proof. $\square$

2.2.17 Corollary *Let I be some nonempty set. Let us assume given a family $(E_i, \mathcal{T}_i)_{i\in I}$ of topological spaces. Let $\mathcal{T}$ be the product topology on $F = \prod_{i\in I} E_i$. Let $(D, \mathcal{T}_D)$ be a topological space. Then $f : D \to F$ is continuous if and only if $p_i \circ f$ is continuous from $(D, \mathcal{T}_D)$ to $(E_i, \mathcal{T}_{E_i})$ for all $i \in I$, where p_i is the canonical surjection on E_i for all $i \in I$.*

Proof. We simply applied the fundamental property of initial topologies. $\square$

2.2.18 Remarks (a) The *box topology* on the cartesian product $\prod_{i\in I} X_i$ is the smallest topology containing all possible cartesian products of open sets $U_i \subset X_i$, $i \in I$. The box topology is strictly finer than the product topology when the index set is infinite and infintely many of the X_i carry a topology strictly finer than the indiscrete topology. Of course, the box and product topologies coincide otherwise, in particular when the product is finite.

(b) Since the product topology is the coarsest topology which makes the canonical projections continuous, it is the preferred and default one on cartesian products.

The metric topology

2.2.19 Definition Let X be a set. A function $d : X \times X \to \mathbb{R}_{\geqslant 0}$ is a *distance* or *metric* on X when:

(M1) For all $x, y \in X$ the relation $d(x, y) = 0$ holds true if and only if $x = y$.

(M2) The map d is *symmetric* that is one has $d(x, y) = d(y, x)$ for all $x, y \in X$.

(M3) For all $x, y, z \in X$ the *triangle inequality* $d(x, y) \leqslant d(x, z) + d(z, y)$ is satisfied.

If instead of (M1) the axiom (M1)' below is fulfilled while (M2) and (M3) are still valid, then d is called a *pseudometric* on X.

(M1)' For all $x \in X$ the equality $d(x, x) = 0$ holds true.

A pair (X, d) is a *metric space* when X is a set and d a distance on X. If d is only a pseudometric on X, one calls the pair (X, d) a *pseudometric space*.

The following is often useful.

2.2.20 Lemma *Let (X, d) be a pseudometric space. Let $x, y, z \in X$. Then*

$$|d(x,y) - d(x,z)| \leqslant d(y,z) \ .$$

Proof. Since $d(x,y) \leqslant d(x,z) + d(z,y)$ we have $d(x,y) - d(x,z) \leqslant d(z,y) = d(y,z)$. Since $d(x,z) \leqslant d(x,y) + d(y,z)$ we have $d(x,z) - d(x,y) \leqslant d(y,z)$. Hence the claim holds. □

2.2.21 Definition Let (X, d) be a pseudometric space. Let $x \in E$ and $r \in \mathbb{R}_{>0}$. The *open ball* with *center* x and *radius* r in (X, d) is the set

$$\mathbb{B}(x,r) = \mathbb{B}_r(x) = \{y \in X \mid d(x,y) < r\} \ .$$

The *closed ball* with *center* x and *radius* r is defined by

$$\overline{\mathbb{B}}(x,r) = \overline{\mathbb{B}}_r(x) = \{y \in X \mid d(x,y) \leqslant r\} \ .$$

2.2.22 Definition Let (X, d) be a pseudometric space. The *metric topology* on X induced by d is the smallest topology containing all the open balls of X.

2.2.23 Theorem *Let (X, d) be a pseudometric space. The set of all open balls on X is a basis for the metric topology on X induced by d.*

Proof. It is enough to show that the set of all open balls is a topological basis. By definition, $X = \bigcup_{x \in X} \mathbb{B}(x, 1)$. Now, let us be given $\mathbb{B}(x, r_x)$ and $\mathbb{B}(y, r_y)$ for some $x, y \in X$ and $r_x, r_y > 0$. If the intersection of these two balls is empty, we are done; let us assume that there exists $z \in \mathbb{B}(x, r_x) \cap \mathbb{B}(y, r_y)$. Let r be the smallest of $r_x - d(x,z)$ and $r_y - d(y,z)$. Let $w \in \mathbb{B}(z,r)$. Then

$$d(x,w) \leqslant d(x,z) + d(z,w) < d(x,z) + r_x - d(x,z) = r_x \ ,$$

so $w \in \mathbb{B}(x, r_x)$. Similarly, $w \in \mathbb{B}(y, r_y)$. Hence, $\mathbb{B}(z,r) \subset \mathbb{B}(x, r_x) \cap \mathbb{B}(y, r_y)$ as desired. □

The following result shows that metric topologies are minimal in the sense of making the distance functions continuous.

2.2.24 Proposition *Let (X, d) be a pseudometric space. For all $x \in X$, the function*

$$d_x : X \to \mathbb{R}_{\geqslant 0}, \quad y \mapsto d(x,y)$$

is continuous on X for the metric topology. Moreover, the metric topology is the smallest topology such that all the functions d_x, $x \in X$ are continuous.

Proof. Fix $x \in X$. To verify continuity of the maps d_x it is sufficient to show that the preimages of $[0, r)$ and (r, ∞) by d_x are open in the metric topology of X, where $r > 0$ is arbitrary. Indeed, these intervals form a subbasis for the topology of $[0, \infty)$ which we assume to carry the subspace topology induced by the order topology on $\mathbb{R}$. Let $r > 0$ be given. Then $d_x^{-1}([0,r)) = \mathbb{B}(x,r)$ by definition, so it is open. Now, let $y \in X$ such that $d(x,y) > r$. Let $\varrho_y = d(x,y) - r > 0$. If $d(w,y) < \varrho_y$ for some $w \in X$, then $d(x,y) - d(w,y) \leqslant d(x,w)$, so $d(x,w) > r$. Hence

$$\mathbb{B}(y, \rho_y) \subset d_x^{-1}((r,\infty)) \quad \text{for all } y \in d_x^{-1}((r,\infty)) \ .$$

Therefore, $d_x^{-1}((r,\infty))$ is open.

Finally, since $d_x^{-1}([0,r)) = \mathbb{B}(x,r)$ for all $x \in X$ and $r > 0$ the minimal topology making all the maps d_x continuous must indeed contain the metric topology as desired, and our proposition is proven. □

2.2.25 Remark The metric topology is the default topology on a pseudometric space.

2.2.26 There are more examples of continuous functions between metric spaces. More precisely, a natural category for metric spaces consists of metric spaces and Lipschitz maps as arrows, defined as follows.

2.2.27 Definition Let (X, d_X), (Y, d_Y) be pseudometric spaces. A function $f : X \to Y$ for which there exists an $L > 0$ such that

$$d_Y(f(x), f(y)) \leqslant L\, d_X(x,y) \quad \text{for all } x, y \in X$$

is called *Lipschitz*.

2.2.28 Definition Let (X, d_X), (Y, d_Y) be pseudometric spaces. Let $f : X \to Y$ be a Lipschitz function. Then the *Lipschitz constant* of f is defined as

$$\mathrm{Lip}(f) = \sup \left\{ \frac{d_Y(f(x), f(y))}{d_X(x,y)} \;\middle|\; x, y \in X, d(x,y) \neq 0 \right\} .$$

A Lipschitz function with Lipschitz constant $L \leqslant 1$ is called a *metric map*. If its Lipschitz constant is < 1, then the Lipschitz function is called a *contraction*.

2.2.29 Examples (a) A constant map $f : X \to Y$ between pseudometric spaces is Lipschitz with Lipschitz constant 0. If both X and Y are metric spaces and $f : X \to Y$ is Lipschitz, then $\mathrm{Lip}(f) = 0$ if and only if f is constant.

(b) The identity map $\mathrm{id}_X : X \to X$ on a pseudometric space (X, d) is Lipschitz. If d is not the zero pseudometric on X, then $\mathrm{Lip}(id_X) = 0$.

2.2.30 Proposition *Let (X, d_X), (Y, d_Y) be pseudometric spaces. If $f : X \to Y$ is a Lipschitz function, then it is continuous.*

Proof. Let L be the Lipschitz constant for f. Let $y \in Y$ and $\varepsilon > 0$. Let $x \in f^{-1}(\mathbb{B}(y,\varepsilon))$. Put $\delta_x = \frac{\varepsilon - d(f(x),y)}{L+1}$ and observe that $\delta_x > 0$. Then, for $z \in \mathbb{B}(x, \delta_x)$,

$$\begin{aligned} d_Y(f(z), y) \leqslant d_Y(f(z), f(x)) + d_Y(f(x), y) \leqslant L d_X(z,x) + d_Y(f(x), y) < \\ < \varepsilon - d_Y(f(x), y) + d_Y(f(x), y) = \varepsilon. \end{aligned}$$

Hence $f^{-1}(\mathbb{B}(y,\varepsilon))$ is open and f is continuous. □

2.2.31 Proposition and Definition *Pseudometric spaces as objects together with metric maps between them form a category* PMet *which is called the* category of pseudometric spaces. *Changing the morphism class to Lipschitz maps between pseudometric spaces gives another category which we denote* $\mathsf{PMetLip}$ *and call the* category of pseudometric spaces and Lipschitz functions. *Metric spaces together with metric or Lipschitz maps between them form full subcategories* Met *and* MetLip *of* PMet *and* $\mathsf{PMetLip}$*, respectively. They are called the* category of metric spaces *respectively the* category of metric spaces and Lipschitz functions.

Proof. The claim immediately follows from the observation that the identity map on a pseudometric space is metric and that the composition of two metric respectively Lipschitz maps is again metric respectively Lipschitz. □

2.2.32 Remark Using metric or Lipschitz maps as morphisms for categories of metric or pseudometric spaces is natural. Other, more general type of morphisms, would be uniform continuous maps, which are discussed in later sections.

Co-Finite Topologies

A potential source for counter-examples, the family of cofinite topologies is easily defined:

2.2.33 Proposition *Let E be a set. Let:*

$$\mathcal{T}_{\mathrm{cof}}(E) = \{\varnothing\} \cup \{U \subset E : \complement_E U \text{ is finite } \}.$$

Then $\mathcal{T}_{\mathrm{cof}}(E)$ is a topology on E.

Proof. By definition, $\varnothing \in \mathcal{T}_{\mathrm{cof}}(E)$. Moreover, $\complement_E E = \varnothing$ which is finite, so $E \in \mathcal{T}_{\mathrm{cof}}(E)$. Let $U, V \in \mathcal{T}_{\mathrm{cof}}(E)$. If U or V is empty then $U \cap V = \varnothing$ so $U \cap V \in \mathcal{T}_{\mathrm{cof}}(E)$. Otherwise, $\complement_E(U \cap V) = \complement_E U \cup \complement V$ which is finite, since by definition $\complement_E U$ and $\complement_E V$ are finite. Hence $U \cap V \in \mathcal{T}_{\mathrm{cof}}(E)$. Last, let $\mathcal{U} \subset \mathcal{T}_{\mathrm{cof}}(E)$. Again, if $\mathcal{U} = \{\varnothing\}$ then $\bigcup \mathcal{U} = \varnothing \in \mathcal{T}_{\mathrm{cof}}(E)$. Let us now assume that $\mathcal{U}$ contains at least one nonempty set V. Then:

$$\complement_E \bigcup \mathcal{U} = \bigcap \{\complement_E U : U \in \mathcal{U}\} \subset \complement_E V.$$

Since $\complement_E V$ is finite by definition, so is $\bigcup \mathcal{U}$, which is therefore in $\mathcal{T}_{\mathrm{cof}}(E)$. This completes our proof. □

The one-point compactification of $\mathbb{N}$

Limits of sequences is a central tool in topology and this section introduces the natural topology for this concept. The general notion of limit is the subject of the next chapter.

2.2.34 Definition Let ∞ be some symbol not found in $\mathbb{N}$. We define $\bar{\mathbb{N}}$ to be $\mathbb{N} \cup \{\infty\}$.

2.2.35 Proposition *The set:*

$$\mathcal{T}_{\bar{\mathbb{N}}} = \{U \subset \bar{\mathbb{N}} : (U \subset \mathbb{N}) \vee (\infty \in U \wedge \complement_{\mathbb{N}} U \text{ is finite})\}$$

is a topology on $\bar{\mathbb{N}}$.

Proof. By definition, $\varnothing \subset \mathbb{N}$ so $\varnothing \in \mathcal{T}_{\bar{\mathbb{N}}}$. Moreover $\complement_{\bar{\mathbb{N}}} \bar{\mathbb{N}} = \varnothing$ which has cardinal 0 so $\bar{\mathbb{N}} \in \mathcal{T}_{\bar{\mathbb{N}}}$. Let $U, V \in \mathcal{T}_{\bar{\mathbb{N}}}$. If either U or V is a subset of $\mathbb{N}$ then $U \cap V$ is a subset of $\mathbb{N}$ so $U \cap V \in \mathcal{T}_{\bar{\mathbb{N}}}$. Othwiwse, $\infty \in U \cap V$. Yet $\complement_{\bar{\mathbb{N}}}(U \cap V) = \complement_{\bar{\mathbb{N}}} U \cup \complement_{\bar{\mathbb{N}}} V$ which is finite as a finite union of finite sets. Hence $U \cap V \in \mathcal{T}_{\bar{\mathbb{N}}}$ again.

Last, assume that $\mathcal{U} \subset \mathcal{T}_{\bar{\mathbb{N}}}$. Of course, $\infty \in \bigcup \mathcal{U}$ if and only if $\infty \in U$ for some $U \in \mathcal{U}$. So, if $\infty \notin \bigcup \mathcal{U}$ then $\bigcup \mathcal{U} \in \mathcal{T}_{\bar{\mathbb{N}}}$ by definition. If, on the other hand, $\infty \in \bigcup \mathcal{U}$, then there exists $U \in \mathcal{U}$ with $\complement_{\bar{\mathbb{N}}} U$ finite. Now, $\complement_{\bar{\mathbb{N}}} \bigcup \mathcal{U} = \bigcap \{\complement_{\bar{\mathbb{N}}} V : V \in \mathcal{U}\} \subset \complement_{\bar{\mathbb{N}}} U$ so it is finite, and thus again $\bigcup \mathcal{U} \in \mathcal{T}_{\bar{\mathbb{N}}}$. □

2.3. Separation properties

2.3.1 The general definition of a topology allows for examples where elements of a topological space, seen as a set, can not be distinguished from each other by open sets (for instance if the topology is indiscrete). When points can be topologically differentiated, a topology is in some sense separated. The standard separation axioms allow to subsume topological spaces with certain separability properties in particular classes. One then studies the properties of these clases, often with a view to particular applications, and attempts to create counter examples, meaning examples not satsifying the corresponding separation axioms. The most important separability property goes back to the founder of set-theoretic topology, Felix Hausdorff, who introduced it in 1914. The first full presentation of the separation axioms as we know them today appeared in the classic book *Topologie* by Alexandroff and Hopf (1965) under their German name *Trennungsaxiome.*

Let us note that the literature on separation axioms is not uniform when it comes to the axioms (T3) to (T6) below, so one needs to always check which convention an author follows. Here, we follow the convention by (Steen and Seebach, 1995, Part I, Chap. 2) which coincides with the one of

2.3.2 Definition (The Separation Axioms) Recall that two subsets A, B of a topological space $(X, \mathcal{T})$ are called *disjoint* if $A \cap B = \varnothing$. The two sets are called *separated* if $\bar{A} \cap B = A \cap \bar{B} = \varnothing$. The topological space $(X, \mathcal{T})$ now is said to be

(T0) or *Kolmogorov* if for each pair of distinct points $x, y \in X$ there is an open $U \subset X$ such that $x \in U$ and $y \notin U$ holds true, or $y \in U$ and $x \notin U$,

(T1) or *Fréchet* if for each pair of distinct points $x, y \in X$ there is an open $U \subset X$ such that $x \in U$ and $y \notin U$,

(T2) or *Hausdorff* if for each pair of distinct points $x, y \in X$ there exist disjoint open sets $U, V \subset X$ such that $x \in U$ and $y \in V$,

(T$2\frac{1}{2}$) or *Uryson* or *completely Hausdorff* if for each pair of distinct points $x, y \in X$ there exist distinct closed neigborhoods U of x and V of y,

(T3) if for each point $x \in X$ and closed subset $A \subset X$ with $x \notin A$ there exist disjoint open sets $U, V \subset X$ such that $x \in U$ and $A \subset V$,

(T$3\frac{1}{2}$) if for each point $x \in X$ and closed subset $A \subset X$ with $x \notin A$ there exists a continuous function $f : X \to \mathbb{R}$ such that $f(x) = 0$ and $f(A) = \{1\}$,

(T4) if for each pair of closed disjoint subsets $A, B \subset X$ there exist disjoint open sets $U, V \subset X$ such that $A \subset U$ and $B \subset V$,

(T5) if for each pair of separated subsets $A, B \subset X$ there exist disjoint open sets $U, V \subset X$ such that $A \subset U$ and $B \subset V$,

(T6) if for each pair of disjoint closed subsets $A, B \subset X$ there exists a continuous function $f : X \to \mathbb{R}$ such that $A = f^{-1}(0)$ and $B = f^{-1}(0)$.

A Hausdorff space will be called *regular* if it fulfills (T3) , *completely regular*, if it satisfies ($\mathsf{T3}_{\frac{1}{2}}$), and *normal* if (T4) holds true. Finally we call a Hausdorff space *completely normal* if it is (T5) and *perfectly normal* if it is (T6) .

2.4. Filters and convergence

Filters and ultrafilters

2.4.1 Definition Let X be a set. A subset $\mathcal{F}$ of the powerset $\mathcal{P}(X)$ is called a *filter* on X if it satisfies the following axioms:

(Fil1) The empty set $\varnothing$ is not an element of $\mathcal{F}$.

(Fil2) The set X is an element of $\mathcal{F}$.

(Fil3) If $A \in \mathcal{F}$ and if $B \in \mathcal{P}(X)$ satisfies $A \subset B$, then $B \in \mathcal{F}$.

(Fil4) If $A \in \mathcal{F}$ and $B \in \mathcal{F}$, then the intersection $A \cap B$ is an element of $\mathcal{F}$ as well.

If $\mathcal{F}_1$ and $\mathcal{F}_2$ are two filters on X such that $\mathcal{F}_1 \subset \mathcal{F}_2$, then one calls $\mathcal{F}_1$ a *subfilter* of $\mathcal{F}_2$ or says that $\mathcal{F}_2$ is *finer* than $\mathcal{F}_1$. Sometimes one expresses this by saying that $\mathcal{F}_2$ *refines* $\mathcal{F}_1$. Filters maximal with respect to set inclusion are called *ultrafilters*. A filter $\mathcal{F}$ is called *free* if $\bigcap_{A\in\mathcal{F}} A = \varnothing$ otherwise it is called *fixed*.

2.4.2 Examples (a) For every set X, the set $\{X\}$ is a filter. It is the smallest of all filters on X.

(b) Given an element $x \in X$ the set $\mathcal{F}_x := \{A \in \mathcal{P}(X) \mid x \in A\}$ is an ultrafilter on X. More generally, if $Y \subset X$ is a non-empty subset, then $\mathcal{F}_Y := \{A \in \mathcal{P}(X) \mid Y \subset A\}$ is a filter on X. It is an ultrafilter if and only if Y has exactly one element.

(c) If $(X, \mathcal{T})$ is a topological space and $x \in X$ an element, then the *neigborhood filter* $\mathcal{U}_x := \{V \in \mathcal{P}(X) \mid \exists U \in \mathcal{T} : \ x \in U \subset V\}$ is a filter contained in $\mathcal{F}_x$. The filters $\mathcal{U}_x$ and $\mathcal{F}_x$ coincide if and only if x is an isolated point.

(d) Now consider the reals and let $\mathcal{F} = \{A \in \mathcal{P}(\mathbb{R}) \mid \exists\, \varepsilon > 0 : \ [0, \varepsilon) \subset A\}$. Then $\mathcal{F}$ is a filter on $\mathbb{R}$ which is properly contained in the ultrafilter $\mathcal{F}_0$ and which properly contains the neighborhood filter $\mathcal{U}_0$ (where $\mathbb{R}$ carries the standard topology).

2.4.3 Proposition *Let $\mathcal{A} \subset \mathcal{P}(X)$ be a non-empty set of subset of X which has the* finite intersection property *that is that $A_1 \cap \ldots \cap A_n$ is non-empty for all $n \in \mathbb{N}^*$ and all $A_1, \ldots, A_n \in \mathcal{A}$. Then there is an ultrafilter $\mathcal{F}$ containing $\mathcal{A}$.*

Proof. Let P be the set of all $\mathcal{J} \subset \mathcal{P}(X)$ having the finite intersection property and containing $\mathcal{A}$. Then P is non-empty, as it contains at least $\mathcal{A}$, and is ordered by set inclusion. If $C \subset P$ is a chain, then $\mathcal{M} := \bigcup_{\mathcal{J}\in C} \mathcal{J}$ contains $\mathcal{A}$ and fulfills the finite intersection property. To verify the latter let $Y_1, \ldots, Y_n \in \mathcal{M}$. Then there exist $\mathcal{J}_1, \ldots, \mathcal{J}_n \in C$ such that $Y_i \in \mathcal{J}_i$ for $i = 1, \ldots, n$. Hence all Y_i lie in the maximum $\mathcal{J}_{\mathrm{m}}$ of the sets $\mathcal{J}_1, \ldots, \mathcal{J}_n$. But $\mathcal{J}_{\mathrm{m}}$ has the finite intersection property, hence $Y_1 \cap \ldots \cap Y_n \neq \varnothing$. So $\mathcal{M}$ is an upper bound of the chain C. By Zorn's Lemma,

P has a maximal element $\mathcal{F}$. It contains $\mathcal{A}$ and has the finite intersection property. Moreover, if $A \in \mathcal{F}$ and $B \in \mathcal{P}(X)$ contains A as a subset, then $\mathcal{F} \cup \{B\}$ also satisfies the finite intersection property, hence by maximality of $\mathcal{F}$ one concludes $B \in \mathcal{F}$. Again by maximality $\mathcal{F}$ has to be an ultrafilter. □

2.4.4 Corollary *Every filter on X is contained in an ultrafilter.*

Proof. This follows from the preceding proposition since a filter has the finite intersection property. □

2.4.5 Theorem *Let $\mathcal{F}$ be a filter on a set X. Then the following are equivalent:*

(i) *$\mathcal{F}$ is an ultrafilter.*

(ii) *If A is a subset of X and A has non-empty intersection with every element of $\mathcal{F}$, then $A \in \mathcal{F}$.*

(iii) *For all $A \subset X$ either $A \in \mathcal{F}$ or $X \backslash A \in \mathcal{F}$.*

Convergence of filters

2.4.6 Definition

2.5. Nets

Directed sets

Let us first recall that by a *preordered set* one understands a set P together with a binary relation $\leqslant$ which is reflexive and transitive, see **??**.

2.5.1 Definition (Directed sets) By a directed set one understands a preordered set $(P, \leqslant)$ such that the binary relation $\leqslant$ is *directed* which means that

(Dir) for all $x, y \in D$ there exists an element $z \in D$ with $x \leqslant z$ and $y \leqslant z$.

2.5.2 Remark The property that $(P, \leqslant)$ is directed is the same as saying that any two elements of the preordered set P have an upper bound.

2.6. Compactness

Quasi-compact topological spaces

2.6.1 Before we come to defining quasi-compactness let us recall some relevant notation. By a *cover* (or *covering*) of a set X one understands a family $\mathcal{U} = (U_i)_{i \in I}$ of subsets $U_i \subset X$ such that $X \subset \bigcup_{i \in I} U_i$. This terminology also holds for a subset $Y \subset X$. That is a family $\mathcal{U} = (U_i)_{i \in I}$ of subsets $U_i \subset X$ is called a *cover* of Y if $Y \subset \bigcup_{i \in I} U_i$. A *subcover* of a cover $\mathcal{U} = (U_i)_{i \in I}$ of Y or shortly a subcover of $\mathcal{U}$ then is a subfamily $(U_i)_{i\in J}$ which also covers Y which means that $J \subset I$ and $Y \subset \bigcup_{i\in J} U_i$. If J is finite, one calls the subcover $(U_i)_{i\in J}$ a *finite subcover.* If $(X, \mathcal{T})$ is a topological space and all elements U_i of a cover $\mathcal{U} = (U_i)_{i\in I}$ of some $Y \subset X$ are open sets, the cover is called an *open cover* of Y.

2.6.2 Proposition *Let be a topological spaces $(X, \mathcal{T})$. Then the following are equivalent:*

(i) *Every open cover of X has a finite subcover.*

(ii) *For every family $(A_i)_{i\in I}$ of closed subset $A_i \subset X$ such that $\bigcap_{i\in I} A_i = \varnothing$ there exist finitely many elements $A_{i_1}, \ldots, A_{i_n}$ such that $A_{i_1} \cap \ldots \cap A_{i_n} = \varnothing$.*

(iii) *Every filter on X has an accummulation point.*

(iv) *Every ultrafilter on X converges.*

Proof. Assume that (i) holds true and let $(A_i)_{i\in I}$ be a family of closed subset $A_i \subset X$ such that $\bigcap_{i\in I} A_i = \varnothing$. Put $U_i := X\backslash A_i$ for all $i \in I$. Then $(U_i)_{i\in I}$ is an open covering of X, hence by assumption there exist $i_1, \ldots, i_n \in I$ such that $X = U_{i_1} \cup \ldots \cup U_{i_n}$. By de Morgan's laws the relation $A_{i_1} \cap \ldots \cap A_{i_n} = \varnothing$ the follows, hence (ii) follows.

Next assume (ii), and let $\mathcal{F}$ be a filter on X. Then $\overline{A_1} \cap \ldots \cap \overline{A_n} \neq \varnothing$ for all $n \in \mathbb{N}^*$ and $A_1, \ldots, A_n \in \mathcal{F}$, since $\mathcal{F}$ is a filter. Hence $\bigcap_{A\in\mathcal{F}} \bar{A} \neq \varnothing$ by (ii). Every element of $\bigcap_{A\in\mathcal{F}} \bar{A}$ now is an accummulation point of $\mathcal{F}$, so (iii) follows.

By **??**, (iii) implies (iv).

Finally assume that every ultrafilter on X converges, and let $\mathcal{U} = (U_i)_{i\in I}$ be an open cover of X. Assume that $\mathcal{U}$ has no finite subcover. For each finite subset $J \subset I$ the set $B_J := X\backslash \bigcup_{i\in J} U_i$ then is non-empty, hence $\mathcal{B} := \{B_J \in \mathcal{P}(X) \mid J \subset I \;\&\; \#J < \infty\}$ is a filter base. Let $\mathcal{F}$ be an ultrafilter containing $\mathcal{B}$. By assumption $\mathcal{F}$ converges to some $x \in X$. Since $\mathcal{U}$ is an open covering of X there is some U_i with $x \in U_i$, hence U_i since $\mathcal{F}$ converges to x. On the other hand $X\backslash U_i \in \mathcal{B} \subset \mathcal{F}$ by construction. This is a contradiction, so $\mathcal{U}$ must have a finite subcover. □

2.6.3 Definition (Bourbaki (1998)[I.§9.1.)]A topological space $(X, \mathcal{T})$ is called *quasi-compact*, if every filter on X has an accummulation point.

2.6.4 Theorem (Alexander Subbase Theorem) *Let $(X, \mathcal{T})$ be a topological space, and $\mathcal{S}$ an* adequate *subbase of the topology that is a subbase of $\mathcal{T}$ such that $X = \bigcup_{S\in\mathcal{S}} S$. If every cover of X by elements of $\mathcal{S}$ has a finite subcover, the topological space $(X, \mathcal{T})$ is quasi-compact.*

Compact topological spaces

2.7. The compact-open topology on function spaces

Let X and Y be topological spaces. We denote the set of all functions from Y to X by X^Y. This is the same thing as the direct product $\prod_Y X$ of X over Y. The space of continuous functions $\mathcal{C}(Y, X)$ sits in X^Y so we can give $\mathcal{C}(Y, X)$ the product topology induced by X^Y. This is the topology of *pointwise convergence* and will not be useful for studying most function spaces.

We will instead be interested in the *compact open topology* which is the topology of *uniform convergence on compact sets.*

2.7.1 Definition Let X and Y be topological spaces. The *compact open topology* on $\mathcal{C}(Y,X)$ is the topology with subbasis given by the sets $\mathcal{V}(K,U) = \{f \in \mathcal{C}(Y,X) | f(K) \subset U\}$ for $K \subset Y$ compact and $U \subset X$ open.

2.7.2 Definition A topology $\mathcal{T}$ on $\mathcal{C}(Y,X)$ is called *admissable* if the evaluation map $e : \mathcal{C}(Y,X) \times Y \to X$, $(f,y) \mapsto f(y)$ is continuous.

2.7.3 Proposition *The compact open topology is coarser than any admissable topology on* $\mathcal{C}(Y,X)$.

Proof. Let $\mathcal{T}$ be an admissable topology on $\mathcal{C}(Y,X)$ so that the evaluation map $e : \mathcal{C}(Y,X) \times Y \to X$ is continuous. Let $K \subset Y$ be compact, $U \subset X$ be open and $f \in T(K,U)$. We have to find $V \in O$ such that $f \in V \subset T(K,U)$. Let $k \in K$. Since e is continuous and U is an open neighborhood of $f(x)$, then there are open sets $W_k \subset Y$ and $V_k \subset C_O(Y,X)$ such that $k \in W_k$, $f(k) \in V_k$ amd $e(V_k \times W_k) \subset U$. Since K is compact, there are $k_1, k_2, ..., k_l \in K$ such that $K \subset \bigcup_{i=1}^{l} W_{k_i}$. Put $V := \bigcap_{i=1}^{l} V_{k_i}$ so that $f \in V$ and V is open in O. Now take $g \in V$ and let $k \in K$. Choose i such that $k \in W_{k_i}$ and observe that $g \in W_{k_i}$ so that

$$g(k) = e(g,k) \in e(V_{k_i} \times W_{k_i} \subset U$$

Hense $g \in T(K,U)$ □

2.7.4 Theorem *If* Y *is locally compact, then the compact open topology on* $\mathcal{C}(Y,X)$ *is admissable, and it is the coarsets topology on* $\mathcal{C}(Y,X)$ *with that property.*

Proof. We have to show that

$$e : \mathcal{C}(Y,X) \times Y \to X(f,y) \mapsto f(y)$$

is continuous. Since sets of the form $T(K,U)$ form a subbasis for the compact open topology, it suffices to show that for an open neighborhood $W \subset X$ of some $e(f,y)$, there is compact $K \subset Y$, open $U \subset X$ and open $V \subset Y$ such that $e(T(K,U) \times V) \subset W$ with $f \in T(K,U)$ and $y \in V$. By assumption, and since f is continuous, there is an open neighborhood $\tilde{W}$ of y such that $f(\tilde{W}) \subset W$. By local compactness, there is an open neighborhood $V \subset Y$ of Y such that $y \in V \subset \bar{V} \subset \tilde{W}$ and $\bar{V}$ is compact. If we put $K := \bar{V}$ and $U = W$, then $e(T(K,U) \times V) \subset W$ since for $f' \in T(K,U)$ and $y' \in V$, we have $e(f',y') = f'(y') \subset W$. □

Let X, Y, Z be topological spaces. As sets, it is always true that $Z^{X \times Y} \cong Z^{Y^X}$ via the maps

$$\Phi : Z^{X \times Y} \to Z^{Y^X} f \mapsto (x \mapsto (y \mapsto f(x,y)))$$

and

$$\Psi : Z^{Y^X} \to Z^{X \times Y} g \mapsto ((x,y) \mapsto g(x)(y))$$

2.7.5 Theorem (The exponential law) *If* Y *is locally compact, then*

$$\Phi(\mathcal{C}(X \times Y), Z) \subset \mathcal{C}(X, \mathcal{C}(Y,Z))$$

and

$$\Psi(\mathcal{C}(X, \mathcal{C}(Y,Z))) \subset (\mathcal{C}(X \times Y), Z)$$

Proof. For $f \in \mathcal{C}(X \times Y, Z)$ and $x \in X$, we have to show that $\Phi(f)(x) \in \mathcal{C}(Y, Z)$ and $\Phi(f) \in \mathcal{C}(X, \mathcal{C}(Y, Z))$. $\Phi(f)(x)(y) = f \circ i_x(y) = f(x, y)$. Consider $T(K, U)$ for $K \subset Y$ compact and $U \subset X$ open. We need ot prove that the preimage $\Phi(f)^{-1}(T(K, U))$ is open in X. Let $x \in \Phi(f)^{-1}(T(K, U))$ so that $f(x,_{)} \in T(K, U)$. Hence for all $y \in K$, we have $f(x, y) \in U$. By the continuity of f, there are open neighborhoods W_y of x and V_y of y such that $f(W_y \times V_y) \subset U$. Since K us compact, there are open sets $y_1, y_2, \ldots y_k \subset Y$ such that $K \subset V_{y_1} \cup V_{y_2} \cup \ldots \cup V_{y_k}$. Put $W = W_{y_1} \cap W_{y_2} \cap \ldots \cap W_{y_k}$ so that W is a neighborhood of x and $\Phi(f)(W) \subset T(K, U)$.

Now we need to show for $g \in \mathcal{C}(X, \mathcal{C}(Y, Z))$ that $\Psi(g) \in \mathcal{C}(X \times Y, Z)$. Let $g : X \times \mathcal{C}(Y, Z)$ be continuous and assume that $U \subset Z$ be open. We have to show that $\Psi(g)^{-1}(U)$ is open. Take $(x, y) \in \Psi(g)^{-1}(U)$. Since g is continuous, there is an open neighborhood W of y such that $g(x)(W) \subset U$. Since Y is locally compact, there is an open $V \subset Y$ such that $y \in V \subset \bar{V} \subset W$ with $\bar{V}$ compact. Hence $g(x)(V) \subset g(x)(\bar{V}) \subset U$. Thus $g(x) \in T(K, U)$ so there is an open neighborhood $O \subset X$ of x such that $g(O) \subset T(\bar{V}, U)$. Therefore

$$\Psi(g)(O \times V) \subset g(O)(V) \subset g(O)(\bar{V}) \subset U$$

2.7.6 Lemma *The sets $(U^L)^K = T(K, T(L, U))$ with $K \subset X$ and $L \subset Y$ compact and $U \subset Z$ open form a subbasis for the compact open topology on $\mathcal{C}(X, \mathcal{C}(Y, Z))$.*

Proof. Let I be an index set $W_i \subset \mathcal{C}(Y, Z)$ be open and $K \subset X$ be compact.

$$T\left(K, \bigcup_I W_i\right) = \bigcup_{n \in \mathbb{N}^+} \bigcup_{\substack{K_1 \times \ldots \times K_n \subset K^n \\ K_1 \cup \ldots \cup K_n = K \\ K_i = \bar{K}_i \forall i}} \bigcup_{(i_1, \ldots, i_n) \in I^n} \bigcap_{l=1}^{n} T(K_{i_l}, W_{i_l})$$

Suppose J is a finite set. then $T\left(K, \bigcap_{j \in J} W_j\right) = \bigcap_{j \in J} T(K, W_j)$. Sets of the form $T(L, U)$ with $L \subset Y$ compact and $U \subset Z$ open form a subbasis of $\mathcal{C}(Y, Z)$, so if $W \subset \mathcal{C}(Y, Z)$ is open, we have $W = \bigcup_{i \in I} \bigcap_{j \in J_i} T(L_{i_j}, U_{i_j})$ so that

$$T(K, W) = \bigcup_{n \in \mathbb{N}^+} \bigcup_{\substack{K_1 \times \ldots \times K_n \subset K^n \\ K_1 \cup \ldots \cup K_n = K \\ K_i = \bar{K}_i \forall i}} \bigcup_{(i_1, \ldots, i_n) \in J^n} \bigcap_{l=1}^{n} \bigcap_{j \in J_{i_l}} T(K_{i_l}, T(L_{i_l j}, U_{i_l j}))$$

2.7.7 Theorem *Let X, Y, Z be topological spaces with X and Y Hausdorff and Y locally compact. Then the natural isomorphism*

$$\bar{\Phi} : \mathcal{C}(X \times Y, Z) \to \mathcal{C}(X, \mathcal{C}(Y, Z))$$

is a homeomorphism.

Proof. Let $f \in \mathcal{C}(X \times Y, Z)$ and let $W \in \mathcal{C}(X, \mathcal{C}(Y, Z))$ be an open neighborhood of $\bar{\Phi}(f)$. By 2.7.6, there is an open $U \subset Z$ and compact subsets $L \subset Y$ and $K \subset X$ such that $p\bar{h}i(f) \in T(K, T(L, U)) \subset W$. $T(K \times L, U)$ is open in $\mathcal{C}(X \times Y, Z)$ and note that $f \in T(K \times L, U)$ since for $(x, y) \in K \times L$, $\bar{\Phi}(f)(x) \in T(L, U)$ and $f(x, y) = \bar{\Phi}(f)(x)(y) \in U$.

Assume that $g \in T(K \times L, U)$. The $\bar{\Phi}(g)(x)(y) = g(x, y) =\in U$ so $\bar{\Phi}(g)(x) \in T(L, U)$ so $\bar{\Phi}(g) \in T(K, T(L, U))$, hence $\bar{\Phi}$ is continuous. Rest of proo email 9/27/10

I.3. Measure and Integration theory

3.1. The category of measurable spaces

The natural domains of measures are the so-called σ-algebras. Similarly to a topology, a σ-algebra is a particular kind of set of subsets of some given ambient space Ω. By definition, a σ-algebra contains the ambient space and is stable under taking complements and countable unions. To construct a measure one usually starts with defining it on a generating set of the σ-algebra which fulfills some weaker properties like for example is only a ring on Ω. In this section we introduce algebras and σ-algebras on sets and the related concepts of rings on sets and Dynkin systems. Crucial is the observation that together with their structure preserving maps, the measurable funtions, sets endowed with σ-algebras form a category, the category of measurable spaces.

σ-Algebras

3.1.1 Definition Let Ω be a set. A *ring on* Ω is a set $\mathcal{R}$ of subset of Ω or in other words an element $\mathcal{R} \in \mathcal{P}(\mathcal{P}(\Omega))$ which satisfies the following axioms:

(Rng1) $\varnothing \in \mathcal{R}$.

(Rng2) For all $A, B \in \mathcal{R}$ the complement $A \backslash B$ belongs to $\mathcal{R}$.

(Rng3) For all $A, B \in \mathcal{R}$ the union $A \cup B$ lies $\mathcal{R}$.

If in addition

(Rng4) $\Omega \in \mathcal{A}$,

then one calls $\mathcal{R}$ an *algebra on* Ω.

3.1.2 Proposition *Let Ω be a set and $\mathcal{A}$ a set of subsets of Ω. Then $\mathcal{A}$ is an algebra on Ω if and only if $\mathcal{A}$ has following properties:*

(Alg1) $\Omega \in \mathcal{A}$.

(Alg2) *For all $A \in \mathcal{A}$ the complement $\complement A = \Omega \backslash A$ belongs to $\mathcal{A}$.*

(Alg3) *For each finite sequence $(A_k)_{k=1}^n$ of elements of $\mathcal{A}$ the union $A = \bigcup\limits_{k=1}^n A_k$ belongs to $\mathcal{A}$.*

Proof. Assume that $\mathcal{A}$ is an algebra on Ω. Then (Alg1) holds true by definition, and (Alg2) by (Rng2) and (Alg1). Property (Alg3) follows from (Rng3) by induction.

Conversely, assume now that $\mathcal{A}$ satisfies (Alg1) to (Alg3). Properties (Rng3) and (Rng4) are immediate. Axiom (Rng1) follows by (Alg2) since $\varnothing = \complement\Omega$. Finally, (Rng2) is true since $A\backslash B$ can be written as $A \cap \complement B$ and since $\mathcal{A}$ is stable under finite intersections by de Morgan's laws. □

3.1.3 Remark Obviously, the set of rings on a set Ω, the set of algebras on Ω and the later defined set of σ-algebras on Ω are all ordered by set-theoretic inclusion. Therefore, when talking about a "smaller" ring or a "largest" σ-algebra we always implicitely mean in regard to set-theoretic inclusion as underlying order relation.

3.1.4 Examples (a) Trivial examples of rings and algebras on a set Ω are the power set $\mathcal{P}(\Omega)$ and the set $\{\varnothing, \Omega\}$. These are the largest and the smallest ring on Ω, respectively.

(b) Of fundamental importance for Lebesgue integration theory is the ring on euclidean space $\mathbb{R}^n$ generated by the n-dimensional right half-open intervals. Let us explain the construction of that ring in some detail. To be precise we mention that the dimension n is assumed to be positive. Now define for any pair of elements $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ and $b = (b_1, \ldots, b_n) \in \mathbb{R}^n$ the right half-open interval $[a, b)$ by

$$[a, b) := \left\{x = (x_1, \ldots, x_n) \in \mathbb{R}^n \;\middle|\; \forall i \in \{1, \ldots, n\} : a_i \leqslant x_i < b_i\right\} .$$

Denote by $\mathcal{I}^n$ the set of right half-open intervals in $\mathbb{R}^n$. Since for $a, b, c, d \in \mathbb{R}^n$ the intersection $[a, b) \cap [c, d)$ coincides with

$$\left\{x \in \mathbb{R}^n \;\middle|\; \forall i \in \{1, \ldots, n\} : \max\{a_i, b_i\} \leqslant x_i < \min\{b_i, d_i\}\right\} ,$$

the set $\mathcal{I}^n$ ist stable under finite intersections. The empty set is an element of $\mathcal{I}^n$ since for example for $a \in \mathbb{R}^n$ and $b = a$ the relation $[a, a) = \varnothing$ holds true. But $\mathcal{I}^n$ is not a ring since the union of finitely many right half-open intervals is in general not a right half-open interval. But one can minimally enlarge $\mathcal{I}^n$ to obtain a ring. Define $\mathcal{R}^n$ as the space of all subsets of $\mathbb{R}^n$ which can be written as the finite union of elements of $\mathcal{I}^n$. Obviously, $\mathcal{I}^n \subset \mathcal{R}^n$ which entails that $\varnothing \in \mathcal{R}^n$. Hence (Rng1) holds for $\mathcal{R}^n$. By definition, the union $A \cup B$ of two elements $A, B \in \mathcal{R}^n$ lies in $\mathcal{R}^n$ which means that (Rng3) is fulfilled. It remains to show (Rng2). To this end we proceed in steps and first prove that for elements $A, B \in \mathcal{R}^n$ the intersection $A \cap B$ is also an element of $\mathcal{R}^n$. By assumption, one can represent the two sets in the form $A = \bigcup_{i=1}^k I_i$ and $B = \bigcup_{j=1}^l J_j$ where $k, l \in \mathbb{N}^+$ and $I_1, \ldots, I_k, J_1, \ldots, J_l \in \mathcal{I}^n$. The distributivity law for sets now entails

$$A \cap B = \bigcup_{i=1}^k I_i \;\cap\; \bigcup_{j=1}^l J_j = \bigcup_{i=1}^k \bigcup_{j=1}^l I_i \cap J_j .$$

Since $I_i \cap J_j \in \mathcal{I}^n$, the intersection $A \cap B$ therefore lies in $\mathcal{R}^n$. By induction one concludes that $\mathcal{R}^n$ is stable under finite intersections. In the next step we show that $I\backslash J \in \mathcal{R}^n$ for all right half-open intervals $I = [a, b)$ and $J = [c, d)$. To avoid trivial cases we assume $a \leqslant b$ and $c \leqslant d$ that is $a_i \leqslant b_i$ and $c_i \leqslant d_i$ for all $i \in \{1, \ldots, n\}$. Note that then

$$[a_i, b_i) \setminus [c_i, d_i) = \begin{cases} [a_i, b_i) & \text{if } b_i \leqslant c_i \text{ or } d_i \leqslant a_i \;, \\ [a_i, c_i) \cup [d_i, b_i) & \text{if } c_i < b_i \text{ and } a_i < d_i \;. \end{cases}$$

Hence $[a_i, b_i) \setminus [c_i, d_i) \in \mathcal{R}^1$. By the formula

$$[a, b) \setminus [c, d) = \prod_{i=1}^{n} [a_i, b_i) \setminus [c_i, d_i)$$

and since the cartesian product distributes over union, the complement $[a, b) \setminus [c, d)$ coincides with the union of finitely many right half-open intervals, hence is in $\mathcal{R}^n$. In the final step we consider the complement $A \backslash B$ for $A, B \in \mathcal{R}^n$. Representing A, B as before one obtains by de Morgan's laws

$$A \setminus B = \bigcup_{i=1}^{k} I_i \setminus \bigcup_{j=1}^{l} J_j = \bigcap_{j=1}^{l} \Big(\bigcup_{i=1}^{k} I_i \Big) \setminus J_j = \bigcap_{j=1}^{l} \bigcup_{i=1}^{k} (I_i \setminus J_j) \ .$$

Since $\mathcal{R}^n$ is stable under finite unions and finite intersections the complement $A \backslash B$ must be in $\mathcal{R}^n$ again. This proves that $\mathcal{R}^n$ is a ring on $\mathbb{R}^n$ as claimed. By construction, $\mathcal{R}^n$ is the smallest ring on $\mathbb{R}^n$ containing the set $\mathcal{I}^n$ of right half-open intervals.

3.1.5 Definition A ring $\mathcal{R}$ on a set Ω is called a *σ-ring on* Ω if it satisfies the following condition:

(σRng3) For each sequence $(A_k)_{k \in \mathbb{N}}$ of elements of $\mathcal{R}$ the union $A = \bigcup_{k \in \mathbb{N}} A_k$ belongs to $\mathcal{R}$.

In case a set $\mathcal{A}$ of subsets of Ω is both a σ-ring and an algebra on Ω, then one calls $\mathcal{A}$ a *σ-algebra* (*on* Ω). A set Ω endowed with a σ-algebra $\mathcal{A}$ on it is referred to as a *measurable space*. The elements of the σ-algebra $\mathcal{A}$ are termed the *measurable subsets* of Ω. We will often denote measurable spaces as pairs $(\Omega, \mathcal{A})$ or $(E, \mathfrak{M})$, where the first component always denotes the underlying set and the second component the σ-algebra on it.

3.1.6 Remark It is immediate by Proposition 3.1.2 that a set $\mathcal{A}$ of subsets of some Ω is a σ-algebra on Ω if and only if $\mathcal{A}$ satisfies the following conditions:

(σAlg1) $\Omega \in \mathcal{A}$.

(σAlg2) For all $A \in \mathcal{A}$ the complement $\complement A = \Omega \backslash A$ belongs to $\mathcal{A}$.

(σAlg3) For each sequence $(A_k)_{k \in \mathbb{N}}$ of elements of $\mathcal{A}$ the union $A = \bigcup_{k \in \mathbb{N}} A_k$ belongs to $\mathcal{A}$.

This is the standard list of axioms defining a σ-algebra and we will use it from now on.

3.1.7 Proposition *If $\mathcal{A}$ is a σ-algebra on Ω, then the intersection of countably many measurable sets is also measurable.*

Proof. This follows immediately from the axioms and the set-theoretic de Morgan's laws. □

3.1.8 Examples (a) Let Ω be any set. Then the power set of Ω is a σ-algebra. The set $\{\emptyset, \Omega\}$ is also a σ-algebra. These are the largest and smallest σ-algebra on Ω, respectively.

(b) Let Ω be any set. Let $\mathcal{A}$ be the set of all sets $A \subset \Omega$ such that A or $\Omega \backslash A$ is a countable set. Then $\mathcal{A}$ is a σ-algebra.

(c) Let E be a set, $(\Omega, \mathcal{A})$ a measurable space, and $f : E \to \Omega$ a map. Then the set

$$f^{-1}(\mathcal{A}) := \{M \in \mathcal{P}(E) \mid \exists A \in \mathcal{A} : \ M = f^{-1}(A)\}$$

is a σ-algebra on E called the *preimage* of $\mathcal{A}$ under f.

The following two results are extremely useful when constructing examples.

3.1.9 Proposition *Let $\mathfrak{S}$ be a non-empty set of algebras on a set Ω. Then the intersection*

$$\mathcal{A}_{\mathfrak{S}} = \bigcap_{\mathcal{A} \in \mathfrak{S}} \mathcal{A} = \{A \in \mathcal{P}(\Omega) \mid A \in \mathcal{A} \text{ for all } \mathcal{A} \in \mathfrak{S}\}$$

is an algebra on Ω. If each of the elements $\mathcal{A} \in \mathfrak{S}$ is a σ-algebra, then $\mathcal{A}_{\mathfrak{S}}$ is so, too.

Proof. Assume first that each $\mathcal{A} \in \mathfrak{S}$ is an algebra on Ω. Obviously, $\Omega \in \mathcal{A}_{\mathfrak{S}}$ because $\Omega \in \mathcal{A}$ for all $\mathcal{A} \in \mathfrak{S}$. Similarly, if $A \in \mathcal{A}_{\mathfrak{S}}$, then $A \in \mathcal{A}$, hence $\complement A \in \mathcal{A}$ for all $\mathcal{A} \in \mathfrak{S}$. Therefore, $\complement A$ lies in the intersection $\mathcal{A}_{\mathfrak{S}} = \bigcap\limits_{\mathcal{A} \in \mathfrak{S}} \mathcal{A}$. Now assume that $(A_k)_{k=1}^n$ is a finite sequence of sets belonging to $\mathcal{A}_{\mathfrak{S}}$, hence to all $\mathcal{A} \in \mathfrak{S}$. Then the union $\bigcup\limits_{k=1}^{n} A_k$ is in each of the $\mathcal{A} \in \mathfrak{S}$, hence in $\mathcal{A}_{\mathfrak{S}}$. The latter argument also works under the condition that every $\mathcal{A} \in \mathfrak{S}$ is a σ-algebra to verify that for a sequence $(A_k)_{k \in \mathbb{N}}$ in $\mathcal{A}(\mathfrak{S})$ the union $\bigcup\limits_{k \in \mathbb{N}} A_k$ lies in $\mathcal{A}_{\mathfrak{S}}$. So the proposition is proved. □

3.1.10 Corollary *Let $\mathcal{E}$ be a collection of subsets of a set Ω. Then there exists a smallest σ-algebra on Ω containing $\mathcal{E}$. It is called the* σ-algebra generated by $\mathcal{E}$ *and will be denoted by $\mathcal{A}(\mathcal{E})$.*

Proof. Let $\mathfrak{S}$ be the set of all σ-algebras on Ω which contain $\mathcal{E}$. The set of all subsets of Ω is certainly a σ-algebra, so $\mathfrak{S} \neq \emptyset$. Let $\mathcal{A}(\mathcal{E})$ be the intersection of all σ-algebras in the set $\mathfrak{S}$. By the preceding proposition $\mathcal{A}(\mathcal{E})$ is a σ-algebra. Since every element of $\mathfrak{S}$ contains $\mathcal{E}$, the intersection $\mathcal{A}(\mathcal{E}) = \bigcap\limits_{\mathcal{A} \in \mathfrak{S}} \mathcal{A}$ contains $\mathcal{E}$ as well. By construction, $\mathcal{A}(\mathcal{E})$ is the smallest σ-algebra with that propery. □

3.1.11 Remark Obviously, given a collection $\mathcal{E}$ of subsets of Ω there also exist a smallest ring and a smallest algebra containing $\mathcal{E}$. They are constructed analogously to the σ-algebra case and are called the *ring generated by* $\mathcal{E}$ and *algebra generated by* $\mathcal{E}$, respectively. Note that the ring $\mathcal{R}^n$ constructed in Examples 3.1.4 (b) is generated in exactly that sense by the set $\mathcal{J}^n$ of right half-open ideals in $\mathbb{R}^n$.

3.1.12 Example Let X be a topological space. The σ-algebra generated by the topology on X is called the *Borel σ-algebra* on X. Its elements are the *Borel measurable sets* or simply the *Borel sets* of X. Obviously, all open and all closed sets of X are Borel measurable, as are all countable unions of closed sets and countable intersections of open sets. We will denote the Borel σ-algebra on X by $\mathcal{A}_{\mathrm{Borel}}(X)$ or shortly by $\mathcal{A}_{\mathrm{Borel}}$.

3.1.13 Example (a) All intervals including the half-open intervals $[a, b)$ and $(a, b]$ with $a < b$ are Borel subsets of $\mathbb{R}$.

(b) The elements of the ring $\mathcal{R}^n$ constrcuted in Examples 3.1.4 (b) are Borel measurable subsets of the euclidean space $\mathbb{R}^n$.

(c) If X is a topological space with the discrete topology, then every subset of X is Borel measurable.

(d) If X is a topological space carrying the indiscrete topology $\{X, \varnothing\}$, then the σ-algebra of Borel sets coincides with the topology $\{X, \varnothing\}$.

Dynkin systems

In measure theory, one often faces the problem to check whether a system of subsets of some given set is a σ-algebra. To address that problem the following concept going back to EUGENE DYNKIN is often useful.

3.1.14 Definition A system $\mathcal{D}$ of subset of a set Ω is called a *Dynkin system* (*in* Ω) if it has the following properties:

(Dyn1) $\Omega \in \mathcal{D}$,

(Dyn2) for all $D \in \mathcal{D}$ the complement $\complement D = \Omega \backslash D$ belongs to $\mathcal{D}$,

(Dyn3) for each sequence $(D_k)_{k\in\mathbb{N}}$ of pairwise disjoint elements of $\mathcal{D}$ the union $D = \bigcup\limits_{k\in\mathbb{N}} D_k$ belongs to $\mathcal{D}$.

3.1.15 Remark By (Dyn1) and (Dyn2), the empty set is contained in every Dynkin system $\mathcal{D}$. Moreover, (Dyn3) then entails that every Dynkin system is stable under finite unions of pairwise disjoints elements.

3.1.16 Dynkin systems on a set Ω are ordered by set-theoretic inclusion. Analogously to the case of σ-algebras one can use this observation to construct the Dynkin system generated by a collection of subsets of Ω.

3.1.17 Proposition *Let $\mathfrak{S}$ be a non-empty set of Dynkin systems on a set Ω. Then the intersection*

$$\mathcal{D}_{\mathfrak{S}} = \bigcap_{\mathcal{D}\in\mathfrak{S}} \mathcal{D} = \big\{ D \in \mathcal{P}(\Omega) \mid D \in \mathcal{D} \text{ for all } \mathcal{D} \in \mathfrak{S} \big\}$$

is a Dynkin system on Ω.

Proof. The claim follows immediately from the proof of Proposition 3.1.9. □

3.1.18 Corollary *Let $\mathcal{E}$ be a collection of subsets of a set Ω. Then there exists a smallest Dynkin system on Ω containing $\mathcal{E}$. It is called the* Dynkin system generated by $\mathcal{E}$ *and will be denoted by $\mathcal{D}(\mathcal{E})$.*

Proof. Analogously as in the proof of Corollary 3.1.10 the claim can be derived from the preceding proposition. □

3.1.19 Lemma *Let $\mathcal{D}$ be a Dynkin system on a set Ω. Then $\mathcal{D}$ is stable with respect to complements of subsets that means*

$$A \backslash B \in \mathcal{D} \quad \textit{for all } A, B \in \mathcal{D} \textit{ with } B \subset A \ .$$

Proof. By assumption, $\complement A$ and B are disjoint and elements of the Dynkin system, hence there union is also in $\mathcal{D}$. Therefore, by de Morgan's laws

$$A \backslash B = A \cap \complement B = \complement(\complement A \cup B) \in \mathcal{D} \ . \qquad \square$$

The next two results are central for the application of Dynkin system in measure theory.

3.1.20 Theorem *A Dynkin system $\mathcal{D}$ on a set Ω is a σ-algebra if and only if it is stable under binary intersections which in other words means if and only if for two elements $D_1, D_2 \in \mathcal{D}$ the intersection $D_1 \cap D_2$ belongs to $\mathcal{D}$.*

Proof. If the Dynkin system $\mathcal{D}$ is a σ-algebra, then it is obviously stable under finite intersections. So let us prove the converse and assume that $\mathcal{D}$ contains with two elements also their intersection. Axioms (σAlg1) and (σAlg2) hold trivially by definition of a Dynkin system. So it remains to show (σAlg3). By assumption on $\mathcal{D}$ and Lemma 3.1.19, the complement $A \backslash B = A \backslash (A \cap B)$ lies again in $\mathcal{D}$ whenever $A, B \in \mathcal{D}$. Moreover, since $A \cup B$ can be written as the disjoint union $(A \backslash B) \cup B$, the Dynkin system $\mathcal{D}$ therefore is stable under finite unions. Now let $(A_k)_{k \in \mathbb{N}}$ be a sequence of elements of $\mathcal{D}$. Define a new sequence $(A'_k)_{k \in \mathbb{N}}$ of elements of $\mathcal{D}$ by

$$A'_0 := A_0, \quad A'_{k+1} = A_{k+1} \backslash \bigcup_{l=0}^{k} A_l \ .$$

Note that by our previous observations the sets A'_k are all elements of $\mathcal{D}$, indeed. By induction one checks that

$$\bigcup_{l=0}^{k} A'_l = \bigcup_{l=0}^{k} A_l \quad \text{for all } k \in \mathbb{N}. \tag{3.1.1}$$

The elements of the sequence $(A'_k)_{k \in \mathbb{N}}$ are pairwise disjoint by construction, hence the union $\bigcup_{k \in \mathbb{N}} A'_k$ is an element of $\mathcal{D}$. Since by (3.1.1) the set $\bigcup_{k \in \mathbb{N}} A_k$ coincides with $\bigcup_{k \in \mathbb{N}} A'_k$, the union of the family $(A_k)_{k \in \mathbb{N}}$ lies again in $\mathcal{D}$. This proves (σAlg3), and $\mathcal{D}$ is a σ-algebra. $\square$

3.1.21 Theorem *Assume that $\mathcal{E}$ is a set of subsets of Ω which contains with each pair of elements also their intersection. Then the Dynkin system and the σ-algebra generated by $\mathcal{E}$ coincide that is*

$$\mathcal{D}(\mathcal{E}) = \mathcal{A}(\mathcal{E}) \ .$$

Proof. Since every σ-algebra is a Dynkin system and since $\mathcal{A}(\mathcal{E})$ contains $\mathcal{E}$, the inclusion $\mathcal{D}(\mathcal{E}) \subset \mathcal{A}(\mathcal{E})$ is clear by the minimality assumption of $\mathcal{D}(\mathcal{E})$. So it remains to show that $\mathcal{A}(\mathcal{E}) \subset \mathcal{D}(\mathcal{E})$. This relation follows if we can verify that $\mathcal{D}(\mathcal{E})$ is a σ-algebra. Hence by Theorem 3.1.20 it suffices to show that $\mathcal{D}(\mathcal{E})$ contains with any two elements also their intersection. To this end put for $D \in \mathcal{D}(\mathcal{E})$

$$\mathcal{D}_D := \{A \in \mathcal{P}(\Omega) \mid A \cap D \in \mathcal{D}(\mathcal{E}\} \ .$$

Let us show that $\mathcal{D}_D$ is a Dynkin system. Obviously, $\Omega \in \mathcal{D}_D$. If $A \in \mathcal{D}_D$, then $\complement A \cap D = D \backslash A = D \backslash (A \cap D) \in \mathcal{D}(\mathcal{E})$ by Lemma 3.1.19. Hence $\complement A \in \mathcal{D}_D$. If $(A_k)_{k\in\mathbb{N}}$ is a family of pairwise disjoint elements of $\mathcal{D}_D$, then $\bigcup_{k\in\mathbb{N}} A_k \in \mathcal{D}_D$ since $A_k \cap D \in \mathcal{D}(\mathcal{E})$ for all $k \in \mathbb{N}$ and therefore

$$\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) \cap D = \bigcup_{k\in\mathbb{N}} (A_k \cap D) \in \mathcal{D}(\mathcal{E}) \ .$$

By assumption on $\mathcal{E}$ and because $\mathcal{E} \subset \mathcal{D}(\mathcal{E})$ the relation $\mathcal{E} \subset \mathcal{D}_E$ holds true for all $E \in \mathcal{E}$. Since $\mathcal{D}_E$ is a Dynkin system, this entails $\mathcal{D}(\mathcal{E}) \subset \mathcal{D}_E$ for all $E \in \mathcal{E}$. Given $D \in \mathcal{D}(\mathcal{E})$ one concludes that $E \cap D \in \mathcal{D}(\mathcal{E})$ for all $E \in \mathcal{E}$, hence $\mathcal{E} \subset \mathcal{D}_D$ and $\mathcal{D}(\mathcal{E}) \subset \mathcal{D}_D$. But that means that $\mathcal{D}(\mathcal{E})$ is stable under binary intersections and the claim is proved. □

Application of this results leads to an important observation about the σ-algebra generated by the right half-open intervals in euclidean space.

3.1.22 Proposition *The Dynkin system generated by the set $\mathcal{J}^n$ of right half-open intervals in $\mathbb{R}^n$ is a σ-algebra and coincides with the Borel σ-algebra $\mathcal{A}_{\mathrm{Borel}}(\mathbb{R}^n)$. More precisely,*

$$\mathcal{J}^n \subset \mathcal{R}^n \subset \mathcal{D}(\mathcal{J}^n) = \mathcal{A}(\mathcal{J}^n) = \mathcal{A}_{\mathrm{Borel}}(\mathbb{R}^n) \ ,$$

where $\mathcal{R}^n$ denotes the ring on $\mathbb{R}^n$ generated by $\mathcal{J}^n$.

Proof. Obviously $\mathcal{J}^n \subset \mathcal{A}_{\mathrm{Borel}}(\mathbb{R}^n)$. Moreover, $\mathcal{J}^n$ is stable under finite intersections as shown in Examples 3.1.4 (b) . By Theorem 3.1.21, the claim now is proved when we can show that the Borel σ-algebra on $\mathbb{R}^n$ is generated by $\mathcal{J}^n$. But that is clear since for all $a, b \in \mathbb{R}^n$ the open interval

$$(a, b) := \big\{x \in \mathbb{R}^n \bigm| \forall i \in \{1, \ldots, n\} \colon a_i < x_i < b_i\big\}$$

can be written as the union of the countable family $\big([a - \frac{1}{k}, b)\big)_{k\in\mathbb{N}^+}$ of elements of $\mathcal{J}^n$, and since the set of all open intervals $(a, b) \subset \mathbb{R}^n$ with $a, b \in \mathbb{Q}^n$ is a countable basis of the topology of $\mathbb{R}^n$. □

Measurable functions

3.1.23 Definition Let $(\Omega, \mathcal{A})$ and $(E, \mathfrak{M})$ be two measurable spaces. A map $f \colon \Omega \to E$ is termed *measurable* if the set $f^{-1}(M)$ is measurable for every measurable set $M \subset E$ that is if $f^{-1}(M) \in \mathcal{A}$ for all $M \in \mathfrak{M}$.

3.1.24 Remark Let $(\Omega, \mathcal{A})$ be a measurable space. A real or complex valued function $f : \Omega \to \mathbb{K}$ with $K = \mathbb{R}$ or $= \mathbb{C}$ is understood to be measurable if it is measurable when $\mathbb{K}$ is equipped with the Borel σ-algebra $\mathcal{A}_{\mathrm{Borel}}(\mathbb{K})$.

3.1.25 Example Let $(\Omega, \mathcal{A})$ be a measurable space, and let $A \subset \Omega$ be a measurable set. Then the function $\chi_A \colon \Omega \to \mathbb{R}$ given by the formula

$$\chi_A(x) = \begin{cases} 1 & \text{for } x \in A, \\ 0 & \text{for } x \notin A, \end{cases}$$

is measurable since for each Borel set $B \subset \mathbb{R}$ the preimage $\chi_A^{-1}(B)$ is either $\emptyset$, A, $\complement A$ or $\mathbb{R}$, depending on whether $0, 1 \notin B$, $1 \in B$ but $0 \notin B$, $0 \in B$ but $1 \notin B$ or $0, 1 \in B$, respectively. The function χ_A is called the *characteristic function* of A.

3.1.26 Theorem and Definition (a) *The identity map* id_Ω *on a measurable space* $(\Omega, \mathcal{A})$ *is measurable.*

(b) *Let* $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$ *and* $(\Omega_3, \mathcal{A}_3)$ *be measurable spaces. Assume that* $f : \Omega_1 \to \Omega_2$ *and* $g : \Omega_2 \to \Omega_3$ *are maps. If* f *and* g *are both measurable, so is* $g \circ f : \Omega_1 \to \Omega_3$.

(c) *Measurable spaces as objects together with measurable maps as morphisms form a category. It is called the* category of measurable spaces *and will be denoted by* $\mathsf{Measble}$.

Proof. By definition the identity map id_Ω is measurable. Under the assumption that f and g are measurable let $A \in \mathcal{A}_3$. Then $g^{-1}(A) \in \mathcal{A}_2$ since g is measurable. Hence $f^{-1}(g^{-1}(A)) \in \mathcal{A}_1$ since f is measurable. Therefore, the composition $g \circ f$ is measurable. The claim follows. $\square$

3.1.27 Proposition *Let* $(\Omega_i, \mathcal{A}_i)$ *for* $i = 1, 2$ *be measurable spaces and assume that the* σ*-algebra* $\mathcal{A}_2$ *is generated by the set* $\mathcal{E} \subset \mathcal{A}_2$. *Then a map* $f : \Omega_1 \to \Omega_2$ *is measurable if and only for all* $E \in \mathcal{E}$ *the preimage* $f^{-1}(E)$ *is measurable.*

Proof. If f is measurable, then $f^{-1}(E) \in \mathcal{A}_1$ for all $E \in \mathcal{E}$ by definition of measurability. It remains to prove the converse. So assume that $f^{-1}(E) \in \mathcal{A}_1$ for all $E \in \mathcal{E}$. Then the set

$$\mathfrak{M} = \left\{ A \in \mathcal{P}(\Omega_2) \mid f^{-1}(A) \in \mathcal{A}_1 \right\}$$

is a σ-algebra since it contains Ω_2 and is stable under complements and countable unions. Moreover, $\mathcal{E} \subset \mathfrak{M}$ by assumption. Since $\mathcal{A}_2$ is generated by $\mathcal{E}$, the relation $\mathcal{A}_2 \subset \mathfrak{M}$ follows and the claim is proved. $\square$

3.1.28 Corollary *Let* $(\Omega, \mathcal{A})$ *be a measurable space and* $f : X \to \mathbb{R}$ *a function. Then the following are equivalent:*

(i) f *is measurable with respect to the Borel* σ*-algebra on* $\mathbb{R}$.

(ii) *The preimage* $f^{-1}(O)$ *of any open subset* $O \subset \mathbb{R}$ *is measurable.*

(iii) *The preimage* $f^{-1}(A)$ *of any closed subset* $A \subset \mathbb{R}$ *is measurable.*

(iv) *The preimage* $f^{-1}((a, b))$ *of any open interval* $(a, b) \subset \mathbb{R}$ *with* $a, b \in \mathbb{R}$ *is measurable.*

(v) *The preimage* $f^{-1}([a, b])$ *of any closed interval* $[a, b] \subset \mathbb{R}$ *with* $a, b \in \mathbb{R}$ *is measurable.*

(vi) *The preimage* $f^{-1}([a, b))$ *of any right half-open interval* $[a, b) \subset \mathbb{R}$ *with* $a, b \in \mathbb{R}$ *is measurable.*

(vii) *The preimage* $f^{-1}((a, b])$ *of any left half-open interval* $(a, b] \subset \mathbb{R}$ *with* $a, b \in \mathbb{R}$ *is measurable.*

Proof. The equivalence of (i) and (ii) follows from the preceding proposition since the open sets generate the Borel σ-algebra on $\mathbb{R}$. Likewise (i) and (iii) are equivalent because the closed subsets of $\mathbb{R}$ also generate the Borel σ-algebra. For the other equivalences it suffices to show that the sets of open intervals, of closed intervals and of right respectively of left half-open intervals each generate the Borel σ-algebra on $\mathbb{R}$. Since every open set in $\mathbb{R}$ is the countable union of open intervals, this is clear for the set of open intervals. An open interval of the form (a, b) can be written as the countable union $\bigcup_{n=1}^\infty [a + \frac{1}{n}, b - \frac{1}{n}]$, which implies that the closed bounded intervals generate the Borel σ-algebra. Similarly, $(a, b) = \bigcup_{n=1}^\infty [a + \frac{1}{n}, b) = \bigcup_{n=1}^\infty (a, b - \frac{1}{n}]$, which entails that the set of right half-open intervals and the set of left half-open intervals each generate the Borel σ-algebra. $\square$

3.1.29 Definition Let $f\colon X \to Y$ be a mapping between topological spaces. If f is measurable with respect to the Borel σ-algebras on X and Y, respectively, then one calls f *Borel measurable* or a *Borel function*.

3.1.30 Example By Proposition 3.1.27, a continuous function $f : X \to Y$ between topological spaces is Borel measurable.

Categorical constructions

3.1.31 Proposition and Definition *Let $(\Omega_i, \mathcal{A}_i)$, $i \in I := \{1,2\}$ be two measurable spaces. Denote by $\Omega_1 \sqcup \Omega_2 = \bigcup_{i\in I}\{(p,i) \in (\Omega_1 \cup \Omega_2) \times I \mid p \in \Omega_i\}$ their disjoint union and by $\Omega_1 \times \Omega_2$ their cartesian product. Let $\mathcal{A}_1 \coprod \mathcal{A}_2$ be the σ-algebra generated by the collection of disjoint unions $A_1 \sqcup A_2 \subset \Omega_1 \sqcup \Omega_2$ and $\mathcal{A}_1 \prod \mathcal{A}_2$ the σ-algebra generated by the set of cartesian products $A_1 \times A_2 \subset \Omega_1 \times \Omega_2$, where in both cases A_i runs through the elements of $\mathcal{A}_i$ for both $i \in I$. Then the pairs $(\Omega_1, \mathcal{A}_1) \coprod (\Omega_2, \mathcal{A}_2) := \big(\Omega_1 \sqcup \Omega_2, \mathcal{A}_1 \coprod \mathcal{A}_2\big)$ and $(\Omega_1, \mathcal{A}_1) \prod (\Omega_2, \mathcal{A}_2) := \big(\Omega_1 \times \Omega_2, \mathcal{A}_1 \prod \mathcal{A}_2\big)$ form the categorical coproduct and product, respectively, of $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ within the category of measurable spaces.*

Proof. By construction, $(\Omega_1, \mathcal{A}_1) \coprod (\Omega_2, \mathcal{A}_2)$ and $(\Omega_1, \mathcal{A}_1) \prod (\Omega_2, \mathcal{A}_2)$ are measurable spaces, so it remains to show that they fulfill the universal properties of the coproduct and product, respectively. To this end observe first that for $i \in I$ we have natural maps

$$\iota_{\Omega_i} = \iota_i \colon\ : \Omega_i \hookrightarrow \Omega_1 \sqcup \Omega_2, \quad p \mapsto (p,i)$$

and

$$\pi_{\Omega_i} = \pi_i \colon\ \Omega_1 \times \Omega_2 \to \Omega_i, \quad (p_1,p_2) \mapsto p_i\ .$$

The injections ι_1 and ι_2 are measurable by Proposition 3.1.27 and the construction of $\mathcal{A}_1 \coprod \mathcal{A}_2$. The projections π_1 and π_2 are measurable by definition of $\mathcal{A}_1 \times \mathcal{A}_2$.

Now assume that $(E, \mathfrak{M})$ is a measurable space and that there are measurable maps $g_i : \Omega_i \to E$. Define $g : \Omega_1 \sqcup \Omega_2 \to E$ by $(p,i) \mapsto g(p,i) = g_i(p)$. Then g is measurable by Proposition 3.1.27 since for $A_i \in \mathcal{A}_i$ the preimage $g^{-1}(A_i \times \{i\}) = g_i^{-1}(A_i)$ is measurable. Moreover, $g \circ \iota_i = g_i$ for $i \in I$, and g is the only map with that property. Hence, $(\Omega_1, \mathcal{A}_1) \coprod (\Omega_2, \mathcal{A}_2)$ together with the maps ι_i, $i \in I$ is the categorical coproduct in the category Measbl.

Finally assume that we are given measurable maps $f_i : E \to \Omega_i$. Define $f : E \to \Omega_1 \times \Omega_2$ by $e \mapsto (f_1(e), f_2(e))$. Since for all $A_1 \in \mathcal{A}$ and $A_2 \in \mathcal{A}_2$ the preimage $f^{-1}(A_1 \times A_2)$ coincides with the intersection $f_1^{-1}(A_1) \cap f_2^{-1}(A_2)$, the map f is measurable by measurability of the f_i and by Proposition 3.1.27. Clearly, $\pi_i \circ f = f_i$ for $i \in I$, and f is the only map having that property. Altogethr this proves that $(\Omega_1, \mathcal{A}_1) \prod (\Omega_2, \mathcal{A}_2)$ together with the maps π_i, $i \in I$ is the categorical product in the category Measbl. $\square$

3.1.32 Remarks (a) The unique map g associated to the measurable maps $g_i : \Omega_i \to E$, $i = 1,2$ in the universal property of the coproduct will sometimes be denoted by $[g_1, g_2] : \Omega_1 \sqcup \Omega_2 \to E$. The unique map f associated to the measurable maps $f_i : E \to \Omega_i$, $i = 1,2$ in the universal property of the product will often be written as a pair (f_1, f_2) or sometimes as $\langle f_1, f_2 \rangle$.

(b) Assume to be given two measurable functions $f_i : (\Omega_i, \mathcal{A}_i) \to (E_i, \mathfrak{M}_i)$ with $i = 1, 2$. By the universal properties of the product and coproduct there exist uniquely determined measurable functions $f_1 \sqcup f_2 : \Omega_1 \sqcup \Omega_2 \to E_1 \sqcup E_2$ and $f_1 \times f_2 : \Omega_1 \times \Omega_2 \to E_1 \times E_2$ making the following diagrams for $i = 1, 2$ commute:

$$\begin{tikzcd} \Omega_i \arrow[r, "\iota_{\Omega_i}"] \arrow[d, "f_i"'] & \Omega_1 \sqcup \Omega_2 \arrow[d, "f_1 \sqcup f_2"] \\ E_i \arrow[r, "\iota_{E_i}"'] & E_1 \sqcup E_2 \end{tikzcd} \qquad \begin{tikzcd} \Omega_1 \times \Omega_2 \arrow[r, "\pi_{\Omega_i}"] \arrow[d, "f_1 \times f_2"'] & \Omega_i \arrow[d, "f_i"] \\ E_1 \times E_2 \arrow[r, "\pi_{E_i}"'] & E_i \end{tikzcd}$$

These diagrams can be understood as functoriality relations for the coproduct and product in Measbl, respectively.

(c) If X and Y are topological spaces, the product σ-algebra $\mathcal{A}_{\text{Borel}}(X \times Y)$ of the product topological space $X \times Y$ coincides with the product σ-algebra of the Borel σ-algebras $\mathcal{A}_{\text{Borel}}(X)$ and $\mathcal{A}_{\text{Borel}}(Y)$ since the product sets $U \times V$ form a basis of the topology on $X \times Y$ and a generating system of the product σ-algebra $\mathcal{A}_{\text{Borel}}(X) \prod \mathcal{A}_{\text{Borel}}(Y)$ when U and V run through the open sets of X and Y, respectively.

Algebras of real and complex valued measurable functions

3.1.33 Proposition *Let $\mathbb{K}$ be the field of real or complex numbers and $(\Omega, \mathcal{A})$ a measurable space. Endow $\mathbb{K}$ with the Borel σ-algebra. Then the set $\mathcal{M}(\Omega, \mathbb{K})$ of measurable $\mathbb{K}$-valued functions becomes an algebra over $\mathbb{K}$ with pointwise addition, poitwise scalar multiplication and pointwise multiplication of functions as structure operations.*

Proof. It suffices to show that $\mathcal{M}(\Omega, \mathbb{K})$ is a subalgebra of the algebra $\mathcal{F}(\Omega, \mathbb{K})$ of $\mathbb{K}$-valued functions on Ω. More precisely, we therefore only need to show that for $f, f_1, f_2 \in \mathcal{M}(\Omega, \mathbb{K})$ and $\lambda \in \mathbb{K}$ the functions $f_1 + f_2$, λf and $f_1 \cdot f_2$ are measurable again. To this end recall that the functions

$$\begin{aligned} \alpha\colon\ & \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \quad (x, y) \mapsto x + y\ , \\ \mu\colon\ & \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \quad (x, y) \mapsto x \cdot y\ , \text{ and} \\ \mu_\lambda\colon\ & \mathbb{K} \to \mathbb{K}, \quad x \mapsto \lambda \cdot y \end{aligned}$$

are continuous, hence Borel measurable. Since the map f is measurable by assumption and $(f_1, f_2) : \Omega \to \mathbb{K} \times \mathbb{K}$, $p \mapsto (f_1(p), f_2(p))$ by the universal property of the product in Measbl, the compositions $\lambda f = \mu_\lambda \circ f$, $f_1 + f_2 = \alpha \circ (f_1, f_2)$ and $f_1 \cdot f_2 = \mu \circ (f_1, f_2)$ are all measurable. □

3.1.34 Proposition *Let $f\colon \Omega \to \mathbb{C}$ be a function on a measurable space $(\Omega, \mathcal{A})$. Then the function f is measurable if and only if the functions $\mathfrak{Re}(f)$ and $\mathfrak{Im}(f)$ are measurable.*

Proof. Since the maps $\mathfrak{Re} : \mathbb{C} \cong \mathbb{R}^2 \to \mathbb{R}$ and $\mathfrak{Im} : \mathbb{C} \cong \mathbb{R}^2 \to \mathbb{R}$ are the projections onto the first and second coordinate, respectively, and since f can be identified with the pair $(\mathfrak{Re}(f), \mathfrak{Im}(f))$, the claim is essentially a consequence of the universal property of the product in the category Measbl. □

3.1.35 Proposition *Let $f\colon \Omega \to \mathbb{C}$ be a measurable function on the measurable space $(\Omega, \mathcal{A})$. Then the function $|f|$ is measurable, and there is a measurable function $\alpha\colon \Omega \to \mathbb{C}$ having image in $\mathbb{S}^1$ such that $f = \alpha|f|$.*

Proof. Since the absolut value $|\cdot|\colon \mathbb{C} \to \mathbb{R}_{\geqslant 0}$ is continuous, hence Borel measurable, the composition $|f|$ is measurable by assumption on f. Let $E = \{p \in \Omega \mid f(p) = 0\}$. Then the set E is the inverse image of a closed subset, and so measurable. We can define a continuous function $\varphi\colon \mathbb{C}\backslash\{0\} \to \mathbb{C}$ by the formula $\varphi(z) = z/|z|$. By measurability of φ it follows from Proposition 3.1.33 that the function $\alpha\colon \Omega \to \mathbb{C}$ defined by the formula $\alpha = \varphi \circ (f + \chi_E)$ is measurable. The formulae $|\alpha(p)| = 1$ for all $p \in \Omega$ and $f = \alpha|f|$ are immediate by construction. □

Measurable numerical functions

3.1.36 Definition Let (a_n) be a sequence of real numbers. Then we define

$$\limsup_{n\to\infty} a_n = \limsup_{n\to\infty}\{a_n, a_{n+1}, a_{n+2}, \ldots\}$$

and

$$\liminf_{n\to\infty} a_n = \liminf_{n\to\infty}\{a_n, a_{n+1}, a_{n+2}, \ldots\}$$

We can pass from results about lim sup to results about lim inf, or conversely, by the observation

$$\limsup_{n\to\infty} a_n = -\liminf_{n\to\infty}(-a_n)$$

It will occasionally be convenient to us to allow ∞ and $-\infty$ as values of limits and functions. This is a safe enough option provided we do not attempt to do arithmetic with these symbols; for example, expressions such as '$\infty - \infty$' are completely meaningless.

However, we can form 'intervals'

$$[a, \infty] = [a, \infty) \cup \{\infty\} \qquad [\infty, b] = (\infty, b] \cup \{\infty\}$$

and so on. These intervals are topological spaces. We can also allow ourselves the inequality

$$-\infty < a < \infty$$

for all $a \in \mathbb{R}$. The standard result about lim sup and lim inf can now be expressed quite simply; although a number of special cases need to be examined in the proof.

3.1.37 Theorem *Let (a_n) be a real-valued sequence. Then the limits*

$$\liminf_{n\to\infty} a_n \in [-\infty, \infty) \qquad \limsup_{n\to\infty} a_n \in (-\infty, \infty]$$

exist and satisfy the inequality

$$\liminf_{n\to\infty} a_n \leqslant \limsup_{n\to\infty} a_n$$

Further, the equality

$$\liminf_{n\to\infty} a_n = a = \limsup_{n\to\infty} a_n$$

holds precisely when the sequence (a_n) converges to the real number a. ***proof to be filled in!***

Note that the number a in the above result must be finite.

3.1.38 Proposition *Let Ω be a measurable space, and let $f\colon \Omega \to [\infty, \infty]$ be any map. Suppose that the inverse image $f^{-1}((\alpha, \infty])$ is measurable for every point $\alpha \in \mathbb{R}$. Then the function f is measurable.*

Proof. Let

$$\mathcal{M} = \{E \subseteq [-\infty, \infty] \mid f^{-1}[E] \text{ is measurable } \}$$

By proposition **??** the set $\mathcal{M}$ is a σ-algebra. Choose points $\alpha \in \mathbb{R}$ and $\alpha_n < \alpha$ such that $\lim_{n\to\infty} \alpha_n = \alpha$. Since the set $(\alpha_n, \infty]$ is measurable by hypothesis, and

$$[-\infty, \alpha) = \bigcup_{n=1}^{\infty} [-\infty, \alpha_n] = \bigcup_{n=1}^{\infty} [-\infty, \infty] \backslash (\alpha_n, \infty]$$

it follows that $[-\infty, \alpha) \in \Omega$. Hence

$$(\alpha, \beta) = [-\infty, \beta) \cap (\alpha, \infty] \in \Omega$$

for every point $\alpha, \beta \in \mathbb{R}$. Since every open set in $[-\infty, \infty]$ is a countable union of such open intervals, the collection $\mathcal{M}$ contains every open set. Thus the map f is measurable. □

3.1.39 Corollary *Let $f_n\colon X \to [-\infty, \infty]$ be measurable functions for $n \in \mathbb{N}$. Then the functions*

$$\sup\{f_n\} \quad \limsup_{n\to\infty} f_n \quad \inf\{f_n\} \quad \liminf_{n\to\infty} f_n$$

are measurable.

Proof. Let $a \in \mathbb{R}$. Observe that the set

$$(\sup\{f_n\})^{-1}(a, \infty] = \bigcup_{n=1}^{\infty} f_n^{-1}(a, \infty]$$

is measurable. Hence by the above proposition, the function $\sup\{f_n\}$ is measurable. The formula $\inf\{f_n\} = -\sup\{-f_n\}$ tells us that the function $\inf\{f_n\}$ is also measurable.

Now, for each point $x \in \Omega$, the sequence of numbers

$$g_n(x) = \sup\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\}$$

is monotonic increasing. It follows that

$$\limsup_{n\to\infty} f_n(x) = \inf\{g_n(x)\}$$

We know that each function f_n is measurable. The above argument tells us that each function g_n is measurable, and that the function $\limsup_{n\to\infty} f_n$ is measurable. A similar argument tells us that the function $\liminf_{n\to\infty} g_n$ is measurable. □

3.1.40 Corollary *If $f, g\colon X \to [-\infty, \infty]$ are measurable functions, then so are the functions $\max\{f, g\}$ and $\min\{f, g\}$.* ***proof to be filled in!***

3.1.41 Corollary *The limit of a pointwise-convergent sequence of meaurable functions is measurable.* ***proof to be filled in!***

3.2. Measure Spaces

3.2.1 Definition Let Ω be a measurable space, equipped with a σ-algebra $\mathcal{A}$. A *measure* on Ω is a function $\mu\colon \mathcal{A} \to [0, \infty]$ such that:

(M1) The function μ is σ-*additive*, i.e.

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) \ ,$$

whenever $(A_n)_{n\in\mathbb{N}}$ is a sequence of disjoint mesaurable sets.

(M2) There is a measurable set A such that $\mu(A) < \infty$.

The number $\mu(A)$ is called the *measure* of a set A. A measurable space equipped with some measure is called a *measure space*.

For the above definition to make sense, we need to make a convention concerning our 'number' ∞, namely that $a + \infty = \infty$ whenever $a \in [0, \infty]$.

3.2.2 Example Let Ω be a measurable space. For any measurable set $E \subseteq \Omega$, let us define $\mu(E) = |E|$, where $|E|$ denotes the number of elements of E. Then μ is a measure on Ω, called the *counting measure*.

3.2.3 Example Let Ω be a measurable space, and let $x_0 \in \Omega$. For any measurable set $E \subseteq \Omega$, let us define

$$\mu(E) = \begin{cases} 1 & x_0 \in E \\ 0 & x_0 \notin E \end{cases}$$

Then μ is a measure on Ω, called the *Dirac measure*.

3.2.4 Proposition *Let Ω be a measure space, with measure μ. Then $\mu(\varnothing) = 0$.*

Proof. Choose a measurable set A such that $\mu(A) < \infty$. Then

$$\mu(A) = \mu(A) + \mu(\varnothing) + \mu(\varnothing) + \cdots$$

Hence $\mu(\varnothing) = 0$. □

3.2.5 Corollary *Let $A_1, \ldots, A_n$ be disjoint measurable sets. Then*

$$\mu(A_1 \cup \cdots \cup A_n) = \mu(A_1) + \cdots + \mu(A_n)$$

proof to be filled in!

3.2.6 Corollary *Let A and B be measurable set where $A \subseteq B$. Then $\mu(A) \leqslant \mu(B)$.*

Proof. The set $B \backslash A = B \cap (\Omega \backslash A)$ is measurable, the sets A and $B \backslash A$ are disjoint, and $B = A \cup B \backslash A$. By the above corollary

$$\mu(B) = \mu(A) + \mu(B \backslash A)$$

The inequality $\mu(A) \leqslant \mu(B)$ follows since $\mu(B \backslash A) \geqslant 0$. □

3.2.7 Proposition *Let (A_n) be a sequence of measurable sets such that $A_n \subseteq A_{n+1}$ for all n. Let $A = \bigcup_{n=1}^{\infty} A_n$. Then $\lim_{n \to \infty} \mu(A_n) = \mu(A)$.*

Proof. Let $B_1 = A_1$, and $B_n = A_n \backslash A_{n-1}$ when $n \geqslant 2$. Then the sets B_n are measurable and disjoint. Further

$$A_n = B_1 \cup \cdots \cup A_n \qquad A = \bigcup_{n=1}^{\infty} B_n$$

Hence

$$\mu(A) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{N \to} \sum_{n=1}^{N} \mu(B_n) = \lim_{N \to \infty} \mu(A_N)$$

3.2.8 Corollary *Let (A_n) be a sequence of measurable sets such that $\mu(A_1) < \infty$ and $A_{n+1} \subseteq A_n$ for all n. Let $A = \bigcap_{n=1}^{\infty} A_n$. Then $\lim_{n \to \infty} \mu(A_n) = \mu(A)$.*

Proof. Let $C_n = A_1 \backslash A_n$. Then the set C_n is measurable, $C_n \subseteq C_{n+1}$ for all n, and $\bigcup_{n=1}^{\infty} C_n = A_1 \backslash A$. Hence, by the above proposition

$$\lim_{n \to \infty} \mu(C_n) = \mu(A_1 \backslash A)$$

We know that the measure $\mu(A_1)$ is finite, and that we have disjoint unions

$$A_1 = A_n \cup C_n \qquad A_1 = A_1 \backslash A \cup A$$

Hence

$$\mu(A_1) - \lim_{n \to \infty} \mu(A_n) = \mu(A_1) - \mu(A)$$

and

$$\lim_{n \to \infty} \mu(A_n) = \mu(A)$$

The above corollary is false if we omit the assumption that $\mu(A_1) < \infty$.

3.3. Lebesgue integration

Simple Functions

3.3.1 Definition A function $s\colon \Omega \to \mathbb{C}$ on a measurable space Ω is called *simple* if the range of s is a finite set of points.

Let $s\colon \Omega \to \mathbb{C}$ be a simple function, with image $s[X] = \{0\} \cup \{\alpha_1, \dots, \alpha_n\}$. Write $A_i = s^{-1}(\alpha_i)$. Then clearly

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

and the function s is measurable if and only if each set A_i is measurable.

3.3.2 Proposition *Let $f\colon \Omega \to [0,\infty]$ be a measurable function. Then there are simple measurable functions $s_n\colon X \to [0,\infty)$ such that the sequence $(s_n(x))$ is monotonically increasing, with limit $f(x)$ for each point $x \in X$.*

Proof. Let $n \in \mathbb{N}$, and $t \in [0,\infty]$. Then there is a unique integer $k_n(t)$ such that

$$k_n(T)2^{-n} \leqslant t \leqslant (k_n(t)+1)2^{-n}$$

Define

$$\varphi_n(t) = \begin{cases} k_n(t)2^{-n} & 0 \leqslant t < n \\ n & n \leqslant t \leqslant \infty \end{cases}$$

The function $\varphi_n\colon [0,\infty] \to [0,\infty]$ is a Borel function, and

$$t - 2^{-n} \leqslant \varphi_n(t) \leqslant t$$

if $0 \leqslant t \leqslant n$. Thus we have a monotonically increasing sequence $(\varphi_n(t))$ with limit t. If we write $s_n = \varphi_n \circ f$, then (s_n) is a monotonically increasing sequence of simple measurable functions, with pointwise limit f as required. $\square$

We now come to the first of our definitions of the integral.

3.3.3 Definition Let Ω be a measure space, with measure μ. Let $s\colon \Omega \to \mathbb{C}$ be a measurable simple function, with set of non-zero values $\{\alpha_1, \dots, \alpha_n\}$. Write

$$s = \sum_{k=1}^{n} \alpha_k \chi_{A_k}$$

Let $E \subseteq \Omega$ be a measurable subset of Ω. Then we define the *integral* of s over E to be the complex number

$$\int_E s\ d\mu = \sum_{k=1}^{n} \alpha_k \mu(A_k \cap E)$$

There are several simple computations we can do immediately with integrals. For example, with s as above:

$$\int_\Omega s\chi_E \, d\mu = \sum_{k=1}^\infty \alpha_k \mu(A_k \cap E) = \int_E s \, d\mu$$

3.3.4 Lemma *Let Ω be a measure space, with measure μ. Let $s\colon \Omega \to [0,\infty)$ be a measurable simple function. Then we can define a new measure φ on Ω by the formula*

$$\varphi(E) = \int_E s \, d\mu$$

Proof. To begin with, observe that $\varphi(E) \geqslant 0$ for every measurable set E, and that if $\mu(E) < \infty$, then $\varphi(E) < \infty$, so there is at least one measurable set with finite measure. We need to test σ-additivity.

Let (E_n) be a sequence of disjoint measurable sets. We know that

$$\mu(\bigcup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \mu(E_i)$$

Let $\{\alpha_1, \ldots, \alpha_k\}$ be the set of non-zero values of the simple function s. Then

$$\varphi(\bigcup_{i=1}^\infty E_i) = \sum_{k=1}^n \sum_{i=1}^\infty \alpha_k \mu(A_k \cap E_i)$$

Exchanging the summation signs is possible since all of the numbers involved in the above equation are positive. Therefore

$$\varphi(\bigcup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \sum_{k=1}^n \alpha_k \mu(A_k \cap E_i) = \sum_{i=1}^\infty \varphi(E_i)$$

and we are done. $\square$

3.3.5 Proposition *Let $s, t\colon \Omega \to [0,\infty]$ be simple functions. Then*

$$\int_\Omega s + t \, d\mu = \int_\Omega s \, d\mu + \int_\Omega t \, d\mu$$

Proof. Write as usual

$$s = \sum_{i=1}^m \alpha_k \chi_{A_i} \qquad t = \sum_{j=1}^n \beta_j \chi_{B_j}$$

Let $E_{ij} = A_i \cap B_j$. Then certainly

$$int_{E_{ij}}(s+t) \, d\mu = (\alpha_i + \beta_j)\mu(E_{ij}) = \int_{E_{ij}} s \, d\mu + \int_{E_{ij}} t \, d\mu$$

Now the sets $\{0, \alpha_1, \ldots, \alpha_m\}$ and $\{0, \beta_1, \ldots, \beta_n\}$ are the ranges of the functions s and t respectively. Let $A_0 = s^{-1}[0]$ and $B_0 = t^{-1}[0]$. Then

$$\Omega = \bigcup_{i=0}^{m} A_i = \bigcup_{j=0}^{n} B_j$$

Hence

$$\Omega = \bigcup_{i,j=0}^{m,n} E_{ij}$$

The sets E_{ij} are certainly disjoint. Hence by the above lemma, we know that

$$\int_\Omega s + t \; d\mu = int_\Omega s \; d\mu + \int_\Omega t \; d\mu$$

and we are done. □

If s is a step function, and $\alpha \in \mathbb{C}$, then clearly

$$\int_\Omega \alpha s \; d\mu = \alpha \int_\Omega s \; d\mu$$

Hence we have proven linearity for integrals of positive-valued step functions.

3.4. Integration of Positive-Valued Functions

3.4.1 Definition Let Ω be a measure space, with measure μ. Let $f\colon \Omega \to [0, \infty]$ be a measurable function, and let $E \subseteq \Omega$ be a measurable set. Let S be the set of simple functions, $s\colon \Omega \to [0, \infty)$, such that $s(x) \leqslant f(x)$ for all $x \in \Omega$. Then we define the *integral* of f over E:

$$\int_E f \; d\mu = \sup\{\int_E s \; d\mu \mid s \in S\}$$

A few properties of the integral are easy to prove. For example:

- Let $f\colon \Omega \to [0, \infty]$ and $E \subseteq \Omega$ be measurable. Then

$$\int_\Omega f \; d\mu = \int_\Omega f\chi_E \; d\mu$$

- Let $f, g\colon \Omega \to [0, \infty]$ be measurable functions such that $f \leqslant g$, that is to say $f(x) \leqslant g(x)$ for all $x \in \Omega$. Then

$$\int_E f \leqslant \int_E g$$

whenever $E \subseteq \Omega$ is a measurable subset.

3.4.2 Theorem (The Monotone Convergence Theorem) *Let $f_n \colon \Omega \to [0,\infty]$ be a sequence of measurable functions, such that for each point $x \in \Omega$ the sequence $(f_n(x))$ is monotonically increasing, with limit $f(x)$. Then the function $f \colon \Omega \to [0,\infty]$ is measurable, and*

$$\int_\Omega f \ d\mu = \lim_{n\to\infty} \int_\Omega f_n \ d\mu$$

Proof. As the limit of a sequence of measurable functions, the function f is measurable. Since the seqyebce $(f_n(x))$ is monotonic increasing, with limit $f(x)$, we know that $F_n \leqslant f_{n+1} \leqslant f$ for all n. Therefore the sequence of integrals $\left(\int_\Omega f_n\right)$ is monotonic increasing, and

$$\int_\Omega f_n \ d\mu \leqslant \int_\Omega f \ d\mu$$

for all n.

Choose a simple function s such that $0 \leqslant s \leqslant f$. Let $0 < \alpha < 1$, and write

$$E_n = \{x \in \Omega \mid f_n(x) \geqslant \alpha s(x)\}$$

Each set E_n is measurable, and $E_n \subseteq E_{n+1}$ for all n since the sequence (f_n) is monotonic increasing. Since the sequence (f_n) has pointwise limit f, we see that

$$\Omega = \bigcup_{n=1}^{\infty} E_n$$

Further

$$\int_\Omega f_n \ d\mu \geqslant \int_{E_n} f_n \ d\mu \geqslant \alpha \int_{E_n} s \ d\mu \qquad (*)$$

By lemma 3.3.4 we can define a measure on the set Ω by the formula

$$\varphi(E) = \int_E s \ d\mu$$

Hence

$$\int_\Omega s \ d\mu = \varphi(\Omega) = \lim_{n\to\infty} \varphi(E_n) = \lim_{n\to\infty} \int_{E_n} s \ d\mu$$

by proposition 3.2.7.

Taking limits in inequality $(*)$, we see that

$$\lim_{n\to\infty} \int_\Omega f_n \ d\mu \geqslant \alpha \int_\Omega s \ d\mu$$

In particular, this inequality holds whenever $0 < \alpha < 1$ and $s \leqslant f$. By the definition of the integral, it follows that

$$\lim_{n\to\infty} \int_\Omega f_n \ d\mu \geqslant \int_\Omega f \ d\mu$$

and we are done. $\square$

Let $f\colon \Omega \to [0,\infty]$ be a measurable function. By proposition 3.3.2, there is a monotonically increasing sequence of simple functions $s\colon \Omega \to [0,\infty)$ with pointwise limit f.

The monotone convergence theorem tells us that

$$\int_\Omega f = \lim_{n\to\infty} \int_\Omega s_n$$

and so gives us a new way of viewing the definition of the integral. Using this viewpoint, the following result follows immediately from proposition 3.3.5

3.4.3 Corollary *Let $f, g\colon \Omega \to [0,\infty]$ be measurable functions, and let $\alpha, \beta \in [0,\infty)$. Then*

$$\int_\Omega (\alpha f + \beta g)\ d\mu = \alpha \int_\Omega f\ d\mu + \beta \int_\Omega g\ d\mu$$

proof to be filled in!

We can also immediately deduce the following result from the monotone convergence theorem.

3.4.4 Corollary *Consider a sequence of measurable functions $f_n\colon \Omega \to [0,\infty]$. Then for any measurable subset $E \subseteq \Omega$ we have the formula*

$$\sum_{n=1}^{\infty} \int_E f_n\ d\mu = \int_E \left(\sum_{n=1}^{\infty} f_n \right)\ d\mu$$

proof to be filled in!

3.4.5 Theorem (Fatou's lemma) *Let $f_n\colon \Omega \to [0,\infty]$ be a sequence of measurable functions. Then*

$$\int_\Omega \liminf_{n\to\infty} f_n \leqslant \liminf_{n\to\infty} \int_\Omega f_n$$

Proof. Let

$$g_n(x) = \inf\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\}$$

Then the function g_n is measurable, the sequence (g_n) is monotonic increasing, and the inequality $g_n \leqslant f_n$ holds for all n.

We know that

$$\lim_{n\to\infty} g_n(x) = \liminf_{n\to\infty} f_n(x)$$

Hence, by the monotone convergence theorem

$$\int_\Omega \liminf_{n\to\infty} f_n = \lim_{n\to\infty} \int_\Omega g_n \leqslant \liminf_{n\to\infty} \int_\Omega f_n$$

and we are done. □

The inequality

$$\int_\Omega \limsup_{n\to\infty} f_n \geqslant \limsup_{n\to\infty} \int_\Omega f_n$$

is easily deduced from Fatou's lemma.

3.5. Integration of Complex-Valued Functions

3.5.1 Definition Let Ω be a measure space, with measure μ. We call a measurable function $f: \Omega \to \mathbb{C}$ *integrable* if

$$\int_\Omega |f| \, d\mu < \infty$$

We write $L^1(\Omega)$ to denote the set of all integrable functions.

Suppose we have a measurable function f and a positive-valued integrable function g such that $|f| \leqslant g$. Then it follows by the above definition that the function f is integrable. This integrability criterion is often used.

3.5.2 Definition Let $f: \Omega \to \mathbb{R}$ be any real-valued function. Then we define functions $f^+, f^-: \Omega \to [0, \infty)$ by the formulae

$$f^+(x) = \max(f(x), 0)) \qquad f^-(x) = \max(-f(x), 0)$$

respectively.

Observe that $f = f^+ - f^-$. If the function f is measurable, then so are the functions f^+ and f^-.

3.5.3 Proposition *Let $f: \Omega \to \mathbb{R}$ be an integrable function. Then the functions f^+ and f^- are also integrable.*

Proof. The functions f^+ and $|f|$ are positive-valued, and $f^+ \leqslant |f|$. We know that $\int_\Omega |f| < \infty$, so $\int_\Omega f^+ < \infty$.

The proof that the function f^- is integrable is identical to the above. $\square$

3.5.4 Definition Let $f: \Omega \to \mathbb{R}$ be an integrable function. Then we define we define the integral

$$\int_\Omega f \, d\mu := \int_\Omega f^+ \, d\mu - \int_\Omega f^- \, d\mu$$

It is easy to see that definition agrees with the previous definition when the function f is positive-valued. Further, the equation

$$\int_\Omega (\alpha f + \beta g) \, d\mu = \alpha \int_\Omega f \, d\mu + \beta \int_\Omega g \, d\mu$$

holds for all real numbers $\alpha, \beta \in \mathbb{R}$ and integrable functions $f, g: \Omega \to \mathbb{R}$.

3.5.5 Definition Let $f, g: \Omega \to \mathbb{C}$ be integrable functions. Then we define the integral

$$\int_\Omega f \, d\mu := \int_\Omega \mathfrak{Re}(f) \, d\mu + i \int_\Omega \mathfrak{Im}(f) \, d\mu$$

An argument similar to that made above tells us that this integral is well-defined, agrees with the previous definition for real-valued functions, and is linear.

3.5.6 Proposition *Let $f\colon \Omega \to \mathbb{C}$ be an integrable function. Then*

$$\left|\inf_{\Omega} f \; d\mu\right| \leqslant \int_{\Omega} |f| \; d\mu$$

Proof. Choose $\alpha \in \mathbb{C}$ such that $|\alpha| = 1$ and

$$\left|\inf_{\Omega} f \; d\mu\right| = \alpha \int_{\Omega} f \; d\mu = \int_{\Omega} \alpha f \; d\mu$$

Let $g = \mathfrak{Re}(\alpha f)$ and $h = \mathfrak{Im}(\alpha f)$. Then

$$\left|\inf_{\Omega} f \; d\mu\right| = \int_{\Omega} g \; d\mu + i \int_{\Omega} h \; d\mu$$

Certainly, $|\inf_{\Omega} f \; d\mu| \in \mathbb{R}$, so

$$\int_{\Omega} h \; d\mu$$

and

$$\left|\inf_{\Omega} f \; d\mu\right| = \int_{\Omega} g \; d\mu$$

However

$$g \leqslant |g| \leqslant |\alpha f| = |f|$$

It follows that

$$\left|\inf_{\Omega} f \; d\mu\right| \leqslant \int_{\Omega} |f| \; d\mu$$

and we are done. □

Observe that the proof of the above result uses only positivity and linearity of the integral.

3.5.7 Theorem (The Dominated Convergence Theorem) *Let (f_n) be a sequence of measurable functions $f_n\colon \Omega \to \mathbb{C}$ such that:*

- *The limit*

$$f(x) = \lim_{n\to\infty} f_n(x)$$

  exists for all $x \in \Omega$.
- *There is an integrable function $g \in L^1(\Omega)$ such that $|f_n(x)| \leqslant g(x)$ for all $x \in \Omega$ and $n \in \mathbb{N}$.*

Then $f \in L^1(\Omega)$, and

$$\lim_{n\to\infty} \int_{\Omega} |f_n - f| \; d\mu = 0$$

Proof. Since each fucntion f_n is measurable, the limit function f is also measurable. We know that $|f_n| \leqslant g$ for all n. Therefore $|f| \leqslant g$. It follows that $f \in L^1(\Omega)$.

Now, let

$$h_n = 2g - |f_n - f|$$

Observe that $h_n \geqslant 0$ for all n. Hence by Fatou's lemma

$$\int_\Omega \liminf_{n\to\infty} h_n \leqslant \liminf_{n\to\infty} \int_\Omega h_n$$

that is

$$\int_\Omega 2g \; d\mu \leqslant \int_\Omega 2g \; d\mu - \liminf_{n\to\infty} \int_\Omega |f_n - f| \; d\mu$$

and so

$$\liminf_{n\to\infty} \int_\Omega |f_n - f| \; d\mu \leqslant 0$$

Since $|f_n - f| \geqslant 0$ for all n, we deduce that

$$\lim_{n\to\infty} \int_\Omega |f_n - f| \; d\mu = 0$$

as required. □

Combining the dominated convergence theorem with proposition 3.5.6 we obtain the following corollary, also referred to as the dominated convergence theorem.

3.5.8 Corollary (The Dominated Convergence Theorem) *Let (f_n) be a sequence of measurable functions $f_n \colon \Omega \to \mathbb{C}$ such that:*

- *The limit*

$$f(x) = \lim_{n\to\infty} f_n(x)$$

  exists for all $x \in \Omega$.

- *There is an integrable function $g \in L^1(\Omega)$ such that $|f_n(x)| \leqslant g(x)$ for all $x \in \Omega$ and $n \in \mathbb{N}$.*

Then $f \in L^1(\Omega)$, and

$$\lim_{n\to\infty} \int_\Omega f_n \; d\mu = \int_\Omega f \; d\mu$$

proof to be filled in!

3.6. Null Sets

3.6.1 Definition Let Ω be a measure space, with measure μ. Then a set $E \subseteq \Omega$ is called a *null set* if E is measurable, and $\mu(E) = 0$.

The measure space Ω is called *complete* if every subspace of a null set is measurable.

The usual manipulations of the axioms tell us that every measure space is contained in a unique smallest complete measure space. To be more precise, we have the following result.

3.6.2 Proposition *Let Ω be a measure space, equipped with σ-algbebra $\mathcal{M}$, and measure μ. Let us define*

$$\mathcal{M}^\star := \{E \subseteq \Omega \mid A \subseteq E \subseteq B,\ A, B \in \Omega,\ \mu(B\backslash A) = 0\}$$

Then the set $\mathcal{M}^\star$ is a σ-algebra. We can define a measure $\mu^\star$ on the set $\mathcal{M}^\star$ by writing

$$\mu^\star(E) = \mu(A) \qquad A \subseteq E \subseteq B,\ \ A, B \in \Omega,\ \ \mu(B\backslash A) = 0$$

proof to be filled in!

As we might expect from the terminology, null sets are irrelevant from the point of view of integration theory.

3.6.3 Theorem *Let $f \colon \Omega \to [0, \infty]$ be a measurable function. Then the integral of f is zero if and only if the function f is equal to zero except on a null set.*

Proof. Suppose that the set

$$N = \{x \in \Omega \mid f(x) \neq 0\}$$

is a null set. Let $s \colon \Omega \to [0, \infty]$ be a simple function such that $s \leqslant f$. Then $s(x) = 0$ when $x \notin N$. The definition of the integral of a simple function tells us that

$$\int_\Omega s \ d\mu = 0$$

The definition of the integral of a non-negative function now implies that

$$\int_\Omega f \ d\mu = 0$$

Conversely, suppose that the integral of the function f is zero. Let

$$A_n = \{x \in \Omega \mid f(x) > 1/n\}$$

Then clearly

$$\frac{1}{n}\mu(A_n) \leqslant \int_{A_n} f \ d\mu \leqslant \int_\Omega f d\mu = 0$$

so $\mu(A_n) = 0$. But

$$\{x \in \Omega \mid f(x) > 0\} = \bigcup_{n=1}^{\infty} A_n$$

Thus σ-additivity implies that the set of all points $x \in \Omega$ such that $f(x) \neq 0$ has measure zero.□

Given two functions $f, g\colon \Omega \to \mathbb{C}$, let us say that f and g are equal *almost everywhere* if they are equal outside of some set of measure zero.

3.6.4 Corollary *Let $f, g\colon \Omega \to \mathbb{C}$ be integrable functions that are equal almost everywhere. Then*

$$\int_\Omega f = \int_\Omega g$$

proof to be filled in!

3.6.5 Corollary *Let $f\colon \Omega \to \mathbb{C}$ be an integrable function. Suppose that*

$$\int_E f = 0$$

whenever the subset $E \subseteq \Omega$ is measurable. Then the function f is equal to zero almost everywhere.

Proof. Let us write

$$f(x) = u(x) + iv(x) = (u^+(x) - u^-(x)) + i(v^+(x) - v^-(x))$$

where the functions u and v are real and integrable, and the functions $u^\pm$ and $v^\pm$ are integrable and non-negative.

Let

$$E = \{x \in \Omega \mid u(x) \geqslant 0\}$$

Then

$$\mathfrak{Re}\left(\int_E f\right) = \int_E u^+ = 0$$

By the above theorem, it follows that $u^+ = 0$ except on a null set. Similarly, it follows that $u^- = 0$ except on a null set. Since the union of two null sets is also a null set, we have shown that $u = 0$ almost everywhere.

A similar argument tells us that $v = 0$ almost everywhere. We conclude that $f = 0$ almost everywhere. □

3.7. The Riesz Representation Theorem

Before we are ready to state the Riesz representation theorem, we need some terminology from point-set topology.

3.7.1 Definition Let X be a topological space. Then we define the *support* of a continuous function $f\colon X \to \mathbb{C}$ to be the closure

$$\operatorname{supp}(f) := \overline{\{x \in X \mid f(x) \neq 0\}}$$

We write $C_c(X)$ to denote the set of all continuous compactly supported functions $f\colon X \to \mathbb{C}$. The set $C_c(X)$ is a vector space under the operations of pointwise addition and scalar multiplication.

3.7.2 Definition A linear map $\Lambda\colon C_c(X) \to \mathbb{C}$ is said to be a *positive functional* if $\Lambda(f) \geqslant 0$ whenever $f \geqslant 0$.

Let X be a topological space equipped with a Borel measure μ such that $\mu(K) < \infty$ whenever $K \subseteq X$ is a compact subspace. Then the integration map

$$f \mapsto \int_X f$$

defines a positive linear functional.

The Riesz representation theorem is essentially a converse of the above observation.

3.7.3 Theorem *Let X be a locally compact Hausdorff space, and let $\Lambda\colon C_c(X) \to \mathbb{C}$ be a positive linear functional.*

Then the set X has a σ-algebra Ω containing all Borel sets, and a unique measure μ on Ω such that

$$\Lambda(f) = \int_X f \; d\mu$$

whenever $f \in C_c(X)$.

The proof of this theorem is in a series of lemmas; the proof is quite long. Before we begin the proof, let us note a theorem from general topology which we shall need.

3.7.4 Theorem *Let X be a locally compact Hausdorff space, and let $\mathcal{U} = \{U_\alpha \mid \alpha \in A\}$ be an open cover of the space X. Then there is a* partition of unity *subordinate to the cover $\mathcal{U}$, that is to say a set of continuous functions $u_\alpha\colon X \to [0,1]$ such that* $\operatorname{supp} u_\alpha \subseteq U_\alpha$ *and*

$$\sum_{\alpha \in A} u_\alpha(x) = 1$$

whenever $x \in X$. ***proof to be filled in!***

The following corollary is known as *Urysohn's lemma.*

3.7.5 Corollary *Let X be a locally compact Hausdorff space, and let $K \subseteq X$ be a compact set, and let U be an open set containing K. Then there is a continuous function $f\colon X \to [0,1]$ such that*

$$\chi_K(x) \leqslant f(x) \leqslant \chi_U(x)$$

Proof. The collection $\{U, X\backslash K\}$ is an open cover of the space X. There is therefore a partition of unity $\{f, g\}$ subordinate to this open cover.

The definition of a partition of unity gives us the required inequality for the function f. □

We now begin our proof of the Riesz representation theorem with the definition of the measure we are looking for.

3.7.6 Definition Let $\Lambda\colon C_c(X) \to \mathbb{C}$ be a positive linear functional. Let $U \subseteq X$ be open. Then we define

$$\mu(U) := \sup\{\Lambda f \mid f \leqslant \chi_U\}$$

In general, for a subset, $E \subset X$, we define

$$\mu(E) = \inf\{\{\mu(U) \mid U \text{ open }, E \subseteq U\}$$

3.7.7 Proposition *Let $f, g \in C_c(X)$, and let $f \leqslant g$. Then $\Lambda f \leqslant \Lambda g$.*

Proof. Observe $g - f \geqslant 0$. The result follows from positivity and linearity of the function Λ. □

3.7.8 Corollary *Let A and B be subsets of the space X where $A \subseteq B$. Then $\mu(A) \leqslant \mu(B)$.* ***proof to be filled in!***

Although we have defined a function μ for every subset of E, the definition is only sensible for a certain σ-algebra.

3.7.9 Definition We define Ω_F to be the sollection of all subsets $E \subseteq X$ such that $\mu(E) < \infty$ and

$$\mu(E) = \sup\{\mu(K) \mid K \subseteq E, K \text{ compact}\}$$

We define Ω to be the collection of all subsets $E \subseteq X$ such that $E \cap K \in \Omega_F$ whenever K is compact.

We need to prove that the set Ω is a σ-algebra which contains all Borel sets; this statement is not obvious.

3.7.10 Proposition *Let $V \subseteq X$ be an open subset such that $\mu(V) < \infty$. Then $V \in \Omega_F$.*

Proof. Choose a number $a < \mu(V)$. By the definition of μ, there is a function $f \in C_c(X)$ such that $f \leqslant \chi_V$ and $a < \Lambda f$. Write $K = \operatorname{supp}(f)$, and let W be an open set that contains K. Then $\Lambda f \leqslant \mu(W)$, so $\Lambda f \leqslant \mu(K)$, using the above proposition and corollary, and the definition of the function μ.

Thus $K \subseteq V$ and $\mu(K) > a$. It follows that

$$\mu(V) = \sup\{\mu(K) \mid K \subseteq E, K \text{ compact}\}$$

and we are done. □

3.7.11 Proposition *Let $U_1, \ldots, U_N \subseteq X$ be open sets. Then $\mu(U_1 \cup \cdots \cup U_N) \leqslant \mu(U_1) + \cdots + \mu(U_N)$.*

Proof. Let $N = 2$. Choose a function $g \in C_c(X)$ such that $g \leqslant \chi_{U_1 \cup U_2}$. By theorem 3.7.4 there are functions $u_1, u_2 \in C_c(X)$ such that $u_1 \leqslant \chi_{U_1}$, $u_2 \leqslant \chi_{u_2}$, and $u_1(x) + u_2(x) = 1$ whenever $x \in U_1 \cup U_2$. It follows that

$$u_1 g \leqslant \chi_{U_1},\ u_2 g \leqslant \chi_{U_2} \qquad g = u_1 g + u_2 g$$

and therefore

$$\Lambda g = \Lambda(u_1 g) + \Lambda(u_2 g) \leqslant \mu(U_1) + \mu(U_2)$$

Since the above inequality holds for every function $g \in C_c(X)$ such that $g \leqslant \chi_{U_1 \cup U_2}$, the result follows from the definition of μ when $N = 2$. The general result follows by induction. □

3.7.12 Lemma *Let $E_1, E_2, E_3, \ldots$ be subsets of the space X. Write*

$$E = \bigcup_{n=1}^{\infty} E_n$$

Then

$$\mu(E) \leqslant \sum_{n=1}^{\infty} \mu(E_n)$$

Proof. If $\mu(E_n) = \infty$ for some n, then the result is obviously true. Thus, let us suppose that $\mu(E_n) < \infty$ for all n. Choose $\varepsilon > 0$. By definition of the function μ, there are open sets $U_n \supseteq V_n$ such that

$$\mu(V_n) < \mu(E_n) + 2^{-n}\varepsilon$$

for all n.

Let $U = \bigcup_{n=1}^{\infty} U$, and choose $f \in C_c(X)$ such that $f \leqslant \chi_U$. The support of the function f is covered by the collection of sets $\{U_n \mid n = 1, 2, 3, \ldots\}$. Since the function f has compact support, it follows that it has a finite subcovering, and so

$$f \leqslant \chi_{U_1 \cup \cdots \cup U_N}$$

for some N. By the above proposition, we see that

$$\Lambda f \leqslant \mu(U_1 \cup \cdots \cup U_N) \leqslant \mu(V_1) + \cdots + \mu(V_N) \leqslant \sum_{n=1}^{\infty} \mu(E_n) + \varepsilon$$

Since the above inueqality holds for every funtion $f \subseteq \chi_U$, and $E \subseteq U$, we see that

$$\mu(E) \leqslant \sum_{n=1}^{\infty} \mu(E_n) + \varepsilon$$

But this inequality holds whenever $\varepsilon > 0$, so the result follows. □

3.7.13 Proposition *Let $K \subseteq X$ be compact. Then $\mu(K) \leqslant \Lambda f$ whenever $f \geqslant \chi_K$, and $K \in \Omega_F$.*

Proof. Let $0 < a < 1$, and choose $f \in C_c(X)$ such that $f \geqslant \chi_K$. Write

$$V_a = \{x \in X \mid f(x) > a\}$$

Then $K \subseteq V_a$, and $ag \leqslant f$ whenever $f \leqslant \chi_{V_a}$. Therefore

$$\mu(K) \leqslant \mu(V_a) = \sup\{\Lambda g \mid g \leqslant \chi_{V_a}\} \leqslant a^{-1}\Lambda f$$

Since this inequaltity holds whenever $0 < a < 1$, it follows that $\mu(K) \leqslant \Lambda f$. It follows that $\mu(K) < \infty$, and so $K \in \Omega_F$. □

3.7.14 Lemma *Let $K \subseteq X$ be compact. Then*

$$\mu(K) = \inf\{\Lambda f \mid \chi_K \leqslant f\}$$

Proof. Let $\varepsilon > 0$. Then there is an open set $U \supseteq K$ such that $\mu(U) < \mu(K) + \varepsilon$. By Urysohn's lemma there is a continuous function $f\colon [0,1] \to X$ such that $\chi_K \leqslant f \leqslant \chi_U$. It follows that

$$\Lambda f \leqslant \mu(U) < \mu(K) + \varepsilon$$

The result follows from the above inequality combined with the previous proposition. □

3.7.15 Proposition *Let $K_1, \ldots K_N$ be disjoint compact sets. Then*

$$\mu(K_1 \cup \cdots \cup K_N) \leqslant \mu(K_1) + \cdots + \mu(K_N)$$

Proof. Let $N = 2$. We can find an open set U such that $U \supseteq K_1$ and $U \cap K_2 = \varnothing$. It follows by Urysohn's lemma that we can find a compactly supported function $u\colon X \to [0,1]$ such that $u(x) = 1$ whenever $x \in K_1$, and $u(x) = 0$ whenever $x \in K_2$.

Let $\varepsilon > 0$. By lemma 3.7.14 there is a function $g \in C_c(X)$ such that

$$\chi_{K_1 \cup K_2} \leqslant g \qquad \Lambda g \leqslant \mu(K_1 + K_2) + \varepsilon$$

Observe that

$$\chi_{K_1} \leqslant fg \qquad \chi_{K_2} \leqslant (1 - f)g$$

Hence

$$\mu(K_1) + \mu(K_2) \leqslant \Lambda(fg) + \Lambda(g - fg) \leqslant \mu(K_1 \cup K_2) + \varepsilon$$

Since the above inequality holds whenever $\varepsilon > 0$, the desired result follows when $N = 2$. The general result follows by induction. □

3.7.16 Lemma *Let $E_1, E_2, E_3, \ldots$ be pairwise disjoint members of the collection Ω_F. Write*

$$E = \bigcup_{n=1}^{\infty} E_n$$

Then

$$\mu(E) = \sum_{n=1}^{\infty} \mu(E_n)$$

Further, if $\mu(E) < \infty$, then $E \in \Omega_F$.

Proof. Observe that the result follows from lemma 3.7.12 when $\mu(E) = \infty$. Let us therefore assume that $\mu(E) < \infty$. Choose $\varepsilon > 0$. Since $E_n \in \Omega_F$, we can find a compact set $K_n \subseteq E_n$ such that

$$\mu(K_n) > \mu(E_n) - 2^{-n}\varepsilon$$

for each n. Let $H_N = K_1 \cup \cdots \cup K_N$. Then by the above propositiion:

$$\mu E) \geqslant \mu(H_N) = \sum_{n=1}^{N} \mu(K_n) > \sum_{n=1}^{N} \mu(E_n) - \varepsilon$$

Since the above inequality holds whenever $\varepsilon > 0$, combining it with the inequality in lemma 3.7.12, we see that

$$\mu(E) = \sum_{n=1}^{\infty} \mu(E_n)$$

Now, if $\mu(E) < \infty$, and $\varepsilon > 0$, then we can find N such that

$$\mu(E) \leqslant \sum_{n=1}^{N} \mu(E_n) + \varepsilon$$

It follows that $\mu(E) \leqslant \mu(H_N) + 2\varepsilon$, and so $E \in \Omega_F$. □

3.7.17 Proposition *Let $E \subseteq \Omega_F$, and let $\varepsilon > 0$. Then there is a compact set K and an open set V such that $K \subseteq E \subseteq V$, and $\mu(V \backslash K) < \varepsilon$.*

Proof. by definition of the collection Ω_F, we can find a compact set $K \subseteq E$ and an open set $U \supseteq E$ such that

$$\mu(V) - \frac{\varepsilon}{2} < \mu(E) < \mu(K) + \frac{\varepsilon}{2}$$

By lemma 3.7.10, we see that $V \backslash K \in \Omega_F$. By lemma 3.7.16, we see

$$\mu(K) + \mu(U \backslash K) = \mu(U) < \mu(K) + \varepsilon$$

and we are done. □

3.7.18 Proposition *Let $A, B \in \Omega_F$. Then the sets $A\backslash B$, $A \cup B$, and $A \cap B$ belong to the collection Ω_F.*

Proof. By the above proposition, there are compact sets K and K', and open sets U and U' such that

$$K \subseteq A \subseteq U \qquad K' \subseteq B \subseteq U'$$

and

$$\mu(U\backslash K) < \varepsilon \qquad \mu(U'\backslash K') < \varepsilon$$

Observe

$$A\backslash B \subseteq U\backslash K' \subseteq U\backslash K \cup K\backslash U' \cup U'\backslash K'$$

Hence by lemme 3.7.12:

$$\mu(A\backslash B) \subseteq \mu(K\backslash V') + 2\varepsilon$$

Further, the set $K\backslash V'$ is compact, so the above inequality tells us that $A\backslash B \in \Omega_F$.

But $A \cup B = (A\backslash B) \cup B$, so $A \cup B \in \Omega_F$ by lemma 3.7.16. Finally, $A \cap B = A\backslash(A\backslash B)$, so $A \cap B \in \Omega_F$ by the above calculation. □

We are now nearly done, and can prove a slightly less technical result.

3.7.19 Theorem *The set Ω is a σ-algebra containing all Borel sets.*

Proof. Let $K \subseteq X$ be compact. If $A \in \Omega$, then $X\backslash A \cap K = K\backslash(A \cap K)$, so $X\backslash A \cap K \in \Omega_F$ by the above proposition, and $X\backslash A \in \Omega$.

Suppose that

$$A = \bigcup_{n=1}^{\infty} A_n \qquad A_n \in \Omega$$

Let $B_1 = A_1 \cap K$, and

$$B_n = (A_n \cap K)\backslash(B_1 \cup \cdots \cup B_n) \qquad n \geqslant 2$$

Then the collection $\{B_n \mid n = 1, 2, \ldots\}$ is a pairwise disjoint, and $B_n \in \Omega_F$ for all n by the above lemma. But $A \cap K = \bigcup_{n=1}^{\infty} B_n$, so $A \cap K \in \Omega_F$ by lemma 3.7.16. It follows that $A \in \Omega$.

We have proved that the collection Ω is a σ-algebra. If $C \subseteq X$ is a closed subset, then the intersection $K \cap C$ is compact. Thus $C \cap K \in \Omega_F$, and so $C \in \Omega$. Thus every closed set belongs to the collection Ω. It follows that the σ-algebra Ω contains all Borel sets. □

3.7.20 Lemma

$$\Omega_F = \{E \in \Omega \mid \mu(E) < \infty\}$$

Proof. Let $E \in \Omega_F$. Then by lemmas 3.7.14 and 3.7.16, we see $E \cap K \in \Omega_F$ whenever $K \subseteq X$ is compact. Then $E \in \Omega$. By definition of the set Ω_F, $\mu(E) < \infty$.

Conversely, suppose that $E \in \Omega$ and $\mu(E) < \infty$. Let $\varepsilon > 0$. We can certainly find an open set $U \supseteq E$ such that $\mu(E) < \infty$. By propositions 3.7.10 and 3.7.17, there is a compact set $K \subseteq U$ such that $\mu(U \backslash K) < \varepsilon$.

We know that $E \cap K \in \Omega_F$. There is therefore a compact set $H \subseteq E \cap K$ such that

$$\mu(E \cap K) < \mu(H) + \varepsilon$$

But $E \subseteq (E \cap K) \cup (U \backslash K)$. Therefore

$$\mu(E) \subseteq \mu(E \cap K) + \mu(V \backslash K) < \mu(H) + \varepsilon$$

and we see that $E \in \Omega_F$. □

We can now prove our main result.

3.7.21 Theorem *The function μ is a measure on the σ-algebra Ω. It is the unique measure with the property*

$$\Lambda f = \int_X f(x)\ d\mu(x)$$

for all $f \in C_c(X)$.

Proof. It follows immediately that μ is a measure from lemmas 3.7.16 and 3.7.20. Our next step is to prove the inequality

$$\Lambda f \leqslant \int_X f(x)\ d\mu(x)$$

for every real-valued compactly supported function f. To do this, let $K = \operatorname{supp}(f)$, and choose $a, b \in \mathbb{R}$ such that $f[K] \subseteq [a, b]$. Let $\varepsilon > 0$, and choose $y_0, \ldots, y_N$ such that

$$a = y_0 < \cdots < y_N \qquad y_n - y_{n-1} < \varepsilon \text{ for all } n$$

We can form Borel sets

$$E_n := \{x \in X \mid y_{n-1} < f(x) \leqslant y_n\}$$

The sets E_n are pairwise disjoint with union K. We can find open sets $U_n \supseteq E_n$ such that

$$\mu(U_k) < \mu(E_k) + \frac{\varepsilon}{n} \qquad f(x) < y_n + \varepsilon$$

whenever $x \in U_n$.

By theorem 3.7.4, we can choose a partition of unity $\{u_1, \ldots, u_N\}$ subordinate to the open cover $\{U_1, \ldots, U_N\}$. It follows that

$$f = \sum_{n=1}^{N} u_n f$$

and by lemma 3.7.14

$$\mu(K) \leqslant \Lambda\left(\sum_{n=1}^{N} u_n\right) = \sum_{n=1}^{N} \Lambda(u_n)$$

But by construction $u_n f \leqslant (y + n + \varepsilon)u_n$, and $y_n - \varepsilon < f(x)$ for all $x \in E_n$, so

$$\Lambda f \leqslant \sum_{n=1}^{N} (y_k + \varepsilon)\Lambda(u_n) = \sum_{n=1}^{N} (|a| + y_k + \varepsilon)\Lambda(u_n) - |a| \sum_{n=1}^{N} \Lambda(u_n)$$

and

$$\Lambda f \leqslant \sum_{n=1}^{N} (|a| + y_k + \varepsilon)(\mu(E_n) + \varepsilon/n) - |a|\mu(K)$$

Multiplying out, we see that

$$\Lambda f \leqslant \sum_{n=1}^{N} (y_n - \varepsilon)\mu(E_n) + 2\varepsilon\mu(K)\frac{\varepsilon}{n} \sum_{n=1}^{N} (|a| + y_n + \varepsilon)$$

so by construction of the integral

$$\Lambda f \leqslant \int_X f \; d\mu + \varepsilon(2\mu(K) + |a| + b + \varepsilon)$$

Since the above inequality must hold for every choice of $\varepsilon > 0$, we see that

$$\Lambda f \leqslant \int_X f(x) \; d\mu(x)$$

as required.

Now, if we replace the function f by the function $-f$, we see that

$$-\Lambda f \leqslant -\int_X f(x) \; d\mu(x)$$

Combining the above two inequalities, we have the equation

$$\Lambda f = \int_X f(x) \; d\mu(x)$$

for every real-valued compactly supported function f. The proof of the above equation for complex-valued functions follows by splitting such a function into real and imaginary parts, and using linearity.

All that remains is to show uniqueness. Let μ' be a measure such that the eqation

$$\Lambda f = \int_X f(x) \; d\mu'(x)$$

holds for every compactly supported function f. Let K be a compact set. By theorem 3.7.4, given an open set $U \supseteq K$, there is a compactly supported function g such that $\chi_K \leqslant g \leqslant \chi_U$. Hence

$$\mu'(K) \leqslant \int_X f d\mu' \leqslant \mu'(U)$$

and

$$\mu'(U) = \sup\{\Lambda f \mid f \leqslant \chi_U\} = \mu(U)$$

It follows that $\mu(B) = \mu'(B)$ whenever B is a Borel set, and we are done. □

3.8. Integration of Continuous Functions

We would like to use the Riesz representation theorem to define a measure on the real line $\mathbb{R}$ that gives the usual integral expected from elementary calculus. To apply the Reisz representation theorem, we need a sensible definition of the integral of a continuous compactly supported function.

Let us consider a continuous function $f\colon [a,b] \to \mathbb{R}$. Let n be a positive integer. Then the interval $[a,b]$ can be divided into S^n equal-sized pieces:

$$a < a + 2^{-n}(b-a) < a + 2(2^{-n})(b-a) < \cdots < a + (2^n - 1)(2^{-n})(b-a) < b$$

Let us define

$$\mu_{n,r} = \inf\{f(x) \mid a + r2^{-n}(b-a) \leqslant f(x) < a + (r+1)2^{-n}(b-a)$$

and

$$I_n(f) = \sum_{r=0}^{2^n-1} 2^{-n}(b-a)\mu_{n,r}$$

The following observations are clear.

- The sequence $(I_n(f))$ is monotonically increasing
- Since the interval $[a,b]$ is compact, and the function f is continuous, there is a constant C such that $f(x) \leqslant C$ for all $x \in [a,b]$. Hence $I_n(f) \leqslant C(b-a)$ for all n.

It follows that we have a well-defined limit

$$\Lambda(f) := \lim_{n\to\infty} I_n(f)$$

We would like to extend the definition of the function Λ. There are two stages to this extension.

- Let $f\colon [a,b] \to \mathbb{C}$ be a continuous function. Write $f(x) = u(x) + iv(x)$, where $u, v\colon [a,b] \to \mathbb{R}$, and define

$$\Lambda(f) = \Lambda(u) + i\Lambda(v)$$

- Let $f \in C_c(\mathbb{R})$. Let $[a, b] \supseteq \mathrm{supp}(f)$. Then we define

$$\Lambda(f) = \Lambda(f|_{[a,b]})$$

The following result is straightforward to check.

3.8.1 Proposition *The map Λ is a positive linear functional on the space $C_c(\mathbb{R})$.* ***proof to be filled in!***

3.8.2 Definition Let $f \in C_c(\mathbb{R})$. Then the number $\Lambda(f)$ is called the *Riemann integral* of f.

3.9. The Lebesgue Measure on $\mathbb{R}$

3.9.1 Definition Let $\Lambda \colon C_c(\mathbb{R}) \to \mathbb{C}$ be the Riemann integral. Then the *Lebesgue measure* on $\mathbb{R}$ is the unique measure such that

$$\int_{\mathbb{R}} f \; d\mu = \Lambda(f)$$

whenever $f \in C_c(\mathbb{R})$.

By the Riesz representation, the Lebesgue measure exists and is unique on the collection of all Borel sets. The integral of a Borel measurable function with respect to the Lebesgue measure is termed the *Lebesgue integral*. We will normally write

$$\int_a^b f \; d\mu := \int_{\mathbb{R}} f \chi_{(a,b)} \; d\mu$$

3.9.2 Proposition *Let $a < b$ be real numbers. Then $\mu(a, b) = b - a$.*

Proof. Let $[c, d] \subseteq (a, b)$ be a compact interval. By Urysohn's lemma, there is a function $f \in C_c(\mathbb{R})$ such that $\chi_{[c,d]} \leqslant f \leqslant \chi_{(a,b)}$.

By definition of the Riemann integral:

$$d - c \leqslant \int_{\mathbb{R}} f \leqslant b - a$$

Let $c \to a$ and $d \to b$. Then $f \to \chi_{(a,b)}$ and by the dominated convergence theroem,

$$\int_{\mathbb{R}} f \to \mu(a, b)$$

It follows that $\mu(a, b) = b - a$, and we are done. $\square$

A similar computation tells us that

$$\mu[a, b] = \mu[a, b) = \mu(a, b] = b - a$$

whenever $a < b$.

The next fundamental property of the Lebesgue measure follows from a topological property of the real line, which we will state without proof.

3.9.3 Proposition *Every open subset of the real line $\mathbb{R}$ is a countable disjoint union of open intervals.* ***proof to be filled in!***

3.9.4 Corollary *Let $E \subseteq \mathbb{R}$ be a Borel set. Then $\mu(X + E) = \mu(E)$ whenever $x \in \mathbb{R}$.* ***proof to be filled in!***

We conclude with a general characterisation of sets of measure zero, or *null sets.*

3.9.5 Theorem *Let $E \subseteq \mathbb{R}$ be a set such that every subset of A is measurable. Then $\mu(A) = 0$.*

Proof. The set $\mathbb{R}$ is an Abelian group under the operation of addition, and the set $\mathbb{Q}$ is a subgroup. Let E be a set of real numbers containing precisely one element of each coset $x + \mathbb{Q} \in \mathbb{R}/\mathbb{Q}$.

We claim:

- $(r + E) \cap (s + E) = \varnothing$ whenever $r, s \in \mathbb{Q}$, $r \neq s$.
- Let $x \in \mathbb{R}$. Then we can find an element $r \in \mathbb{Q}$ such that $x \in r + E$.

To see the first claim, suppose that $x \in (r+E) \cap (s+E)$, where $r, s \in \mathbb{Q}$. Then there are elements $y, z \in E$ such that $r + y = s + z$, and so $y - z \in \mathbb{Q}$. But the definition of the set E means that $r = s$.

As for the second claim, let $x \in \mathbb{R}$. Construction of the set E means that we can find a point $y \in E$ such that $x - y \in \mathbb{Q}$. But $x = y + (x - y)$ so the claim is established.

We now use the above to claims to prove the theorem. Let $t \in \mathbb{Q}$, and define $A_t := A \cap (t + E)$. The set A_t is measurable since it is a subset of the set A. Consider a compact subset $K \subseteq A_t$, and let

$$H = \bigcup_{r \in \mathbb{Q} \cap [0,1]} (r + K)$$

Then the set H is bounded and measurable, so $\mu(H) < \infty$. The first of the above claims tells us that the sets $r + K$ are pair-wise disjoint, so

$$\mu(H) = \sum_{r \in \mathbb{Q} \cap [0,1]} \mu(r + K) = \sum_{r \in \mathbb{Q} \cap [0,1]} \mu(K)$$

by corollary 3.9.4. It follows that $\mu(K) = 0$ whenever $K \subseteq A_t$ is compact.

So $\mu(A_t) = 0$. But

$$A = \bigcup_{t \in \mathbb{Q}} A_t$$

and it follows that $\mu(A) = 0$. $\square$

3.9.6 Corollary *Any countable subset of the space $\mathbb{R}$ has measure zero.* ***proof to be filled in!***

3.9.7 Corollary *There are non-measurable subsets of the space $\mathbb{R}$.* ***proof to be filled in!***

3.10. The Fundamental Theorem of Calculus

By convention, when $a < b$ are real numbers, and μ is the Lebesgue measure on the space $\mathbb{R}$, we simplfy our notation slightly and write just

$$\int_a^b f(x)\ dx := \int_a^b f\ d\mu$$

If $b < a$, we write

$$\int_a^b f(x)\ dx := -\int_b^a f(x)\ dx$$

Linearity of the integral gives us the equation

$$\int_a^b f(x)\ dx = \int_a^c f(x)\ dx + \int_c^b f(x)\ dx$$

whenever $a, b, c \in \mathbb{R}$.

This new notation is convenient when integating a concrete function given by some definite formula.

In this section we will focus on one major result, which is of absolutely vital importance when trying to calculate integrals. This result is termed the em fundamental theorem of calculus.

3.10.1 Theorem *Let $f\colon [a,b] \to \mathbb{C}$ be a continuous function. Define a function $F\colon [a,b] \to \mathbb{C}$ by the formula*

$$F(x) = \int_a^x f(y)\ dy$$

Then the function F is differentiable on the open interval (a,b), and has a one-sided derivative at the end-points a and b. In all cases, the derivative is given by the formula

$$F'(x) = f(x)$$

Proof. Let $\varepsilon > 0$, and let $x \in [a,b]$. Since the function f is continuous, we can choose $\delta > 0$ such that $|f(x+h) - f(x)| < \varepsilon$ whenever $|h| < \delta$ and $x + h \in [a,b]$.

Let $x \in [a,b]$, and $x + h \in [a,b]$. Observe:

$$F(x+h) - F(x) = \int_x^{h+h} f(y)\ dy$$

and

$$hf(x) = f(x)\mu(x, x+h) = \int_x^{x+h} f(x)\ dy$$

Suppose that $|h| < \delta$. Then $|f(y) - f(x)| < \varepsilon$ whenever $y \in [x, x+h]$, and so:

$$\left| \int_x^{x+h} f(y) - f(x) \; dy \right| \leqslant \int_x^{x+h} |f(y) - f(x)| \; dy \leqslant \varepsilon |h|$$

Thus:

$$|F(x+h) - F(x) - hf(x)| \leqslant \varepsilon |h|$$

whenver $|h| < \delta$. It follows that the function F is differentiable, and $F'(x) = f(x)$ as claimed.□

In actual fact, the more useful form of the fundmantal theorem of calculus is a variation of the above formula.

3.10.2 Corollary *Let $F\colon [a,b] \to \mathbb{C}$ be a function with a continuous derivative f. Then*

$$\int_a^b f(x) \; dx = F(b) - F(a)$$

Proof. Define

$$F_0(x) = \int_a^x f(y) \; dy$$

Then by the above version of the fundamental theorem of calculus, $F_0'(x) = f(x)$ whenever $x \in [a,b]$. Hence $F_0'(x) = F'(x)$ whenever $x \in [a,b]$, so there is a constant C such that $F_0(x) = F(x) + C$ for all $x \in [a,b]$.

We know that $F_0(a) = 0$. Therefore $C = -F(a)$. We see that

$$int_a^b f(x) \; dx = F_0(b) = F(b) - F(a)$$

as claimed. □

The various integration formulae, such as integration by parts and the change of variable formula, come from the fundamental theorem of calculus along with the corresponding formulae for differentives, such as the derivative of a product and the derivative of a composition.

3.11. Product Measures

Let Ω_1 and Ω_2 be measure spaces, with measures μ_1 and μ_2 on σ-algebras $\mathcal{M}_1$ and $\mathcal{M}_2$ respectively.

3.11.1 Definition We call a subset of the form $A \times B \subseteq X \times Y$, where $A \in \mathcal{M}_1$ and $B \in \mathcal{M}_2$ a *measurable rectangle*. A finite union of measurable rectangles is called an *elementary set*.

We write $\mathcal{M}_{12}$ to denote the smallest σ-algebra in the set $\Omega_1 \times \Omega_2$ that contains every measurable rectangle.

We want to define a measure on the σ-algebra $\mathcal{M}_{12}$. Before we can do this, we need some technical constructions.

3.11.2 Definition Let $\mathcal{C}$ be a collection of subsets of some set. Suppose that the following two conditions hold:

- Let (A_n) be a sequence of sets in the collection $\mathcal{C}$ such that $A_n \subseteq A_{n+1}$ for all n. Then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{C}$.
- Let (B_n) be a sequence of sets in the collection $\mathcal{C}$ such that $B_n \supseteq B_{n+1}$ for all n. Then $\bigcup_{n=1}^{\infty} B_n \in \mathcal{C}$.

Then we call the collection $\mathcal{C}$ a *monotone class*.

The proof of the following lemma is elementary, but rather abstract. We omit it.

3.11.3 Lemma *The σ-algebra $\mathcal{M}_{12}$ is the smallest monotone class in the product $\Omega_1 \times \Omega_2$ which contains all elementary sets.* ***proof to be filled in!***

Given a subset $E \subseteq \Omega_1 \times \Omega_2$, and points $x \in Omega_1$ and $y \in \Omega_2$, let us write

$$E_x = \{y \in \Omega_2 \mid (x,y) \in E\} \qquad E^y = \{x \in \Omega_1 \mid (x,y) \in E\}$$

3.11.4 Proposition *Let $E \in \mathcal{M}_{12}$. Then $E_x\ in \mathcal{M}_1$ and $E^y \in \mathcal{M}_2$ whenever $x \in \Omega_1$ and $y \in \Omega_2$.*

Proof. Let $x \in \Omega_1$. Let $\mathcal{M}$ be the collection of all elements $E \in \Omega_1 \times \Omega_2$ such that $E_x \in \Omega_2$. It is straightforward to check that $\mathcal{M}$ is a σ-algebra that contains every measurable rectangle. Therefore $\mathcal{M}_{12} \subseteq \mathcal{M}$, and we see that $E_x \in \mathcal{M}_2$ for every measurable set $E \subseteq \Omega_1 \times \Omega_2$ and point $x \in \Omega_2$.

The corresponding statement concerning sets of the form E^y is proved in the same way. □

3.11.5 Corollary *Let X be a topological space, and let $f : \Omega_1 \times \Omega_2 \to X$ be a measurable function. Choose points $x \in \Omega_1$ and $y \in \Omega_2$. Then the functions*

$$f(x,-) : \Omega_2 \to X \qquad f(-,y) : \Omega_1 \to X$$

are measurable. ***proof to be filled in!***

3.11.6 Definition A measure space Ω is called *σ-finite* if it is a countable union of spaces of finite measure.

3.11.7 Example The space $\mathbb{R}$, equipped with the standard Lebesgue measure, is σ-finite.

The following result lets us define measures on products of σ-finite measure spaces.

3.11.8 Theorem *Let Ω_1 and Ω_2 be σ-finite measure spaces. Let $E \subseteq \Omega_1 \times \Omega_2$ be a measurable subset. Then we can define measurable functions $f\colon \Omega_1 \to [0,\infty]$ and $g\colon \Omega_1 \to [0,\infty]$ by the formulae*

$$f_E(x) = \mu_2(E_x) \qquad g_E(y) = \mu_1(E^y)$$

respectively. Further,

$$\int_{\Omega_1} f_E = \int_{\Omega_2} g_E$$

Proof. Measurability of the functions f_E and g_E associated as above to a measurable set $E \subseteq X \times Y$ follows from the above proposition and corollary; all that remains it to prove the main equation.

Let $\mathcal{M}$ be the set of all measurable subsets $E \subseteq \Omega_1 \times \Omega_2$ such that the equation

$$\int_{\Omega_1} f_E = \int_{\Omega_2} g_E$$

holds.

Let $E = A \times B$ be a measurable rectangle. Then $f_E = \mu_2(B)\chi_A$ and $g_E = \mu_1(A)\chi_B$. It follows that

$$\int_{\Omega_1} f_E = \int_A \mu_2(B) = \mu_1(A)\mu_2(B) \qquad \int_{\Omega_1} g_E = \int_B \mu_1(A) = \mu_1(A)\mu_2(B)$$

so $E \in \mathcal{M}$.

Let (E_n) be a sequence of sets in the collection $\mathcal{M}$ such that $E_n \subseteq E_{n+1}$ for all n. Write

$$E = \bigcup_{n=1}^{\infty} E_n$$

Then the sequences of functions (f_{E_n}) and (g_{E_n}) are monotonic increasing, with limits f_E and g_E respectively. We know that $E_n \in \mathcal{M}$ for all n, so that the equation

$$\int_{\Omega_1} f_{E_n} = \int_{\Omega_2} g_{E_n}$$

holds for all n. The monotone convergence theorem gives us the equation

$$\int_{\Omega_1} f_E = \int_{\Omega_2} g_E$$

and so tells us that $E \in \mathcal{M}$.

As a consequence of the above calculation, we can easily show that the union of a discrete sequence of measurable sets in the set $\mathcal{M}$ also belongs to the set $\mathcal{M}$. Let (E_n) be a sequence of sets in the collection $\mathcal{M}$ such that $E_1 \subseteq A \times B$, where $\mu_1(A) < \infty$, $\mu_2(B) < \infty$, and $E_n \supseteq E_{n+1}$ for all n. Write

$$E = \bigcup_{n=1}^{\infty} E_n$$

Then an argument similar to the above one, only using the dominated convergence theorem rather than the monotone convergence theorem, tells us that the set E belongs to the collection $\mathcal{M}$.

Now, let $\Omega_1 = \cup_{n=1}^{\infty}\Omega_1^{(n)}$ and $\Omega_2 = \cup_{n=1}^{\infty}\Omega_2^{(n)}$, where $\mu_1(\Omega_1^{(m)}) < \infty$ and $\mu_2(\Omega_2^{(n)}) < \infty$ for all $m, n \in \mathbb{N}$. Given a set $E \subseteq \Omega_1 \times \Omega_2$, let us write

$$E_{mn} = E \cap (\Omega_1^{(m)} \times \Omega_2^{(n)}$$

Let $\mathcal{C}$ be the collection of all measurable sets $E \subseteq \Omega_1 \times \Omega_2$ such that $E_{mn} \in \mathcal{M}$ for all natural numbers m and n. Then the above calculations tell us that the collection $\mathcal{C}$ is a monotone class that contains every elementary rectangle. It follows from lemma 3.11.3 $\mathcal{M}_{12} \subseteq \mathcal{C}$, and we are done. $\square$

To paraphrase the above theorem, the equation

$$\int_{\Omega_1}\left(\int_{\Omega_2}\chi_E(x,y)\ d\mu_2(y)\right)\ d\mu_1(x) = \int_{\Omega_2}\left(\int_{\Omega_1}\chi_E(x,y)\ d\mu_1(x)\right)\ d\mu_2(y)$$

holds for every measurable set $E \subseteq \Omega_1 \times \Omega_2$.

3.11.9 Definition Let Ω_1 and Ω_2 be σ-finite measure sets. Then we define a measure μ on the product $\Omega_1 \times \Omega_2$ by writing

$$\mu(E) := \int_{\Omega_1}\left(\int_{\Omega_2}\chi_E(x,y)\ d\mu_2(y)\right)\ d\mu_1(x) = \int_{\Omega_2}\left(\int_{\Omega_1}\chi_E(x,y)\ d\mu_1(x)\right)\ d\mu_2(y)$$

whenever the set $E \subseteq \Omega_1 \times \Omega_2$ is measurable.

It is easy to check that the above definition satisfies the axioms required of a measure. As a special case of the above definition, we can now define a Lebesgue measure on the space $\mathbb{R}^n$ by viewing it as a product of copies of the space $\mathbb{R}$. This measure is defined on every Borel set, and the measure of the n-dimensional cuboid

$$[a_1, b_1] \times \cdots \times [a_n, b_n]$$

is the product

$$(b_1 - a_1)(b_2 - a_2)\cdots(b_n - a_n)$$

3.12. Fubini's Theorem

In the previous section, we saw how to define measures on products of σ-finite measure spaces. We can therefore integrate on such spaces. The purpose of this section is two state two results on the integrability of such functions, and how they are integrated. These results are usually put together, and referred to in one piece as *Fubini's theorem*.

3.12.1 Theorem *Let Ω_1 and Ω_2 be σ-finite measure spaces, and let $f\colon \Omega_1 \times \Omega_2 \to \mathbb{C}$ be an integrable function. Then the functions $f(x,-)$ and $f(-,y)$ are integrable almost everywhere, and the functions*

$$x \mapsto \int_{\Omega_2} f(x,y)d\mu_2(y) \qquad y \mapsto \int_{\Omega_2} f(x,y)d\mu_2(y)$$

are integrable. Moreover,

$$\int_{\Omega_1\times\Omega_2} f(x,y)d\mu(x,y) = \int_{\Omega_1}\left(\int_{\Omega_2} f(x,y)\ d\mu_2(y)\right)\ d\mu_1(x) = \int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)\ d\mu_1(x)\right)\ d\mu_2(y)$$

Proof. Let $s\colon \Omega_1 \times \Omega_2 \to \mathbb{C}$ be a simple function. Then the functions $s(x,-)$ and $s(-,y)$ are integrable almost everywhere, the functions

$$x \mapsto \int_{\Omega_2} s(x,y)d\mu_2(y) \qquad y \mapsto \int_{\Omega_2} s(x,y)d\mu_2(y)$$

are integrable, and the equation

$$\int_{\Omega_1\times\Omega_2} s(x,y)d\mu(x,y) = \int_{\Omega_1}\left(\int_{\Omega_2} s(x,y)\ d\mu_2(y)\right)\ d\mu_1(x) = \int_{\Omega_2}\left(\int_{\Omega_1} s(x,y)\ d\mu_1(x)\right)\ d\mu_2(y)$$

holds by theorem 3.11.8 and the definition of the product measure.

Now, suppose that $f(x,y) \geqslant 0$ for all points $(x,y) \in \Omega_1 \times \Omega_2$. Since the function f is measurable, by proposition 3.3.2 there is a monotonically increasing sequence, (s_n), of simple functions, with point-wise limit f. The result therefore follows in this case by the monotone convergence theorem.

By splitting a real-valued function into positive and negative parts, we see that the result holds for all real-valued functions. We can deduce the result for complex-valued functions by splitting such a function into real and imaginary parts. □

For the above theorem to be useful, we would like a criterion for a function $f\colon \Omega_1 \times \Omega_2 \to \mathbb{C}$ to be integrable. Fortunately, such a condition forms the second half of Fubini's theorem, which is also sometimes referred to as Tonelli's theorem.

3.12.2 Theorem *Let Ω_1 and Ω_2 be σ-finite measure spaces, and let $f\colon \Omega_1 \times \Omega_2 \to \mathbb{C}$ be an integrable function. Suppose that*

$$\int_{\Omega_1}\left(\int_{\Omega_2} |f(x,y)|\ d\mu_2(y)\right)\ d\mu_1(x) < \infty$$

or

$$\int_{\Omega_2}\left(\int_{\Omega_1} |f(x,y)|\ d\mu_1(x)\right)\ d\mu_2(y) < \infty$$

Then the function $f\colon \Omega_1 \times \Omega_2 \to \mathbb{C}$ is integrable.

Proof. The result is obvious if the function f is simple. A similar argument to the proof of Fubini's theorem gives us the result in general. □

Combining the two theorems in this section (ie: the two halves of Fubini's theorem), we have the following handy result on swapping the order of integration.

3.12.3 Corollary *Let Ω_1 and Ω_2 be σ-finite measure spaces, and let $f\colon \Omega_1 \times \Omega_2 \to \mathbb{C}$ be an integrable function. Suppose that*

$$\int_{\Omega_1} \left(\int_{\Omega_2} |f(x,y)| \; d\mu_2(y) \right) \; d\mu_1(x) < \infty$$

Then

$$\int_{\Omega_2} \left(\int_{\Omega_1} f(x,y) \; d\mu_1(x) \right) \; d\mu_2(y) < \infty = \int_{\Omega_1} \left(\int_{\Omega_2} f(x,y) \; d\mu_2(y) \right) \; d\mu_1(x) < \infty$$

proof to be filled in!

Part II.

Functional Analysis

II.1. Topological Vector Spaces

1.1. Topological division rings and fields

1.1.1 Vector spaces with a compatible topology can not only defined for vector spaces over the ground fields $\mathbb{R}$ and $\mathbb{C}$ but also over fields $\mathbb{K}$ carrying an absolute value $|\cdot| : \mathbb{K} \to \mathbb{R}_{\geqslant 0}$. This endows the ground field with a topology which will be needed in the definition of a topological vector space. We therefore give here a brief introduction to topological division rings and fields first.

1.1.2 Definition Let R be a division ring. By an *absolute value* on R one understands a map $|\cdot| : R \to \mathbb{R}_{\geqslant 0}$ such that the following axioms hold true.

(VDR1) The function $|\cdot|$ is multiplicative that is

$$|xy| = |x|\,|y| \quad \text{for all } x, y \in R \ .$$

(VDR2) The triangle inequality is satisfied which means that

$$|x+y| \leqslant |x| + |y| \quad \text{for all } x, y \in R \ .$$

(VDR3) For all $x \in R$ the relation $|x| = 0$ holds true if and only if $x = 0$.

A division ring or field endowed with an absolute value is called a *valued division ring* respectively a *valued field*. An absolute value $|\cdot|$ on a division ring R and the corresponding valued division ring $(R, |\cdot|)$ are called *non-archimedean* if the *strong triangle inequality* is satisfied that is if

(VDR4) $|x+y| \leqslant \max\{|x|, |y|\}$ for all $x, y \in R$.

Otherwise $|\cdot|$ and $(R, |\cdot|)$ are called *archimedean*.

1.1.3 Lemma *Let $(R, |\cdot|)$ be a valued division ring. Then*

(i) $|1| = 1$,

(ii) $|-x| = |x|$ *for all* $x \in R$, *and*

(iii) $\big||x| - |y|\big| \leqslant |x - y| \leqslant |x| + |y|$ *for all* $x, y \in R$.

Proof. (i) holds true since $|1| = |1^2| = |1|^2$ and $|1| \neq 0$ by $1 \neq 0$. To verify (ii) it suffices to show that $|-1| = 1$. But that holds true since $|-1|^2 = |(-1)^2| = 1$ and $|-1| \geqslant 0$. The last claim follows by

$$-|x-y| = |x| - (|y-x| + |x|) \leqslant |x| - |y| \leqslant |x-y| + |y| - |y| = |x-y|$$

and

$$|x-y| = |x+(-y)| \leqslant |x| + |-y| = |x| + |y| \ . \qquad \square$$

1.1.4 Examples (a) Obviously, the *standard absolute values*

$$|\cdot|_\infty : \mathbb{Q}, \mathbb{R} \to \mathbb{R}_{\geqslant 0}, \ x \mapsto \begin{cases} x & \text{if } x \geqslant 0 \\ -x & \text{if } x < 0 \end{cases} \quad \text{and} \quad |\cdot|_\infty : \mathbb{C} \to \mathbb{R}_{\geqslant 0}, \ z \mapsto \sqrt{z\overline{z}}$$

are absolute values on the fields $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, respectively. These absolute values are all archimedean since $|1+1|_\infty = 2 > 1$. Unless mentioned differently, we always assume $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ to be equipped with the standard absolute values. If no confusion can arise we usually write $|\cdot|$ instead of $|\cdot|_\infty$.

(b) The *standard absolute value* on the quaternions

$$|\cdot|_\infty : \mathbb{H} \to \mathbb{R}_{\geqslant 0}, \ q = a + b\,\mathsf{i} + c\,\mathsf{j} + d\,\mathsf{k} \mapsto \sqrt{\overline{q}q} = \sqrt{a^2 + b^2 + c^2 + d^2} \ ,$$

where a, b, c, d are real, is an archimedean absolute value. Usually it is briefly denoted $|\cdot|$.

(c) For every division ring R the map

$$|\cdot| : R \to \mathbb{R}, \ x \mapsto \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{else} \end{cases}$$

is a non-archimedean absolute value. It is called the *trivial absolute value* on R.

(d) An absolute value $|\cdot| : \mathbb{F} \to \mathbb{R}_{\geqslant 0}$ defined on a finite field $\mathbb{F}$ has to be trivial. To see this observe that for each $x \in \mathbb{K}^\times$ there exists an $n \in \mathbb{N}$ such that $x^n = 1$. This entails $|x|^n = 1$, hence $|x| = 1$ for all $x \in \mathbb{K}^\times$. So $|\cdot|$ is trivial.

(e) The field of formal Laurent power series $\mathbb{K}((X))$ over a field $\mathbb{K}$ can be equipped with an absolute value as follows. Choose $0 < \varepsilon < 1$ and define the absolute value $\left|\sum_{k\in\mathbb{Z}} a_k X^k\right|$ of an element $\sum_{n\in\mathbb{Z}} a_n X^n \in \mathbb{K}((X))$ as ε^n, where n is the minimal integer such that $a_n \neq 0$.

(f) Let p be prime number. For every integer $m \neq 0$ let $\nu_p(m)$ be the exponent of p in the prime factor decomposition of m that is $m = p^{\nu_p(n)} n$ where n is relatively prime to p. For $m \in \mathbb{Z}$ and $n \in \mathbb{N}_{>0}$ one defines the *p-adic absolute value* of the rational number $x = \frac{m}{n}$ by

$$|x|_p = \begin{cases} 0 & \text{if } m = 0 \ , \\ p^{-\nu_p(m) + \nu_p(n)} & \text{else} \ . \end{cases}$$

Note that $|x|_p$ does not depend on the particular representation of x as the quotient of integers m and n. By definition it is immediately clear that the p-adic absolute value is an absolute value on $\mathbb{Q}$ indeed. It is non-archimedean.

1.1.5 Proposition *A valued division ring $(R, |\cdot|)$ is non-archimedean if and only if the image of $\mathbb{Z}$ under the canonical map $\mathbb{Z} \to R$ is bounded.*

Proof. Assume that $(R, |\cdot|)$ is a non-archimedean valued division ring. Then, $|0 \cdot 1| = |0| = 0$ and, under the assumption that $|(n-1) \cdot 1| \leqslant 1$ for some $n \in \mathbb{N}_{>0}$, $|n \cdot 1| = |(n-1) \cdot 1 + 1| = \max\{|(n-1) \cdot 1|, 1\} = 1$. Hence by induction and since $|-1| = 1$ one obtains that $|n \cdot 1| \leqslant 1$ for all $n \in \mathbb{Z}$, and the image of $\mathbb{Z}$ in R is bounded.

To show the converse assume that the image of $\mathbb{Z}$ in R is bounded by some constant $C > 0$. Then, for all $x, y \in R$ and $n \in \mathbb{N}_{>0}$ by the binomial formula and the triangle inequality

$$|x+y|^n = \left| \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \right| \leqslant (n+1)\, C \max\{|x|, |y|\}^n \ .$$

Taking the n-th root gives $|x + y| \leqslant \big((n+1)C\big)^{1/n} \max\{|x|, |y|\}$ which after passing to the limit $n \to \infty$ entails $|x + y| \leqslant \max\{|x|, |y|\}$ since $\lim\limits_{n\to\infty} \big((n+1)C\big)^{1/n} = 1$. Hence $(R, |\cdot|)$ is non-archimedean. $\square$

1.1.6 Proposition *Let $|\cdot|$ be an absolute value on the division ring R. Then for every $\tau > 0$ with $\tau \leqslant 1$ the map $|\cdot|^\tau : R \to \mathbb{R}_{\geqslant 0}$ is an absolute value on R as well. It is archimedean if and only if $|\cdot|$ is archimedean.*

Proof. To prove that $|\cdot|^\tau$ is an absolute value it suffices to show that $(a + b)^\tau \leqslant a^\tau + b^\tau$ for all $a, b \geqslant 0$. Without loss of generality we may assume $a \geqslant b > 0$. By dividing through b^τ one sees that the claim is equivalent to $(t + 1)^\tau \leqslant t^\tau + 1$ for all $t \geqslant 1$. For $t = 1$ this is certainly true. The derivative of the function $h : [1, \infty) \to \mathbb{R}$, $t \mapsto (t + 1)^\tau - t^\tau$ now is given by $h'(t) = \tau\big((t + 1)^{\tau-1} - t^{\tau-1}\big)$ which is negative since $\tau - 1 \leqslant 0$ and $1 + t > t \geqslant 1$. Hence h is monotone decreasing and $(t + 1)^\tau - t^\tau \leqslant 1$ for all $t \geqslant 1$.

Since $(0, \infty) \to \mathbb{R}$, $t \mapsto t^\tau$ is strictly increasing and unbounded, the image of $\mathbb{Z}$ in R is unbounded with respect to $|\cdot|$ if and only if it is with respect to $|\cdot|^\tau$. $\square$

1.1.7 An absolute value $|\cdot| : R \to \mathbb{R}_{\geqslant 0}$ on a division ring R induces the metric $d : R \times R \to \mathbb{R}_{\geqslant 0}$, $(x, y) \mapsto |x - y|$ which then gives rise to a topology on R. This topology has the following properties:

(TDR1) Addition $+ : R \times R \to R$ is continuous.

(TDR2) Multiplication $\cdot : R \times R \to R$ is continuous.

(TDR3) Inversion $(\,\cdot\,)^{-1} : R^\times \to R^\times$ is continuous, where $R^\times$ denotes the set of units in R i.e. $R^\times = R\backslash\{0\}$.

Proof. Addition is continuous since for all $a, b, x, y \in R$ by the triangle inequality

$$d(x + y, a + b) = |x + y - (a + b)| \leqslant |x - a| + |y - b| = d(x, a) + d(y, b) \ .$$

Actually, this even shows that addition is Lipschitz continuous. Now fix $a, b \in R$ and let $C = \max\{|a|, |b|\} + 1$. Then for all $x, y \in R$ with $d(y, b) < 1$

$$d(x \cdot y, a \cdot b) = |(x \cdot y - a \cdot y) + (a \cdot y - a \cdot b)| \leqslant |x - a|\,|y| + |a|\,|y - b| \leqslant C\big(d(x, a) + d(y, b)\big) \ .$$

Hence multiplication is continuous. Finally, fix $a \in R^\times$ and let $x \in R^\times$ with $d(x,a) < \frac{|a|}{2}$. Then $|x| \geqslant |a| - d(x,a) > \frac{|a|}{2} > 0$ and

$$d\left(x^{-1}, a^{-1}\right) = \left|x^{-1} - a^{-1}\right| = \left|x^{-1} \cdot a^{-1}\right| |x-a| = \frac{1}{|x|\,|a|} d(x,a) < \frac{2}{|a|^2} d(x,a) \ .$$

So inversion is also continuous. □

1.1.8 Definition A division ring or field R which is equipped with a topology so that (TDR1), (TDR2) and (TDR3) are satisfied is called a *topological division ring* or a *topological field*, respectively.

1.1.9 Lemma *If $|\cdot|$ is a non-trivial absolute value on the division ring R, then there exists an element $t \in R^\times$ such that the sequence $(t^n)_{n\in\mathbb{N}}$ converges to 0. Furthermore in this case every 0-neighborhood in R contains infinitely many elements.*

Proof. By non-triviality of $|\cdot|$ there exists $t \in R^\times$ such that $|t| \neq 1$. By possibly passing to t^{-1} we can assume $|t| < 1$. Since then $\lim\limits_{n\to\infty} |t|^n = 0$, the sequence $(t^n)_{n\in\mathbb{N}}$ converges to 0. This implies in particular that for every $\varepsilon > 0$ the open ball $\mathbb{B}(0,\varepsilon) = \{t \in R \mid |t| < \varepsilon\}$ contains infinitely many elements. So the lemma is proved. □

1.1.10 Definition Two absolute values $|\cdot|$ and $|\cdot|'$ on a division ring R are called *equivalent* if they induce the same topology on R.

1.1.11 Theorem *Let $|\cdot|$ and $|\cdot|'$ be two absolute values on the division ring R. Then they are equivalent if and only if there exists $e > 0$ such that $|\cdot|' = |\cdot|^\tau$. In particular the trivial absolute value is the only one inducing the discrete topology on R.*

Proof. Let us first show the following proposition.

(A) If $|\cdot|$ and $|\cdot|'$ are equivalent, then the relation $|x| < 1$ holds true for $x \in R^\times$ if and only if $|x|' < 1$.

Since $\left|x^{-1}\right| = \frac{1}{|x|}$ and $\left|x^{-1}\right|' = \frac{1}{|x|'}$ for all $x \in R^\times$, (A) implies that $|x| > 1$ if and only if $|x|' > 1$ and that $|x| = 1$ if and only if $|x|' = 1$. To verify claim (A) assume now that $0 < |x| < 1$. Then $\lim\limits_{n\to\infty} |x^n| = 0$, hence $(x^n)_{n\in\mathbb{N}}$ converges to 0. By assumption, $\lim\limits_{n\to\infty} |x^n|' = 0$ then holds as well which implies that $|x|' < 1$. By switching $|\cdot|$ and $|\cdot|'$ the converse holds true, so (A) is proved.

Next we show that $|\cdot|$ is trivial if and only if the induced topology on R is discrete. Namely, if $|\cdot|$ is non-trivial, then there exists $x \in R^\times$ such that $|x| \neq 1$. After possibly passing to $\frac{1}{x}$ we can achieve that $|x| < 1$. So $\lim\limits_{n\to\infty} |x^n| = 0$, which means that $(x^n)_{n\in\mathbb{N}}$ is a sequence of non-zero elements of R converging to 0. But this implies that the singleton $\{0\}$ is not open in the topology induced by $|\cdot|$, hence this topology is non-discrete. Since obviously the trivial absolute value induces the discrete topology on R the second claim of the theorem is proved.

Now assume that $|\cdot|' = |\cdot|^\tau$ for some $\tau > 0$. Then a subset $B \subset R$ is a metric open ball with respect to $|\cdot|$ if and only if it is one with respect to $|\cdot|'$ since for $x \in R$ and $\varepsilon > 0$

$$\left\{y \in R \,\middle|\, |y-x| < \varepsilon\right\} = \left\{y \in R \,\middle|\, |y-x|' < \varepsilon^\tau\right\} \text{ and}$$
$$\left\{y \in R \,\middle|\, |y-x|' < \varepsilon\right\} = \left\{y \in R \,\middle|\, |y-x| < \varepsilon^{1/\tau}\right\} \ .$$

Hence the open sets with respect to the metric defined by $|\cdot|$ coincide with those defined by $|\cdot|'$ and the two absolute values are equivalent.

Let us finally show the other direction and assume that $|\cdot|$ and $|\cdot|'$ are equivalent. By the already proven second claim of the theorem we can restrict to the case where the induced topology is non-discrete which means to the case where both $|\cdot|$ and $|\cdot|'$ are non-trivial. We show that there exists $\tau > 0$ such that $|x|' = |x|^\tau$ for all $x \in R^\times$ with $|x| > 1$. This is sufficient, since if $|x| = 1$, then $|x|' = 1 = |x|^\sigma$ for any $\sigma > 0$ by (A), and since if $x \in R^\times$ with $|x| < 1$ then $|x^{-1}| > 1$ and

$$|x|' = \frac{1}{|x^{-1}|'} = \frac{1}{|x^{-1}|^\tau} = |x|^\tau \ .$$

The existence of a $\tau > 0$ with the claimed property is equivalent to the function

$$R^\times \to \mathbb{R},\ x \mapsto \frac{\ln|x|'}{\ln|x|}$$

being constant. Assume that that is not the case. Then there exist $x, y \in R^\times$ with $|x|, |y| > 1$ such that $\frac{\ln|x|'}{\ln|x|} \neq \frac{\ln|y|'}{\ln|y|}$. By possibly switching x and y we can assume $\frac{\ln|x|'}{\ln|x|} < \frac{\ln|y|'}{\ln|y|}$. But that implies $\frac{\ln|x|'}{\ln|y|'} < \frac{\ln|x|}{\ln|y|}$ since the logarithms are positive by assumptions on x and y and (A). Hence there exists a rational number $\frac{p}{q}$ with $p, q \in \mathbb{N}_{>0}$ such that

$$\frac{\ln|x|'}{\ln|y|'} < \frac{p}{q} < \frac{\ln|x|}{\ln|y|} \ .$$

Then $|x^q|' < |y^p|'$ and $|y^p| < |x^q|$ which entails

$$\left|\frac{x^q}{y^p}\right|' < 1 \text{ and } \left|\frac{x^q}{y^p}\right| > 1 \ .$$

This contradicts (A) and the theorem is proved. □

1.1.12 Remarks (a) By Ostrowski's theorem (Ostrowski, 1916, p. 276), see also (Gouvêa, 1997, Thm. 3.1.3), every non-trivial absolute value on the field $\mathbb{Q}$ of rational numbers is either equivalent to the standard absolute value $|\cdot|_\infty$ or to a p-adic absolute value $|\cdot|_p$ for some prime number p. Observe that for different primes p and q the absolute values $|\cdot|_p$ and $|\cdot|_q$ are not equivalent.

(b) Another theorem of Ostrowski (Ostrowski, 1916, p. 284), sometimes called big Ostrowski's theorem, tells that for every archimedean valued field $(\mathbb{K}, |\cdot|)$ there exists an embedding $\iota : \mathbb{K} \hookrightarrow \mathbb{C}$ into the field of complex numbers with its standard absolute value and a positive real number $\tau \leqslant 1$ such that

$$|x| = |\iota(x)|_\infty^\tau \quad \text{for all } x \in \mathbb{K} \ .$$

In particular this means that every complete archimedean valued field is isomorphic to either $(\mathbb{R}, |\cdot|_\infty^\tau)$ or $(\mathbb{C}, |\cdot|_\infty^\tau)$ for some positive $\tau \leqslant 1$.

(c) The p-adic absolute values on $\mathbb{Q}$ have extensions to $\mathbb{R}$ by (Lang, 2002, XII, §4, Thm. 4.1). This is a highly non-obvious result. To prove it one has to check first that $|\cdot|_p$ can be extended to an absolute value $|\cdot|$ on the field $\Bbbk$ of real numbers algebraic over $\mathbb{Q}$. This extended absolute value is, and that turns out to be crucial, again non-archimedean. Now one observes that $|\cdot|$

can be extended to the polynomial ring $\Bbbk[X]$ by the *Gauß norm* $|p(X)| = \max_{0\leqslant i\leqslant n}\{a_i\}$ where $p(X) = a_nX^n + \ldots + a_1X + a_0 \in \Bbbk[X]$. The Gauß norm obviously extends to an absolute value on the fraction field $\Bbbk(X)$. Again, this extension is non-archimedean. Now one recalls that $\mathbb{R}$ is a purely transcendental field extension of $\Bbbk$ and uses a transfinite induction type argument involving the just constructed Gauß norm to extend $|\cdot|$ from $\mathbb{K}$ to $\mathbb{R}$. The thus obtained extension of the p-adic absolute value to $\mathbb{R}$ is not unique. In its construction, the axiom of choice is used, so one can not even give an explicit formula for such an extension.

1.2. The category of topological vector spaces

Vector space topologies

1.2.1 Definition Let R be a topological division ring. A topology $\mathcal{T}$ on a vector space E over R is called a *vector space topology* if the following axioms hold true:

(TVS1) Addition $+ : \mathrm{E} \times \mathrm{E} \to \mathrm{E}$ is continuous.

(TVS2) Multiplication by scalars $\cdot : R \times \mathrm{E} \to \mathrm{E}$ is continuous.

The topology $\mathcal{T}$ on E is called *translation invariant* if for every $w \in \mathrm{E}$ the linear map $\ell_w : \mathrm{E} \to \mathrm{E}$, $v \mapsto v + w$ is a homeomorphism.

A vector space E endowed with a vector space topology on it is called a *topological vector space* (*over* R), for short a tvs

1.2.2 Remark Let us recall at this point some notation from linear algebra. Assume that V is a left vector space over the divison ring R. If $A, B \subset \mathrm{V}$ are two non-empty subsets, then $A + B$ is the set of all $v \in \mathrm{V}$ for which there exist $x \in A$ and $y \in B$ such that $v = x + y$. If A or B is empty, then $A + B$ is defined as the empty set. In case A is a singleton that is if $A = \{x\}$, then we often write $x + B$ instead of $\{x\} + B$. If $\mathcal{B} \subset \mathcal{P}(\mathrm{V})$ is a non-empty set of subsets of V, then we denote by $A + \mathcal{B}$ and $x + \mathcal{B}$ the sets $\{A + B \in \mathcal{P}(\mathrm{V}) \mid B \in \mathcal{B}\}$ and $\{x + B \in \mathcal{P}(\mathrm{V}) \mid B \in \mathcal{B}\}$, respectively. If $\mathcal{A} \subset \mathcal{P}(\mathrm{V})$ is a second non-empty set of subsets of V, then $\mathcal{A} + \mathcal{B}$ stands for the set of all sets of the form $A + B$, where $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

In case C is a subset of the ground ring R, then $C \cdot A$ is defined as the set of all $v \in \mathrm{V}$ for which there exist $r \in C$ and $x \in A$ such that $v = r \cdot x$. If $r \in R$ we write $r \cdot A$ for $\{r\} \cdot A$. Likewise, if $x \in \mathrm{V}$, $C \cdot x$ stands for $C \cdot \{x\}$. Analogously as for addition the sets $\mathcal{C} \cdot A$, $C \cdot \mathcal{A}$ and $\mathcal{C} \cdot \mathcal{A}$ are defined when $\mathcal{C} \subset \mathcal{P}(R)$ and $\mathcal{A} \subset \mathcal{P}(\mathrm{V})$ are non-empty.

1.2.3 Proposition *Let* E *be a tvs over a topological division ring* R. *Then the following holds true:*

(i) *For every* $r \in R^\times$ *and* $w \in \mathrm{E}$ *the homothety* $\ell_{r,w} : \mathrm{E} \to \mathrm{E}$, $v \mapsto rv + w$ *is a homeomorphism with inverse* $\ell_{r^{-1},-r^{-1}w}$.

(ii) *Let* w *be an element of* E *and* $r \in R^\times$. *A filter base* $\mathcal{B}$ *on* E *then is a filter base for the zero neighborhoods if and only if* $w + r\mathcal{B}$ *is a filter base for the neighborhoods of* w.

(iii) *If $\mathcal{B}$ is a filter base of the filter of zero neighborhoods, then the closure of any non-empty $A \subset \mathrm{E}$ is given by*

$$\bar{A} = \bigcap_{U \in \mathcal{B}} A + U \ .$$

(iv) *Let $A \subset \mathrm{E}$ be open and $B \subset \mathrm{E}$. Then the set $A + B$ is open.*

(v) *Let $A, B \subset \mathrm{E}$ be closed and assume that A is quasi-compact that is that any filter on A has a cluster point. Then the set $A + B$ is closed.*

(vi) *The space* E *is* (T3) *or, equivalently, each point of* E *possesses a neighborhood base consisting of closed subsets.*

Proof. ad (i). The homothety $\ell_{r,w}$ is continuous since addition and multiplication by a scalar are continuous maps on a tvs Since for all $v \in V$

$$\ell_{r^{-1},-r^{-1}w} \circ \ell_{r,w}(v) = r^{-1}(rv + w) - r^{-1}w = v, \text{ and}$$
$$\ell_{r,w} \circ \ell_{r^{-1},-r^{-1}w}(v) = r(r^{-1}v - r^{-1}w) + w = v$$

the homothety $\ell_{r,w}$ is invertible, and its inverse is $\ell_{r^{-1},-r^{-1}w}$.

ad (ii). This follows since $\ell_{r,w}$ is a homeomorphism.

ad (iii). Let $B = \bigcap_{U \in \mathcal{B}} A + U$. Let v be an element of the closure of A. Then, for $U \in \mathcal{B}$, there exists an element $a \in A \cap v - U$ by (ii) and since $-U$ is a zero neighborhood. Hence $v \in a + U$, and $\bar{A} \subset B$ follows. Now let $v \in B$ and V be a neighborhood of v. Then there exists $U \in \mathcal{B}$ such that $v - U \subset V$. By definition of B there exists an element $a \in A$ such that $v \in a + U$. Hence $a \in v - U \subset V$ which implies that $v \in \bar{A}$. So $B \subset \bar{A}$.

ad (iv). The set $A + B$ is either empty or coincides with the union $\bigcup_{v \in B} v + A$. In the latter case, each of the sets $v + A$ is non-empty and open by continuity of addition. So $A + B$ is open under the assumptions made.

ad (v). We can assume that A and B are non-empty because the claim is trivial otherwise. Assume that $A + B$ is not closed. Then there exists an element $v \in \mathrm{E} \backslash (A + B)$ such that each neighborhood of v meets $A + B$. This means in particular that the restriction of the neighborhood filter $\mathcal{U}$ of v to $A + B$ is a filter base. Consequently, $(-B + \mathcal{U}) \cap A$ is a filter base on A, hence possesses an accummulation point $x \in A$. For each neighborhood $V \in \mathcal{U}$ the point x is then contained in the closure of $-B + V$. Hence, by (iii), x is contained in $v - B + U + U$ for every zero neighborhood U. Since by continuity of addition $U + U$ runs through a base of zero neighborhoods when U runs through the zero neighborhoods, $x \in v - \bar{B} = v - B$ follows. Since $x \in A$ this contradicts the assumption $v \in A + B$ and $A + B$ has to be closed.

ad (vi). Let $v \in \mathrm{E}$, $A \subset \mathrm{E}$ closed, and assume $v \notin A$. Choose an open neighborhood V of v such that $V \cap A = \varnothing$. Then there exists an open zero neighborhood U such that $v + U + U \subset V$. By possibly passing to $U \cap (-U)$ we can assume that $U = -U$. Now $v + U$ is an open neighborhood of v and $A + U$ one of A. These neighborhoods are disjoint because if the intersection $v + U \cap A + U$ is non-empty, then there exists an element $w \in v + U + U \cap A$ since $-U = U$. This contradicts $V \cap A = \varnothing$, so $v + U$ and $A + U$ are disjoint neighborhoods of v and A, respectively. Hence E satisfies (T3) . □

1.2.4 Corollary *Every vector space topology on a vector space over a topological division ring is translation invariant.*

Proof. This follows immediately by Proposition 1.2.3 (i). □

1.2.5 Definition A subset C of a vector space E over a valued division ring $(R, |\cdot|)$ is called

(i) *symmetric* if $-v \in C$ for all $v \in C$,

(ii) *circled* or *balanced* if $rv \in C$ for all $v \in C$ and $r \in R$ with $|r| \leqslant 1$.

1.2.6 Remark Symmetry of a subset of a vector space of a division ring is even defined when the underlying division ring does not carry an absolute value.

1.2.7 Lemma *Let C be a subset of a topological vector space E over a valued division ring $(R, |\cdot|)$ and $r \in R$.*

(i) *If C is symmetric, then the closure $\bar{C}$ and the interior $\mathring{C}$ are symmetric.*

(ii) *If C is circled, then the closure $\bar{C}$ and the union $\mathring{C} \cup \{0\}$ are circled.*

(iii) *The set rC is symmetric (respectively circled) if C has that property.*

Proof. Without loss of generality we can assume $C \neq \emptyset$. Claim (i) then follows immediately since multiplication by -1 is a homeomorphism. To prove claim (ii) assume that C is circled. Let $s \in R$ with $|s| \leqslant 1$. Assume $v \in \bar{C}$ and consider sv. We have to show that $sv \in \bar{C}$. If $s = 0$ then $sv = 0 \in C \subset \bar{C}$ since C is circled. So we can assume $s \neq 0$ and need to show that for every neighborhood V of sv the intersection $C \cap V$ is non-empty. Since $|s| > 0$, the homothety $\ell_s : \mathrm{E} \to \mathrm{E}$, $w \mapsto sw$ is a homeomorphism with inverse $\ell_{s^{-1}}$. Hence $s^{-1}V$ is a neighborhood of v. Since v lies in the closure of C there exists an element $w \in C \cap s^{-1}V$. Hence $sw \in C \cap V$ by assumption on C and $\bar{C}$ is circled.

If $v \in \mathring{C} \cup \{0\}$ then $0 = 0 \cdot v \in \mathring{C} \cup \{0\}$. It remains to show that $sv \in \mathring{C} \cup \{0\}$ for $s \in R$ with $0 < |s| \leqslant 1$ and $v \in \mathring{C} \backslash \{0\}$. Under this assumption the homothety ℓ_s is a homeomorphism, so $s\mathring{C}$ is an open subset of C since C is circled. Hence $sv \in s\mathring{C} \subset \mathring{C}$, and $\mathring{C} \cup \{0\}$ is circled as well.

Claim (iii) follows immediately from the observation that for $v \in C$ and $s \in R$ the relation $srv \in rC$ holds true if $sv \in C$. □

1.2.8 Proposition and Definition *The intersection of a non-empty family $(C_i)_{i \in I}$ of symmetric (respectively circled) subsets $C_i \subset \mathrm{E}$, $i \in I$ of a topological vector space E over a valued division ring $(R, |\cdot|)$ is symmetric (respectively circled). In particular, if $A \subset \mathrm{E}$ is a subset, then the sets*

$$\operatorname{Sym} A = \bigcap_{\substack{A \subset B \subset \mathrm{E} \\ B \text{ is symmetric}}} B \quad \text{and} \quad \operatorname{Circ} A = \bigcap_{\substack{A \subset B \subset \mathrm{E} \\ B \text{ is circled}}} B$$

are symmetric and circled, respectively. They have the property that $\operatorname{Sym} A$ *is the smallest symmetric and* $\operatorname{Circ} A$ *the smallest circled subsets of* E *containing A. They are called the* symmetric *and the* circled hull *of A, respectively. Analogously,*

$$\overline{\operatorname{Sym}} A = \bigcap_{\substack{A \subset B = \bar{B} \subset \mathrm{E} \\ B \text{ is symmetric}}} B \quad \text{and} \quad \overline{\operatorname{Circ}} A = \bigcap_{\substack{A \subset B = \bar{B} \subset \mathrm{E} \\ B \text{ is circled}}} B$$

are called the closed symmetric *and the* closed circled hull *of A, respectively. They have the property that* $\overline{\mathrm{Sym}}\, A$ *is the smallest closed symmetric and* $\overline{\mathrm{Circ}}\, A$ *the smallest closed circled subset of* E *containing* A*.*

Proof. Note first that all the hulls in the proposition are well-defined since E is closed and circled. Let C denote the intersection of the family $(C_i)_{i\in I}$. Assume that for some $r \in R$ with $|r| \leqslant 1$ the inclusion $rC_i \subset C$ holds true for all $i \in I$. Then $rC \subset C$, hence if all C_i are symmetric (respectively circled), so is C. This observation now entails that $\mathrm{Sym}\, A$ is symmetric, Circ is circled, $\overline{\mathrm{Sym}}\, A$ is closed and symmetric, and finally that $\overline{\mathrm{Circ}}\, A$ is closed and circled. Moreover, all those sets contain A. The minimality properties of these sets are clear by construction. □

1.2.9 Remark Observe that by the proposition A is symmetric if and only if $\mathrm{Sym}\, A = A$ and circled if and only if $\mathrm{Circ}\, A = A$. Analogously, $\overline{\mathrm{Sym}}\, A = A$ if and only if A is closed symmetric and $\overline{\mathrm{Circ}}\, A = A$ if and only if A is closed and circled.

1.2.10 Lemma *Let* E *be a topological vector space over the valued division ring* $(R, |\cdot|)$ *and* $A \subset \mathrm{E}$ *non-empty. Then*

$$\mathrm{Sym}\, A = A \cup -A \quad \textit{and} \quad \mathrm{Circ}\, A = \bigcup_{r\in R,\, |r|\leqslant 1} rA \ .$$

For the closed hulls one has

$$\overline{\mathrm{Sym}}\, A = \overline{\mathrm{Sym}\, A} \quad \textit{and} \quad \overline{\mathrm{Circ}}\, A = \overline{\mathrm{Circ}\, A} \ .$$

Proof. Since $A \cup -A$ is symmetric by definition, contains A, and is contained in $\mathrm{Sym}\, A$, the equality $\mathrm{Sym}\, A = A \cup -A$ holds true. Similarly, $\bigcup_{r\in R,\, |r|\leqslant 1} rA$ is circled by definition, contains A, and is contained in $\mathrm{Circ}\, A$ by definition of the circled hull. Hence $\mathrm{Circ}\, A = \bigcup_{r\in R,\, |r|\leqslant 1} rA$. The remainder of the claim follows from Lemma 1.2.7. □

1.2.11 Definition Assume that B, C are subsets of a vector space E over the valued division ring $(R, |\cdot|)$. Then one says that

(i) C *absorbes* B if there exists a real number $t \in \mathbb{R}_{\geqslant 0}$ such that $B \subset rC$ for all $r \in R$ with $|r| \geqslant t$,

(ii) C is *absorbing* or *absorbent* if C absorbes every one-point set of E that is if for every $v \in \mathrm{E}$ there exists $t \in \mathbb{R}_{\geqslant 0}$ such that $v \in rC$ for all $r \in R$ with $|r| \geqslant t$.

If the vector space E carries in addition a vector space topology, then one says that

(iii) the subset $B \subset \mathrm{E}$ is *bounded* if it is absorbed by every zero neighborhood.

1.2.12 Lemma *Let* E *be a vector space over the valued division ring* $(R, |\cdot|)$*. Then the following holds true.*

(i) *If* $C_1, \ldots, C_n$ *are absorbing subset of* E*, then the intersection* $C_1 \cap \ldots \cap C_n$ *is absorbing.*

(ii) *If* C *is an absorbing subset of* E*, then* rC *is absorbing for every* $r \in R^\times$*.*

Proof. ad (i). Let $v \in \mathrm{E}$ and choose $t_1, \ldots, t_n \in \mathbb{R}_{\geqslant 0}$ such that $v \in rC_i$ for $|r| \geqslant t_i$. Put $t = \max\{t_1, \ldots, t_n\}$. Then $v \in r(C_1 \cap \ldots \cap C_n)$ for $|r| \geqslant t$, hence $C_1 \cap \ldots \cap C_n$ is absorbing.

ad (*ii*). Choose $t \in \mathbb{R}_{\geqslant 0}$ such that $v \in sC$ for all $s \in R$ with $|s| \geqslant t$. Then one has $|sr| \geqslant t$ for all $s \in R$ with $|s| \geqslant \frac{t}{|r|}$, hence $v \in s(rC)$ for all such s. Therefore rC is absorbing. □

1.2.13 Proposition *The filter of zero neighborhoods of a topological vector space* E *over* $(R, |\cdot|)$ *has a filter base* $\mathcal{B}$ *with the following properties:*

(i) *For each* $V \in \mathcal{B}$ *there exists* $U \in \mathcal{B}$ *such that* $U + U \subset V$.

(ii) *Every element* $V \in \mathcal{B}$ *is circled and absorbing.*

(iii) *There exists an element* $r \in R^\times$ *with* $0 < |r| < 1$ *such that* $V \in \mathcal{B}$ *implies* $rV \in \mathcal{B}$.

Conversely, if $\mathcal{B}$ *is a filter base on an* R*-vector space* E *such that* (i) *to* (iii) *hold true, then there exists a unique vector space topology on* E *such that* $\mathcal{B}$ *is a neighborhood base at the origin. In case the ground ring* R *is archimedean, a filter base on* E *which satisfies* (i) *and* (ii) *already induces a unique vector space topology having* $\mathcal{B}$ *as a neighborhood base at* 0. *In either of these two cases, the thus constructed topology coincides with the coarsest translation invariant topology for which* $\mathcal{B}$ *is a set of zero neighborhoods.*

Proof. Assume that E is a tvs Let $\mathcal{B}$ be the set of circled neighborhoods of 0. We show first that $\mathcal{B}$ is a base of the filter $\mathcal{U}_0$ of zero neighborhoods. Let $W \in \mathcal{U}_0$. By Axiom (TVS2) there exists an $\varepsilon > 0$ and an open zero neighborhood U such that $sU \subset W$ for all $s \in R$ with $|s| < \varepsilon$. Then $V = \bigcup_{s \in R^\times \,\&\, |s| < \varepsilon} sU$ is a zero neighborhood since by Lemma 1.1.9 the set of $s \in R^\times$ with $|s| < \varepsilon$ is non-empty. By construction V is contained in W and circled, so $V \in \mathcal{B}$. Hence $\mathcal{B}$ is a filter base of $\mathcal{U}_0$.

Next recall that there exists $r \in R^\times$ with $0 < |r| < 1$ since the absolute value $|\cdot|$ is non-trivial. Let $V \in \mathcal{B}$. Then $sV \subset V$ for all $s \in R$ with $|s| \leqslant 1$ which entails $srV \subset rV$ for all such s. Hence rV is circled and an element of $\mathcal{B}$ as well. This proves (iii). Since addition on E is continuous, there exist for given $V \in \mathcal{B}$ open neighborhoods U_1, U_2 of the origin such that $U_1 + U_2 \subset V$. Choose $U \in \mathcal{B}$ such that $U \subset U_1 \cap U_2$. Then $U + U \subset V$ and (i) is proved. To show that any $V \in \mathcal{B}$ is absorbing let $v \in$ E. By continuity of scalar multiplication there exists $\varepsilon > 0$ such that $sv \in V$ for all $s \in R$ with $|s| < \varepsilon$. By Proposition 1.2.3 (i) this entails $v \in sV$ for all $s \in R$ with $|s| > \varepsilon$ and V is absorbing.

Now assume that E is an R-vector space and that $\mathcal{B}$ is a filter base that satisfies (i), (ii) and, if $|\cdot|$ is non-archimedean, (iii). Since $\mathcal{B}$ consists of non-empty circled sets, $0 \in V$ for all $V \in \mathcal{B}$. Let $\mathcal{T} \subset \mathcal{P}(\mathrm{E})$ be the set of all $U \subset$ E such that for each $v \in U$ there exists $V \in \mathcal{B}$ with $v + V \subset U$. By definition and since $\mathcal{B}$ is a filter base, $\mathcal{T}$ is a topology on E. By construction, $\mathcal{T}$ is also the coarsest translation invariant topology for which $\mathcal{B}$ is a set of zero neighborhoods. We show that $\mathcal{B}$ is a base of the filter $\mathcal{U}_0$ of zero neighborhoods. By definition of $\mathcal{T}$ there exists for each $U \in \mathcal{U}_0$ a $V \in \mathcal{B}$ such that $V \subset U$. So it remains to show that each $V \in \mathcal{B}$ is a zero neighborhood. To this end let U be the set of all $v \in V$ for which there exists a $W \in \mathcal{B}$ with $v + W \subset V$. Since $0 + V \subset V$ one has $0 \in U$. The relation $U \subset V$ holds because $0 \in W$ for all $W \in \mathcal{B}$. Now let $v \in U$. By (i) there exists W' such that $v + W' + W' \subset V$ which entails $v + W' \subset U$. Hence $U \in \mathcal{T}$ and V is a zero neighborhood. Next we verify that $\mathcal{T}$ is a vector space topology. We start with continuity of addition. Let W be an open neighborhood of $v + w$, where $v, w \in$ E. Then there exists $V \in \mathcal{B}$ such that $v + w + V \subset W$. Choose $U \in \mathcal{B}$ such that $U + U \subset V$. Then $v + U$ and

$w+U$ are neighborhoods of v and w, respectively, and $(v+U)+(w+U) \subset v+w+V \subset W$. So addition is continuous. We continue with scalar multiplication. Let W be an open neighborhood of rv, where $r \in R$ and $v \in \mathrm{E}$. Then there exists $V \in \mathcal{B}$ such that $rv+V+V \subset W$. Since V is absorbing by (ii) there exists $\varepsilon > 0$ such that $(s-r)v \in V$ for all $s \in R$ with $|s-r| < \varepsilon$. Now if $|\cdot|$ is non-archimedean choose $t \in R^\times$ according to (iii), and put $V_n = t^n V$ for all $n \in \mathbb{N}$. In the archimedean case let $t = \frac{1}{2}$ and use (i) to construct recursively a sequence $(V_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{B}$ such that $2^n V_n = V_n + \ldots + V_n \subset V$, where the sum has 2^n summands. In either of these cases, choose $N \in \mathbb{N}$ large enough so that $|t|^N < \frac{1}{|r|+\varepsilon}$. Then $V_N \in \mathcal{B}$ and $v + V_N$ is a neighborhood of v. Moreover, for $w \in v + V_N$ there exists an element $x \in V$ such that $w - v = t^N x$. Then the relation $s(w-v) = st^N x \in V$ holds whenever $|s-r| < \varepsilon$ since V_N is circled. Hence for such w and s

$$sw = rv + s(w-v) + (s-r)v \in rv + V + V \subset W \ .$$

This means that scalar multiplication is continuous, and the proof is finished. □

Morphisms of topological vector spaces

1.2.14 Definition By a *morphism* of topological vector spaces over the topological division ring R one understands a continuous R-linear map $f : \mathrm{E} \to \mathrm{F}$ between two topological vector spaces E and F over R. The space of morphisms between E and F will be denoted $\mathrm{Hom}_{R\text{-}\mathsf{TVS}}(\mathrm{E},\mathrm{F})$ or just $\mathrm{Hom}_R(\mathrm{E},\mathrm{F})$ or $\mathrm{Hom}(\mathrm{E},\mathrm{F})$ if now confusion can arrise.

1.2.15 Theorem *The topological vector spaces over a topological division ring R as objects together with their morphisms form an additive category which we denote by R-*TVS*. More precisely, R-*TVS* is a category enriched over the category of R-vector spaces where addition and scalar multiplication on the hom-spaces* $\mathrm{Hom}(\mathrm{E},\mathrm{F})$ *are given by*

$$+ : \mathrm{Hom}(\mathrm{E},\mathrm{F}) \times \mathrm{Hom}(\mathrm{E},\mathrm{F}) \to \mathrm{Hom}(\mathrm{E},\mathrm{F}),\ (f,g) \mapsto f+g = (\mathrm{E} \ni v \mapsto f(v)+g(v) \in \mathrm{F}) \ ,$$
$$\cdot : R \times \mathrm{Hom}(\mathrm{E},\mathrm{F}) \to \mathrm{Hom}(\mathrm{E},\mathrm{F}),\ (r,f) \mapsto r\cdot f = (\mathrm{E} \ni v \mapsto r\cdot f(v) \in \mathrm{F}) \ .$$

Proof. Observe first that the identity map id_E on a topological vector space E is linear and continuous and so is the composition $g \circ f$ of two morphisms of topological vector spaces $f : \mathrm{E} \to \mathrm{F}$ and $g : \mathrm{F} \to \mathrm{G}$. Hence topological vector spaces over R together with linear and continuous maps between them form a category.

Next check that the hom-space $\mathrm{Hom}(\mathrm{E},\mathrm{F})$ is an abelian group. Associativity and commutativity of addition follow from the respective properties on F. The zero element is the constant map $\mathrm{E} \to \mathrm{F}$, $v \mapsto 0$ and the inverse of a morphism $f : \mathrm{E} \to \mathrm{F}$ is given by $-f : \mathrm{E} \to \mathrm{F}$, $v \mapsto -f(v)$. Similarly one checks that multiplication by scalars on $\mathrm{Hom}(\mathrm{E},\mathrm{F})$ is associative and distributes from the left and from the right over addition since scalar multiplication on F has these properties. Finally, the unit of R acts as identity on $\mathrm{Hom}(\mathrm{E},\mathrm{F})$ since it does so on F. Hence $\mathrm{Hom}(\mathrm{E},\mathrm{F})$ carries the structure of an R left vector space.

Composition of morphisms $\mathrm{Hom}(\mathrm{E},\mathrm{F}) \times \mathrm{Hom}(\mathrm{F},\mathrm{G}) \to \mathrm{Hom}(\mathrm{E},\mathrm{G})$, $(f,g) \to g \circ f$ is an R-bilinear map as the following equalities for $f, f_1, f_2 \in \mathrm{Hom}(\mathrm{E},\mathrm{F})$, $g, g_1, g_2 \in \mathrm{Hom}(\mathrm{F},\mathrm{G})$, $r \in R$, and $v \in \mathrm{E}$

show:

$$\begin{aligned}
&(f \circ (g_1 + g_2))(v) = f((g_1 + g_2)(v)) = f(g_1(v) + g_2(v)) = \\
&\quad = f \circ g_1(v) + f \circ g_2(v) = (f \circ g_1 + f \circ g_2)(v) \ , \\
&(f \circ (rg))(v) = f((rg)(v)) = f(rg(v)) = rf(g(v)) = (r(f \circ g))(v) \ , \\
&((f_1 + f_2) \circ g)(v) = (f_1 + f_2)(g(v)) = f_1(g(v)) + f_2(g(v)) = \\
&\quad = f_1 \circ g(v) + f_2 \circ g(v) = (f_1 \circ g + f_2 \circ g)(v) \ , \\
&((rf) \circ g)(v) = (rf)(g(v)) = r(f(g(v))) = r(f \circ g(v)) = (r(f \circ g))(v) \ .
\end{aligned}$$

Hence R-TVS is a category enriched over the category of R-vector spaces. In particular, R-TVS then is an additive category. □

1.2.16 Example For every tvs E and non-zero element t of the ground ring R the map $\ell_t : \mathrm{E} \to \mathrm{E}$, $v \mapsto tv$ is an isomorphism of topological vector spaces by Proposition 1.2.3 (i).

1.2.17 Proposition and Definition *A linear map $f : \mathrm{E} \to \mathrm{F}$ between topological vector spaces over a valued division ring $(R, |\cdot|)$ maps symmetric sets to symmetric sets and circled sets to circled sets. If in addition f is continuous, then f is* bounded *that means it maps bounded subsets of* E *to bounded subsets of* F.

Proof. Since by linearity $f(tv) = tf(v)$ for all $v \in \mathrm{E}$ and $t \in R$, $f(C)$ is symmetric (respectively circled) if the subset $C \subset \mathrm{E}$ is.

To verify the second claim let $B \subset \mathrm{E}$ be bounded and $V \subset \mathrm{F}$ a zero neighborhood. Then $f^{-1}(V)$ is a zero neighborhood in E by continuity of f. Hence there exists an $r \in \mathbb{R}_{\geqslant 0}$ such that $B \subset tf^{-1}(V)$ for all $t \in R$ with $|t| \geqslant r$. By linearity of f one obtains $f(B) \subset tV$ for all such t, so f is bounded. □

1.2.18 Remark By the proposition continuity of a linear map between topological vector spaces implies the map to be bounded. As we will see later in this monograph, the converse does in general not hold true unless the underlying topological vector spaces are for example normable.

Normed real division algebras and local convexity

1.2.19 The major class of topological divison rings over which topological vector spaces are defined is formed by valued division rings $(R, |\cdot|)$ which carry the structure of an $\mathbb{R}$-algebra such that for all $r \in \mathbb{R}$ and $x \in R$ the equality

$$|rx| = |r|_\infty \cdot |x|$$

holds true. We will therefore given them a particular name and call them *normed real division algebras*. Note that the field of real numbers can be embedded into a normed real division algebra R by the natural map $\mathbb{R} \mapsto R$, $r \mapsto r \cdot 1$. Since $\mathbb{R}$ with its standard absolute value is archimedean, so is every normed real division algebra. By the Frobenius theorem, Frobenius (1878), there exist only three finite dimensional real division algebras, namely the field of real numbers $\mathbb{R}$, the field of complex numbers $\mathbb{C}$, and the quaternions $\mathbb{H}$.

1.2.20 Definition Under the assumption that R is a normed real division algebra one calls a subset $C \subset \mathrm{E}$ of an R-vector space

(i) *convex* if $tv + (1-t)w \in C$ for all $v, w \in C$ and $t \in \mathbb{R}$ with $0 \leqslant t \leqslant 1$,

(ii) *absolutely convex* if $rv + sw \in C$ for all $v, w \in C$ and $r, s \in R$ such that $|r| + |s| \leqslant 1$,

(iii) a *cone* if $tv \in C$ for all $v \in C$ and $t \in \mathbb{R}$ with $0 \leqslant t \leqslant 1$.

1.2.21 Lemma *Let R be a normed real division algebra. A subset C of an R-vector space* E *then is absolutely convex if and only if it is circled and convex.*

Proof. The claim is trivial when $C = \varnothing$, so we assume that C is nonempty.

Let C be absolutely convex. Since C contains at least one element v one has $0 = 0 \cdot v + 0 \cdot v \in C$. Hence $rv = (1 - |r|) \cdot 0 + rv \in C$ for all $v \in C$ and $r \in R$ with $|r| \leqslant 1$. So C is circled. By definition of absolute convexity C is convex.

If C is circled and convex, then it contains with elements v, w also $rv + sw$ if $|r| + |s| \leqslant 1$. To see this observe first that $\varrho v \in C$ and $\sigma w \in C$ where the elements $\varrho, \sigma \in R$ have been chosen so that $|\varrho| = |\sigma| = 1$, $r = |r| \cdot \varrho$ and $s = |s| \cdot \sigma$. Now if $|r| + |s| = 0$, then $rv + sw = 0 \in C$ since C is circled. If $|r| + |s| > 0$, then

$$rv + sw = (|r| + |s|) \left(\frac{|r|}{|r| + |s|} \varrho v + \frac{|s|}{|r| + |s|} \sigma w \right) \in C$$

since C is convex and circled. Hence C is absolutely convex. □

1.2.22 Lemma *A linear map $f : \mathrm{E} \to \mathrm{F}$ between vector spaces over a normed real divison algebra R maps convex sets to convex sets, absolutely convex sets to absolutely convex sets, and cones to cones.*

Proof. This an immediate consequence of the linearity of f. □

1.2.23 Lemma *Let* E *be a tvs over a normed real division algebra R, let $C, D \subset \mathrm{E}$ be convex and $r \in R$. Then the following holds true.*

(i) *The closure $\bar{C}$ and the interior $\mathring{C}$ are convex.*

(ii) *The sets $C + D$ and rC are convex.*

(iii) *If C is absolutely convex, then so are $\bar{C}$ and $\mathring{C}$.*

(iv) *If C is absolutely convex, then so is rC for each $r \in R^{\times}$.*

Proof. We consider only the cases $C, D \neq \varnothing$ because otherwise the claim is trivial.

ad (i). Let $t \in (0, 1)$. Then $t\bar{C} + (1-t)\bar{C} \subset \bar{C}$ by continuity of the map $\mathrm{E} \times \mathrm{E} \to \mathrm{E}$, $(v, w) \mapsto tv + (1-t)w$. Hence $\bar{C}$ is convex. Now let v, w be points of the interior of C and $z = tv + (1-t)w$. Then $z \in C$, and there exists a zero neighborhood U such that $v + U \subset C$ and $w + U \subset C$. Let $u \in U$ and compute

$$z + u = tv + (1-t)w + tu + (1-t)u = t(v + u) + (1-t)(w + u) \ .$$

Since both $v + u$ and $w + u$ are elements of C so is $z + u$ by convexity of C. Hence $z + U \subset C$ and z lies in the interior of C.

ad (ii). If $v, w \in C$, $x, y \in D$ and $t \in (0,1)$, then by convexity of C and D

$$t(v+x) + (1-t)(w+y) = \big(tv + (1-t)w\big) + \big(tx + (1-t)y\big) \in C + D \ .$$

Hence $C + D$ is convex. Similarly,

$$t(rv) + (1-t)(rw) = r\big(tv + (1-t)w\big) \in rC \ ,$$

so rC is convex as well.

ad (iii). Let C be absolutely convex. If $\mathring{C} \neq \emptyset$, then $0 \in \frac{1}{2}\mathring{C} - \frac{1}{2}\mathring{C} \subset C$, hence $0 \in \mathring{C}$. By Lemma 1.2.7 and (i) the claim now follows.

ad (iv). By (ii), rC is convex, so it remains to show that rC is circled. Assume that $v \in rC$. Then $v = rw$ for a unique $w \in C$. Since C is circled, $tw \in C$ for every $t \in R$ with $|t| \leqslant 1$. Hence $tv = r(tw) \in rC$ for such t and rC is circled. □

1.2.24 Proposition and Definition *The intersection of a non-empty family $(C_i)_{i\in I}$ of convex (respectively absolutely convex) subsets $C_i \subset \mathrm{E}$, $i \in I$ of a topological vector space* E *over a normed real division algebra R is convex (respectively absolutely convex). In particular, if $A \subset \mathrm{E}$ is a subset, then the sets*

$$\operatorname{Conv} A = \bigcap_{\substack{A \subset B \subset \mathrm{E} \\ B \text{ is convex}}} B \quad \text{and} \quad \operatorname{AConv} A = \bigcap_{\substack{A \subset B \subset \mathrm{E} \\ B \text{ is absolutely convex}}} B$$

are convex and absolutely convex, respectively. The set Conv A *is called the* convex hull *of A and is the smallest convex set containing A. Similarly,* AConv A *is the smallest absolutely convex set containing A. It is called the* absolutely convex hull *of A. The* closed convex hull $\overline{\operatorname{Conv}}\, A$ *and the* closed absolutely convex hull $\overline{\operatorname{AConv}}\, A$ *of A are defined by*

$$\overline{\operatorname{Conv}}\, A = \bigcap_{\substack{A \subset B = \bar{B} \subset \mathrm{E} \\ B \text{ is convex}}} B \quad \text{and} \quad \overline{\operatorname{AConv}}\, A = \bigcap_{\substack{A \subset B = \bar{B} \subset \mathrm{E} \\ B \text{ is absolutely convex}}} B \ .$$

These sets have the property that $\overline{\operatorname{Conv}}\, A$ is the smallest closed convex subset and $\overline{\operatorname{AConv}}\, A$ the smallest closed absolutely convex subset of E *containing A.*

Proof. Let C be the intersection $\bigcap\limits_{i\in I} C_i$ and assume that each C_i is absolutely convex. Let $v, w \in C$ and $r, s \in R$ with $|r| + |s| \leqslant 1$. Then $v, w \in C_i$, hence $rv + sw \in C_i$ for all $i \in I$. Therefore $rv + sw \in C$ and C is absolutely convex. This argument also shows that C is convex if all C_i are convex. The rest of the claim follows as in the proof of Proposition and Definition 1.2.8. □

1.2.25 Remark The proposition in particular entails that A is convex if and only if $\operatorname{Conv} A = A$ and absolutely convex if and only if $\operatorname{AConv} A = A$. Analogously, $\overline{\operatorname{Conv}}\, A = A$ if and only if A is closed and convex, and $\overline{\operatorname{AConv}}\, A = A$ if and only if A is closed and absolutely convex.

1.2.26 Lemma *Let $A \subset \mathrm{E}$ be a non-empty subset of a tvs E over a normed real division algebra R. Then*

$$\mathrm{Conv}\, A = \left\{ \sum_{i=1}^{k} t_i v_i \in \mathrm{E} \;\Big|\; k \in \mathbb{N}_{>0},\ v_1, \ldots v_k \in A,\ t_1 \ldots, t_k \in \mathbb{R}_{\geqslant 0},\ \sum_{i=1}^{k} t_i = 1 \right\}, \tag{1.2.1}$$

$$\mathrm{AConv}\, A = \left\{ \sum_{i=1}^{k} r_i v_i \in \mathrm{E} \;\Big|\; k \in \mathbb{N}_{>0},\ v_1, \ldots v_k \in A,\ r_1 \ldots, r_k \in R,\ \sum_{i=1}^{k} |r_i| \leqslant 1 \right\}. \tag{1.2.2}$$

For the closed hulls one has

$$\overline{\mathrm{Conv}}\, A = \overline{\mathrm{Conv}\, A} \quad \text{and} \quad \overline{\mathrm{AConv}}\, A = \overline{\mathrm{AConv}\, A} \; .$$

Finally, if A is circled, then

$$\mathrm{AConv}\, A = \mathrm{Conv}\, A \; .$$

Proof. By definition, the right hand side of Eq. (1.2.1) is convex and contains A, hence it contains $\mathrm{Conv}\, A$. Conversely, one shows by induction on $k \in \mathbb{N}_{>0}$ and convexity of $\mathrm{Conv}\, A$ that each element of the form $\sum_{i=1}^{k} t_i v_i$ with $v_1, \ldots, v_k \in A$ and $t_1, \ldots, t_k \in \mathbb{R}_{\geqslant 0}$ such that $\sum_{i=1}^{k} t_i = 1$ is in $\mathrm{Conv}\, A$. This proves Eq. (1.2.1). The proof of Eq. (1.2.2) is similar. Observe that the right hand side of Eq. (1.2.2) is absolutely convex and contains A. Hence it contains $\mathrm{AConv}\, A$. An argument using induction on $k \in \mathbb{N}_{>0}$ and absolute convexity of $\mathrm{AConv}\, A$ shows that each element of the form $\sum_{i=1}^{k} r_i v_i$ with $v_1, \ldots v_k \in A$ and $r_1 \ldots, r_k \in R$ such that $\sum_{i=1}^{k} |r_i| \leqslant 1$ is in $\mathrm{Conv}\, A$. So Eq. (1.2.2) holds true as well. The claim about the closed hulls is a consequence of Lemma 1.2.23. For the proof of the last claim it suffices to show that $\mathrm{Conv}\, A$ is circled if A is. To this end let $v \in \mathrm{Conv}\, A$ and $r \in R$ with $|r| \leqslant 1$. Then one can write v in the form $v = \sum_{i=1}^{k} t_i v_i$ with $v_1, \ldots, v_k \in A$ and $t_1, \ldots, t_k \in \mathbb{R}_{\geqslant 0}$, where $\sum_{i=1}^{k} t_i = 1$. Hence $rv = \sum_{i=1}^{k} t_i (r v_i)$, which is in $\mathrm{Conv}\, A$, since $r v_i \in A$ for all i because A is circled. □

1.2.27 Lemma *Let $A \subset \mathrm{E}$ be a non-empty subset of a tvs E over a normed real division algebra R.*

(i) *If A is convex and $t_1, \ldots, t_k \in \mathbb{R}_{\geqslant 0}$ with $k \in \mathbb{N}_{>0}$, then*

$$\sum_{i=1}^{k} t_i A = \left(\sum_{i=1}^{k} t_i \right) A \; .$$

(ii) *If A is absolutely convex and $r_1, \ldots, r_k \in R$ with $k \in \mathbb{N}_{>0}$, then*

$$\sum_{i=1}^{k} r_i A = \left(\sum_{i=1}^{k} |r_i| \right) A \; .$$

Proof. ad (i). Obviously $\sum_{i=1}^{k} t_i A \supset \left(\sum_{i=1}^{k} t_i \right) A$. Let us show the converse inclusion. Without loss of generality we can assume that $t_i > 0$ for all i. Then $t = \sum_{i=1}^{k} t_i > 0$, so, after division by t, we can reduce the claim to showing that $\sum_{i=1}^{k} t_i A \subset A$ for $t_1, \ldots, t_k \in \mathbb{R}_{>0}$ such that $\sum_{i=1}^{k} t_i = 1$. But $\sum_{i=1}^{k} t_i A \subset \mathrm{Conv}\, A = A$ by Lemma 1.2.26 and convexity of A.

ad (*ii*). Since by absolute convexity $r_i A = |r_i| A$ for $i = 1, \ldots, k$, the claim follows from (i). $\square$

1.2.28 Lemma *Let* $\mathbb{K}$ *be one of the division rings* $\mathbb{C}$ *or* $\mathbb{H}$ *with their standard absolute values and let* E *be a vector space over* $\mathbb{K}$. *Then a convex subset* $C \subset \mathrm{E}$ *is absorbent in* E *if and only if it is absorbent in the realification* $\mathrm{E}^{\mathbb{R}}$.

Proof. It suffices to show the non-trivial direction. So assume that C is convex and absorbent in the realification $\mathrm{E}^{\mathbb{R}}$. Denote by $u_1, \ldots, u_n$ the standard basis of $\mathbb{K}$ over $\mathbb{R}$ with $n = 2$ or $n = 4$ depending on $\mathbb{K}$. In particular this means $u_1 = 1$. For given $v \in \mathrm{E}$ there now exists $t \in \mathbb{R}_{\geqslant 0}$ such that

$$\pm \frac{1}{u_1} v, \ldots, \pm \frac{1}{u_n} v \in rC \quad \text{for all } r \geqslant t \ .$$

Without loss of generality we can assume $t \geqslant 1$. Let $z \in \mathbb{K}$ with $|z| \geqslant nt$. Then the vectors $c_1 = \operatorname{sgn} z_1 \frac{n}{|z|\, u_1} v, \ldots, c_n = \operatorname{sgn} z_n \frac{n}{|z|\, u_n} v$ are elements of C. By convexity of C and since $0 \in C$ one has $\frac{|z_1|}{|z|} c_1, \ldots, \frac{|z_n|}{|z|} c_n \in C$. Again by convexity one concludes

$$\frac{1}{z} v = \sum_{i=1}^{n} \frac{z_i}{|z|^2\, u_i} v = \sum_{i=1}^{n} \frac{|z_i|}{n|z|} c_i \in C \ .$$

Hence C is absorbing and the claim is proved. $\square$

1.2.29 Definition A topological vector space E over a normed real division algebra R for which Axiom LCVS below holds true is called a *locally convex topological vector space*, a *locally convex vector space* or shortly a *locally convex tvs.*

(LCVS) The vector space topology on E has a base consisting of convex sets.

1.2.30 Remark For better readability, we often say *locally convex topology* instead of *locally convex vector space topology.*

1.2.31 Proposition *The locally convex topological vector spaces over a normed real division algebra* R *together with the continuous linear maps between them form a full subcategory of the category* R-TVS *of topological* R*-vector spaces. It is denoted* R-LCVS.

Proof. This is clear by definition. $\square$

1.2.32 Proposition and Definition *The filter of zero neighborhoods of a locally convex topological vector space* E *over a normed real divison algebra* R *has a filter base* $\mathcal{B}$ *with the following properties:*

(i) *For each* $V \in \mathcal{B}$ *there exists* $U \in \mathcal{B}$ *such that* $U + U \subset V$.

(ii) *Every element of* $\mathcal{B}$ *is a* barrel *that means is absolutely convex, closed and absorbing.*

(iii) *Let* $r \in R^{\times}$. *Then* $V \in \mathcal{B}$ *if and only if* $rV \in \mathcal{B}$.

Conversely, if $\mathcal{B}$ *is a filter base on an* R*-vector space* E *such that* (i) *holds true and such that each element of* $\mathcal{B}$ *is absolutely convex and absorbing, then there exists a unique locally convex topology on* E *such that* $\mathcal{B}$ *is a neighborhood base of the origin. It is the coarsest among all translation invariant topologies for which* $\mathcal{B}$ *is a set of zero neighborhoods and is called the* locally convex topology generated *or* induced by $\mathcal{B}$.

Proof. Let E be a locally convex tvs. Let $\mathcal{B}$ be the collection of all barrels which are at the same time zero neighborhoods. Let V be an element of $\mathcal{U}_0$, the filter of zero neighborhoods. Since E is (T3) by Proposition 1.2.3, there exists a closed zero neighborhood V_a such that $V_a \subset V$. By local convexity of E there exists a convex zero neighborhood V_b with $V_b \subset V_a$. By Proposition 1.2.13 there exists a circled zero neighborhood V_c with $V_c \subset V_b$. The closed convex hull $U = \overline{\operatorname{Conv}} V_c$ then is a barrel contained in V. Since it is a zero neighborhood it is an element of $\mathcal{B}$, and $\mathcal{B}$ is a filter base of $\mathcal{U}_0$. This proves (ii).

To verify (i), let $V \in \mathcal{B}$ and observe that by continuity of addition there exist zero neighborhoods U_1 and U_2 such that $U_1 + U_2 \subset V$. Choose $U \in \mathcal{B}$ such that $U \subset U_1 \cap U_2$. Then $U + U \subset V$.

Claim (iii) holds true since multiplication by an element $r \in R^\times$ is a homeomorphism which preserves circled and convex sets.

The remaining claim follows immediately from Proposition 1.2.13 and the observation that a real division algebra is archimedean. □

1.2.33 Corollary *Let $\mathcal{S}$ be a non-empty set of absolutely convex and absorbent subsets of a vector space* E *over a normed real divison algebra R. Then the set*

$$\mathcal{B} = \Big\{ r \bigcap_{B \in \mathcal{F}} B \in \mathcal{P}(\mathrm{E}) \ \Big|\ \mathcal{F} \in \mathcal{P}_{\mathrm{fin}}(\mathcal{S}),\ \mathcal{F} \neq \varnothing\ \&\ r \in R^\times \Big\}$$

consists of absolutely convex and absorbent subsets of V *and is a base of the filter of zero neighborhoods of a locally convex topology $\mathcal{T}$ on* E *uniquely determined by that property. This topology is the coarsest among all vector space topologies for which $\mathcal{S}$ is a set of zero neighborhoods. The topology $\mathcal{T}$ is called the* locally convex topology generated *or* induced *by $\mathcal{S}$.*

Proof. The intersection of finitely many absolutely convex and absorbing sets is non-empty and again absolutely convex and absorbing by Lemma 1.2.12 (i) and Proposition and Definition 1.2.24. By Lemma 1.2.12 (ii) and Lemma 1.2.23, the scalar multiple of an absolutely convex and absorbing set again has these properties whenever the scalar is invertible. Hence each element of $\mathcal{B}$ is absolutely convex and absorbing. Given two elements $C, D \in \mathcal{B}$ there exist non-empty $\mathcal{F}, \mathcal{G} \in \mathcal{P}_{\mathrm{fin}}(\mathcal{S})$ and $r, s \in R^\times$ such that $C = r \bigcap_{B \in \mathcal{F}} B$ and $D = s \bigcap_{B \in \mathcal{G}} B$. Without loss of generality one can assume that $|r| \leqslant |s|$. Then $A = r \bigcap_{B \in \mathcal{F} \cup \mathcal{G}} B \in \mathcal{B}$ and $A = C \cap rs^{-1} D \subset C \cap D$ since D is balanced and $|rs^{-1}| \leqslant 1$. Hence $\mathcal{B}$ is a filter base consisting of absolutely convex and absorbent sets. Moreover, $\frac{1}{2}C + \frac{1}{2}C \subset C$ for every $C \in \mathcal{B}$ by absolut convexity. By Proposition 1.2.32 the filter base $\mathcal{B}$ therefore generates a unique locally convex topology $\mathcal{T}$ for which $\mathcal{B}$ is a base of the filter of zero neighborhoods. Moreover, $\mathcal{T}$ is the coarsest translation invariant topology so that $\mathcal{B}$ is a set of zero neighborhoods. This implies in particular that $\mathcal{S}$ is a set of zero neighborhoods for $\mathcal{T}$. Now let $\mathcal{T}'$ be a vector topology such that each element of $\mathcal{S}$ is a zero neighborhood. Then finite intersections of elements of $\mathcal{S}$ are zero neighborhoods with respect to $\mathcal{T}'$ and therefore also all elements of $\mathcal{B}$. Since $\mathcal{T}'$ is translation invariant one concludes that $\mathcal{T}$ is coarser than $\mathcal{T}'$ and the claim is proved. □

1.3. Seminorms and gauge functionals

1.3.1 Throughout the rest of this chapter the symbol $\mathbb{K}$ will always stand for the field of real numbers $\mathbb{R}$, the field of complex numbers $\mathbb{C}$ or the division algebra of quaternions $\mathbb{H}$. We assume these division algebras to be equipped with their standard absolute values $|\cdot|$. Moreover, vector spaces are assumed to be defined over the ground field $\mathbb{K}$ unless mentioned differently and are always assumed to be left vector spaces.

Seminorms and induced vector space topologies

1.3.2 Definition By a *seminorm* on a vector space E one understands a map $p : \mathrm{E} \to \mathbb{R}$ with the following properties:

(N0) The map p is *positive* that is $p(v) \geqslant 0$ for all $v \in \mathrm{E}$.

(N1) The map p is *absolutely homogeneous* that means

$$p(rv) = |r|\, p(v) \quad \text{for all } v \in \mathrm{E} \text{ and } r \in \mathbb{K}.$$

(N2) The map p is *subadditive* or in other words satisfies the *triangle inequality*

$$p(v + w) \leqslant p(v) + p(w) \quad \text{for all } v, w \in \mathrm{E}.$$

A seminorm is called a *norm* if in addition the following axiom is satisfied:

(N3) For all $v \in \mathrm{E}$ the relation $p(v) = 0$ holds true if and only if $v = 0$.

A vector space E equipped with a norm $\|\cdot\| : \mathrm{E} \to \mathbb{R}_{\geqslant 0}$ is called a *normed vector space.*

1.3.3 Let us introduce some useful further properties a map $p : \mathrm{E} \to \mathbb{R}$ can have. One calls such a map p

(1) *positively homogeneous* if $p(tv) = t\, p(v)$ for all $t \in \mathbb{R}_{>0}$ and all $v \in \mathrm{E}$,

(2) *sublinear* if $p(tv + sw) \leqslant t\, p(v) + s\, p(w)$ for all $t, s \in \mathbb{R}_{\geqslant 0}$ and all $v, w \in \mathrm{E}$, and

(3) *convex* if $p(tv + sw) \leqslant t\, p(v) + s\, p(w)$ for all $t, s \in \mathbb{R}_{\geqslant 0}$ with $t + s = 1$ and all $v, w \in \mathrm{E}$.

1.3.4 Lemma *For a real-valued map $p : \mathrm{E} \to \mathbb{R}$ on a vector space E the following are equivalent:*

(i) *p is sublinear.*

(ii) *p is positively homogeneous and convex.*

(iii) *p is positively homogeneous and subadditive.*

Proof. Let p be sublinear. Then p is subadditive by definition. Subadditivity implies $p(0) \leqslant p(0) + p(0)$, hence $p(0) \geqslant 0$. By sublinearity

$$p(0) = p(0 \cdot 0 + 0 \cdot 0) \leqslant 0 \cdot p(0) + 0 \cdot p(0) = 0 \ ,$$

so $p(0) = 0$. We show that p is positively homogeneous. Applying sublinearity again one checks for $v \in \mathrm{E}$ and $t \geqslant 0$ that

$$p(tv) = p(tv + 0 \cdot 0) \leqslant tp(v) + 0 \cdot p(0) = tp(v) \ ,$$

so p is positively homogeneous and the implication (i) $\implies$ (iii) follows. If p is positively homogeneous and subadditive, then for $v, w \in \mathrm{E}$ and $t, s > 0$ with $t + s = 1$

$$p(tv + sw) \leqslant p(tv) + p(sw) \leqslant tp(v) + sp(w),$$

so p is convex. This gives the implication (iii) $\implies$ (ii). If p is positively homogeneous and convex, then one computes for $v, w \in \mathrm{E}$ and $t, s \geqslant 0$ with $t + s > 0$

$$p(tv + sw) = (t+s)\, p\left(\frac{t}{t+s}v + \frac{s}{t+s}w\right) \leqslant (t+s)\left(\frac{t}{t+s}p(v) + \frac{s}{t+s}p(w)\right) = tp(v) + sp(w) \ .$$

Since $p(0) = \lim\limits_{t \searrow 0} p(t0) = \lim\limits_{t \searrow 0} t\, p(0) = 0$ by positive homogeneity, p then has to be sublinear and one obtains the implication (ii) $\implies$ (i). $\square$

1.3.5 Lemma *Let $p : \mathrm{E} \to \mathbb{R}$ be a real-valued map defined on a vector space E over $\mathbb{K}$.*

(i) *If $p : \mathrm{E} \to \mathbb{R}$ is positively homogeneous, then $p(0) = 0$.*

(ii) *If $p : \mathrm{E} \to \mathbb{R}$ is subadditive, then $p(0) \geqslant 0$ and for all $v, w \in \mathrm{E}$*

$$|p(v) - p(w)| \leqslant \max\{p(v - w), p(w - v)\} \ .$$

(iii) *If $p : \mathrm{E} \to \mathbb{R}$ is convex, then the sets $\mathbb{B}_{p,\varepsilon} := \{v \in \mathrm{E} \mid p(v) < \varepsilon\}$ and $\overline{\mathbb{B}}_{p,\varepsilon} := \{v \in \mathrm{E} \mid p(v) \leqslant \varepsilon\}$ are convex for all $\varepsilon > 0$.*

(iv) *If p is sublinear, then $\mathbb{B}_{p,\varepsilon}$ and $\overline{\mathbb{B}}_{p,\varepsilon}$ are convex and absorbent for all $\varepsilon > 0$.*

Proof. *ad (i).* As already observed, $p(0) = \lim\limits_{t \searrow 0} p(t0) = \lim\limits_{t \searrow 0} t\, p(0) = 0$.

ad (ii). Note that by subadditivity

$$p(0) \leqslant p(0) + p(0), \quad p(v) - p(w) \leqslant p(v - w), \quad \text{and} \quad p(w) - p(v) \leqslant p(w - v) \ .$$

This entails (ii).

ad (iii). Let $v, w \in \{v \in \mathrm{E} \mid p(v) < \varepsilon\}$ and $0 \leqslant t \leqslant 1$. Then, by convexity of p,

$$p\big(tv + (1-t)w\big) \leqslant tp(v) + (1-t)p(w) < t\varepsilon + (1-t)\varepsilon = \varepsilon \ .$$

Hence $tv + (1-t)w \in \{v \in \mathrm{E} \mid p(v) < \varepsilon\}$. The proof for $\{v \in \mathrm{E} \mid p(v) \leqslant \varepsilon\}$ is analogous.

ad (iv). Convexity of the sets $\mathbb{B}_{p,\varepsilon}$ and $\overline{\mathbb{B}}_{p,\varepsilon}$ holds by (iii). Moreover, $\mathbb{B}_{p,\varepsilon} \subset \overline{\mathbb{B}}_{p,\varepsilon}$ by definition. Hence it suffices by Lemma 1.2.28 to show that $\mathbb{B}_{p,\varepsilon}$ is absorbent in the realification $\mathrm{E}^{\mathbb{R}}$. Since p is positively homogenous by Lemma 1.3.4 and $0 \leqslant p(v) + p(-v)$ for all $v \in \mathrm{E}$, one concludes that for all $t \in \mathbb{R}$ and $v \in \mathrm{E}$

$$|p(tv)| \leqslant |t| \max\{p(v), p(-v)\} .$$

Hence $tv \in \mathbb{B}_{p,\varepsilon}$ if $0 < t < \frac{\varepsilon}{\max\{p(v),p(-v)\}+1}$, and $\mathbb{B}_{p,\varepsilon}$ is absorbent in $\mathrm{E}^{\mathbb{R}}$. □

1.3.6 Definition If $p : \mathrm{E} \to \mathbb{R}$ is a seminorm on a vector space E, we denote for every $v \in \mathrm{E}$ and $\varepsilon > 0$ by $\mathbb{B}_{p,\varepsilon}(v)$ the (*open*) *ε-ball associated with* p and *with center* v that is the set

$$\mathbb{B}_{p,\varepsilon}(v) = \left\{ w \in \mathrm{E} \;\middle|\; p(w - v) < \varepsilon \right\} .$$

The *closed ε-ball associated with* p and *with center* v is defined as

$$\overline{\mathbb{B}}_{p,\varepsilon}(v) = \left\{ w \in \mathrm{E} \;\middle|\; p(w - v) \leqslant \varepsilon \right\} .$$

The positive number ε is called the *radius* of the ball. In case the center of the ball is the origin, we often write $\mathbb{B}_{p,\varepsilon}$ and $\overline{\mathbb{B}}_{p,\varepsilon}$ for $\mathbb{B}_{p,\varepsilon}(0)$ and $\overline{\mathbb{B}}_{p,\varepsilon}(0)$, respectively. If in addition the radius equals 1, then we usually write only $\mathbb{B}_p$ and $\overline{\mathbb{B}}_p$ and call these sets the *open* respectively the *closed unit ball*. More generally, for the particular radius 1 we denote the corresponding balls by $\mathbb{B}_p(v)$ and $\overline{\mathbb{B}}_p(v)$ and call them the *open* respectively *closed unit balls with center v*. When by the context it is clear which seminorm p a ball is associated with we often do not mention p explicitely. This is in particular the case when the underlying vector space is a normed vector space.

If P is a finite set or a finite family of seminorms on E we define the *open* and *closed ε-multiballs with center v* by

$$\mathbb{B}_{P,\varepsilon}(v) = \left\{ w \in \mathrm{E} \;\middle|\; p(w - v) < \varepsilon \text{ for all } p \in P \right\}$$

and

$$\overline{\mathbb{B}}_{P,\varepsilon}(v) = \left\{ w \in \mathrm{E} \;\middle|\; p(w - v) \leqslant \varepsilon \text{ for all } p \in P \right\} ,$$

respectively. As before, we abbreviate $\mathbb{B}_{P,\varepsilon} = \mathbb{B}_{P,\varepsilon}(0)$ and $\overline{\mathbb{B}}_{P,\varepsilon} = \overline{\mathbb{B}}_{P,\varepsilon}(0)$.

1.3.7 Remark For convenience, we will also use the symbols $\mathbb{B}_{p,\varepsilon}$ and $\overline{\mathbb{B}}_{p,\varepsilon}$ to denote the sets $\left\{ v \in \mathrm{E} \;\middle|\; p(v) < \varepsilon \right\}$ and $\left\{ v \in \mathrm{E} \;\middle|\; p(v) \leqslant \varepsilon \right\}$, respectively, when $p : \mathrm{E} \to \mathbb{R}$ is just a real-valued convex map on the vector space E. Note that for such a p the set $\left\{ v \in \mathrm{E} \;\middle|\; p(v) < 0 \right\}$ might be non-empty. But as we have shown in Lemma 1.3.5 the sets $\mathbb{B}_{p,\varepsilon}$ and $\overline{\mathbb{B}}_{p,\varepsilon}$ associated to a convex p share with the the balls associated to a seminorm several nice properties like convexity.

1.3.8 Proposition *Let* E *be a* $\mathbb{K}$*-vector space, and* P *a finite set of seminorms on* E*. Then, for every* $\varepsilon > 0$ *and* $v \in \mathrm{E}$*, the ε-multiballs* $\mathbb{B}_{P,\varepsilon}(v)$ *and* $\overline{\mathbb{B}}_{P,\varepsilon}(v)$ *are convex. The ε-multiballs* $\mathbb{B}_{P,\varepsilon}$ *and* $\overline{\mathbb{B}}_{P,\varepsilon}$ *centered at the origin are absolutely convex and absorbent.*

Proof. Axiom (N1) immediately entails that $\mathbb{B}_{P,\varepsilon}$ and $\overline{\mathbb{B}}_{P,\varepsilon}$ are circled. Axiom (N2) together with (N1) entails that the sets $\mathbb{B}_{P,\varepsilon}(v)$ and $\overline{\mathbb{B}}_{P,\varepsilon}(v)$ are convex. Namely, if $w_1, w_2 \in \mathbb{B}_{P,\varepsilon}(v)$ and $t \in [0, 1]$, then one has for all seminorms $p \in P$

$$p\,(tw_1 + (1 - t)w_2 - v) \leqslant t\,p\,(w_1 - v) + (1 - t)\,p\,(w_2 - v) < t\,\varepsilon + (1 - t)\,\varepsilon = \varepsilon$$

and likewise $p\left(tw_1 + (1-t)w_2 - v\right) \leqslant \varepsilon$ for all $w_1, w_2 \in \overline{\mathbb{B}}_{P,\varepsilon}(v)$ and $p \in P$.

Now let $v \in \mathrm{E}$ and $\varepsilon > 0$ be given. Put $t_p = \frac{p(v)+1}{\varepsilon}$ for every $p \in P$ and $t_0 = \max\{t_p \mid p \in P\}$. Then one has for all $t \in \mathbb{K}$ with $|t| \geqslant t_0$ and for all $p \in P$

$$p\left(\frac{1}{t}v\right) \leqslant \frac{\varepsilon}{p(v)+1}\, p(v) < \varepsilon \ ,$$

hence $v \in t\,\mathbb{B}_{P,\varepsilon}$. So $\mathbb{B}_{P,\varepsilon}$ is absorbing. Since $\overline{\mathbb{B}}_{P,\varepsilon}$ contains the absorbing set $\mathbb{B}_{P,\varepsilon}$, it is absorbing as well. □

1.3.9 Proposition and Definition *Assume to be given a set Q of seminorms on a vector space* E. *Let $\mathcal{P}_{\mathrm{fin}}(Q)$ be the collection of all finite subsets of Q. A base of a topology on* E *then is given by*

$$\mathcal{B} = \left\{\mathbb{B}_{P,\varepsilon}(v) \;\middle|\; P \in \mathcal{P}_{\mathrm{fin}}(Q),\ v \in \mathrm{E},\ \varepsilon > 0\right\} .$$

The topology $\mathcal{T}$ generated by $\mathcal{B}$ is called the topology generated, induced *or* defined *by Q. Moreover, $\mathcal{T}$ is a locally convex vector space topology on* E. *It coincides with the coarsest translation invariant topology on* E *such that each seminorm in Q is continuous.*

Proof. Consider the set $\mathcal{B}_0$ of all multiballs $\mathbb{B}_{P,\varepsilon}$ with $P \in \mathcal{P}_{\mathrm{fin}}(Q)$ and $\varepsilon > 0$ centered at the origin. Clearly, $\mathcal{B}_0$ is a filter base since for $P_1, P_2 \in \mathcal{P}_{\mathrm{fin}}(Q)$ and $\varepsilon_1, \varepsilon_2 > 0$ the multiball $\mathbb{B}_{P_1 \cup P_2, \min\{\varepsilon_1,\varepsilon_2\}}$ is contained in $\mathbb{B}_{P_1,\varepsilon_1} \cap \mathbb{B}_{P_2,\varepsilon_2}$. Moreover it consists of absolutely convex and absorbing sets by Proposition 1.3.8.

By a similar argument one shows that $\mathcal{B}$ is base of a topology. Let $\mathbb{B}_{P_1,\varepsilon_1}(v_1), \mathbb{B}_{P_2,\varepsilon_2}(v_2) \in \mathcal{B}$ and $v \in \mathbb{B}_{P_1,\varepsilon_1}(v_1) \cap \mathbb{B}_{P_2,\varepsilon_2}(v_2)$. Let ε be the minium of the numbers $\varepsilon_1 - p_1(v - v_1)$ and $\varepsilon_2 - p_2(v - v_2)$, where p_1 runs through the elements of P_1 and p_2 through the ones of P_2. Then $\varepsilon > 0$ and $\mathbb{B}_{P_1 \cup P_2,\varepsilon}(v) \subset \mathbb{B}_{P_1,\varepsilon_1}(v_1) \cap \mathbb{B}_{P_2,\varepsilon_2}(v_2)$, and $\mathcal{B}$ is a base for a topology $\mathcal{T}$ indeed. By construction, $\mathcal{B}_0$ then is a base for the filter of zero neighborhoods and each element of $\mathcal{B}_0$ is open in $\mathcal{T}$. Moreover, each closed multiball $\overline{\mathbb{B}}_{P,\varepsilon}(v)$ is closed in $\mathcal{T}$ since the complement $\mathrm{E}\backslash\overline{\mathbb{B}}_{P,\varepsilon}(v)$ contains with w also the open multiball $\mathbb{B}_{P,\delta}(w)$, where $\delta = \min\{p(v-w) - \varepsilon | p \in P\}$.

We now prove continuity of addition with respect to $\mathcal{T}$. Let $v_1, v_2 \in \mathrm{E}$, $P \in \mathcal{P}_{\mathrm{fin}}(Q)$, and $\varepsilon > 0$. Since the triangle inequality holds for every seminorm in F, one has

$$\mathbb{B}_{P,\frac{\varepsilon}{2}}(v_1) + \mathbb{B}_{P,\frac{\varepsilon}{2}}(v_2) \subset \mathbb{B}_{P,\varepsilon}(v_1 + v_2) \ ,$$

which entails continuity of addition at each $(v_1, v_2) \in \mathrm{E} \times \mathrm{E}$. Next consider multiplication by scalars and let $\lambda \in \mathbb{K}$ and $v \in \mathrm{E}$. Again let $P = \{p_1, \ldots, p_n\} \in \mathcal{P}_{\mathrm{fin}}(Q)$ and $\varepsilon > 0$. Let $C_1 = \sup\{p_j(v) \mid 1 \leqslant j \leqslant n\} + 1$, $C_2 = |\lambda| + 1$ and put $\delta_1 = \min\{1, \frac{\varepsilon}{2\,C_1}\}$ and $\delta_2 = \frac{\varepsilon}{2\,C_2}$. Then one obtains by absolute homogeneity and subadditivity of each seminorm

$$p_j(\mu w - \lambda v) \leqslant |\mu|\, p_j(w - v) + |\mu - \lambda|\, p_j(v) \quad \text{for all } \mu \in \mathbb{K} \text{ and } w \in \mathrm{E},$$

hence

$$\mathbb{B}_{\delta_1}(\lambda) \cdot \mathbb{B}_{P,\delta_2}(v) \subset \mathbb{B}_{P,\varepsilon}(\lambda \cdot v) \ ,$$

where $\mathbb{B}_{\delta_1}(\lambda) = \{\mu \in \mathbb{K} \mid |\mu - \lambda| < \delta_1\}$. This shows continuity of scalar multiplication at each $(\lambda, v) \in \mathbb{K} \times \mathrm{E}$, and $\mathcal{T}$ is a vector space topology.

Since each of the base elements $\mathbb{B}_{P,\varepsilon} \in \mathcal{B}_0$ is convex, Axiom LCVS holds true as well and the topology $\mathcal{T}$ is locally convex.

Every seminorm $p \in Q$ is continuous with respect to the topology $\mathcal{T}$ since for all $a < b$ the preimage $p^{-1}((a,b)) = \mathbb{B}_{p,b} \backslash \overline{\mathbb{B}}_{p,a}$ is open in $\mathcal{T}$. Now let $\mathcal{T}'$ be a translation invariant topology on E for which every seminorm $p \in Q$ is continuous. In that topology $\mathcal{B}_0$ is a set of zero neighborhoods. As shown before, every element $B \in \mathcal{B}_0$ is absolutely convex, absorbing and satisfies $\frac{1}{2}B + \frac{1}{2}B \subset B$. Hence by Proposition and Definition 1.2.32 the topology $\mathcal{T}'$ is finer than the locally convex topology generated by $\mathcal{B}_0$. But the latter topology coincides with $\mathcal{T}$ by construction. This shows the last part of the claim and the proof is finished. □

Gauge functionals and induced seminorms

1.3.10 As we have seen, any vector space with a topology defined by a family of seminorms on it is a locally convex topological vector space. The converse also holds true. The fundamental notion needed for the proof of this is the following.

1.3.11 Definition Let E be a vector space and $A \subset \mathrm{E}$ absorbent. Then the map

$$p_A : \mathrm{E} \to \mathbb{R}_{\geqslant 0},\ v \mapsto p_A(v) = \inf \left\{ t \in \mathbb{R}_{>0} \;\middle|\; v \in tA \right\}$$

is called the *gauge functional*, the *Minkowski functional* or the *Minkowski gauge* of A.

1.3.12 Remark By definition of an absorbent set, $\left\{ t \in \mathbb{R}_{>0} \;\middle|\; v \in tA \right\}$ is non-empty whenever $A \subset \mathrm{E}$ is absorbent. Hence p_A is well-defined for such A.

1.3.13 Proposition *The Minkowski gauge $p_A : \mathrm{E} \to \mathbb{R}_{\geqslant 0}$ of an absorbent subset A of a vector space* E *has the following properties.*

(i) *The gauge functional is positively homogeneous that is $p_A(tv) = t\, p_A(v)$ for all $t \in \mathbb{R}_{>0}$ and all $v \in \mathrm{E}$.*

(ii) *If A is convex, then p_A is subadditive and*

$$\mathbb{B}_p(v) = \bigcup_{0<t<1} tA \subset A \subset \bigcap_{1<t} tA = \overline{\mathbb{B}}_p(v)\ .$$

(iii) *If A is absolutely convex, then p_A is a seminorm on* E.

Proof. If $t > 0$, then $tv \in sA$ for some $s > 0$ if and only if $v \in \frac{s}{t}A$. Hence $\left\{ s \in \mathbb{R}_{>0} \;\middle|\; tv \in sA \right\}$ and $t\left\{ s \in \mathbb{R}_{>0} \;\middle|\; v \in sA \right\}$ coincide for all $t > 0$, so (i) follows.

Assume that A is convex. Let $v, w \in \mathrm{E}$ and $\varepsilon > 0$. Then there exist $t > p_A(v)$ and $s > p_A(w)$ such that $v \in tA$, $w \in sA$, $t < p_A(v) + \frac{\varepsilon}{2}$ and $s < p_A(w) + \frac{\varepsilon}{2}$. By convexity of A and Lemma 1.2.27, $v + w \in tA + sA = (t+s)A$. Hence $p_A(v+w) \leqslant (t+s) < p_A(v) + p_A(w) + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, $p_A(v+w) \leqslant p_A(v) + p_A(w)$ and p_A is subadditive. If $v \in tA$ for some t with $0 < t < 1$, then $p_A(v) \leqslant t < 1$ by definition. Conversely, if $p_A(v) < 1$, then there exists a $t > 0$ such that $t < 1$ and $v \in tA$. Hence the equality $\mathbb{B}_p(v) = \bigcup_{0<t<1} tA$ follows. Since A is absorbing, 0 is an element of A. By convexity of A one therefore concludes $tA = (1-t)\{0\} + tA \subset A$ whenever $0 <$

$t < 1$. For $t > 1$ this shows $\frac{1}{t}A \subset A$, hence $A \subset tA$. So the relation $\bigcup_{0<t<1} tA \subset A \subset \bigcap_{1<t} tA$ is proved. Now assume that $v \in tA$ for all $t > 1$. Then $p_A(v) \leqslant 1$ by definition. If conversely $p_A(v) \leqslant 1$, then there exists for each $\varepsilon > 0$ an $s \geqslant 0$ such that $p_A(v) \leqslant s$, $v \in sA$ and $s < 1 + \varepsilon$. Hence, for $t \geqslant 1 + \varepsilon$ by Lemma 1.2.27 and $0 \in A$,

$$v \in sA = sA + (t-s)\{0\} \subset sA + (t-s)A = tA \ .$$

Since $\varepsilon > 0$ was arbitrary, $v \in tA$ for all $t > 1$ follows. So one obtains the equality $\bigcap_{1<t} tA = \overline{\mathbb{B}}_p(v)$, and (ii) is proved.

To verify (iii) recall that A is circled whenever A is absolutely convex. This entails for $r \in \mathbb{K}$, $v \in \mathrm{E}$ and absolutely convex A

$$p_A(rv) = \inf\left\{t \in \mathbb{R}_{>0} \;\middle|\; rv \in tA\right\} = \inf\left\{t \in \mathbb{R}_{>0} \;\middle|\; |r|v \in tA\right\} = p_A(|r|v) = |r|p_A(v) \ ,$$

where for the last equality we have used (i). □

1.3.14 Lemma *Let A and B be absorbent subsets of a vector space* E*. Then the following holds true.*

(i) $p_{tA}(v) = p_A(t^{-1}v)$ *for all* $t \in \mathbb{K}^\times$ *and* $v \in \mathrm{E}$.

(ii) *If* $B \subset A$, *then* $p_A \leqslant p_B$.

(iii) *If A is convex, then* $v \in tA$ *for all* $v \in \mathrm{E}$ *and* $t > p_A(v)$.

(iv) *If A and B are convex, then the intersection $A \cap B$ is absorbent and convex and* $p_{A\cap B} = \sup\{p_A, p_B\}$, *where* $\sup\{p_A, p_B\}(v) = \sup\{p_A(v), p_B(v)\}$ *for all* $v \in \mathrm{E}$.

Proof. ad (i). If $t \in \mathbb{K}$ is invertible, then $v \in tA$ if and only if $t^{-1}v \in A$.

ad (ii). Let $v \in \mathrm{E}$ and $\varepsilon > 0$. Then there exists t with $p_B(v) \leqslant t < p_B(v) + \varepsilon$ such that $v \in tB$. By $B \subset A$ this implies $v \in tA$, hence $p_A(v) \leqslant t < p_B(v) + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, the estimate $p_A \leqslant p_B$ follows.

ad (iii). By definition of the Minkowski gauge there exists $s \in \mathbb{R}$ such that $p_A(v) < s < t$ and $v \in sA$. By convexity of A one concludes $\frac{s}{t}v = \frac{s}{t}v + \left(1 - \frac{s}{t}\right) \cdot 0 \in sA$, hence $v \in tA$.

ad (iv). The intersection of convex sets is convex, so $A \cap B$ is convex. Let $v \in \mathrm{E}$ and choose $r_A \geqslant 0$ and $r_B \geqslant 0$ such that $v \in tA$ for all $t \geqslant r_A$ and $v \in sB$ for all $s \geqslant r_B$. Then $v \in (tA) \cap (tB) = t(A \cap B)$ for all $t \geqslant \max\{r_A, r_B\}$, so $A \cap B$ is absorbent. One has $p_{A\cap B} \geqslant \sup\{p_A, p_B\}$ by (ii). To show the converse inequality assume that $v \in \mathrm{E}$ and $t > \sup\{p_A(v), p_B(v)\}$. Then $v \in tA \cap tB = t(A \cap B)$, which implies $p_{A\cap B}(v) \leqslant t$. Hence $p_{A\cap B}(v) \leqslant \sup\{p_A(v), p_B(v)\}$ since $t > \sup\{p_A(v), p_B(v)\}$ was arbitrary. □

1.3.15 Lemma *Let* $p : \mathrm{E} \to \mathbb{R}$ *be a sublinear map on a vector space* E *and* $A \subset \mathrm{E}$ *convex. If*

$$\mathbb{B}_p \subset A \subset \overline{\mathbb{B}}_p \ ,$$

then the gauge functional p_A coincides with $\sup\{p, 0\}$. *If p is even a seminorm, then* $p = p_A$.

Proof. Let $p : \mathrm{E} \to \mathbb{R}$ be sublinear. Observe that then $\mathbb{B}_p$ is absorbent by Lemma 1.3.5 (iv). Hence A must also be absorbent by assumption, so the associated Minkowski gauge p_A is positively homogeneous by Proposition 1.3.13 (i).

Assume now that there exists $v \in \mathrm{E}$ such that $\max\{p(v), 0\} < p_A(v)$. By positive homogeneity of p and p_A one can achive by possibly multiplying v by a positive real number that $\max\{p(v), 0\} < 1 < p_A(x)$. The first inequality entails $v \in \mathbb{B}_p$, the second $v \notin \overline{\mathbb{B}}_p$ which is a contradiction. Next assume that there exists $v \in \mathrm{E}$ with $p_A(v) < \max\{p(v), 0\}$. As before one can then achieve that $p_A(v) < 1 < \max\{p(v), 0\}$ for some $v \in \mathrm{E}$. By the first inequality one concludes $v \in A$, by the second $v \notin A$. This is a contradiction. So the equality $\max\{p(v), 0\} = p_A(v)$ holds for all $v \in \mathrm{E}$.

In case p is a seminorm, then $p(v) \geqslant 0$ for all $v \in \mathrm{E}$ and the second claim follows by the first. □

1.3.16 Proposition *Let* E *be a topological vector space, and* $p : \mathrm{E} \to \mathbb{R}$ *be sublinear. Then the following are equivalent.*

(i) *The map p is continuous in the origin.*

(ii) *The map p is uniformly continuous.*

(iii) *The map p is continuous.*

(iv) *The unit ball* $\mathbb{B}_p$ *is a zero neighborhood.*

Proof. Let us first show (i) $\Longrightarrow$ (ii). To this end fix $\varepsilon > 0$. By assumption there exists a zero neighborhood $V \subset \mathrm{E}$ such that $|p(v)| < \varepsilon$ for all $v \in V$. By possibly passing to $V \cap (-V)$ one can assume that V is symmetric. Lemma 1.3.5 (ii) now implies

$$|p(v) - p(w)| < \varepsilon \quad \text{for all } v, w \in V \ .$$

Hence p is uniformly continuous. The implications (ii) $\Longrightarrow$ (iii) and (iii) $\Longrightarrow$ (iv) are trivial. It remains to prove (iv) $\Longrightarrow$ (i). Assume that $\mathbb{B}_p(0,1)$ is a zero neighborhood. Then there exists a symmetric zero neighborhood V contained in $\mathbb{B}_p(0,1)$. Since $p(0) = 0$ one concludes by Lemma 1.3.5 (ii)

$$|p(v)| < \max\{p(v), p(-v)\} < 1 \quad \text{for all } v \in V \ .$$

But this implies $|p(v)| < \varepsilon$ for all $v \in \varepsilon V$ and $\varepsilon > 0$, so p is continuous at the origin. □

Normability

1.3.17 Definition A topological vector space E is called *seminormable* if its topology is generated by a single seminorm $p : \mathrm{E} \to \mathbb{R}_{\geqslant 0}$. If the topology on E coincides with the vector space topology generated by a norm $\| \cdot \|$, then one calls E *normable.*

1.3.18 Theorem (Kolmogorov's normability criterion) *A topological vector space* E *is normable if and only if it is a* (T1) *space and possesses a bounded convex neighborhood of the origin.*

1.4. Cauchy filters and completeness

1.5. Function spaces and their topologies

1.5.1 Proposition *Let X be a topological space and (Y,d) a metric space. Then the following holds true.*

(i) *The space*

$$\mathcal{B}(X,Y) = \left\{ f : X \to Y \;\middle|\; \exists y_0 \in Y\, \exists C > 0\, \forall x \in X : \; d\big(f(x), y_0\big) \leqslant C \right\}$$

of bounded functions from X to Y is a metric space with metric

$$\varrho : \mathcal{B}(X,Y) \times \mathcal{B}(X,Y) \to \mathbb{R}_{\geqslant 0},\; (f,g) \mapsto \sup_{x \in X} d\big(f(x), g(x)\big) \; .$$

(ii) *If (Y,d) is complete, then $(\mathcal{B}(X,Y), \varrho)$ is so, too.*

(iii) *The space*

$$\mathcal{C}_{\mathrm{b}}(X,Y) = \mathcal{C}(X,Y) \cap \mathcal{B}(X,Y)$$

of continuous bounded functions from X to Y is a closed subspace of $\mathcal{B}(X,Y)$.

Proof. Note first that by the triangle inequality there exists for every $f \in \mathcal{B}(X,Y)$ and $y \in Y$ a real number $C_{f,y} > 0$ such that

$$d\big(f(x), y\big) \leqslant C_{f,x} \quad \text{for all } x \in X \; .$$

ad (*i*). Before verifying the axioms of a metric for ϱ we need to show that ϱ is well-defined meaning that $\sup_{x \in X} d\big(f(x), g(x)\big) < \infty$ for all $f, g \in \mathcal{B}(X,Y)$. To this end fix some $y \in Y$ and observe using the triangle inequality that

$$d\big(f(x), g(x)\big) \leqslant d\big(f(x), y\big) + d\big(y, g(x)\big) \leqslant C_{f,y} + C_{g,y} \quad \text{for all } x \in X \; .$$

Since furthermore $d\big(f(x), g(x)\big) \geqslant 0$ for all $x \in X$, the map ϱ is well-defined indeed with image in $\mathbb{R}_{\geqslant 0}$. If $\varrho(f,g) = 0$, then $d\big(f(x), g(x)\big) = 0$ for all $x \in X$, hence $f = g$. Obviously, ϱ is symmetric since d is symmetric. Finally, let $f, g, h \in \mathcal{B}(X,Y)$ and check using the triangle inequality for d:

$$\begin{aligned} \varrho(f,g) &= \sup_{x \in X} d\big(f(x), g(x)\big) \leqslant \sup_{x \in X} \Big(d\big(f(x), h(x)\big) + d\big(h(x), g(x)\big) \Big) \leqslant \\ &\leqslant \sup_{x \in X} d\big(f(x), h(x)\big) + \sup_{x \in X} d\big(h(x), g(x)\big) = d(f,h) + d(h,g) \; . \end{aligned}$$

Hence ϱ is a metric.

ad (*ii*). Assume (Y,d) to be complete and let $(f_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in $\mathcal{B}(X,Y)$. Let $\varepsilon > 0$ and choose $N_\varepsilon \in \mathbb{N}$ so that

$$\varrho(f_n, f_m) < \varepsilon \quad \text{for all } n, m \geqslant N \; .$$

Then for every $x \in X$ the relation

$$d\big(f_n(x), f_m(x)\big) < \varepsilon \quad \text{for all } n, m \geqslant N_\varepsilon \tag{1.5.1}$$

holds true, so $(f_n(x))_{n\in\mathbb{N}}$ is a Cauchy sequence in Y. By completeness of (Y, d) it has a limit which we denote by $f(x)$. By passing to the limit $m \to \infty$ in (1.5.1) one obtains that

$$d\big(f(x), f_n(x)\big) \leqslant \varepsilon \quad \text{for all } x \in X \text{ and } n \geqslant N_\varepsilon \ . \tag{1.5.2}$$

Using the triangle inequality one infers from this for an element $y \in Y$ which we now fix that

$$d\big(f(x), y)\big) \leqslant d\big(f(x), f_{N_1}(x))\big) + d\big(f_{N_1}(x), y\big) \leqslant 1 + C_{f_{N_1}, y} \ .$$

Hence f is a bounded function. Moreover, (1.5.2) entails that

$$\varrho(f, f_n) = \sup_{x\in X} d\big(f(x), f_n(x)\big) \leqslant \varepsilon \quad \text{for all } n \geqslant N_\varepsilon \ ,$$

so $(f_n)_{n\in\mathbb{N}}$ converges to f.

ad (iii). We have to show that the limit f of a sequence $(f_n)_{n\in}$ of functions $f_n \in \mathcal{C}_{\mathrm{b}}(X, Y)$ which converges in $(\mathcal{B}(X, Y), \varrho)$ has to be continuous. To this end let $\varepsilon > 0$ and choose $N_\varepsilon \in \mathbb{N}$ so that

$$\varrho(f_n, f) < \frac{\varepsilon}{3} \quad \text{for all } n \geqslant N_\varepsilon \ .$$

Let $x_0 \in X$. By continuity of f_{N_ε} there exists a neighborhood $U \subset X$ of x so that

$$d\big(f_{N_\varepsilon}(x), f_{N_\varepsilon}(x_0)\big) < \frac{\varepsilon}{3} \quad \text{for all } x \in U \ .$$

By the triangle inequality one concludes that

$$d\big(f(x), f(x_0)\big) \leqslant d\big(f(x), f_{N_\varepsilon}(x)\big) + d\big(f_{N_\varepsilon}(x), f_{N_\varepsilon}(x_0)\big) + d\big(f_{N_\varepsilon}(x_0), f(x_0)\big) < \varepsilon$$

for all $x \in U$. Hence f is continuous at x_0. Since $x_0 \in X$ was arbitrary f, is a continuous map, hence an elemnt of $\mathcal{C}_{\mathrm{b}}(X, Y)$.

□

1.5.2 Proposition *Let X be a topological space and $\mathbb{K}$ the division algebra of real or complex numbers or of quaternions. Then the following holds true.*

(i) *The space $\mathcal{B}(X, \mathbb{K})$ of bounded $\mathbb{K}$-valued functions on X can be expressed as*

$$\mathcal{B}(X, \mathbb{K}) = \big\{ f : X \to \mathbb{K} \ \big| \ \exists C > 0 \, \forall x \in X : \ |f(x)| \leqslant C \big\} \ . \tag{1.5.3}$$

It carries the structure of a $\mathbb{K}$-algebra by pointwise addition and multiplication of functions and becomes a Banach algebra when equipped with the supremums-norm

$$\| \cdot \|_\infty : \ \mathcal{B}(X, \mathbb{K}) \to \mathbb{K}, \quad f \mapsto \sup_{x\in X} |f(x)| \ .$$

(ii) *The subspace $\mathcal{C}_{\mathrm{b}}(X,\mathbb{K}) \subset \mathcal{B}(X,\mathbb{K})$ of bounded continuous $\mathbb{K}$-valued functions on X is a closed subalgebra of $\big(\mathcal{B}(X,\mathbb{K}), \|\cdot\|_\infty\big)$, so a Banach algebra as well when endowed with the supremums-norm. For X compact this means in particular that the algebra $\big(\mathcal{C}(X,\mathbb{K}), \|\cdot\|_\infty\big)$ is a Banach algebra.*

Proof. Eq. (1.5.3) is obvious since the distance of two elements $a, b \in \mathbb{K}$ is given by $d(a,b) = |a-b|$, so in particular $d(a,0) = |a|$. Let $f, g \in \mathcal{B}(X,\mathbb{K})$ and choose $C_f, C_g \geqslant 0$ so that $|f(x)| \leqslant C_f$ and $|g(x)| \leqslant C_g$ for all $x \in X$. Then, by the triangle inequality and absolute homogeneity of the absolute value,

$$|f(x) + g(x)| \leqslant C_f + C_g, \quad |a\, f(x)| \leqslant |a|\, C_f, \quad \text{and} \quad |f(x) \cdot g(x)| \leqslant C_f \cdot C_g \ .$$

Hence the sum and the product of two bounded functions are bounded and so is any scalar multiple of a bounded function. Therefore, $\mathcal{B}(X,\mathbb{K})$ is an algebra over $\mathbb{K}$. Using the triangle inequality and absolute homogeneity of the absolute value again one verifies that $\|f\|_\infty$ is a norm on $\mathcal{B}(X,\mathbb{K})$ indeed and that it fulfills $\|fg\|_\infty \leqslant \|f\|_\infty \cdot \|g\|_\infty$ for all $f, g \in \mathcal{B}(X,\mathbb{K})$. Furthermore, by definition, $\|f\|_\infty = \varrho(f,0)$ for all $f \in \mathcal{B}(X,\mathbb{K})$, where ϱ is defined as in Proposition 1.5.1. Since $(\mathcal{B}(X,\mathbb{K}), \varrho)$ is a complete metric space, $(\mathcal{B}(X,\mathbb{K}), \|\cdot\|_\infty)$ therefore is a Banach algebra. This proves the first claim.

For the second observe that for $f, g \in \mathcal{C}_{\mathrm{b}}(X,\mathbb{K})$ and $a \in \mathbb{K}$ the sum $f+g$, the scalar multiple af, and the product $f \cdot g$ are elements of $\mathcal{C}_{\mathrm{b}}(X,\mathbb{K})$ again. To verify this let $x \in X$ and $\varepsilon > 0$. Choose neighborhoods U_1 and U_2 of x so that

$$|f(y) - f(x)| < \min\left\{\frac{\varepsilon}{2}, \frac{\varepsilon}{|a|+1}, \frac{\varepsilon}{2(|g(x)|+1)}\right\} \quad \text{for } y \in U_1$$

and

$$|g(y) - g(x)| < \left\{1, \frac{\varepsilon}{2}, \frac{\varepsilon}{2(|f(x)|+1)}\right\} \quad \text{for } y \in U_2 \ .$$

Then for all $y \in U_1 \cap U_2$

$$\begin{aligned}
|(f+g)(y) - (f+g)(x)| &\leqslant |f(y) - f(x)| + |g(y) - g(x)| < \varepsilon \ , \\
|(af)(y) - (af)(x)| &\leqslant |a| \cdot |f(y) - f(x)| < \varepsilon \ , \\
|(f \cdot g)(y) - (f \cdot g)(x)| &\leqslant |g(y)| \cdot |f(y) - f(x)| + |f(x)| \cdot |(g(y) - g(x)| < \varepsilon \ .
\end{aligned}$$

This means that $f+g$, af and fg are continuous in x, hence elements of $\mathcal{C}_{\mathrm{b}}(X,\mathbb{K})$ since $x \in X$ was arbitrary. So $\mathcal{C}_{\mathrm{b}}(X,\mathbb{K})$ is a subalgebra of $\mathcal{B}(X,\mathbb{K})$. By Proposition 1.5.1 one knows that $\mathcal{C}_{\mathrm{b}}(X,\mathbb{K})$ is a closed subspace of $\mathcal{B}(X,\mathbb{K})$. The rest of the claim is obvious. □

1.5.3 As the next step, we introduce seminorms and their topologies on spaces of differentiable functions defined over an open set $\Omega \subset \mathbb{R}^n$. We agree that from now on Ω will always denote in this section an open subset of $\mathbb{R}^n$. For any differentiability order $m \in \mathbb{N} \cup \{\infty\}$ the symbol $\mathcal{C}^m(\Omega)$ stands for the space of m-times continuously differentiable complex valued functions on Ω. For $i = 1, \ldots, n$ we denote by $x^i : \mathbb{R}^n \to \mathbb{R}$ the i-th coordinate function and, if $m \geqslant 1$, by $\partial_i : \mathcal{C}^m(\Omega) \to \mathcal{C}^{m-1}(\Omega)$ the operator which maps $f \in \mathcal{C}^m(\Omega)$ to the partial derivative $\frac{\partial f}{\partial x^i}$. More generally, if $\alpha \in \mathbb{N}^n$ is a multiindex satisfying $|\alpha| = \alpha_1 + \ldots \alpha_n \leqslant m$, then we write

$\partial^\alpha : \mathcal{C}^m(\Omega) \to \mathcal{C}^{m-|\alpha|}(\Omega)$ for the higher order partial derivative which maps $f \in \mathcal{C}^m(\Omega)$ to $\frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdot \ldots \cdot \partial x_n^{\alpha_n}}$. Recall that the sum and the product of two m-times differentiable functions and scalar multiples of m-times differentiable functions are again m-times differentiable, hence $\mathcal{C}^m(\Omega)$ forms a $\mathbb{C}$-algebra. Now we define $\bar{\mathcal{C}}^m(\Omega)$ to be the space of continuous functions on the closure $\bar{\Omega}$ which are m-times continuosly differentiable on Ω so that each of its partial derivatives of order $\leqslant m$ has a continuos extension to $\bar{\Omega}$. Since the operators ∂_i are linear and also derivations by the Leibniz rule, $\bar{\mathcal{C}}^m(\Omega)$ is a subalgebra of $\mathcal{C}^m(\Omega)$. In general, these algebras do not coincide as for example the function $\frac{1}{x}$ on $\mathbb{R}_{>0}$ shows. It is an element of $\mathcal{C}^\infty(\mathbb{R}_{>0})$ but can not be extended to a continuous function on $\mathbb{R}_{\geqslant 0}$, so is not an element of $\bar{\mathcal{C}}^\infty(\mathbb{R}_{>0})$.

If $X \subset \mathbb{R}^n$ is locally closed which means that X is the intersection of an open and a closed susbet of $\mathbb{R}^n$, then define $\mathcal{C}^m(X)$ as the quotient space $\mathcal{C}^m(\Omega)/\mathcal{J}_X(\Omega)$, where $\Omega \subset \mathbb{R}^n$ open is chosen so that $X = \bar{X} \cap \Omega$ and where $\mathcal{J}_X$ denotes the ideal sheaf of all m-times continuously differentiable functions vanishing on X that is

$$\mathcal{J}_X(\Omega) = \left\{ f \in \mathcal{C}^m(\Omega) \;\middle|\; f|_X = 0 \right\} .$$

Using a smooth partition of unity type of argument one shows that $\mathcal{C}^m(X)$ does not depend on the particular choice of the neighborhood Ω in which X is relatively closed and that $\mathcal{C}^m(X)$ can be naturally identified with the space of continuous functions on X which have an extension to an element of $\mathcal{C}^m(\Omega)$.

1.5.4 Proposition *Let $\Omega \subset \mathbb{R}^n$ be open and bounded and $m \in \mathbb{N}_{>0}$. Then $\bar{\mathcal{C}}^m(\Omega)$ equipped with the norm*

$$\| \cdot \|_{\Omega,m} : \bar{\mathcal{C}}^m(\Omega) \to \mathbb{R}_{\geqslant 0}, \quad f \mapsto$$

1.6. Summability

1.6.1 Definition Assume to be given a locally convex topological vector space V over the field $\mathbb{K}$ of real or complex numbers. Let $(v_i)_{i \in I}$ be a family of elements of V. Let $\mathcal{F}(I)$ be the set of finite subsets of I and note that it is filtered by set-theoretic inclusion. The family $(v_i)_{i \in I}$ then gives rise to the net $\left(\sum_{i \in J} v_i \right)_{J \in \mathcal{F}(I)}$. One calls the family $(v_i)_{i \in I}$ *summable* to an element $v \in$ V if the net $\left(\sum_{i \in J} v_i \right)_{J \in \mathcal{F}(I)}$ converges to v. In other words this means that for every convex zero neighborhood $U \subset$ V and $\varepsilon > 0$ there exists an element $J_{U,\varepsilon} \in \mathcal{F}(I)$ such that for all finite sets J with $J_{U,\varepsilon} \subset J \subset I$

$$p_U \left(v - \sum_{i \in J} v_i \right) < \varepsilon .$$

As before, p_U denotes here the gauge of U. If V is Hausdorff, the limit v of a summable family $(v_i)_{i \in I}$ is uniquely determined, and one writes in this situation

$$v = \sum_{i \in I} v_i .$$

We denote the space of summable families in V over the given index set I by $\ell^1(I, \mathrm{V})$. For $E = \mathbb{C}$ we just write $\ell^1(I)$ instead of $\ell^1(I, \mathbb{C})$. If in addition the index set coincides with $\mathbb{N}$, we briefly denote $\ell^1(\mathbb{N})$ by ℓ^1.

1.6.2 Proposition (Cauchy criterion for summability) *Let* V *be a complete locally convex topological vector space. A family* $(v_i)_{i\in I}$ *of elements of* V *then is summable to some* $v \in \mathrm{V}$ *if and only if it satisfies the following Cauchy condition:*

(C) *For every convex zero neighborhood* $U \subset \mathrm{V}$ *and* $\varepsilon > 0$ *there exists an element* $J_{U,\varepsilon} \in \mathcal{F}(I)$ *such that for all* $K \in \mathcal{F}(I)$ *with* $K \cap J_{U,\varepsilon} = \varnothing$ *the relation*

$$p_U\left(\sum_{i\in K} v_i\right) < \varepsilon$$

holds true.

Proof. By completeness of V it suffices to verify that the net $\left(\sum_{i\in J} v_i\right)_{J\in\mathcal{F}(I)}$ is a Cauchy net if and only if condition (C) is satisfied. Recall that one calls $\left(\sum_{i\in J} v_i\right)_{J\in\mathcal{F}(I)}$ a Cauchy net if for every convex zero neighborhood $U \subset \mathrm{V}$ all $\varepsilon > 0$ there exists an element $J_{U,\varepsilon} \in \mathcal{F}(I)$ such that for all $J, J' \in \mathcal{F}(I)$ containing $J_{U,\varepsilon}$ as a subset the relation

$$p_U\left(\sum_{i\in J} v_i - \sum_{i\in J'} v_i\right) < \varepsilon$$

holds true. But that is clearly equivalent to condition (C). □

1.6.3 Several other notions of summability have been introduced in the analysis and functional analysis literature. These are mainly either used to establish summability criteria or are used in the study of topological tensor products and nuclearity of locally convex topological vector spaces, see Grothendieck (1955); Pietsch (1972). In the following we define these further notions of summability and study their properties. The symbol V hereby always stands for a locally convex tvs, I always denotes a nonempty index set, and $\mathcal{F}(I)$ the set of its finite subsets.

1.6.4 Definition A family $(v_i)_{i\in I}$ in V is called *weakly summable* to $v \in \mathrm{V}$ if for every continuous linear form $\alpha : \mathrm{V} \to \mathbb{K}$ the net $\left(\sum_{i\in J} \alpha(v_i)\right)_{J\in\mathcal{F}(I)}$ converges in $\mathbb{K}$ to $\alpha(v)$. In other words this means that for every $\alpha \in \mathrm{V}'$ and $\varepsilon > 0$ there exists a finite set $J_{\alpha,\varepsilon} \subset I$ such that for all finite sets J with $J_{\alpha,\varepsilon} \subset J \subset I$

$$\left|\alpha(v) - \sum_{j\in J} \alpha(v_i)\right| < \varepsilon \ .$$

The set of all weakly summable families in V with index set I is denoted $\ell^1[I, \mathrm{V}]$.

1.6.5 Definition A family $(v_i)_{i\in I}$ in V is called *absolutely summable* if for every circled convex zero neighborhood $U \subset \mathrm{V}$ there exists some $C \geqslant 0$ such that

$$\sum_{i\in J} p_U(v_i) \leqslant C \quad \text{for all } J \in \mathcal{F}(I) \ .$$

We denote the set of all absolutely summable families in V by $\ell^1\{I, \mathrm{V}\}$.

1.6.6 Proposition *A family $(v_i)_{i\in I} \subset \mathrm{V}$ is absolutely summable if and only if for every element U of a basis of circled convex zero neighborhoods there exists a $C \geqslant 0$ such that*

$$\sum_{i\in J} p_U(v_i) \leqslant C \quad \text{for all } J \in \mathcal{F}(I) \ .$$

Proof.

1.6.7 Definition A family $(v_i)_{i\in I}$ in V is called *totally summable* if there exists a bounded absolutely convex subset $B \subset \mathrm{V}$ and a $C \geqslant 0$ such that

$$\sum_{i\in J} p_B(v_i) \leqslant C \quad \text{for all } J \in \mathcal{F}(I) \ .$$

We write $\ell^1\langle I, \mathrm{V}\rangle$ for the set of all totally summable families in V.

Summable families of complex numbers

1.6.8 Lemma (cf. (Pietsch, 1972, Lem. 1.1.2)) *Let $(z_i)_{i\in I}$ be a family of complex numbers for which there exists a positive real number $C > 0$ such that*

$$\left|\sum_{i\in J} z_i\right| \leqslant C \quad \text{for all } J \in \mathcal{F}(I) \ .$$

Then one has the estimate

$$\sum_{i\in J} |z_i| \leqslant 4C \quad \text{for all } J \in \mathcal{F}(I) \ .$$

Proof. We assume first that all z_i are real. Then let I^+ the set of all indices $i \in I$ such that $z_i \geqslant 0$, and I^- the set of all $i \in I$ such that $z_i < 0$. Then, for all finite $J \subset I$

$$\sum_{i\in J} |z_i| = \sum_{i\in J\cap I^+} |z_i| + \sum_{i\in J\cap I^-} |z_i| = \left|\sum_{i\in J\cap I^+} z_i\right| + \left|\sum_{i\in J\cap I^-} z_i\right| \leqslant 2C \ .$$

In the general case decompose z_i into real and imaginary parts $x_i = \mathfrak{Re} z_i$ and $y_i = \mathfrak{Im} z_i$. By the triangle inequality one obtains for all finite $J \subset I$

$$\sum_{i\in J} |z_i| \leqslant \sum_{i\in J} |x_i| + \sum_{i\in J} |y_i| \leqslant 4C \ .$$

□

1.6.9 Proposition *For a family $(z_i)_{i\in I}$ of complex numbers the following are equivalent.*

(i) *The family $(z_i)_{i\in I}$ is summable.*

(ii) *The family $(|z_i|)_{i\in I}$ is summable.*

(iii) *The family $(z_i)_{i\in I}$ is absolutely summable.*

(iv) *There exists some $C > 0$ such that $\sum_{i\in J} |z_i| \leqslant C$ for all $J \in \mathcal{F}(I)$.*

In case that one hence all of the conditions are fulfilled, the estimate

$$\left| \sum_{i \in I} z_i \right| \leqslant \sum_{i \in I} |z_i|$$

holds true.

Proof. Assume that $(z_i)_{i \in I}$ is absolutely summable. Since $\mathbb{C}$ is normed with norm given by the absolut value this just means that there exists some $C > 0$ such that $\sum_{i \in J} |z_i| \leqslant C$ for all $J \in \mathcal{F}(I)$. Hence the supremum $c = \sup \{ \sum_{i \in J} |z_i| \mid J \in \mathcal{F}(I) \}$ exists and is $\leqslant C$. For given $\varepsilon > 0$ choose $J_\varepsilon \in \mathcal{F}(I)$ such that

$$c - \varepsilon \leqslant \sum_{i \in J_\varepsilon} |z_i| \leqslant c \ .$$

Then one has for all $K \in \mathcal{F}(I)$ with $K \cap J_\varepsilon = \varnothing$

$$\left| \sum_{i \in K} z_i \right| \leqslant \sum_{i \in K} |z_i| \leqslant \varepsilon \ .$$

Hence $\left(\sum_{i \in J} z_i \right)_{J \in \mathcal{F}(I)}$ is a Cauchy net, so has to converges by completeness of $\mathbb{C}$. This proves summability of $(z_i)_{i \in I}$.

Vice versa, assume now that $(z_i)_{i \in I}$ is summable. Then $\left(\sum_{i \in J} z_i \right)_{J \in \mathcal{F}(I)}$ is a Cauchy net. Hence there exists an element $J_1 \in \mathcal{F}(I)$ such that for all $K \in \mathcal{F}(I)$ with $K \cap J_1 = \varnothing$ the inequality

$$\left| \sum_{i \in K} z_i \right| < 1$$

holds true. Let $C = \sum_{i \in J_1} |z_i|$. Then one has for all $J \in \mathcal{F}(I)$

$$\left| \sum_{i \in J} z_i \right| \leqslant \left| \sum_{i \in J \setminus J_1} z_i \right| + \left| \sum_{i \in J \cap J_1} z_i \right| \leqslant 1 + C \ .$$

By the preceding lemma the set of partial sums $\sum_{i \in J} |z_i|$, where J runs through the finite subsets of I, is then bounded by $4 + 4C$, hence $(z_i)_{i \in I}$ is absolutely summable. □

Summability in Banach spaces

1.6.10 Proposition *Let* V *be a normed vector space. For a family* $(v_i)_{i \in I}$ *of elements in* V *the following are equivalent:*

(i) *The family* $(v_i)_{i \in I}$ *is absolutely summable.*

(ii) *The family* $(\|v_i\|)_{i \in I}$ *is summable.*

(iii) *There exists some* $C > 0$ *such that* $\sum_{i \in J} \|v_i\| \leqslant C$ *for all* $J \in \mathcal{F}(I)$.

If V *is even a Banach space, these conditions are all equivalent to*

(iv) *The family* $(v_i)_{i \in I}$ *is summable.*

Proof. (ii) and (iii) are equivalent by Proposition 1.6.9 Assume now that (i) holds true. □

to do: Carl Neumann series

Properties of and relations between the various summability types

1.6.11 Theorem *Let I be a non-empty index set. Then the spaces $\ell^1(I, \mathrm{V})$ of summable families, $\ell^1[I, \mathrm{V}]$ of weakly summable families, $\ell^1\{I, \mathrm{V}\}$ of absolutely summable families and $\ell^1\langle I, \mathrm{V}\rangle$ of totally summable families in E are all subvector spaces of the product vector space $E^I = \Pi_{i \in I} E$. Furthermore one has the following chain of inclusions:*

$$\ell^1\langle I, \mathrm{V}\rangle \subset \ell^1\{I, \mathrm{V}\} \quad \textit{and} \quad \ell^1(I, \mathrm{V}) \subset \ell^1[I, \mathrm{V}] \ .$$

If E is complete, then one even has

$$\ell^1\{I, \mathrm{V}\} \subset \ell^1(I, \mathrm{V})$$

Proof. Now let (v_i) be a summable family and $\alpha : \mathrm{V} \to \mathbb{K}$ a continuous linear form.

Let U be an absolutely convex zero neighborhood. Then U absorbes B, so there exists $r > 0$ such that $B \subset rU$. Hence □

1.7. Topological tensor products

1.7.1 Definition (cf. (Grothendieck, 1955, Chap. I, § 3, nº 3)) Let V and W be two locally convex topological vector spaces over the ground field $\mathbb{K}$. A locally convex vector topology τ on the (algebraic) tensor product $\mathrm{V} \otimes \mathrm{W}$ is called *compatible with the tensor product structure*, an *admissible tensor product topology* or just *admissible* if the following conditions hold true:

(ATPT1) The canonical map $\mathrm{V} \times \mathrm{W} \to \mathrm{V} \otimes_\tau \mathrm{W}$ is seperately continuous that is for each $v \in \mathrm{V}$ and each $w \in \mathrm{W}$ the linear maps

$$\mathrm{W} \to \mathrm{V} \otimes_\tau \mathrm{W}, y \mapsto v \otimes y \quad \text{and} \quad \mathrm{V} \to \mathrm{V} \otimes_\tau \mathrm{W}, x \mapsto x \otimes w$$

are continuous where $\mathrm{V} \otimes_\tau \mathrm{W}$ denotes the vector space $\mathrm{V} \otimes \mathrm{W}$ equipped with τ.

(ATPT2) For all linear maps $\alpha \in \mathrm{V}'$ and $\beta \in \mathrm{W}'$ the canonical linear map map $\alpha \otimes \beta : \mathrm{V} \otimes_\tau \mathrm{W} \to \mathbb{K}$ is continuous.

(ATPT3) For every equicontinuous subset $A \subset \mathrm{V}'$ and equicontinuous subset $B \subset \mathrm{W}'$ the set $\{\alpha \otimes \beta \mid \alpha \in A \,\&\, \beta \in B\}$ is an equicontinuous subset of the topological dual of $\mathrm{V} \otimes_\tau \mathrm{W}$.

The locally convex vector topology τ is called *strongly compatible with the tensor product structure*, a *strongly admissible tensor product topology* or briefly *strongly admissible* if it satisfies:

(sATPT) The canonical map $V \times W \to V \otimes_\tau W$ is continuous where $V \times W$ carries the product topology.

1.7.2 The admissible respectively strongly admissible vector topologies on $V \otimes W$ are obviously partially ordered by set-theoretic inclusion. Therefore, the following definition makes sense.

1.7.3 Definition

II.2. Banach Spaces and Algebras

2.1. Functional calculus

II.3. Hilbert Spaces

3.1. Inner product spaces

3.1.1 Let us first remind the reader that as before $\mathbb{K}$ stands for the field of real or of complex numbers. We will keep this notational agreement throughout the whole chapter.

3.1.2 Definition By a *sesquilinear form* on a $\mathbb{K}$-vector space V one understands a map $\langle\cdot,\cdot\rangle : \mathrm{V}\times\mathrm{V}\to\mathbb{K}$ with the following two properties:

(SF1) The map $\langle\cdot,\cdot\rangle$ is *conjugate-linear* in its first coordinate which means that

$$\langle v_1+v_2,w\rangle = \langle v_1,w\rangle + \langle v_2,w\rangle \quad \text{and} \quad \langle rv,w\rangle = \overline{r}\langle v,w\rangle$$

for all $v,v_1,v_2,w\in\mathrm{V}$ and $r\in\mathbb{K}$.

(SF2) The map $\langle\cdot,\cdot\rangle$ is *linear* in its second coordinate which means that

$$\langle v,w_1+w_2\rangle = \langle v,w_1\rangle + \langle v,w_2\rangle \quad \text{and} \quad \langle v,rw\rangle = r\langle v,w\rangle$$

for all $v,w,w_1,w_2\in\mathrm{V}$ and $r\in\mathbb{K}$.

A *hermitian form* is a sesquilinear form $\langle\cdot,\cdot\rangle$ on V with the following additional property:

(SF3) The map $\langle\cdot,\cdot\rangle$ is *conjugate-symmetric* which means that

$$\langle v,w\rangle = \overline{\langle w,v\rangle} \quad \text{for all } v,w\in\mathrm{V}\ .$$

A sesquilinear form $\langle\cdot,\cdot\rangle$ is called *weakly-nondegenerate* if it satisfies axiom

(SF4w) For every $v\in\mathrm{V}$, the map $\mathrm{V}\to\mathbb{K}$, $w\to\langle w,v\rangle$ is the zero map if and only if $v=0$.

Finally, one calls a hermitian form $\langle\cdot,\cdot\rangle$ on V *positive semidefinite* if

(SF5s) $\langle v,v\rangle\geqslant 0$ for all $v\in\mathrm{V}$.

3.1.3 Remark Recall that a map $\langle\cdot,\cdot\rangle : \mathrm{V}\times\mathrm{V}\to\mathbb{K}$ is called *bilinear* if it satisfies (SF2) and the following condition:

(BF1) The map $\langle\cdot,\cdot\rangle$ is *linear* in its first coordinate which means that

$$\langle v_1+v_2,w\rangle = \langle v_1,w\rangle + \langle v_2,w\rangle \quad \text{and} \quad \langle rv,w\rangle = r\langle v,w\rangle$$

for all $v,v_1,v_2,w\in\mathrm{V}$ and $r\in\mathbb{K}$.

In case the underlying ground field $\mathbb{K}$ coincides with the field of real numbers, a sesquilinear form is by definition the same as a bilinear form, and a hermitian form the same as a symmetric bilinear form.

3.1.4 Given a positive semidefinite hermitian form $\langle\cdot,\cdot\rangle$ on a $\mathbb{K}$-vector space V, one calls two vectors $v, w \in \mathrm{V}$ *orthogonal* if $\langle v, w\rangle = 0$. Since the hermitian form $\langle\cdot,\cdot\rangle$ is assumed to be positive semidefinite, the map

$$\|\cdot\| : \mathrm{V} \to \mathbb{R}_{\geqslant 0},\ v \mapsto \|v\| = \sqrt{\langle v, v\rangle}$$

is well-defined. We will later see that $\|\cdot\|$ is a seminorm on V and therefore call the map $\|\cdot\|$ the *seminorm associated to* $\langle\cdot,\cdot\rangle$. The following formulas are immediate consequences of the properties defining a positive semidefinite hermitian form and the definition of the associated seminorm:

$$\|v + w\|^2 = \|v\|^2 + 2\,\mathfrak{Re}\,\langle v, w\rangle + \|w\|^2 \quad \text{for all } v, w \in \mathrm{V}\ , \tag{3.1.1}$$

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 \quad \text{for all orthogonal } v, w \in \mathrm{V}\ , \tag{3.1.2}$$

$$\|v + w\|^2 + \|v - w\|^2 = 2\big(\|v\|^2 + \|w\|^2\big) \quad \text{for all } v, w \in \mathrm{V}\ , \tag{3.1.3}$$

$$\|rv\| = \sqrt{|r|^2\langle v, v\rangle} = |r|\|v\| \quad \text{for all } v, w \in \mathrm{V} \text{ and } r \in \mathbb{K}\ . \tag{3.1.4}$$

Formula (3.1.2) is an abstract version of the *pythagorean theorem*, Equation (3.1.3) is called the *parallelogram identity*. The triangle inequality for the map $\|\cdot\|$ will turn out to be a consequence of the next result.

3.1.5 Proposition (Cauchy–Schwarz inequality) *Given a positive semidefinite hermitian form $\langle\cdot,\cdot\rangle$ on a $\mathbb{K}$-vector space V the following inequality holds true:*

$$|\langle v, w\rangle| \leqslant \|v\|\|w\| \quad \textit{for all } v, w \in \mathrm{V}. \tag{3.1.5}$$

Equality holds if v and w are linearly dependant. In case $\langle\cdot,\cdot\rangle$ is positive definite, the converse holds true as well.

Proof. First consider the case where $\|v\| = \|w\| = 0$. Note that this does not imply that $v = 0$ (or $w = 0$) unless the hermitian form $\langle\cdot,\cdot\rangle$ is positive definite. Now put $c = -\langle v, w\rangle$ and compute

$$0 \leqslant \|cv + w\|^2 = 2\,\mathfrak{Re}\big(\bar{c}\,\langle v, w\rangle\big) = -2|\langle v, w\rangle|^2\ . \tag{3.1.6}$$

This entails $\langle v, w\rangle = 0$ and the Cauchy–Schwarz inequality is proved for $\|v\| = \|w\| = 0$.

If $\|v\| \neq 0$ or $\|w\| \neq 0$, we can assume without loss of generality that $\|v\| \neq 0$. Under this assumption put

$$c = -\frac{\langle v, w\rangle}{\|v\|^2}$$

and compute

$$\begin{aligned} 0 &\leqslant \|cv + w\|^2 = |c|^2\|v\|^2 + 2\,\mathfrak{Re}\big(\bar{c}\,\langle v, w\rangle\big) + \|w\|^2 = \\ &= \frac{|\langle v, w\rangle|^2}{\|v\|^2} - 2\frac{|\langle v, w\rangle|^2}{\|v\|^2} + \|w\|^2 = \|w\|^2 - \frac{|\langle v, w\rangle|^2}{\|v\|^2}\ . \end{aligned} \tag{3.1.7}$$

Hence the estimate

$$|\langle v, w\rangle|^2 \leqslant \|v\|^2 \|w\|^2$$

holds which entails the Cauchy–Schwarz inequality.

In the case where v, w are linearly dependant nonzero elements of V there exists a nonzero scalar $a \in \mathbb{K}$ such that $v = aw$. Therefore

$$|\langle v, w\rangle| = |a| \, \|w\|^2 = \|v\| \|w\| \, .$$

If one of v or w is 0, then both sides of the Cauchy–Schwarz inequality are 0.

In the positive definite case, equality in (3.1.5) entails by Equation (3.1.7) that $cv + w = 0$ whenever $v \neq 0$. If $v = 0$, then $v = 0 \cdot w$. In either case this means that v and w are linearly dependant. □

3.1.6 Lemma *A positive semidefinite hermitian form* $\langle \cdot, \cdot \rangle$ *on a* $\mathbb{K}$*-vector space* V *is weakly-nondegenerate if and only if it is* positive definite *that is if and only if*

(SF5p) $\langle v, v\rangle > 0$ *for all* $v \in \mathrm{V} \backslash \{0\}$.

Proof. A positive definite real bilinear or complex hermitian form $\langle \cdot, \cdot \rangle$ is weakly-nondegenerate since for every $v \in \mathrm{V}\backslash\{0\}$ the linear form $\langle v, - \rangle : \mathrm{V} \to \mathbb{K}$ is nonzero by $\langle v, v\rangle > 0$.

Conversely, if $\langle v, - \rangle : \mathrm{V} \to \mathbb{K}$ is nonzero for all $v \in \mathrm{V}\backslash\{0\}$, then there exists an element $w \in \mathrm{V}$ such that $\langle w, v\rangle \neq 0$. The Cauchy–Schwarz inequality entails

$$0 < |\langle w, v\rangle|^2 \leqslant \langle w, w\rangle \langle v, v\rangle \, ,$$

which implies $\langle v, v\rangle > 0$. Hence $\langle \cdot, \cdot \rangle$ is positive definite. □

3.1.7 Proposition *The map*

$$\| \cdot \| : V \to \mathbb{R}_{\geqslant 0}, \ v \mapsto \|v\| = \sqrt{\langle v, v\rangle}$$

associated to a positive semidefinite hermitian form $\langle \cdot, \cdot \rangle$ *on a* $\mathbb{K}$*-vector space* V *is a seminorm. If the hermitian form is positive definite, then* $\| \cdot \|$ *is even a norm.*

Proof. Absolute homogeneity (N1) is given by Eq. (3.1.4). The triangle inequality is a consequence of the Cauchy–Schwarz inequality:

$$\|v + w\|^2 = \|v\|^2 + 2\,\mathfrak{Re}\,\langle v, w\rangle + \|w\|^2 \leqslant \|v\|^2 + 2\,\|v\|\,\|w\| + \|w\|^2 = \big(\|v\| + \|w\|\big)^2 \, .$$

Finally, if $\langle \cdot, \cdot \rangle$ is positive definite, then $\|v\| = \sqrt{\langle v, v\rangle} > 0$ for all $v \in \mathrm{V}\backslash\{0\}$, so $\| \cdot \|$ is a norm. □

3.1.8 Definition By an *inner product* or a *scalar product* on a $\mathbb{K}$-vector space $\mathcal{H}$ one understands a positive definite hermitian form on $\mathcal{H}$. A $\mathbb{K}$-vector space $\mathcal{H}$ endowed with an inner product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{K}$ is called an *inner product space* or a *pre-Hilbert space.*

A hermitian form on a $\mathbb{K}$-vector space $\mathcal{H}$ which is only positive semidefinite is called a *semi-inner product* or a *semi-scalar product.*

A *Hilbert space* is an inner product space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ which is complete as a normed vector space. In other words, a Hilbert space is Banach space where the norm on the space is induced by an inner product.

3.1.9 Examples (a) The vector space $\mathbb{R}^n$ with the *euclidean inner product*

$$\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},\ \big((v_1,\ldots,v_n),(w_1,\ldots,w_n)\big) \mapsto \sum_{i=1}^n v_i w_i$$

is a real Hilbert space. Obviously, $\langle\cdot,\cdot\rangle$ is linear in the first argument, symmetric, and positive definite, hence a real inner product. The associated norm is the *euclidean norm*. We have seen before that $\mathbb{R}^n$ with the euclidean norm is complete.

(b) The vector space $\mathbb{C}^n$ together with the hermitian form

$$\langle\cdot,\cdot\rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C},\ \big((v_1,\ldots,v_n),(w_1,\ldots,w_n)\big) \mapsto \sum_{i=1}^n \overline{v}_i w_i$$

is a complex Hilbert space. One immediately verifies that $\langle\cdot,\cdot\rangle$ is linear in the second argument, conjugate-symmetric, and positive definite. Hence $\langle\cdot,\cdot\rangle$ is a complex inner product which we sometimes call the *standard hermitian inner product* on $\mathbb{C}^n$. Its associated norm is again the euclidean norm, so by completeness of $\mathbb{C}^n \cong \mathbb{R}^{2n}$ with respect to the euclidean norm one obtains the claim.

(c) The set

$$\ell^2 = \left\{ (z_k)_{k\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}} \;\middle|\; \sum_{k=0}^\infty |z_k|^2 < \infty \right\}$$

of square summable sequences of complex numbers is a complex Hilbert space with inner product

$$\langle\cdot,\cdot\rangle : \ell^2 \times \ell^2 \to \mathbb{C},\ \big((z_k)_{k\in\mathbb{N}},(w_k)_{k\in\mathbb{N}}\big) \mapsto \sum_{k=0}^\infty \overline{z}_k w_k\ .$$

To prove this one needs to first verify that ℓ^2 is a subvector space of $\mathbb{C}^{\mathbb{N}}$. For $z = (z_k)_{k\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ denote by $\|z\|$ the *extended norm* $\sqrt{\sum_{k=0}^\infty |z_k|^2} = \sup\limits_{K\in\mathbb{N}} \sqrt{\sum_{k=0}^K |z_k|^2} \in [0,\infty]$. Then $z \in \ell^2$ if and only if $\|z\| < \infty$. Now let $a \in \mathbb{C}$ and $z \in \ell^2$ and compute

$$\|az\| = \sqrt{\sum_{k=0}^\infty |az_k|^2} = |a| \sqrt{\sum_{k=0}^\infty |z_k|^2} = |a| \cdot \|z\| < \infty\ .$$

Hence $az \in \ell^2$. If $z, w \in \ell^2$, denote for each $K \in \mathbb{N}$ by $z_{(K)}$ and $w_{(K)}$ the "cut-off" vectors $(z_0,\ldots,z_K) \in \mathbb{C}^{K+1}$ and $(w_0,\ldots,w_K) \in \mathbb{C}^{K+1}$, respectively. By the triangle inequality for the norm on the Hilbert space $\mathbb{C}^{K+1}$ one concludes

$$\sqrt{\sum_{k=0}^K |z_k + w_k|^2} = \|z_{(K)} + w_{(K)}\| \leqslant \|z_{(K)}\| + \|w_{(K)}\| \leqslant \|z\| + \|w\| < \infty\ .$$

Therefore, the sequence of partial sums $\sum_{k=0}^K |z_k + w_k|^2$, $K \in \mathbb{N}$, is bounded, so convergent by the the monotone convergence theorem. One obtains

$$\|z + w\| = \lim_{K\to\infty} \sqrt{\sum_{k=0}^K |z_k + w_k|^2} \leqslant \|z\| + \|w\| < \infty\ .$$

Hence $z+w$ is square summable and ℓ^2 a vector subspace of $\mathbb{C}^{\mathbb{N}}$ indeed. Note that our argument also shows that the restriction of the extended norm to ℓ^2 is a norm.

We need to show that $\langle\cdot,\cdot\rangle$ is well-defined. To this end it suffices to prove that for all $z, w \in \ell^2$ the family $(z_k\overline{w}_k)_{k\in\mathbb{N}}$ is absolutely summable or in other words that $\sum_{k=0}^{\infty} |z_k\overline{w}_k| < \infty$. One concludes by the Hölder inequality for sums

$$\sum_{k=0}^{K} |\overline{z}_k w_k| = \sum_{k=0}^{K} |z_k w_k| \leqslant \|z_{(K)}\| \, \|w_{(K)}\| \leqslant \|z\| \, \|w\| \; .$$

So the left hand side has an upper bound uniform in K which by the monotone convergence theorem entails convergence of the partial sums and the estimate

$$\sum_{k=0}^{\infty} |\overline{z}_k w_k| \leqslant \|z\| \, \|w\| < \infty \; .$$

By definition it is clear that $\langle\cdot,\cdot\rangle$ is linear in the second argument, conjugate-symmetric and positive definite, hence a complex inner product. Note that the norm associated to $\langle\cdot,\cdot\rangle$ coincides with the above defined map $\|\cdot\|$.

It remains to be shown that ℓ^2 is complete. Let $(z^n)_{n\in\mathbb{N}}$ with $z^n = (z^n_k)_{k\in\mathbb{N}} \in \ell^2$ for all $n \in \mathbb{N}$ be a Cauchy sequence in ℓ^2. For $\varepsilon > 0$ choose $N_\varepsilon \in \mathbb{N}$ so that

$$\|z^n - z^m\| < \varepsilon \quad \text{for all } n, m \geqslant N_\varepsilon \; .$$

For each fixed $k \in \mathbb{N}$ one therefore has

$$|z^n_k - z^m_k| \leqslant \|z^n - z^m\| < \varepsilon \quad \text{for all } n, m \geqslant N_\varepsilon \; . \tag{3.1.8}$$

By completeness of $\mathbb{C}$ there exist $z_k \in \mathbb{C}$ such that $\lim_{n\to\infty} z^n_k = z_k$ for all $k \in \mathbb{N}$. We claim that $z = (z_k)_{k\in\mathbb{N}}$ is an element of ℓ^2 and that $(z^n)_{n\in\mathbb{N}}$ converges to z. To verify this observe that for all $\varepsilon > 0$, $K \in \mathbb{N}$ and $n \geqslant N_\varepsilon$

$$\sum_{k=0}^{K} |z_k - z^n_k|^2 = \lim_{m\to\infty} \sum_{k=0}^{K} |z^m_k - z^n_k|^2 \leqslant \sup_{m\geqslant N_\varepsilon} \sum_{k=0}^{K} |z^m_k - z^n_k|^2 \leqslant \sup_{m\geqslant N_\varepsilon} \|z^m - z^n\|^2 \leqslant \varepsilon^2 \; .$$

This implies by the triangle inequality and the fact that the Cauchy sequence $(z^n)_{n\in\mathbb{N}}$ is bounded in norm by some $C > 0$ that for all $K \in \mathbb{N}$ and $N = N_1$

$$\sqrt{\sum_{k=0}^{K} |z_k|^2} = \|z_{(K)}\| \leqslant \|z_{(K)} - z^N_{(K)}\| + \|z^N_{(K)}\| \leqslant \|z_{(K)} - z^N_{(K)}\| + \|z^N\| \leqslant 1 + C \; .$$

Hence $\|z\| = \sqrt{\sum_{k=0}^{\infty} |z_k|^2} \leqslant 1 + C$ and $z \in \ell^2$. In addition one obtains

$$\|z - z^n\| = \lim_{K\to\infty} \sqrt{\sum_{k=0}^{K} |z_k - z^n_k|^2} \leqslant \varepsilon \quad \text{for all } n \geqslant N_\varepsilon \; .$$

This means that z is the limit of the sequence $(z^n)_{n\in\mathbb{N}}$ and ℓ^2 is complete.

(d) Denote by λ the Lebesgue measure and let

$$\mathcal{L}^2(\mathbb{R}^d) = \left\{ f : \mathbb{R}^d \to \mathbb{C} \;\middle|\; f \text{ is Lebesgue measurable and } \|f\|_2 := \sqrt{\int_{\mathbb{R}^d} |f|^2 d\lambda} < \infty \right\}$$

be the space of Lebesgue square integrable functions on $\mathbb{R}^d$. Then $\mathcal{L}^2(\mathbb{R}^d)$ is a linear subspace of the space of all measurable functions by Minkowski's inequality which reads

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p \quad \text{for all measurable } f, g : \mathbb{R}^d \to \mathbb{C} .$$

Hereby, $\|f\|_p$ denotes for $p \in [1, \infty)$ the $\mathcal{L}^p$-seminorm $\left(\int_{\mathbb{R}^d} |f|^p d\lambda\right)^{1/p}$ of a measurable function $f : \mathbb{R}^d \to \mathbb{C}$. Note that $\|f\|_p$ can attain the value ∞, namely when f is not in the space $\mathcal{L}^p(\mathbb{R}^d)$. By Hölder's inequality, the product fg is Lebesgue integrable for $f, g \in \mathcal{L}^2(\mathbb{R}^d)$ and one has the estimate

$$\int_{\mathbb{R}^d} |fg| d\lambda = \|fg\|_1 \leqslant \|f\|_2 \, \|g\|_2 .$$

Hence the map

$$\langle \cdot, \cdot \rangle : \mathcal{L}^2(\mathbb{R}^d) \times \mathcal{L}^2(\mathbb{R}^d) \to \mathbb{C}, \ (f, g) \mapsto \int_{\mathbb{R}^d} \overline{f} g \, d\lambda$$

is well-defined and a positive semidefinite hermitian form on $\mathcal{L}^2(\mathbb{R}^d)$. By construction, the associated seminorm is the $\mathcal{L}^2$-seminorm $\| \cdot \|_2$. Modding out $\mathcal{L}^2(\mathbb{R}^d)$ by the kernel

$$\mathcal{N} := \operatorname{Ker}(\| \cdot \|_2) = \left\{ f \in \mathcal{L}^2(\mathbb{R}^d) \;\middle|\; \int_{\mathbb{R}^d} |f|^2 d\lambda = 0 \right\}$$

gives the Lebesgue space

$$L^2(\mathbb{R}^d) := \mathcal{L}^2(\mathbb{R}^d) / \mathcal{N} .$$

The hermitian form $\langle \cdot, \cdot \rangle$ vanishes on $\mathcal{N} \times \mathcal{L}^2(\mathbb{R}^d)$ and $\mathcal{L}^2(\mathbb{R}^d) \times \mathcal{N}$ by the Cauchy–Schwarz inequality, hence descends to a hermitian form

$$\langle \cdot, \cdot \rangle : L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d) \to \mathbb{C}, \ (f + \mathcal{N}, g + \mathcal{N}) \mapsto \int_{\mathbb{R}^d} \overline{f} g \, d\lambda .$$

That hermitian form is positive definite, since $\langle f + \mathcal{N}, f + \mathcal{N} \rangle = 0$ means $\int_{\mathbb{R}^d} |f|^2 d\lambda = 0$, hence $f \in \mathcal{N}$. Let us show that $L^2(\mathbb{R}^d)$ is complete with respect to the L^2-norm $\| \cdot \|_2$ induced by the inner product. Note that on the quotient space $\| \cdot \|_2$ is a norm indeed by construction. So let $(f_n + \mathcal{N})_{n \in \mathbb{N}}$ be a Cauchy sequence in $L^2(\mathbb{R}^d)$. Choose a subsequence $(f_{n_k})_{k \in \mathbb{N}}$ such that

$$\|f_{n_k} - f_{n_{k-1}}\|_2 < \frac{1}{2^k} \quad \text{for all } k \in \mathbb{N}_{>0}$$

and put

$$g_n(x) = \sum_{k=1}^{n} |f_{n_k}(x) - f_{n_{k-1}}(x)| \quad \text{for } x \in \mathbb{R}^d \text{ and } n \in \mathbb{N} .$$

The limit function

$$g : \mathbb{R}^d \to [0, \infty], \ x \mapsto \lim_{n \to \infty} g_n(x) = \liminf_{n \to \infty} g_n(x)$$

then exists even though it might not be finite everyhwere. Minkowski's inequality for the $\mathcal{L}^2$-norm entails that $\|g_n\|_2 \leqslant 1$ for all $n \in \mathbb{N}$, hence g is measurable and $\|g\|_2 \leqslant \liminf_{n\to\infty} \|g_n\|_2 \leqslant 1$ by Fatou's lemma. Therefore, $g(x)$ is finite for all x up to a set $Z \subset \mathbb{R}^d$ of measure 0, and for those x the series with partial sums $g_n(x)$ converges absolutely. For all $x \in \mathbb{R}^d \backslash Z$ the limit

$$f(x) = \lim_{k\to\infty} f_{n_k}(x) = f_{n_0} + \lim_{k\to\infty} \sum_{j=1}^{k} (f_{n_j}(x) - f_{n_{j-1}}(x))$$

therefore exists in $\mathbb{C}$. Put $f(x) = 0$ for all $x \in Z$, and let $\chi_Z : \mathbb{R}^d \to \mathbb{R}$ be the characteristic function of Z. Then the sequence of functions $(\chi_Z f_{n_k})_{k\in\mathbb{N}}$ converges pointwise to f, and each of the functions $\chi_Z f_n$ is measurable, actually even square integrable. Since

$$|\chi_Z f_{n_k}| \leqslant |\chi_Z f_{n_0}| + g_k \leqslant |\chi_Z f_{n_0}| + g \quad \text{for all } k \in \mathbb{N}$$

and since $|\chi_Z f_{n_0}| + g$ is square integrable by Minkowski's inequality, the pointwise limit f is square integrable by Lebesgue's dominated convergence theorem, and $f + \mathcal{N}$ is in $L^2(\mathbb{R}^d)$. It remains to show that $(f_n + \mathcal{N})_{n\in\mathbb{N}}$ converges to $f + \mathcal{N}$ in the norm $\|\cdot\|_2$. To this end let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that $\|f_n - f_m\|_2 < \varepsilon$ for $n, m \geqslant N$. By Fatou's lemma one obtains

$$\int_{\mathbb{R}^d} |f_n - f|^2 d\lambda \leqslant \liminf_{m\to\infty} \int_{\mathbb{R}^d} |f_n - f_m|^2 d\lambda \leqslant \varepsilon^2 \quad \text{for all } n \geqslant N \ .$$

Hence $\lim_{n\to\infty} \|f_n - f\|_2 = 0$, and $L^2(\mathbb{R}^d)$ endowed with the inner product $\langle\cdot,\cdot\rangle$ is a Hilbert space. It is called the *Hilbert space of square-integrable functions* on $\mathbb{R}^d$. Note that for every complete measure space (Ω, μ) one obtains in the same way the Hilbert space $L^2(\Omega, \mu)$ of square-integrable functions on (Ω, μ).

3.1.10 Theorem *Let* V *be a normed* $\mathbb{K}$*-vector space. Then the norm* $\|\cdot\| : \mathrm{V} \to \mathbb{R}_{\geqslant 0}$ *is associated to an inner product* $\langle\cdot,\cdot\rangle : \mathrm{V} \times \mathrm{V} \to \mathbb{K}$ *if and only if the* parallelogram identity

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2$$

holds true for all $v, w \in \mathrm{V}$. *In this case, the inner product of two elements* $v, w \in \mathrm{V}$ *can be expressed by the* polarization identity for $\mathbb{K} = \mathbb{R}$

$$\langle v, w\rangle = \frac{1}{4}\left(\|v + w\|^2 - \|v - w\|^2\right) = \frac{1}{2}\left(\|v + w\|^2 - \|v\|^2 - \|w\|^2\right) \tag{3.1.9}$$

respectively by the polarization identity for $\mathbb{K} = \mathbb{C}$

$$\langle v, w\rangle = \frac{1}{4}\sum_{k=1}^{4} \mathrm{i}^k \, \|w + \mathrm{i}^k \, v\|^2 \ . \tag{3.1.10}$$

Proof. The forward direction is a consequence of 3.1.4, Eq. 3.1.3. To show the backward direction we consider two cases $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$ separately.

1. Case. Given the norm $\|\cdot\|$ define $\langle\cdot,\cdot\rangle : \mathrm{V} \times \mathrm{V} \to \mathbb{R}$ by real polarization

$$\langle v, w\rangle = \frac{1}{4}\left(\|v + w\|^2 - \|v - w\|^2\right), \quad \text{where } v, w \in \mathrm{V} \ .$$

Note that the parallelogram identity entails

$$\frac{1}{4}\left(\|v+w\|^2-\|v-w\|^2\right)=\frac{1}{2}\left(\|v+w\|^2-\|v\|^2-\|w\|^2\right) .$$

Observe that by definition $\langle v,w\rangle=\langle w,v\rangle$ and $\|v\|=\sqrt{\langle v,v\rangle}$. Let us show additivity in the first variable. Let $v_1,v_2,w\in \mathrm{V}$ and compute using the parallelogram identity

$$\begin{aligned}\|v_1+v_2+w\|^2 &= 2\|v_1+w\|^2+2\|v_2\|^2-\|v_1+w-v_2\|^2 \ ,\\ \|v_1+v_2+w\|^2 &= 2\|v_2+w\|^2+2\|v_1\|^2-\|v_2+w-v_1\|^2 \ .\end{aligned}$$

Hence

$$\|v_1+v_2\pm w\|^2=\|v_1\pm w\|^2+\|v_2\pm w\|^2+\|v_1\|^2+\|v_2\|^2-\|v_1\pm w-v_2\|^2-\|v_2\pm w-v_1\|^2 \ .$$

Subtracting the $-$ version from the $+$ version of this equation entails

$$\begin{aligned}\langle v_1+v_2,w\rangle &= \frac{1}{4}\left(\|v_1+v_2+w\|^2-\|v_1+v_2-w\|^2\right)=\\ &= \frac{1}{4}\left(\|v_1+w\|^2+\|v_2+w\|^2-\|v_1-w\|^2-\|v_2-w\|^2\right)=\langle v_1,w\rangle+\langle v_2,w\rangle \ ,\end{aligned}$$

so additivity in the first variable is proved. By induction one derives from this that for all natural n

$$\langle nv,w\rangle=n\langle v,w\rangle \quad \text{for all } v,w\in \mathrm{V} \ . \tag{3.1.11}$$

Since then $\langle -nv,w\rangle-n\langle v,w\rangle=\langle -nv+nv,w\rangle=0$ for all $n\in\mathbb{N}$, Eq. (3.1.11) also holds for $n\in\mathbb{Z}$. Now let $p\in\mathbb{Z}$ and $q\in\mathbb{N}_{>0}$. Then $q\langle \frac{p}{q}v,w\rangle=\langle pv,w\rangle=p\langle v,w\rangle$, hence one has for rational r

$$\langle rv,w\rangle=r\langle v,w\rangle \quad \text{for all } v,w\in \mathrm{V} \ . \tag{3.1.12}$$

Since addition, multiplication by scalars and the norm are continuous, the function

$$\mathbb{R}\to\mathbb{R},\ r\mapsto\langle rv,w\rangle-r\langle v,w\rangle=\frac{1}{4}\left(\|rv+w\|^2+r\|v-w\|^2-\|rv-w\|^2-r\|v+w\|^2\right)$$

is continuous. Since it vanishes over $\mathbb{Q}$, it has to coincide with the zero map. Therefore, Eq. (3.1.12) holds for all $r\in\mathbb{R}$. So $\langle\cdot,\cdot\rangle$ is linear in the first coordinate. By symmetry, it is so too in the second coordinate. Hence $\langle\cdot,\cdot\rangle$ is a symmetric bilinear form inducing $\|\cdot\|$.

2. Case. In the case $\mathbb{K}=\mathbb{C}$ use complex polarization and put

$$\langle v,w\rangle=\frac{1}{4}\sum_{k=1}^{4}\mathsf{i}^k\,\|w+\mathsf{i}^k\,v\|^2 \quad \text{for all } v,w\in \mathrm{V} \ .$$

Then $\langle\cdot,\cdot\rangle$ is conjugate-symmetric, since

$$\overline{\langle v,w\rangle}=\frac{1}{4}\sum_{k=1}^{4}(-\mathsf{i})^k\,\|w+\mathsf{i}^k\,v\|^2=\frac{1}{4}\sum_{k=1}^{4}(-\mathsf{i})^k\,\|(-\mathsf{i})^k\,w+v\|^2=\langle w,v\rangle \ .$$

Next compute

$$\mathfrak{Re}\,\langle v,w\rangle=\frac{1}{4}\left(\|w+v\|^2-\|w-v\|^2\right)$$

and

$$\mathfrak{Im}\,\langle v, w\rangle = \frac{1}{4}\left(\|w + \mathsf{i}v\|^2 - \|w - \mathsf{i}v\|^2\right) \ .$$

By the first case one concludes that $\mathfrak{Re}\langle\cdot,\cdot\rangle$ and $\mathfrak{Im}\langle\cdot,\cdot\rangle$ are both $\mathbb{R}$-linear in the first and the second coordinate. Moreover,

$$\mathfrak{Re}\,\langle v, \mathsf{i}w\rangle = \frac{1}{4}\left(\|\mathsf{i}w + v\|^2 - \|\mathsf{i}w - v\|^2\right) = \frac{1}{4}\left(\|w - \mathsf{i}v\|^2 - \|w + \mathsf{i}v\|^2\right) = -\mathfrak{Im}\,\langle v, w\rangle = \mathfrak{Re}\,\mathsf{i}\langle v, w\rangle$$

and

$$\mathfrak{Im}\,\langle v, \mathsf{i}w\rangle = \frac{1}{4}\left(\|\mathsf{i}w + \mathsf{i}v\|^2 - \|\mathsf{i}w - \mathsf{i}v\|^2\right) = \mathfrak{Re}\,\langle v, w\rangle = \mathfrak{Im}\,\mathsf{i}\langle v, w\rangle \ ,$$

hence $\langle\cdot,\cdot\rangle$ is complex linear in the second coordinate. Finally,

$$\mathfrak{Re}\,\langle v, v\rangle = \|v\|^2 \quad \text{and} \quad \mathfrak{Im}\,\langle v, v\rangle = \frac{1}{4}\left(\|v + \mathsf{i}v\|^2 - \|v - \mathsf{i}v\|^2\right) = 0 \ .$$

This finishes the proof that $\langle\cdot,\cdot\rangle$ is a complex inner product inducing the norm $\|\cdot\|$. □

3.1.11 Next we will turn Hilbert spaces into a category. To this end one needs to know what morphisms in this category should be. There are two options each giving rise to a category of Hilbert spaces. These categories just differ by their morphism classes. The first one is to have as morphisms linear maps $A : \mathcal{H}_1 \to \mathcal{H}_2$ preserving the inner products which means that they fulfill

$$\langle Av_1, Av_2\rangle = \langle v_1, v_2\rangle \quad \text{for all } v_1, v_2 \in \mathcal{H}_1 \ .$$

By Theorem 3.1.10 this property is equivalent to

$$\|Av\| = \|v\| \quad \text{for all } v \in \mathcal{H}_1 \ ,$$

that is to A being *norm preserving* or *isometric*. Obviously, the identity map on a Hilbert space is isometric and the composition of two composable isometric linear maps is again isometric and linear. Hence Hilbert spaces together with norm preserving linear maps between them form a category which we denote by $\mathsf{Hilb}_{\mathrm{np}}$. The isomorphisms in this category are the surjective isometric linear maps between Hilbert spaces. Such maps are called *unitary*. The condition of a linear map being norm preserving is pretty restrictive, so the category $\mathsf{Hilb}_{\mathrm{np}}$ contains only few morphisms. This can be cured by allowing all *bounded* linear maps between Hilbert spaces to be morphisms that is of all linear $A : \mathcal{H}_1 \to \mathcal{H}_2$ for which there exists a $C \geqslant 0$ such that

$$\|Av\| \leqslant C\|v\| \quad \text{for all } v \in \mathcal{H}_1 \ . \tag{3.1.13}$$

The existence of a smallest such C is guaranteed by the following. It is called the *operator norm* of A and is denoted $\|A\|$.

3.1.12 Lemma *The operator norm of a bounded linear operator $A : \mathcal{H}_1 \to \mathcal{H}_2$ between Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ exists and is given by*

$$\begin{aligned}\|A\| &= \sup\left\{\|Av\| \;\middle|\; v \in \mathcal{H}_1,\ \|v\| = 1\right\} \\ &= \sup\left\{\|Av\| \;\middle|\; v \in \mathcal{H}_1,\ \|v\| \leqslant 1\right\} \\ &= \sup\left\{|\langle w, Av\rangle| \;\middle|\; v \in \mathcal{H}_1,\ w \in \mathcal{H}_2,\ \|v\| = \|w\| = 1\right\} \ .\end{aligned}$$

Proof. If $A : \mathcal{H}_1 \to \mathcal{H}_2$ is bounded, then the set $\{\|Av\| \mid v \in \mathcal{H}_1,\ \|v\| = 1\}$ is bounded, hence has a supremum C_0. This implies that for all non-zero $v \in \mathcal{H}_1$

$$\|Av\| = \|v\| \left\| A\left(\frac{v}{\|v\|}\right)\right\| \leqslant C_0\|v\| \ .$$

Hence the estimate (3.1.13) holds true for $C = C_0$. Moreover, C_0 is the smallest such C because if $0 \leqslant C_1 < C_0$, then there exists $v \in \mathcal{H}_1$ with $\|v\| = 1$ and $\|Av\| > C_1$. This proves that the operator norm of A exists and that it fulfills $\|A\| = C_0$.

By definition of C_0, the estimate $\|A\| = C_0 \leqslant \sup\{\|Av\| \mid v \in \mathcal{H}_1,\ \|v\| \leqslant 1\}$ holds true. By definition of the operator norm, $\|Av\| \leqslant \|A\|$ for all $v \in \mathcal{H}_1$ with $\|v\| \leqslant 1$. The two estimates together entail the equality $\|A\| = \sup\{\|Av\| \mid v \in \mathcal{H}_1,\ \|v\| \leqslant 1\}$.

The Cauchy–Schwarz inequality entails

$$\sup\{|\langle w, Av\rangle| \mid v \in \mathcal{H}_1,\ w \in \mathcal{H}_2,\ \|v\| = \|w\| = 1\} \leqslant \|A\| \ .$$

The converse estimate follows by the observation that

$$\sup\{|\langle w, Av\rangle| \mid w \in \mathcal{H}_2,\ \|w\| = 1\} \geqslant \left|\left\langle \frac{Av}{\|Av\|}, Av\right\rangle\right| = \|Av\|$$

whenever $Av \neq 0$. This proves the last claimed equality. □

Every norm preserving linear map is bounded with operator norm 1. In particular, the identity map on a Hilbert space is bounded. Moreover, if $A : \mathcal{H}_1 \to \mathcal{H}_2$ and $B : \mathcal{H}_2 \to \mathcal{H}_3$ are bounded linear operators between Hilbert spaces, then the composition $BA : \mathcal{H}_1 \to \mathcal{H}_3$ is bounded with operator norm $\leqslant \|B\|\,\|A\|$ since for all $v \in \mathcal{H}_1$ with $\|v\| \leqslant 1$

$$\|BAv\| \leqslant \|B\|\,\|Av\| \leqslant \|B\|\,\|A\| \ .$$

Hence Hilbert spaces as objects together with bounded linear maps as morphisms form a category which we denote by Hilb and call the *category of Hilbert spaces*. Note that the morphisms in this category appear to "forget" the inner product and just preserve the linear and the topological structure. John Baez (Baez, 1997, p. 133) has explained how to heal this apparent defect by showing that Hilb carries a so-called $*$-structure given by the adjoint map on bounded linear operators. We will come back to this point later when we introduce adjoint operators.

As proved already for Banach spaces, a linear map between Hilbert spaces is bounded if and only if it is continuous. For reasons of completeness and convenience we state here the result for inner product spaces.

3.1.13 Proposition *Let $A : \mathcal{H}_1 \to \mathcal{H}_2$ be a linear map between two inner product spaces. Then the following are equivalent.*

(i) *A is bounded.*

(ii) *A is continuous.*

(iii) *A is continuous at* 0.

Proof. (i) $\implies$ (ii). Assume that A is bounded. Let $\|A\| := \sup_{v \in \mathcal{H}_1} \|Av\|$ be its norm. Then, for all $v, w \in \mathcal{H}_1$

$$\|Av - Aw\| \leqslant \|A\| \cdot \|v - w\| \ .$$

Hence A is Lipschitz continuous, so in particular continuous.
(ii) $\implies$ (iii). If the map A is continuous, it is in particular continuous at the origin.
(iii) $\implies$ (i). If A is continuous at the origin, there exists $\delta > 0$ such that for all $v \in \mathcal{H}_1$ the estimate $\|Av\| < 1$ holds whenever $\|v\| < \delta$. This implies that for v with $\|v\| \leqslant 1$

$$\|Av\| = 2\delta \left\| A\left(\frac{1}{2\delta} v\right) \right\| < 2\delta \ .$$

This means that A is bounded. □

3.1.14 Last in this section we will introduce bounded bilinear and sesquilinear maps. We define them for normed vector spaces. Their main application lies in the operator theory on Hilbert spaces, so we introduce them here.

3.1.15 Definition Let V_1 and V_2 be two normed vector spaces over $\mathbb{K}$ and denote the norms on V_1 and V_2 by the same symbol $\|\cdot\|$. Assume that $b : \mathrm{V}_1 \times \mathrm{V}_2 \to \mathbb{K}$ is a *bilinear* or *sesquilinear form* that is b is linear in each argument respectively b is conjugate linear in the first and linear in the second argument. The form $b : \mathrm{V}_1 \times \mathrm{V}_2 \to \mathbb{K}$ then is called *bounded* if there exists a $C > 0$ such that

$$|b(v, w)| \leqslant C \, \|v\| \, \|w\| \quad \text{for all } v \in \mathrm{V}_1, \ w \in \mathrm{V}_2 \ .$$

In this case,

$$\|b\| := \sup \left\{ |b(v, w)| \ \middle| \ v \in \mathrm{V}_1, \ w \in \mathrm{V}_2, \ \|v\| = \|w\| = 1 \right\}$$

exists and is called the *norm* of the form b.

3.1.16 Example The inner product on a (pre-) Hilbert space is bounded by the Cauchy–Schwarz inequality and has norm 1.

3.1.17 Proposition *A bilinear or sesquilinear form* $b : \mathrm{V}_1 \times \mathrm{V}_2 \to \mathbb{K}$ *defined on the cartesian product of two normed vector space* V_1 *and* V_2 *over* $\mathbb{K}$ *is bounded if and only if it is continuous.*

Proof. If b is bounded, then

$$\begin{aligned} \left|b(v, w) - b(v', w')\right| &\leqslant \left|b(v, w) - b(v', w)\right| + \left|b(v', w) - b(v', w')\right| \leqslant \\ &\leqslant \|b\| \, \big(\|w\| \, \|v - v'\| + \|v'\| \, \|w - w'\|\big) \end{aligned}$$

for all $v, v' \in \mathrm{V}_1$ and $w, w' \in \mathrm{V}_2$. Hence b is locally Lipschitz continuous, so in particular continuous.

Conversely, assume now that b is continuous. Then one can find $\delta > 0$ such that for all $v \in \mathrm{V}_1$ and $w \in \mathrm{V}_2$ of norm less than δ the relation $|b(v, w)| < 1$ holds true. But that entails for all non-zero v, w

$$|b(v, w)| = \frac{4 \, \|v\| \, \|w\|}{\delta^2} \cdot b\left(\delta \frac{v}{2\|v\|}, \delta \frac{w}{2\|w\|}\right) \leqslant \frac{4}{\delta^2} \|v\| \, \|w\| \ .$$

Hence b is bounded. □

3.1.18 Remark Given two normed vector spaces or more generally two topological vector spaces V_1 and V_2 one can consider bilinear or sesquilinear forms $b : \mathrm{V}_1 \times \mathrm{V}_2 \to \mathbb{K}$ which are only *separately-continuous*. That means that for all $v \in \mathrm{V}_1$ the map $b_v = b(v, -) : \mathrm{V}_2 \to \mathbb{K}$ and for all $w \in \mathrm{V}_2$ the map $\overline{b}_w = b(-, w) : \mathrm{V}_1 \to \mathbb{K}$ is continuous. In general, separate-continuity is strictly weaker than continuity unless the underlying vector spaces are Banach spaces where the two notions coincide as a consequence of the Banach–Steinhaus theorem. Let us prove this. By continuity of b_v there exist $C_v \geqslant 0$ such that $|b_v(w)| \leqslant C_v \|w\|$ for all $w \in \mathrm{V}_2$ and $\overline{C}_w \geqslant 0$ such that $|\overline{b}_w(v)| \leqslant \overline{C}_w \|v\|$ for all $v \in \mathrm{V}_1$. Hence, for all $w \in \mathrm{V}_2$

$$\sup_{v \in \mathrm{V},\ \|v\| \leqslant 1} |b_v(w)| = \sup_{v \in \mathrm{V},\ \|v\| \leqslant 1} |\overline{b}_w(v)| \leqslant \overline{C}_w < \infty \ .$$

The Banach–Steinhaus theorem now entails

$$\sup_{v,w \in \mathrm{V},\ \|v\|,\ \|w\| \leqslant 1} |b(v,w)| = \sup_{v \in \mathrm{V},\ \|v\| \leqslant 1} \|b_v\| < \infty \ .$$

Therefore, b is bounded, so continuous by the preceding proposition.

3.2. Orthogonal decomposition and the Riesz representation theorem

3.2.1 One of the issues with infinite-dimensional analysis is that a closed subspace of an infinite dimensional Banach space might not have a closed complement. Fortunately, the situation in Hilbert space theory is not so grim because every closed subspace of a Hilbert space admits an orthogonal complement. This is one of the four crucial properties which distinguish Hilbert spaces from Banach spaces and which are stated in the following.

In this section $\mathcal{H}$ will always denote a Hilbert space over the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. The symbol $\langle \cdot, \cdot \rangle$ will stand for the inner product of $\mathcal{H}$.

3.2.2 Theorem (Best approximation theorem) *Every closed convex nonempty subset C of a Hilbert space $\mathcal{H}$ has a unique element of minimal norm.*

Proof. Let $d = \inf\{\|v\| \mid v \in C\}$ which is a non-negative real number. We claim there exists a unique $v_0 \in C$ with $\|v_0\| = d$. For uniqueness, consider two vectors v_0, v_1 satisfying the desired property, and let $v = \frac{1}{2}(v_0 + v_1)$ be their midpoint. Then

$$\|v\| = \frac{1}{2}\|v_0 + v_1\| \leqslant \frac{1}{2}(\|v_0\| + \|v_1\|) = d$$

By minimality of d this entails $\|v\| = d$. By the parallelogram identity

$$\left\|\frac{1}{2}(v_0 + v_1)\right\|^2 + \left\|\frac{1}{2}(v_0 - v_1)\right\|^2 = 2\left\|\frac{v_0}{2}\right\|^2 + 2\left\|\frac{v_1}{2}\right\|^2 = d^2 \ ,$$

hence

$$\left\|\frac{1}{2}(v_0 - v_1)\right\|^2 \leqslant d^2 - \|v\|^2 = 0 \ ,$$

proving $v_0 = v_1$.

For the proof of existence observe that by definition of d there exists a sequence $(v_n)_{n\in\mathbb{N}} \subset C$ such that $\lim_{n\to\infty} \|v_n\| = d$. By convexity

$$\frac{1}{2}(v_n + v_m) \in C$$

for all $n, m \in \mathbb{N}$, hence $\frac{1}{4}\|v_n + v_m\|^2 \geqslant d^2$. The parallelogram equality entails

$$0 \leqslant \|v_n - v_m\|^2 = 2\|v_n\|^2 + 2\|v_m\|^2 - \|v_n + v_m\|^2 \leqslant 2\|v_n\|^2 + 2\|v_m\|^2 - 4d^2 \ .$$

Since $\lim_{n\to\infty} \|v_n\| = d$ there exists for given $\varepsilon > 0$ an $N \in \mathbb{N}$ such that $\|v_n\|^2 - d^2 \leqslant \frac{1}{4}\varepsilon^2$ for all $n \geqslant N$. Hence, for $n, m \geqslant N$

$$0 \leqslant \|v_n - v_m\| \leqslant \varepsilon \ ,$$

and $(v_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence, so convergent by completeness of $\mathcal{H}$. Put $v_0 := \lim_{n\to\infty} v_n$. Then $v_0 \in C$ since C is closed and $\|v_0\| = \lim_{n\to\infty} \|v_n\| = d$. The existence claim follows and the proof is finished. □

3.2.3 Theorem and Definition (Orthogonal decomposition theorem) *Let* $\mathrm{V} \subset \mathcal{H}$ *be a closed subspace of the Hilbert space* $\mathcal{H}$*. Then the* orthogonal complement

$$\mathrm{V}^\perp = \big\{ w \in \mathcal{H} \;\big|\; \langle v, w\rangle = 0 \textit{ for each } v \in \mathrm{V} \big\}$$

is a closed subspace of $\mathcal{H}$ *and* $\mathcal{H} = \mathrm{V} \oplus \mathrm{V}^\perp$*. The map* $\mathrm{pr}_\mathrm{V} : \mathcal{H} \to \mathrm{V}$ *which maps* $w \in \mathcal{H}$ *to the unique* $w_1 \in \mathrm{V}$ *such that* $w - w_1 \in \mathrm{V}^\perp$ *is called the* orthogonal projection *onto* V*. It satisfies* $\|w - \mathrm{pr}_\mathrm{V}(w)\| = d(w, \mathrm{V}) := \inf\big\{\|v - w\| \;\big|\; v \in \mathrm{V}\big\}$ *that is* $\mathrm{pr}_\mathrm{V}(w)$ *is the unique element of* V *having shortest distance from* w*.*

Proof. For $v \in \mathcal{H}$ define $v^\flat : \mathcal{H} \to \mathbb{R}$ by $v^\flat(w) = \langle w, v\rangle$. Recall that this map is continuous and linear. Hence the kernel $(v^\flat)^{-1}(0)$ is a closed linear subspace of $\mathcal{H}$ and

$$\mathrm{V}^\perp = \bigcap_{v\in\mathrm{V}} (v^\flat)^{-1}(0) \tag{3.2.1}$$

is a closed linear subspace. To show $\mathrm{V} \cap \mathrm{V}^\perp = \{0\}$, consider $v \in \mathrm{V} \cap \mathrm{V}^\perp$. Then $\|v\|^2 = \langle v, v\rangle = 0$. Next we want to show that every $w \in \mathcal{H}$ can be written in the form $w = w_1 + w_2$ with $w_1 \in \mathrm{V}$ and $w_2 \in \mathrm{V}^\perp$. To see this put $C = w - \mathrm{V}$. Then C is closed and convex. By the best approximation theorem there exists a unique element $w_2 \in C$ of minimal norm. Let w_1 be the unique element of V such that $w_2 = w - w_1$. It remains to show $w_2 \in \mathrm{V}^\perp$. Since w_2 has minimal norm among the elements of $w - \mathrm{V}$, the following inequality holds for all vectors $v \in \mathrm{V}$:

$$\|w_2\|^2 \leqslant \|w_2 + v\|^2 = \|w_2\|^2 + 2\,\mathfrak{Re}\langle w_2, v\rangle + \|v\|^2 \ .$$

Hence

$$0 \leqslant 2\,\mathfrak{Re}\langle w_2, v\rangle + \|v\|^2 \quad \text{for all } v \in \mathrm{V} \ .$$

Now assume that $\|v\| = 1$ and choose $\varphi \in \mathbb{R}$ such that $e^{i\varphi}\langle w_2, v\rangle \in \mathbb{R}$. Setting $v' = e^{i\varphi}v$, one obtains for all $\lambda \in \mathbb{R}$ by the last inequality

$$0 \leqslant 2\langle w_2, \lambda v'\rangle + \|\lambda v'\|^2 = 2\lambda\langle w_2, v'\rangle + \lambda^2 \ .$$

For $\lambda = -\langle w_2, v' \rangle$ this entails the estimate

$$|\langle w_2, v' \rangle|^2 = -\left(-2|\langle w_2, v' \rangle|^2 + |\langle w_2, v' \rangle|^2\right) = -\left(2\lambda\langle w_2, v' \rangle + \lambda^2\right) \leqslant 0 \ .$$

Hence $\langle w_2, v \rangle = 0$ for all unit vectors $v \in \mathrm{V}$, therefore $w_2 \in \mathrm{V}^\perp$.

The remainder of the claim is now an immediate consequence of the construction of w_1 from the given w and the observation that $\mathrm{pr}_\mathrm{V}(w) = w_1$. □

3.2.4 Corollary *For every subspace* $\mathrm{V} \subset \mathcal{H}$ *of a Hilbert space* $\mathcal{H}$ *the orthogonal complement* $V^\perp$ *is closed, and the relation*

$$V^\perp = \overline{V}^\perp$$

holds true. Moreover,

$$\overline{V} = (V^\perp)^\perp \ .$$

Proof. By Equation (3.2.1), the orthogonal complement $V^\perp$ is closed. Since $V \subset \overline{V}$ the inclusion $\overline{V}^\perp \subset V^\perp$ holds true. The converse inclusion $V^\perp \subset \overline{V}^\perp$ follows from the observation that if $w \in V^\perp$ and $(v_n)_{n\in\mathbb{N}}$ is a sequence in V converging to some $v \in \overline{V}$, then

$$\langle w, v \rangle = \lim_{n\to\infty} \langle w, v_n \rangle = 0 \ .$$

This proves the equality $V^\perp = \overline{V}^\perp$. The inclusion $\overline{V} \subset (\overline{V}^\perp)^\perp = (V^\perp)^\perp$ is immediate by definition of the orthogonal complement. Since

$$\mathcal{H} = \overline{V} \oplus V^\perp = (V^\perp)^\perp \oplus V^\perp$$

by the preceding theorem, the equality $\overline{V} = (V^\perp)^\perp$ follows. □

3.2.5 Theorem (Riesz representation theorem for Hilbert spaces) *Let* $\mathcal{H}$ *be a Hilbert space and* $\mathcal{H}'$ *its topological dual. Then the* musical map

$$^\flat : \mathcal{H} \to \mathcal{H}', \quad v \mapsto v^\flat = (\mathcal{H} \ni w \mapsto \langle v, w \rangle \in \mathbb{K})$$

is an isometric isomorphism which is linear in the real case and conjugate-linear in the complex case.

Proof. Obviously, $^\flat$ is linear if the ground field $\mathbb{K}$ equals $\mathbb{R}$ and conjugate-linear if $\mathbb{K} = \mathbb{C}$. Now observe that for all $v \in \mathcal{H}$ by the Cauchy–Schwarz inequality

$$\|v^\flat\| = \sup\left\{|\langle v, w \rangle| \;\middle|\; w \in \mathcal{H} \ \& \ \|w\| = 1\right\} = \|v\| \ ,$$

hence $^\flat$ is an isometry, so in particular injective. It remains to show surjectivity. So assume that $\alpha : \mathcal{H} \to \mathbb{K}$ is a nontrivial continuous linear form. Let V be its kernel. Then V is a closed linear subspace of $\mathcal{H}$. Since α is nontrivial, the orthogonal complement $\mathrm{V}^\perp$ is nontrivial, too. Hence $\mathrm{V}^\perp \cong \mathcal{H}/\mathrm{V}$ is isomorphic to $\operatorname{im}\alpha = \mathbb{K}$ and there exists a vector $v \in \mathrm{V}^\perp \backslash \{0\}$ such that $\alpha(v) = 1$. Since v spans $\mathrm{V}^\perp$ there exists for every $w \in \mathcal{H}$ a unique $\lambda_w \in \mathbb{K}$ such that $w = \mathrm{pr}_V(w) + \lambda_w v$. Then compute

$$\alpha(w) = \alpha(\lambda_w v) = \lambda_w \quad \text{and} \quad \left(\frac{v}{\|v\|^2}\right)^\flat (w) = \frac{1}{\|v\|^2}\langle v, w \rangle = \frac{\lambda_w}{\|v\|^2}\langle v, v \rangle = \lambda_w \ .$$

This entails $\alpha = \left(\frac{v}{\|v\|^2}\right)^\flat$, and $^\flat$ is surjective. □

3.2.6 Remark Sometimes in the Hilbert space literature the inverse of the musical isomorphism $^\flat : \mathcal{H} \to \mathcal{H}'$ is denoted $^\sharp : \mathcal{H}' \to \mathcal{H}$. We will follow that convention.

3.2.7 Corollary *Every Hilbert space $\mathcal{H}$ is* reflexive *that is the canonical map*

$$H \to H'',\ v \mapsto \big(H' \ni \lambda \mapsto \lambda(v) \in \mathbb{K}\big)$$

is an isometric isomorphism.

Proof. By the Riesz Representation Theorem, the dual $\mathcal{H}'$ is a Hilbert space with inner product

$$\langle\!\langle \cdot, \cdot \rangle\!\rangle : \mathcal{H}' \times \mathcal{H}' \to \mathbb{K},\ (\lambda, \mu) \mapsto \langle\!\langle \lambda, \mu \rangle\!\rangle = \langle \mu^\sharp, \lambda^\sharp \rangle\ .$$

Hence, by applying the Riesz Representation Theorem twice, the map $^\flat \circ {}^\flat : \mathcal{H} \to \mathcal{H}''$ is an isometric linear isomorphism. Now compute for $v \in \mathcal{H}$ and $\lambda \in \mathcal{H}'$

$$(v^\flat)^\flat(\lambda) = \langle\!\langle v^\flat, \lambda \rangle\!\rangle = \langle \lambda^\sharp, v \rangle = \lambda(v)\ .$$

Hence $^\flat \circ {}^\flat$ coincides with the canonical map above and the claim follows. □

3.2.8 Corollary *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces and $b : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathbb{K}$ a bounded sesquilinear form. Then there exists unique bounded linear map $A : \mathcal{H}_2 \to \mathcal{H}_1$ such that*

$$b(v, w) = \langle v, Aw \rangle \quad \text{for all } v \in \mathcal{H}_1,\ w \in \mathcal{H}_2\ . \tag{3.2.2}$$

Moreover, the operator norm $\|A\|$ coincides with $\|b\|$.

Proof. First let us show uniqueness. So let $A, B : \mathcal{H}_2 \to \mathcal{H}_1$ be bounded and linear so that

$$b(v, w) = \langle v, Aw \rangle = \langle v, Bw \rangle \quad \text{for all } v \in \mathcal{H}_1,\ w \in \mathcal{H}_2\ .$$

Then $\|(A - B)w\|^2 = \langle (A - B)w, Aw - Bw \rangle = b((A - B)w, w) - b((A - B)w, w) = 0$ for all $w \in \mathcal{H}_2$ which entails equality of A and B.

To prove existence observe that for every $w \in \mathcal{H}_2$ the map

$$\overline{b}_w : \mathcal{H}_1 \to \mathbb{K},\ v \mapsto \overline{b}(w, v) := \overline{b(v, w)}$$

is bounded and linear, so by the Riesz representation theorem there exists for every w a unique element $Aw \in \mathcal{H}_1$ such that $\langle Aw, v \rangle = \overline{b}(w, v)$ for all $v \in \mathcal{H}_1$. By construction, $Aw = (\overline{b}_w)^\sharp$. Since the maps $\mathcal{H}_2 \to \mathcal{H}_1'$, $w \mapsto \overline{b}_w$ and $^\sharp : \mathcal{H}_1' \to \mathcal{H}_1$ are both conjugate-linear, A is linear. Hence A is the desired linear operator fulfilling Equation (3.2.2).

For the operator norm compute

$$\begin{aligned}\|A\| &= \sup\big\{\, |\langle v, Aw \rangle| \,\big|\, v \in \mathcal{H}_1,\ w \in \mathcal{H}_2,\ \|v\| = \|w\| = 1 \big\} = \\ &= \sup\big\{\, |b(v, w)| \,\big|\, v \in \mathcal{H}_1,\ w \in \mathcal{H}_2,\ \|v\| = \|w\| = 1 \big\} = \|b\|\ .\end{aligned}$$

Hence A is bounded with operator norm equal to $\|b\|$ and the claim is proved. □

3.2.9 Last in this section we will examine the *Hilbert direct sum* or just *Hilbert sum* of a family $(\mathcal{H}_i)_{i\in I}$ of Hilbert spaces. It is defined by

$$\widehat{\bigoplus_{i\in I}}\mathcal{H}_i = \left\{(v_i)_{i\in I} \in \prod_{i\in I}\mathcal{H}_i \;\middle|\; \left(\|v_i\|^2\right)_{i\in I} \text{ is summable}\right\} =$$
$$= \left\{(v_i)_{i\in I} \in \prod_{i\in I}\mathcal{H}_i \;\middle|\; \exists C \geqslant 0\, \forall J \in \mathcal{P}_{\text{fin}}(I) : \sum_{i\in J}\|v_i\|^2 \leqslant C\right\} ,$$

where, as usual, $\mathcal{P}_{\text{fin}}(I)$ denotes the set of all finite subsets of I.

3.2.10 Proposition *Let $(\mathcal{H}_i)_{i\in I}$ be a family of Hilbert spaces. Then the Hilbert direct sum $\widehat{\bigoplus_{i\in I}}\mathcal{H}_i$ is a Hilbert space with inner product given by*

$$\langle -,-\rangle : \widehat{\bigoplus_{i\in I}}\mathcal{H}_i \times \widehat{\bigoplus_{i\in I}}\mathcal{H}_i \to \mathbb{K}, \quad ((v_i)_{i\in I},(w_i)_{i\in I}) \mapsto \sum_{i\in I}\langle v_i, w_i\rangle .$$

Proof. We show first that $\widehat{\bigoplus_{i\in I}}\mathcal{H}_i$ is a subvector space of the direct product $\prod_{i\in I}\mathcal{H}_i$. Let $z \in \mathbb{K}$ and $(v_i)_{i\in I},(w_i)_{i\in I} \in \widehat{\bigoplus_{i\in I}}\mathcal{H}_i$. Choose $C, D \geqslant 0$ such that

$$\sum_{i\in J}\|v_i\|^2 \leqslant C \quad \text{and} \quad \sum_{i\in J}\|w_i\|^2 \leqslant D \quad \text{for all } J \in \mathcal{P}_{\text{fin}}(I) .$$

Then

$$\sum_{i\in J}\|zv_i\|^2 = |z| \sum_{i\in J}\|v_i\|^2 \leqslant |z|\, C \quad \text{for all } J \in \mathcal{P}_{\text{fin}}(I) , \tag{3.2.3}$$

so $(zv_i)_{i\in I} \in \widehat{\bigoplus_{i\in I}}\mathcal{H}_i$. Moreover, by Minkowski's inequality for finite sums,

$$\sum_{i\in J}\|v_i + w_i\|^2 \leqslant \left(\sqrt{\sum_{i\in J}\|v_i\|^2} + \sqrt{\sum_{i\in J}\|w_i\|^2}\right)^2 \leqslant \left(\sqrt{C}+\sqrt{D}\right)^2 \quad \text{for all } J \in \mathcal{P}_{\text{fin}}(I) . \tag{3.2.4}$$

Hence the family $\left(\|v_i + w_i\|^2\right)_{i\in I}$ is summable and $(v_i + w_i)_{i\in I} \in \widehat{\bigoplus_{i\in I}}\mathcal{H}_i$.

Next observe that the map

$$\| - \| : \widehat{\bigoplus_{i\in I}}\mathcal{H}_i \to \mathbb{K},\ (v_i)_{i\in I} \mapsto \left\|(v_i)_{i\in I}\right\| = \sqrt{\sum_{i\in I}\|v_i\|^2}$$

is well-defined by definition of the Hilbert direct sum. It is even a norm by (3.2.3) and (3.2.4). Now we need to show that the inner product on $\widehat{\bigoplus_{i\in I}}\mathcal{H}_i$ is well-defined which means that the family $(\langle v_i, w_i\rangle)_{i\in I}$ is summable for all $(v_i)_{i\in I},(w_i)_{i\in I} \in \widehat{\bigoplus_{i\in I}}\mathcal{H}_i$. To this end let $J \subset I$ be a finite subset.

Then, by the triangle inequality, the Cauchy–Schwarz inequality on the Hilbert spaces $\mathcal{H}_i$ and the Cauchy–Schwarz inequality for finite sums,

$$\left|\sum_{i\in J}\langle v_i, w_i\rangle\right| \leqslant \sum_{i\in J}|\langle v_i, w_i\rangle| \leqslant \sum_{i\in J}\|v_i\|\,\|w_i\| \leqslant \sqrt{\sum_{i\in J}\|v_i\|^2}\cdot\sqrt{\sum_{i\in J}\|w_i\|^2} \leqslant \left\|(v_i)_{i\in I}\right\|\left\|(w_i)_{i\in I}\right\| .$$

Hence the family $(\langle v_i, w_i\rangle)_{i\in I}$ is absolutely summable, so in particular summable, and the inner product is well-defined.

By definition and since all the inner products on the Hilbert spaces $\mathcal{H}_i$ are conjugate symmetric and positive definite, the map $\langle -,-\rangle$ on $\widehat{\bigoplus_{i\in I}}\mathcal{H}_i$ has to be conjugate symmetric and positive definite as well. It remains to show linearity in the second argument. Denote for $(v_i)_{i\in I}, (w_i)_{i\in I} \in \prod_{i\in I}\mathcal{H}_i$ and $J \in \mathcal{P}_{\mathrm{fin}}(I)$ by $\langle (v_i)_{i\in I}, (w_i)_{i\in I}\rangle_J$ the finite sum $\sum_{i\in J}\langle v_i, w_i\rangle$. Observe that the net $\left(\langle (v_i)_{i\in I}, (w_i)_{i\in I}\rangle_J\right)_{J\in\mathcal{P}_{\mathrm{fin}}(I)}$ converges to $\langle (v_i)_{i\in I}, (w_i)_{i\in I}\rangle$ in case both $(v_i)_{i\in I}$ and $(w_i)_{i\in I}$ are in $\widehat{\bigoplus_{i\in I}}\mathcal{H}_i$. Now let $z \in \mathbb{K}$ and $(v_i)_{i\in I}, (w_i)_{i\in I}, (w_i')_{i\in I} \in \widehat{\bigoplus_{i\in I}}\mathcal{H}_i$. Then

$$\begin{aligned}\langle (v_i)_{i\in I}, (w_i)_{i\in I} + (w_i')_{i\in I}\rangle_J &= \langle (v_i)_{i\in I}, (w_i)_{i\in I}\rangle_J + \langle (v_i)_{i\in I}, (w_i')_{i\in I}\rangle_J \quad \text{and} \\ \langle (v_i)_{i\in I}, z(w_i)_{i\in I}\rangle_J &= z\langle (v_i)_{i\in I}, (w_i)_{i\in I}\rangle_J .\end{aligned}$$

By convergence of all the nets $\left(\langle (v_i)_{i\in I}, (w_i)_{i\in I}\rangle_J\right)_{J\in\mathcal{P}_{\mathrm{fin}}(I)}$, linearity in the second argument follows.

By construction, the norm associated to the inner product $\langle -,-\rangle$ on $\widehat{\bigoplus_{i\in I}}\mathcal{H}_i$ coincides with the above defined norm $\|-\|$. It remains to show that $\widehat{\bigoplus_{i\in I}}\mathcal{H}_i$ equipped with the norm $\|-\|$ is complete. To this end observe that for every finite $J \subset I$ the map

$$\|-\|_J : \prod_{i\in I}\mathcal{H}_i \to \mathbb{R}_{\geqslant 0},\ (v_i)_{i\in I} \mapsto \sqrt{\langle (v_i)_{i\in I}, (v_i)_{i\in I}\rangle_J} = \sqrt{\sum_{i\in J}\|v_i\|^2}$$

is a seminorm and that $(v_i)_{i\in I} \in \prod_{i\in I}\mathcal{H}_i$ lies in the Hilbert direct sum $\widehat{\bigoplus_{i\in I}}\mathcal{H}_i$ if and only if the family $\left(\|(v_i)_{i\in I}\|_J\right)_{J\in\mathcal{P}_{\mathrm{fin}}(I)}$ is bounded. Now let $((v_i^n)_{i\in I})_{n\in\mathbb{N}}$ be a Cauchy sequence. Let $\varepsilon > 0$ and choose $N_\varepsilon \in \mathbb{N}$ such that

$$\left\|(v_i^m)_{i\in I} - (v_i^n)_{i\in I}\right\| < \varepsilon \quad \text{for all } n, m \geqslant N_\varepsilon . \tag{3.2.5}$$

Then

$$\left\|(v_i^m)_{i\in I} - (v_i^n)_{i\in I}\right\|_J < \varepsilon \quad \text{for all } J \in \mathcal{P}_{\mathrm{fin}}(I) \text{ and } n, m \geqslant N_\varepsilon . \tag{3.2.6}$$

Taking $J = \{j\}$ for $j \in I$ this implies that the sequence $(v_j^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in the Hilbert space $\mathcal{H}_j$. Let $v_j \in \mathcal{H}_j$ be its limit. The family $(v_i)_{i\in I}$ then is an element of $\widehat{\bigoplus_{i\in I}}\mathcal{H}_i$. To verify this put $N = N_1$ and observe that by (3.2.6) for all finite $J \subset I$

$$\begin{aligned}\left\|(v_i)_{i\in I}\right\|_J &\leqslant \left\|(v_i^N)_{i\in I}\right\|_J + \left\|(v_i)_{i\in I} - (v_i^N)_{i\in I}\right\|_J = \\ &= \left\|(v_i^N)_{i\in I}\right\|_J + \lim_{m\to\infty}\left\|(v_i^m)_{i\in I} - (v_i^N)_{i\in I}\right\|_J \leqslant \left\|(v_i^N)_{i\in I}\right\| + 1.\end{aligned}$$

Hence the family $\left(\left\|(v_i)_{i\in I}\right\|_J\right)_{J\in\mathcal{P}_{\text{fin}}(I)}$ is bounded and $(v_i)_{i\in I}$ lies in the Hilbert direct sum of the spaces $\mathcal{H}_i$, $i \in I$. Moreover, (3.2.6) entails that

$$\left\|(v_i)_{i\in I} - (v_i^n)_{i\in I}\right\|_J = \lim_{m\to\infty} \left\|(v_i^m)_{i\in I} - (v_i^n)_{i\in I}\right\|_J \leqslant \varepsilon \quad \text{for all } J \in \mathcal{P}_{\text{fin}}(I) \text{ and } n \geqslant N_\varepsilon \ .$$

Since $\left\|(v_i)_{i\in I} - (v_i^n)_{i\in I}\right\|$ is the limit of the net $\left(\left\|(v_i)_{i\in I} - (v_i^n)_{i\in I}\right\|_J\right)_{J\in\mathcal{P}_{\text{fin}}(I)}$, the estimate

$$\left\|(v_i)_{i\in I} - (v_i^n)_{i\in I}\right\| \leqslant \varepsilon \quad \text{for all } n \geqslant N_\varepsilon$$

follows, and the sequence $\left((v_i^n)_{i\in I}\right)_{n\in\mathbb{N}}$ convergences to $(v_i)_{i\in I}$. This finishes the proof. □

3.3. Orthonormal bases in Hilbert spaces

3.3.1 Definition A (possibly empty) subset S of a Hilbert space $\mathcal{H}$ is called an *orthogonal system* or just *orthogonal* if for any two different elements $v, w \in S$ the relation $\langle v, w\rangle = 0$ holds true. If in addition $\|v\| = 1$ for all elements $v \in S$, then the set is called *orthonormal* or an *orthonormal system.* A family $(v_i)_{i\in I}$ of vectors in $\mathcal{H}$ is called *orthogonal* if $\langle v_i, v_j\rangle = 0$ for all $i, j \in I$ with $i \neq j$ and *orthonormal* if in addition $\|v_i\| = 1$ for all $i \in I$.

3.3.2 Obviously, the set of orthonormal subsets of a Hilbert space is ordered by set-theoretic inclusion. Therefore, the following definition makes sense.

3.3.3 Definition A maximal orthonormal set in a Hilbert space $\mathcal{H}$ is called an *orthonormal basis* or a *Hilbert basis* of $\mathcal{H}$.

3.3.4 Proposition *Every Hilbert space $\mathcal{H}$ has an orthonormal basis.*

Proof. Wothout loss of generality we can assume that $\mathcal{H} \neq \{0\}$, because $\varnothing$ is a Hilbert basis for $\{0\}$. Let $\mathcal{O}$ denote the set of orthonormal subsets of $\mathcal{H}$. As mentioned before, $\mathcal{O}$ is ordered by set-theoretic inclusion. Let $\mathcal{C} \subset \mathcal{O}$ be a non-empty chain. Put $U = \bigcup_{S\in\mathcal{C}} S$. Then U is an upper bound of $\mathcal{C}$. So by Zorn's lemma $\mathcal{O}$ has a maximal element. □

3.3.5 Remark (a) By slight abuse of language we sometimes call an orthonormal family $(b_i)_{i\in I}$ in a Hilbert space $\mathcal{H}$ an *orthonormal basis* or a *Hilbert basis* of $\mathcal{H}$ if the set $\{b_i \mid i \in I\}$ is an orthornormal basis.

(b) If on an orthonormal basis $B \subset \mathcal{H}$ a total order relation is given, one calls B an *ordered Hilbert basis* of $\mathcal{H}$. Likewise, an orthonormal basis of the form $(b_i)_{i\in I}$ is called *ordered* if the index set I carries a total order.

3.3.6 Proposition (Pythagorean theorem for orthogonal families) *An orthogonal family $(v_i)_{i\in I}$ in a Hilbert space $\mathcal{H}$ is summable if and only if the family of norms $\left(\|v_i\|\right)_{i\in I}$ is square summable. In this case one has*

$$\Big\|\sum_{i\in I} v_i\Big\|^2 = \sum_{i\in I} \|v_i\|^2 \ .$$

Proof. Assume that $(\|v_i\|)_{i\in I}$ is square summable or in other words that the net of partial sums $\left(\sum_{i\in J}\|v_i\|^2\right)_{J\in\mathcal{P}_{\text{fin}}(I)}$ converges to some $s\in\mathbb{R}$. For $\varepsilon>0$ choose a finite $J_\varepsilon\subset I$ such that for all finite J with $J_\varepsilon\subset J\subset I$ the relation

$$\left|s-\sum_{i\in J}\|v_i\|^2\right|<\frac{\varepsilon^2}{2}$$

holds true. For finite $K\subset I$ with $K\cap J_\varepsilon=\varnothing$ one then obtains by the pythagorean theorem for finite orthogonal families, Eq. (3.1.2),

$$\Big\|\sum_{i\in K}v_i\Big\|^2=\sum_{i\in K}\|v_i\|^2\leqslant\Big|s-\sum_{i\in K\cup J_\varepsilon}\|v_i\|^2\Big|+\Big|s-\sum_{i\in J_\varepsilon}\|v_i\|^2\Big|<\varepsilon^2\ .$$

Hence $\left(\sum_{i\in J}v_i\right)_{J\in\mathcal{P}_{\text{fin}}(I)}$ is a Cauchy net in $\mathcal{H}$, so convergent.

Now let $(v_i)_{i\in I}$ be summable to $v\in\mathcal{H}$. Then there exists a $J_1\in\mathcal{P}_{\text{fin}}(I)$ such that for all finite $J\subset I$ containing J_1

$$\Big\|v-\sum_{i\in J}v_i\Big\|\leqslant 1\ .$$

This implies by the pythagorean theorem for finite orthogonal families

$$\sum_{i\in J}\|v_i\|^2=\Big\|\sum_{i\in J}v_i\Big\|^2\leqslant\left(\Big\|v-\sum_{i\in J}v_i\Big\|+\|v\|\right)^2\leqslant(1+\|v\|)^2\ .$$

Therefore, the net of partial sums $\left(\sum_{i\in J}\|v_i\|^2\right)_{J\in\mathcal{P}_{\text{fin}}(I)}$ is bounded, so convergent since each term $\|v_i\|^2$ is non-negative.

By continuity of the inner product and pairwise orthogonality of the v_i one finally obtains in the convergent case

$$\Big\|\sum_{i\in I}v_i\Big\|^2=\Big\langle\sum_{i\in I}v_i,\sum_{j\in I}v_j\Big\rangle=\sum_{i\in I}\Big\langle v_i,\sum_{j\in I}v_j\Big\rangle=\sum_{i\in I}\sum_{j\in I}\langle v_i,v_j\rangle=\sum_{i\in I}\|v_i\|^2\ .$$

□

3.3.7 Proposition *Let $(v_i)_{i\in I}$ be an orthonormal family in a Hilbert space $\mathcal{H}$. Then for every $v\in\mathcal{H}$ the family $(\langle v_i,v\rangle)_{i\in I}$ is square summable and* Bessel's inequality *holds true that is*

$$\sum_{i\in I}|\langle v_i,v\rangle|^2\leqslant\|v\|^2\ .$$

Proof. Let $J\subset I$ be finite. Then, by the pythagorean theorem for finite orthogonal families

$$0\leqslant\Big\|v-\sum_{i\in J}\langle v_i,v\rangle v_i\Big\|^2=\|v\|^2-2\sum_{i\in J}|\langle v_i,v\rangle|^2+\Big\|\sum_{i\in J}\langle v_i,v\rangle v_i\Big\|^2=\|v\|^2-\sum_{i\in J}|\langle v_i,v\rangle|^2\ .$$

Therefore, for all $J\in\mathcal{P}_{\text{fin}}(I)$

$$\sum_{i\in J}|\langle v_i,v\rangle|^2\leqslant\|v\|^2\ . \tag{3.3.1}$$

Hence, by Proposition 1.6.9, the family $(|\langle v_i,v\rangle|)_{i\in I}$ is square summable. Bessel's inequality now follows from the observation that in Equation (3.3.1) one can pass over to the limit of the net $\left(\sum_{i\in J}|\langle v_i,v\rangle|^2\leqslant\|v\|^2\right)_{J\in\mathcal{P}_{\text{fin}}(I)}$. □

3.3.8 Theorem *Let B be an orthonormal system in a Hilbert space $\mathcal{H}$. Then the following are equivalent:*

(1) *The orthonormal system B is maximal, i.e. a Hilbert basis.*

(2) *The orthonormal system B is* total *that is for all $v \in H$ such that $\langle v, b\rangle = 0$ for all $b \in B$ the equality $v = 0$ holds true.*

(3) *For every $b \in B$ let $\mathcal{H}_b = \{rb \in \mathcal{H} \mid r \in \mathbb{K}\}$. Then the canonical map*

$$\iota : \widehat{\bigoplus_{b\in B}}\mathcal{H}_b \to \mathcal{H},\ (v_b)_{b\in B} \mapsto \sum_{b\in B} v_b$$

is an isometric isomorphism.

(4) *The closed linear span of B coincides with $\mathcal{H}$.*

(5) *For all $v \in \mathcal{H}$, one has the* Fourier expansion

$$v = \sum_{b\in B}\langle v, b\rangle b\ .$$

(6) *For all $v, w \in \mathcal{H}$, one has*

$$\langle v, w\rangle = \sum_{b\in B}\langle v, b\rangle\langle b, w\rangle\ .$$

(7) *For all $v \in \mathcal{H}$,* Parseval's identity *holds true that is*

$$\|v\|^2 = \sum_{b\in B} |\langle v, b\rangle|^2\ .$$

Proof. (1) $\Rightarrow$ (2): If $v \neq 0$, then $\frac{v}{\|v\|}$ is a unit vector orthogonal to each v_i. Hence $\{v\} \cup B$ is an orthonormal system which is strictly larger than B, contradicting (1).

(2) $\Rightarrow$ (3). First note that by the pythagorean theorem for infinite families, Proposition 3.3.6, the canonical map $\iota : \widehat{\bigoplus}_{b\in B} H_b \to H$ is well-defined and an isometry. Hence ι is injective. It remains to show that ι is surjective. To this end observe that $\operatorname{im}\iota$ is closed in $\mathcal{H}$ since ι is an isometry (the image is complete). If ι is not surjective, then $\operatorname{im}\iota^\perp$ is not the zero vector space. Choose $v \in \operatorname{im}\iota^\perp\backslash\{0\}$. Then v is orthogonal to each element of B, but $v \neq 0$. This contradicts (2), so $\operatorname{im}\iota = \mathcal{H}$.

(3) $\Rightarrow$ (5): We can represent any $v \in \mathcal{H}$ in the form $v = \iota\left((v_b)_{b\in B}\right) = \sum_{b\in B} v_b$ with $(v_b)_{b\in B} \in \widehat{\bigoplus}_{b\in B} H_b$. Write $v_b = r_b\, b$ for every $b \in B$, where $r_b \in \mathbb{K}$ is uniquely determined by v_b. Then compute using continuity of the inner product

$$\langle v, b\rangle = \langle \sum_{c\in B} v_c, b\rangle = \sum_{c\in B} r_c\langle c, b\rangle = r_b\ .$$

Therefore,

$$v = \sum_{b\in B} r_b\, b = \sum_{b\in B}\langle v, b\rangle b\ .$$

(5) $\Rightarrow$ (6): Fourier expansion of $v, w \in H$ gives $v = \sum_{b \in B} \langle v, b \rangle b$ and $w = \sum_{b \in B} \langle w, b \rangle b$. Then, by continuity of the inner product,

$$\langle v, w \rangle = \sum_{b \in B} \langle v, b \rangle \langle b, w \rangle \ .$$

(5) $\Rightarrow$ (4): Let $v \in \mathcal{H}$. Then $\sum_{b \in J} \langle v, b \rangle b \in \mathrm{Span}(B)$ for all finite $J \subset B$. By Fourier expansion v is the limit of the net $\left(\sum_{b \in J} \langle v, b \rangle b \right)_{J \in \mathcal{P}_{\mathrm{fin}}(B)}$, so v lies in the closure $\overline{\mathrm{Span}}(B)$.

(4) $\Rightarrow$ (2): Assume that $\langle v, b \rangle = 0$ for all $b \in B$. By (4), v can be written as a limit $v = \lim_{n \to \infty} v_n$, where $v_n \in \mathrm{Span}(B)$ for all $n \in \mathbb{N}$. Then $\langle v, v_n \rangle = 0$ for all $n \in \mathbb{N}$ by assumption. By continuity of the inner product this implies

$$\|v\|^2 = \lim_{n \to \infty} \langle v, v_n \rangle = 0 \ ,$$

so $v = 0$.

(6) $\Rightarrow$ (7): Put $v = w$. Then, by assumption,

$$\|v\|^2 = \langle v, v \rangle = \sum_{b \in B} \langle v, b \rangle \langle b, v \rangle = \sum_{b \in B} |\langle v, b \rangle|^2 \ .$$

(7) $\Rightarrow$ (1): Assume (7) and that (1) is not true. Then there exists $v \in H$ with $\|v\| = 1$ and $\langle v, b \rangle = 0$ for all $b \in B$. But then

$$\|v\|^2 = \sum_{b \in B} |\langle v, b \rangle|^2 = 0, \qquad \Box$$

which is a contradiction.

3.4. The monoidal structure of the category of Hilbert spaces

3.4.1 Let $\mathbb{K}$ be the field of real or complex numbers. Hilbert spaces over $\mathbb{K}$ together with bounded $\mathbb{K}$-linear maps between them form a category denoted by $\mathbb{K}$-Hilb or just Hilb if no confusion can arise. This can be seen immediately by observing that the identity map $\mathbb{1}_{\mathcal{H}}$ on a Hilbert space is a bounded linear operator and that the composition $B \circ A : \mathcal{H}_1 \to \mathcal{H}_3$ of two bounded linear operators between Hilbert spaces $A : \mathcal{H}_1 \to \mathcal{H}_2$ and $B : \mathcal{H}_2 \to \mathcal{H}_3$ is again a bounded linear operator. We want to endow the category Hilb with a bifunctor $\widehat{\otimes} : \mathsf{Hilb} \times \mathsf{Hilb} \to \mathsf{Hilb}$ so that it becomes a monoidal category. The (bi)functor $\widehat{\otimes}$ will be called the *Hilbert tensor product.*

Unless mentioned differently, Hilbert spaces, vector spaces and the algebraic tensor product $\otimes$ in this section are assumed to be taken over the ground field $\mathbb{K}$.

3.4.2 Proposition *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces. Then there exists a unique inner product $\langle \cdot, \cdot \rangle : (\mathcal{H}_1 \otimes \mathcal{H}_2) \times (\mathcal{H}_1 \otimes \mathcal{H}_2) \to \mathbb{K}$ on the algebraic tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ such that*

$$\langle v_1 \otimes v_2, w_1 \otimes w_2 \rangle = \langle v_1, w_1 \rangle \cdot \langle v_2, w_2 \rangle \quad \text{for all } v_1, w_1 \in \mathcal{H}_1,\ v_2, w_2 \in \mathcal{H}_2 \ . \tag{3.4.1}$$

Proof. Let us first provide some preliminary constructions. Recall that for every pair of vector spaces V_1 and V_2 the bilinear map

$$\tau : \mathrm{Hom}(\mathrm{V}_1, \mathbb{K}) \times \mathrm{Hom}(\mathrm{V}_2, \mathbb{K}) \to \mathrm{Hom}(\mathrm{V}_1 \otimes \mathrm{V}_2, \mathbb{K}),$$
$$(\lambda_1, \lambda_2) \mapsto \big(\mathrm{V}_1 \otimes \mathrm{V}_2 \to \mathbb{K},\ v_1 \otimes v_2 \mapsto \lambda_1(v_1) \cdot \lambda_2(v_2)\big)$$

induces a linear map

$$\widehat{\tau} : \mathrm{Hom}(\mathrm{V}_1, \mathbb{K}) \otimes \mathrm{Hom}(\mathrm{V}_2, \mathbb{K}) \to \mathrm{Hom}(\mathrm{V}_1 \otimes \mathrm{V}_2, \mathbb{K})$$

by the universal property of the tensor product. This map is an isomorphism. To see this choose a basis $(v_{1i})_{i\in I}$ of V_1 and a basis $(v_{2j})_{j\in J}$ of V_2. Let $(v'_{1i})_{i\in I}$ and $(v'_{2j})_{j\in J}$ denote the respective dual bases of V'_1 and V'_2. Then $\left(v'_{1i} \otimes v'_{2j}\right)_{(i,j)\in I\times J}$ is a basis of $\mathrm{Hom}(\mathrm{V}_1, \mathbb{K}) \otimes \mathrm{Hom}(\mathrm{V}_2, \mathbb{K})$ which under $\widehat{\tau}$ is mapped bijectively to the basis $((v_{1i} \otimes v_{2j})')_{(i,j)\in I\times J}$ of $\mathrm{Hom}(\mathrm{V}_1 \otimes \mathrm{V}_2, \mathbb{K})$ dual to the basis $(v_{1i} \otimes v_{2j})_{(i,j)\in I\times J}$ of $\mathrm{V}_1 \otimes \mathrm{V}_2$. Hence $\widehat{\tau}$ is a linear isomorphism as claimed, and we can identify the tensor product $\lambda_1 \otimes \lambda_2$ of two linear functionals $\lambda_i : \mathrm{V}_i \to \mathbb{K}$, $i = 1, 2$ with its image in $\mathrm{Hom}(\mathrm{V}_1 \otimes \mathrm{V}_2, \mathbb{K})$.

Now observe that for two conjugate-linear maps $\mu_1 : \mathrm{V}_1 \to \mathbb{K}$ and $\mu_2 : \mathrm{V}_2 \to \mathbb{K}$ the map $\tau^*(\mu_1, \mu_2) = \overline{\overline{\mu_1} \otimes \overline{\mu_2}} : \mathrm{V}_1 \otimes \mathrm{V}_2 \to \mathbb{K}$ is conjugate-linear and satisfies

$$\tau^*(\mu_1, \mu_2)\,(v_1 \otimes v_2) = \mu_1(v_1) \cdot \mu_2(v_2) \quad \text{for all } v_1 \in \mathrm{V}_1,\ v_2 \in \mathrm{V}_2 \ . \tag{3.4.2}$$

One obtains a map

$$\tau^* : \mathrm{Hom}^*(\mathrm{V}_1, \mathbb{K}) \times \mathrm{Hom}^*(\mathrm{V}_2, \mathbb{K}) \to \mathrm{Hom}^*(\mathrm{V}_1 \otimes \mathrm{V}_2, \mathbb{K}) \ ,$$

where here the symbol $\mathrm{Hom}^*(\mathrm{V}, \mathbb{K})$ denotes the space of all conjugate linear functionals on a vector space V. Since τ^* is biadditive and since $\tau^*(z\mu_1, \mu_2) = \tau^*(\mu_1, z\mu_2)$ for all $\mu_1 \in \mathrm{Hom}^*(\mathrm{V}_1, \mathbb{K})$, $\mu_2 \in \mathrm{Hom}^*(\mathrm{V}_2, \mathbb{K})$, and $z \in \mathbb{K}$, the map τ^* factors through a linear map

$$\widehat{\tau^*} : \mathrm{Hom}^*(\mathrm{V}_1, \mathbb{K}) \otimes \mathrm{Hom}^*(\mathrm{V}_2, \mathbb{K}) \to \mathrm{Hom}^*(\mathrm{V}_1 \otimes \mathrm{V}_2, \mathbb{K}) \ .$$

Using the above bases $(v_{1i})_{i\in I}$ and $(v_{2j})_{j\in J}$ of V_1 and V_2 respectively, one observes that $\widehat{\tau^*}$ is an isomorphism since it maps the basis $\left(\overline{v'_{1i}} \otimes \overline{v'_{2j}}\right)_{(i,j)\in I\times J}$ of $\mathrm{Hom}^*(\mathrm{V}_1, \mathbb{K}) \otimes \mathrm{Hom}^*(\mathrm{V}_2, \mathbb{K})$ bijectively to the basis $\left(\overline{(v_{1i} \otimes v_{2j})'}\right)_{(i,j)\in I\times J}$ of the space $\mathrm{Hom}^*(\mathrm{V}_1 \otimes \mathrm{V}_2, \mathbb{K})$. So $\widehat{\tau^*}$ is also a linear isomorphism, which allows us to identify the tensor product $\mu_1 \otimes \mu_2$ of two conjugate linear functionals $\mu_i : \mathrm{V}_i \to \mathbb{K}$, $i = 1, 2$ with its image in $\mathrm{Hom}^*(\mathrm{V}_1 \otimes \mathrm{V}_2, \mathbb{K})$.

After these preliminary considerations we consider the map

$$\mathcal{H}_1 \times \mathcal{H}_2 \to \mathrm{Hom}^*(\mathcal{H}_1 \otimes \mathcal{H}_2, \mathbb{K}),\ (w_1, w_2) \mapsto \overline{w_1^\flat} \otimes \overline{w_2^\flat} = \tau^*\left(\overline{w_1^\flat}, \overline{w_2^\flat}\right) = \widehat{\tau^*}\left(\overline{w_1^\flat} \otimes \overline{w_2^\flat}\right) \ ,$$

which is well-defined and bilinear since the musical isomorphisms $^\flat : \mathcal{H}_l \to \mathcal{H}_l'$, $w \mapsto \langle w, - \rangle$, $l = 1, 2$, are conjugate-linear and since τ^* is bilinear. Hence it factors through a linear map

$$\beta : \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathrm{Hom}^*(\mathcal{H}_1 \otimes \mathcal{H}_2, \mathbb{K})$$

such that

$$\beta(w_1 \otimes w_2)(v_1 \otimes v_2) = \langle v_1, w_1 \rangle \cdot \langle v_2, w_2 \rangle \quad \text{for all } v_1, w_1 \in \mathcal{H}_1,\ v_2, w_2 \in \mathcal{H}_2 . \tag{3.4.3}$$

Now put

$$\langle \cdot, \cdot \rangle : (\mathcal{H}_1 \otimes \mathcal{H}_2) \times (\mathcal{H}_1 \otimes \mathcal{H}_2) \to \mathbb{K},\ (v, w) \mapsto \langle v, w \rangle := \beta(w)(v) .$$

Then $\langle \cdot, \cdot \rangle$ is sesquilinear by construction, and (3.4.1) holds true by (3.4.3).

Let us show that $\langle \cdot, \cdot \rangle$ is positive definite. Let $v = \sum_{k=1}^n v_{1k} \otimes v_{2k} \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Choose an orthonormal basis $e_1, \ldots, e_m$ of the linear subspace spanned by $v_{21}, \ldots, v_{2n}$. Expand $v_{2k} = \sum_{i=1}^m c_{ki} e_i$ with $c_{k1}, \ldots, c_{km} \in \mathbb{K}$. Then

$$v = \sum_{k=1}^n v_{1k} \otimes v_{2k} = \sum_{k=1}^n \sum_{i=1}^m v_{1k} \otimes (c_{ki} e_i) = \sum_{i=1}^m \left(\sum_{k=1}^n c_{ki} v_{1k} \right) \otimes e_i = \sum_{i=1}^m w_{1i} \otimes e_i , \tag{3.4.4}$$

where $w_{1i} = \sum_{k=1}^n c_{ki} v_{1k}$. Hence

$$\langle v, v \rangle = \langle \sum_{i=1}^m w_{1i} \otimes e_i, \sum_{j=1}^m w_{1j} \otimes e_j \rangle = \sum_{i=1}^m \sum_{j=1}^m \langle w_{1i}, w_{1j} \rangle \langle e_i, e_j \rangle = \sum_{i=1}^m \|w_{1i}\|^2 \geqslant 0 . \tag{3.4.5}$$

Moreover, if $\langle v, v \rangle = 0$, then $w_{1i} = 0$ for $i = 1, \ldots, m$, which implies $v = \sum_{i=1}^m w_{1i} \otimes e_i = 0$. So $\langle \cdot, \cdot \rangle$ is an inner product on $\mathcal{H}_1 \otimes \mathcal{H}_2$ satisfying (3.4.1). It is uniquely determined by this condition since the vectors $v_1 \otimes v_2$ with $v_1 \in \mathcal{H}_1$ and $v_2 \in \mathcal{H}_2$ span $\mathcal{H}_1 \otimes \mathcal{H}_2$. □

3.4.3 Definition Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. The Hilbert completion of the algebraic tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ equipped with the unique inner product $\langle \cdot, \cdot \rangle$ fulfilling (3.4.1) will be denoted $\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2$, its inner product again by $\langle \cdot, \cdot \rangle$. One calls the Hilbert space $\big(\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2, \langle \cdot, \cdot \rangle\big)$ the *Hilbert tensor product* of $\mathcal{H}_1$ and $\mathcal{H}_2$ or just the *tensor product* of $\mathcal{H}_1$ and $\mathcal{H}_2$ if no confusion can arise.

3.4.4 Proposition *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces.*

(i) *Let $A_1 \subset \mathcal{H}_1$ and $A_2 \subset \mathcal{H}_2$ be subsets which are total $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Then the set of simple vectors $a_1 \otimes a_2$ with $a_1 \in A_1$ and $a_2 \in A_2$ is total in the Hilbert tensor product $\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2$.*

(ii) *If $(e_i)_{i \in I}$ and $(f_j)_{j \in J}$ are orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, then $(e_i \otimes f_j)_{(i,j) \in I \times J}$ is an orthonormal basis of the Hilbert tensor product $\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2$.*

Proof. ad (i). Recall that a subset $A \subset \mathcal{H}$ or a family $A = (a_j)_{j \in J}$ of elements of a Hilbert space $\mathcal{H}$ is called *total* in $\mathcal{H}$ if the linear span of A is dense in $\mathcal{H}$. By density of the algebraic tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ in the Hilbert tensor product $\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2$, the set of simple tensors $v_1 \otimes v_2$ with

$(v_1, v_2) \in \mathcal{H}_1 \times \mathcal{H}_2$ is total in $\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2$. Hence it suffices to find for each such pair (v_1, v_2) and all $\varepsilon > 0$ vectors $w_1 \in \operatorname{Span} A_1$ and $w_2 \in \operatorname{Span} A_2$ such that

$$\|v_1 \otimes v_2 - w_1 \otimes w_2\| < \frac{\varepsilon}{2} .$$

By totality of A_i in $\mathcal{H}_i$ there exist $w_i \in \operatorname{Span} A_i$ for $i = 1, 2$ such that

$$\|v_1 - w_1\| < \min\left\{1, \frac{\varepsilon}{2(\|v_2\| + 1)}\right\} \quad \text{and} \quad \|v_2 - w_2\| < \frac{\varepsilon}{2(\|v_1\| + 1)} .$$

Then

$$\|v_1 \otimes v_2 - w_1 \otimes w_2\| \leqslant \|v_1 - w_1\| \, \|v_2\| + \|v_2 - w_2\| \|w_1\| < \varepsilon .$$

ad (ii). The family $(e_i \otimes f_j)_{(i,j)\in I\times J}$ is orthonormal by definition of the inner product on $\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2$. It is total by (i) and therefore a Hilbert basis. □

3.4.5 Proposition *Assigning to each pair of Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ the Hilbert tensor product $\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2$ and to each pair of bounded linear operators $A_1 : \mathcal{H}_1 \to \mathcal{H}_3$ and $A_2 : \mathcal{H}_2 \to \mathcal{H}_4$ between Hilbert spaces the unique bounded extension $A_1 \hat{\otimes} A_2 : \mathcal{H}_1 \hat{\otimes} \mathcal{H}_2 \to \mathcal{H}_3 \hat{\otimes} \mathcal{H}_4$ of the operator $A_1 \otimes A_2 : \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}_3 \hat{\otimes} \mathcal{H}_4$, $v_1 \otimes v_2 \mapsto A_1(v_1) \otimes A_2(v_2)$ comprises a (covariant) bifunctor*

$$\hat{\otimes} : \mathsf{Hilb} \times \mathsf{Hilb} \to \mathsf{Hilb} .$$

Moreover, $\hat{\otimes}$ is isometric in the sense that

$$\|v_1 \otimes v_2\| = \|v_1\| \, \|v_2\| \quad \textit{for all } v_1 \in \mathcal{H}_1,\ v_2 \in \mathcal{H}_1 \textit{ and} \tag{3.4.6}$$

$$\|A_1 \hat{\otimes} A_2\| = \|A_1\| \, \|A_2\| \quad \textit{for all } A_1 \in \mathfrak{B}(\mathcal{H}_1, \mathcal{H}_3),\ A_2 \in \mathfrak{B}(\mathcal{H}_2, \mathcal{H}_4) . \tag{3.4.7}$$

Proof. We first show that $A_1 \otimes A_2$ is a bounded operator. To this end observe that $A_1 \otimes A_2$ can be written as the composition of the two operators $A_1 \otimes \mathbb{1}_{\mathcal{H}_2}$ and $\mathbb{1}_{\mathcal{H}_3} \otimes A_2$. Hence it suffices to show that each of these linear maps is bounded. Let $v = \sum_{k=1}^n v_{1k} \otimes v_{2k} \in \mathcal{H}_1 \otimes \mathcal{H}_2$ be of norm 1. As in the proof of Proposition 3.4.2 expand $v_{2k} = \sum_{i=1}^m c_{ki} e_i$, $k = 1, \ldots, n$, where $e_1, \ldots, e_m$ is an orthonormal basis of $\operatorname{Span}\{v_{21}, \ldots, v_{2n}\}$ and $c_{k1}, \ldots, c_{km} \in \mathbb{K}$. Equations (3.4.4) and (3.4.5) then entail that

$$v = \sum_{i=1}^m w_{1i} \otimes e_i \quad \text{and} \quad 1 = \langle v, v\rangle = \sum_{i=1}^m \|w_{1i}\|^2$$

where $w_{1i} = \sum_{k=1}^n c_{ki} v_{1k}$ for $i = 1, \ldots, m$. Hence

$$\|(A_1 \otimes \mathbb{1}_{\mathcal{H}_2})v\|^2 = \left\|\sum_{i=1}^m A_1(w_{1i}) \otimes e_i\right\|^2 = \sum_{i=1}^m \|A_1(w_{1i})\|^2 \leqslant \|A_1\|^2 \sum_{i=1}^m \|w_{1i}\|^2 = \|A_1\|^2 ,$$

so $A_1 \otimes \mathbb{1}_{\mathcal{H}_2}$ is bounded with norm $\leqslant \|A_1\|$. By symmetry, $\mathbb{1}_{\mathcal{H}_3} \otimes A_2$ is bounded with norm $\leqslant \|A_2\|$. Hence $A_1 \otimes A_2 = (\mathbb{1}_{\mathcal{H}_3} \otimes A_2) \circ (A_1 \otimes \mathbb{1}_{\mathcal{H}_2})$ is bounded and

$$\|A_1 \otimes A_2\| \leqslant \|A_1\| \, \|A_2\| .$$

Therefore, $A_1 \otimes A_2$ has a unique bounded extension $A_1 \widehat{\otimes} A_2$ of norm

$$\|A_1 \widehat{\otimes} A_2\| = \|A_1 \otimes A_2\| \leqslant \|A_1\| \, \|A_2\| \, .$$

Let us show that the converse inequality holds as well. For given $\varepsilon > 0$ there exist unit vectors $v_i \in \mathcal{H}_i$, $i = 1, 2$ such that $\|A_i v_i\| \geqslant \|A_i\| - \frac{\varepsilon}{2(\|A_1\| + \|A_2\| + 1)}$. Then

$$\|(A_1 \otimes A_2)(v_1 \otimes v_2)\| = \|A_1 v_1\| \, \|A_2 v_2\| \geqslant \|A_1\| \, \|A_2\| - \varepsilon \, .$$

This implies

$$\|A_1 \widehat{\otimes} A_2\| = \|A_1 \otimes A_2\| \geqslant \|A_1\| \, \|A_2\|$$

and (3.4.7) follows. Equality (3.4.6) is clear by construction of the Hilbert tensor product.

Next observe that $\mathbb{1}_{\mathcal{H}_1} \widehat{\otimes} \mathbb{1}_{\mathcal{H}_2} = \mathbb{1}_{\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2}$ by definition. Given Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_6$ and bounded linear operators $A_i : \mathcal{H}_i \to \mathcal{H}_{i+2}$ and $B_i : \mathcal{H}_{i+2} \to \mathcal{H}_{i+4}$ for $i = 1, 2$, the composition $(B_1 \otimes B_2) \circ (A_1 \otimes A_2)$ coincides with $(B_1 \circ A_1) \otimes (B_2 \circ A_2)$ by functoriality of the algebraic tensor product. By continuity of the operators $A_1 \widehat{\otimes} A_2$ and $B_1 \widehat{\otimes} B_2$ and by density of $\mathcal{H}_1 \otimes \mathcal{H}_2$ in $\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2$ the equality

$$(B_1 \widehat{\otimes} B_2) \circ (A_1 \widehat{\otimes} A_2) = (B_1 \circ A_1) \widehat{\otimes} (B_2 \circ A_2)$$

follows. Hence $\widehat{\otimes}$ is a bifunctor as claimed. □

3.4.6 Proposition *For every Hilbert space $\mathcal{H}$ one has two natural isomorphisms*

$$\widehat{u}_{\mathcal{H}} : \mathbb{K} \widehat{\otimes} \mathcal{H} \to \mathcal{H}, \; z \otimes v \to zv \quad \text{and} \quad {}_{\mathcal{H}}\widehat{u} : \mathcal{H} \widehat{\otimes} \mathbb{K} \to \mathcal{H}, \; v \otimes z \to zv$$

called the left *and the* right unit*, respectively. Furthermore, for every triple of Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ there is a natural isomorphism, called* associator

$$\widehat{a}_{\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3} : (\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2) \widehat{\otimes} \mathcal{H}_3 \to \mathcal{H}_1 \widehat{\otimes} (\mathcal{H}_2 \widehat{\otimes} \mathcal{H}_3), \; (v_1 \otimes v_2) \otimes v_3 \mapsto v_1 \otimes (v_2 \otimes v_3) \, .$$

These data fulfill the so-called coherence conditions *that is the* pentagon diagram

$$\begin{tikzcd}
 & ((\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2) \widehat{\otimes} \mathcal{H}_3) \widehat{\otimes} \mathcal{H}_4 \arrow[dl, "\widehat{a}_{\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3} \widehat{\otimes} \mathbb{1}_{\mathcal{H}_4}"'] \arrow[dr, "\widehat{a}_{\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2, \mathcal{H}_3, \mathcal{H}_4}"] & \\
(\mathcal{H}_1 \widehat{\otimes} (\mathcal{H}_2 \widehat{\otimes} \mathcal{H}_3)) \widehat{\otimes} \mathcal{H}_4 \arrow[d, "\widehat{a}_{\mathcal{H}_1, \mathcal{H}_2 \widehat{\otimes} \mathcal{H}_3, \mathcal{H}_4}"'] & & (\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2) \widehat{\otimes} (\mathcal{H}_3 \widehat{\otimes} \mathcal{H}_4) \arrow[d, "\widehat{a}_{\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3 \widehat{\otimes} \mathcal{H}_4}"] \\
\mathcal{H}_1 \widehat{\otimes} ((\mathcal{H}_2 \widehat{\otimes} \mathcal{H}_3) \widehat{\otimes} \mathcal{H}_4) \arrow[rr, "\mathbb{1}_{\mathcal{H}_1} \widehat{\otimes} \widehat{a}_{\mathcal{H}_2, \mathcal{H}_3, \mathcal{H}_4}"'] & & \mathcal{H}_1 \widehat{\otimes} (\mathcal{H}_2 \widehat{\otimes} (\mathcal{H}_3 \widehat{\otimes} \mathcal{H}_4))
\end{tikzcd}$$

and the triangle diagram

$$\begin{tikzcd}
(\mathcal{H}_1 \mathbin{\widehat{\otimes}} \mathbb{K}) \mathbin{\widehat{\otimes}} \mathcal{H}_2 \arrow[rr, "\widehat{a}_{\mathcal{H}_1,\mathbb{K},\mathcal{H}_3}"] \arrow[dr, "{}_{\mathcal{H}_1}\widehat{u} \mathbin{\widehat{\otimes}} \mathbb{1}_{\mathcal{H}_2}"'] & & \mathcal{H}_1 \mathbin{\widehat{\otimes}} (\mathbb{K} \mathbin{\widehat{\otimes}} \mathcal{H}_2) \arrow[dl, "\mathbb{1}_{\mathcal{H}_1} \mathbin{\widehat{\otimes}} \widehat{u}_{\mathcal{H}_2}"] \\
& \mathcal{H}_1 \mathbin{\widehat{\otimes}} \mathcal{H}_2 &
\end{tikzcd}$$

commute for all Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3, \mathcal{H}_4$. *In other words, the category* Hilb *endowed with the Hilbert tensor product* $\widehat{\otimes}$ *is a monoidal category.*

Proof. The category of $\mathbb{K}$-vector spaces with the usual tensor product as tensor functor is monoidal. Denote the corresponding unit isomorphisms and associator by ${}_{-}u$, u_{-}, and $a_{-,-,-}$, respectively. Then observe that by construction $\mathbb{K} \mathbin{\widehat{\otimes}} \mathcal{H} = \mathbb{K} \otimes \mathcal{H}$ and $\mathcal{H} \mathbin{\widehat{\otimes}} \mathbb{K} = \mathcal{H} \otimes \mathbb{K}$ for every Hilbert space $\mathcal{H}$. In particular this means that $\widehat{u}_{\mathcal{H}}$ coincides with the unit $u_{\mathcal{H}}$ and ${}_{\mathcal{H}}\widehat{u}$ with the unit ${}_{\mathcal{H}}u$. Moreover, both units $\widehat{u}_{\mathcal{H}}$ and ${}_{\mathcal{H}}\widehat{u}$ are bounded. Next recall that $\mathcal{H}_1 \otimes \mathcal{H}_2$ is dense in $\mathcal{H}_1 \mathbin{\widehat{\otimes}} \mathcal{H}_2$ which by Proposition 3.4.4 implies density of $(\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathcal{H}_3$ and $\mathcal{H}_1 \otimes (\mathcal{H}_2 \otimes \mathcal{H}_3)$ in $(\mathcal{H}_1 \mathbin{\widehat{\otimes}} \mathcal{H}_2) \mathbin{\widehat{\otimes}} \mathcal{H}_3$ and $\mathcal{H}_1 \mathbin{\widehat{\otimes}} (\mathcal{H}_2 \mathbin{\widehat{\otimes}} \mathcal{H}_3)$, respectively. Similarly one argues that $\mathcal{H}_1 \otimes (\mathcal{H}_2 \otimes (\mathcal{H}_3 \otimes \mathcal{H}_4))$ is dense in $\mathcal{H}_1 \mathbin{\widehat{\otimes}} (\mathcal{H}_2 \mathbin{\widehat{\otimes}} (\mathcal{H}_3 \mathbin{\widehat{\otimes}} \mathcal{H}_4))$, and so on. Since the associator map $a_{\mathcal{H}_1,\mathcal{H}_2,\mathcal{H}_3} : (\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathcal{H}_3 \to \mathcal{H}_1 \otimes (\mathcal{H}_2 \otimes \mathcal{H}_3)$ is bounded, it extends in a unique way to a linear bounded map $\widehat{a}_{\mathcal{H}_1,\mathcal{H}_2,\mathcal{H}_3} : (\mathcal{H}_1 \mathbin{\widehat{\otimes}} \mathcal{H}_2) \mathbin{\widehat{\otimes}} \mathcal{H}_3 \to \mathcal{H}_1 \mathbin{\widehat{\otimes}} (\mathcal{H}_2 \mathbin{\widehat{\otimes}} \mathcal{H}_3)$. Using density, continuity, and commutativity of the pentagon and triangle diagrams for the tensor product functor one concludes that the coherence conditions for $\widehat{\otimes}$ with the unit and associator maps ${}_{-}\widehat{u}$, $\widehat{u}_{-}$, and $\widehat{a}_{-,-,-}$ are satisfied. □

3.5. Adjoints of bounded operators

3.5.1 As before, the symbols $\mathcal{H}$ and $\mathcal{H}_k$ with $k = 1, 2$ always stand for Hilbert spaces over the field $\mathbb{K}$ of real or complex numbers. Several results of this section hold only in the complex case, thouhgh. Therefore we will be quite precise in stating all necessary assumptions, in particular about the ground field.

Let $A \in \mathfrak{B}(\mathcal{H}_1, \mathcal{H}_2)$ that is let $A : \mathcal{H}_1 \to \mathcal{H}_2$ be linear and bounded. Then the map

$$b_A : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathbb{K},\ (v, w) \mapsto \langle Av, w \rangle$$

is sesquilinear and bounded with norm

$$\|b_A\| = \sup \left\{ |b_A(v, w)| \,\middle|\, v \in \mathcal{H}_1,\ w \in \mathcal{H}_2,\ \|w\| = \|v\| = 1 \right\} = \|A\| \ .$$

By Corollary 3.2.8 to the Riesz representation theorem there exists a unique bounded linear operator $A^* : \mathcal{H}_2 \to \mathcal{H}_1$ such that

$$b_A(v, w) = \langle v, A^* w \rangle \quad \text{for all } v \in \mathcal{H}_1,\ w \in \mathcal{H}_2 \ .$$

This operator satisfies

$$\|A^*\| = \|b_A\| = \|A\| \ . \tag{3.5.1}$$

3.5.2 Definition The unique operator $A^* \in \mathfrak{B}(\mathcal{H}_2, \mathcal{H}_1)$ associated to an operator $A \in \mathfrak{B}(\mathcal{H}_1, \mathcal{H}_2)$ such that

$$\langle Av, w\rangle = \langle v, A^*w\rangle \quad \text{for all } v \in \mathcal{H}_1,\ w \in \mathcal{H}_2$$

is called the *adjoint* of A.

The fundamental property of the adjoint operation is given by the following result.

3.5.3 Proposition *The adjoint map* $^* : \mathfrak{B}(\mathcal{H}_1, \mathcal{H}_2) \to \mathfrak{B}(\mathcal{H}_2, \mathcal{H}_1)$ *is a conjugate linear isometry whose square coincides with the identity operation that is* $A^{**} = A$ *for all* $A \in \mathfrak{B}(\mathcal{H}_1, \mathcal{H}_2)$.

Proof. By the proof of Corollary 3.2.8, $A^*w = \langle w, A(-)\rangle^\sharp$ for all $w \in \mathcal{H}_2$. Since the inner product is linear in the second argument and the operator $^\sharp$ conjugate linear, the map $A \mapsto A^*$ is conjugate linear in A. By Equation (3.5.1), the adjoint map is an isometry. The relation $A^{**} = A$ follows by uniqueness of the adjoint and since

$$\langle A^*w, v\rangle = \langle w, Av\rangle \quad \text{for all } v \in \mathcal{H}_1,\ w \in \mathcal{H}_2 \ . \qquad \square$$

3.5.4 Definition An operator $A \in \mathfrak{B}(\mathcal{H})$ is called *self-adjoint* if $A = A^*$, *unitary* if $A^* = A^{-1}$, and *normal* if $[A, A^*] := AA^* - A^*A = 0$.

We note that self-adjoint and unitary operators are always normal, but normal operators do not have to be self-adjoint or unitary. In the remainder of this section, we gather several results on self-adjoint and normal operators.

3.5.5 Proposition *Assume that the ground field* $\mathbb{K}$ *of the Hilbert space* $\mathcal{H}$ *is the field of complex numbers. An operator* $A \in \mathfrak{B}(\mathcal{H})$ *then is self-adjoint if and only if* $\langle Av, v\rangle \in \mathbb{R}$ *for all* $v \in \mathcal{H}$.

Proof. ($\Rightarrow$) If A is self-adjoint, then

$$\langle Av, v\rangle = \langle v, A^*v\rangle = \langle v, Av\rangle = \overline{\langle Av, v\rangle} \ ,$$

which implies that $\langle Av, v\rangle \in \mathbb{R}$.

($\Leftarrow$) Suppose that $\langle Av, v\rangle \in \mathbb{R}$ for all $v \in \mathcal{H}$. We know

$$\langle A(v+w), v+w\rangle = \langle Av, v\rangle + \langle Av, w\rangle + \langle Aw, v\rangle + \langle Aw, w\rangle \ . \tag{3.5.2}$$

By assumption, $\langle A(v+w), v+w\rangle$, $\langle Av, v\rangle$, and $\langle Aw, w\rangle$ are all real. This implies that the sum $\langle Av, w\rangle + \langle Aw, v\rangle$ is real as well, so

$$\mathfrak{Im}\,\langle Av, w\rangle = -\mathfrak{Im}\,\langle Aw, v\rangle = \mathfrak{Im}\,\langle v, Aw\rangle \ .$$

Since this holds for all $w \in \mathcal{H}$, it holds for $\mathsf{i}w$, too. Thus,

$$\mathfrak{Re}\,\langle Av, w\rangle = \mathfrak{Im}\,\mathsf{i}\langle Av, w\rangle = \mathfrak{Im}\,\langle Av, \mathsf{i}w\rangle = \mathfrak{Im}\,\langle v, A(\mathsf{i}w)\rangle = \mathfrak{Im}\,\mathsf{i}\langle v, Aw\rangle = \mathfrak{Re}\,\langle v, Aw\rangle \ .$$

Combining the above two lines yields $\langle Av, w\rangle = \langle v, Aw\rangle$ for all $v, w \in \mathcal{H}$. By uniqueness of the adjoint this implies that $A = A^*$. $\square$

3.5.6 Proposition *Assume that the ground field $\mathbb{K}$ of the Hilbert space $\mathcal{H}$ is the field of complex numbers and let $A \in \mathfrak{B}(\mathcal{H})$. If $\langle Av, v\rangle = 0$ holds for all $v \in \mathcal{H}$, then $A = 0$.*

Proof. Since $\langle Av, v\rangle = 0$ for all $v \in H$, equation (3.5.2) from the proof of Proposition 3.5.5 reduces to

$$\langle Av, w\rangle = -\langle Aw, v\rangle = -\langle w, Av\rangle = -\overline{\langle Av, w\rangle} \quad \text{for all } v, w \in \mathcal{H} \ .$$

That means that $\langle Av, w\rangle$ has no real part for all $v, w \in \mathcal{H}$. But then fixing v and setting $w = Av$ implies $\|Av\|^2 = 0$ for all $v \in \mathcal{H}$, so $A = 0$. □

3.5.7 Example The preceding proposition does not hold in the real case. To see this take rotation by $\frac{\pi}{2}$:

$$R = \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix}$$

Then $\langle Rv, v\rangle = 0$ for all $v \in \mathbb{R}^2$, but R is non-zero. Note that the example of the rotation operator R also shows that the criterion for self-adjointness from Proposition 3.5.5 can not be applied in the real case.

3.5.8 Lemma (cf. (Hirzebruch and Scharlau, 1991, Lem. 22.4)) *Assume that A is a bounded linear operator on the real or complex Hilbert space $\mathcal{H}$ for which there exists a $C \geqslant 0$ such that*

$$|\langle Av, v\rangle| \leqslant C\|v\|^2 \quad \textit{for all } v \in \mathcal{H} \ .$$

Then

$$|\langle Av, w\rangle + \langle v, Aw\rangle| \leqslant 2C\|v\|\|w\| \quad \textit{for all } v, w \in \mathcal{H} \ . \tag{3.5.3}$$

In case $\mathcal{H}$ is a complex Hilbert space one even has the sharper estimate

$$|\langle Av, w\rangle| + |\langle v, Aw\rangle| \leqslant 2C\|v\|\|w\| \quad \textit{for all } v, w \in \mathcal{H} \ . \tag{3.5.4}$$

Proof. We start with the equality

$$\langle A(v+w), v+w\rangle + \langle A(v-w), v-w\rangle = 2(\langle Av, w\rangle + \langle Aw, v\rangle) \ . \tag{3.5.5}$$

By assumption and the parallelogram identity (3.1.3) this entails

$$2|\langle Av, w\rangle + \langle Aw, v\rangle| \leqslant C\left(\|v+w\|^2 + \|v-w\|^2\right) = 2C\left(\|v\|^2 + \|w\|^2\right) \ . \tag{3.5.6}$$

The claim obviously holds for $v = 0$ or $w = 0$, so we assume from now on that both v and w are non-zero. Then put $a = \sqrt{\frac{\|v\|}{\|w\|}}$ and replace in (3.5.6) v by $\frac{v}{a}$ and w by aw. One obtains

$$|\langle Av, w\rangle + \langle Aw, v\rangle| \leqslant C\left(\left\|\frac{v}{a}\right\|^2 + \left\|aw\right\|^2\right) = 2C\|v\|\|w\|$$

which is the claim in the real case. If $\mathcal{H}$ is a complex Hilbert space, let x, y be complex numbers of modulus 1. In the just proven estimate multiply the left side with $|x|$ and replace w with yw. This gives

$$\left|xy\langle Av, w\rangle + x\overline{y}\langle Aw, v\rangle\right| = |x| \cdot |\langle Av, yw\rangle + \langle A(yw), v\rangle| \leqslant 2C\|v\|\|w\| \ . \tag{3.5.7}$$

Now write $\langle Av, w\rangle = re^{\mathrm{i}\varphi}$ and $\langle Aw, v\rangle = se^{\mathrm{i}\psi}$ with $r, s \geqslant 0$ and $\varphi, \psi \in \mathbb{R}$. Then put

$$x = e^{-\mathrm{i}\frac{1}{2}(\varphi+\psi))} \quad \text{and} \quad y = e^{-\mathrm{i}\frac{1}{2}(\varphi-\psi))} \ .$$

With these values, (3.5.7) becomes

$$|\langle Av, w\rangle| + |\langle v, Aw\rangle| \leqslant 2C\|v\|\|w\|$$

which was to be shown. □

3.5.9 Proposition *If $\mathcal{H}$ is a Hilbert space over the field $\mathbb{K}$ of real or complex numbers and $A \in \mathfrak{B}(\mathcal{H})$ is self-adjoint, then*

$$\|A\| = \sup_{\|v\|=1} |\langle Av, v\rangle| \ .$$

Proof. We know

$$\|A\| = \sup_{\|v\|=\|w\|=1} |\langle Av, w\rangle| \ , \tag{3.5.8}$$

so we clearly have

$$\sup_{\|v\|=1} |\langle Av, v\rangle| \leqslant \|A\| \ .$$

The other direction follows from Equation (3.5.8) and Lemma 3.5.8 since A is self-adjoint. □

3.5.10 Proposition *If $\mathcal{H}$ is a real or complex Hilbert space and $A \in \mathfrak{B}(\mathcal{H})$, then A^*A is self-adjoint and $\|A^*A\| = \|A\|^2$.*

Proof. For arbitrary $v, w \in \mathcal{H}$, we have

$$\langle A^*Av, w\rangle = \langle Av, Aw\rangle = \langle v, A^*Aw\rangle$$

so A^*A is self-adjoint. Then

$$\|A^*A\| = \sup_{\|v\|=\|w\|=1} |\langle A^*Av, w\rangle| = \sup_{\|v\|=\|w\|=1} |\langle Av, Aw\rangle| = \|A\|^2 \ ,$$

where the last equality is a consequence of the Cauchy–Schwarz inequality and the observation that for all $\varepsilon > 0$ there exists a unit vector v such that $\langle Av, Av\rangle \geqslant \|A\|^2 - \varepsilon$. □

3.5.11 Proposition *Let $\mathcal{H}$ be a complex Hilbert space $\mathcal{H}$. If $A \in \mathfrak{B}(\mathcal{H})$, then there exist unique self-adjoint $B, C \in \mathfrak{B}(\mathcal{H})$ such that $A = B + \mathrm{i}C$. Furthermore, A is normal if and only if $[B, C] = 0$.*

Proof. We define

$$B = \frac{1}{2}(A + A^*) \quad \text{and} \quad C = \frac{\mathrm{i}}{2}(A^* - A).$$

Clearly $A = B + \mathrm{i}C$. Note also that $A^* = B - \mathrm{i}C$. Furthermore, by Proposition 3.5.3

$$B^* = \frac{1}{2}(A^* + A) = B$$

and

$$C^* = -\frac{\mathsf{i}}{2}(A - A^*) = C \ .$$

Hence B and C are self-adjoint, so fulfill the claim. Let us show uniqueness. Assume that $B', C' \in \mathfrak{B}(\mathcal{H})$ are selfadjoint and satisfy $A = B' + \mathsf{i}C'$. Then

$$B - B' = B^* - B'^* = \big(\mathsf{i}(C' - C)\big)^* = -\mathsf{i}(C' - C) = -(B - B') \ .$$

Hence $B = B'$ and consequently $C = C'$. Finally, we compute

$$[A, A^*] = [B + \mathsf{i}C, B - \mathsf{i}C] = -\mathsf{i}[B, C] + \mathsf{i}[C, B] = -2\mathsf{i}[B, C] \ .$$

This entails that A is normal if and only if $[B, C] = 0$. □

3.5.12 Proposition *If A is a normal operator on a real or complex Hilbert space $\mathcal{H}$, then*

$$\|Av\| = \|A^*v\| \quad \textit{for all } v \in \mathcal{H} \ .$$

Proof. Using the fact that $A^*A = AA^*$, we compute

$$\|Av\|^2 = \langle Av, Av \rangle = \langle v, A^*Av \rangle = \langle v, AA^*v \rangle = \langle A^*v, A^*v \rangle = \|A^*v\|^2 \ .$$

Taking the square root yields the desired result. □

3.6. The lattice of orthogonal projections

3.6.1 In their celebrated article on quantum logic from 1936, Birkhoff and von Neumann showed that the set of closed linear subspaces of a Hilbert space carries the structure of a complete orthocomplemented lattice. In this section, we will describe the Birkhoff–von Neumann lattice structure. We start with the definition of orthogonal projections and the observation that the space of orthogonal projections on a Hilbert space $\mathcal{H}$ is in bijective correspondence with the closed linear subspaces of $\mathcal{H}$.

3.6.2 Definition By an *orthogonal projection* on a Hilbert space $\mathcal{H}$ one understands a bounded self-adjoint operator $P : \mathcal{H} \to \mathcal{H}$ which is an *idempotent* that is it fulfills the relation

$$P^2 = P \ . \tag{3.6.1}$$

3.7. Projection-valued measures and spectral integrals

3.7.1 In this section $\mathcal{H}$ will always denote a fixed complex Hilbert space.

3.7.2 Definition By a *projection-valued measure* or a *spectral measure* on a measurable space $(\Omega, \mathcal{A})$ one understands a map $E : \mathcal{A} \to \mathfrak{B}(\mathcal{H})$ having the following properties:

(SM0) For each $\Delta \in \mathcal{A}$ the operator $E(\Delta)$ is an orthogonal projection that is $E(\Delta)^2 = E(\Delta)$ and $E(\Delta)^* = E(\Delta)$.

(SM1) $E(\Omega) = \mathrm{id}_{\mathcal{H}}$.

(SM2) For every sequence $(\Delta_n)_{n\in\mathbb{N}}$ of pairwise disjoint elements of $\mathcal{A}$ one has

$$E\left(\bigcup_{n\in\mathbb{N}} \Delta_n\right) = \text{s-}\sum_{n=0}^{\infty} E(\Delta_n) \ ,$$

where convergence is with respect to the strong operator toplogy.

3.7.3 Remark Recall that *convergence* of a sequence of operators $(A_n)_{n\in\mathbb{N}} \subset \mathfrak{B}(\mathcal{H})$ in the *strong operator topology* to some A means that for every $v \in \mathcal{H}$ the sequence $(A_n v)_{n\in\mathbb{N}}$ converges in $\mathcal{H}$ to Av. One denotes this by $A = \text{s-}\lim_{n\to\infty} A_n$. Likewise, $B = \text{s-}\sum_{n=0}^{\infty} A_n$ means that the sequence of partial sums $\left(\sum_{k=0}^{n} A_n\right)_{n\in\mathbb{N}}$ converges in the strong operator topology to some $B \in \mathfrak{B}(\mathcal{H})$.

3.7.4 Proposition *A spectral measure $E : \mathcal{A} \to \mathfrak{B}(\mathcal{H})$ has the following properties in addition to the defining axioms:*

(SM1') $E(\varnothing) = 0$.

(SM2') (Finite additivity) *One has for all disjoint $\Delta_1, \Delta_2 \in \mathcal{A}$*

$$E(\Delta_1 \cup \Delta_2) = E(\Delta_1) + E(\Delta_2) \ .$$

(SM3) *One has for all $\Delta_1, \Delta_2 \in \mathcal{A}$*

$$E(\Delta_1 \cap \Delta_2) = E(\Delta_1) \cdot E(\Delta_2) \ .$$

Proof. ad (SM1').

ad (SM2').

ad (SM3). □

3.8. Spectral theory of bounded operators

3.8.1 We now apply the foundations of Hilbert space theory built in the previous sections to spectral theory. For the moment we will sacrifice generality and work only with bounded linear operators. The spectral theory of unbounded linear operators will be treated later.

Let us a recall that a linear map $A : \mathcal{H}_1 \to \mathcal{H}_2$ between Hilbert spaces is continuous if and only if it is bounded, i.e. has finite operator norm, and that $\mathfrak{B}(\mathcal{H}_1, \mathcal{H}_2)$ is a Banach space with the operator norm. For the rest of this section, $\mathcal{H}$, $\mathcal{H}_1$, $\mathcal{H}_2, \ldots$ will always denote complex Hilbert spaces and A, B bounded linear operators. We will also now fix the base field to be complex, i.e. $\mathbb{K} = \mathbb{C}$. Last we agree on writing $I_{\mathcal{H}}$ or just I for the identity operator on a Hilbert space $\mathcal{H}$.

Spectrum and Resolvent

3.8.2 Definition Let $A : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator. A complex number λ is then called an *eigenvalue* of A if there exists a nonzero $v \in H$ such that $Av = \lambda v$. For every $\lambda \in \mathbb{C}$ one defines the λ-*eigenspace* of A as

$$\mathrm{Eig}_\lambda(A) = \left\{ v \in H \;\middle|\; Av = \lambda v \right\} \subset \mathcal{H},$$

which is clearly a linear subspace of $\mathcal{H}$.

3.8.3 By definition it is immediately clear that

$$\mathrm{Eig}_\lambda(A) = \ker(A - \lambda),$$

where the λ on the right stands for the operator λI. In other words this means that $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if $A - \lambda$ is not injective.

3.8.4 Definition Let $A \in \mathfrak{B}(\mathcal{H})$. We make the following definitions.

(i) A *regular value* of A is a complex number λ such that $A - \lambda$ is invertible.

(ii) The set of all regular values is the *resolvent* of A, denoted $\varrho(A)$.

(iii) A *spectral value* of A is a complex number λ such that $A - \lambda$ is not invertible.

(iv) The set of all spectral values is the *spectrum* of A, denoted $\sigma(A)$.

(v) The *point* or *eigenspectrum* of A is the set

$$\sigma_{\mathrm{p}}(A) = \left\{ \lambda \in \mathbb{C} \;\middle|\; \ker(A - \lambda) \neq \{0\} \right\}.$$

(vi) An *approximate eigenvalue* of A is a complex number λ for which there exists a sequence of unit vectors $(v_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ such that

$$\lim_{n\to\infty} (A - \lambda) v_n = 0.$$

The set $\sigma_{\mathrm{ap}}(A)$ is the set of all approximate eigenvalues.

3.8.5 Evidently, $\sigma(A) = \mathbb{C} \backslash \varrho(A)$ and $\sigma_{\mathrm{p}}(A) \subset \sigma_{\mathrm{ap}}(A) \subset \sigma(A)$, and these may all be strict inclusions. Note that $A - \lambda$ is bounded for any $\lambda \in \mathbb{C}$, so the open mapping theorem **??** implies that $(A - \lambda)^{-1} \in \mathfrak{B}(\mathcal{H})$ when $\lambda \in \varrho(A)$. We call the map

$$R_\bullet(A) : \varrho(A) \to \mathfrak{B}(\mathcal{H}), \quad R_\lambda(A) = (A - \lambda)^{-1}$$

the *resolvent* of A, not to be confused with the resolvent set $\varrho(A)$. To keep the notation clean, we often briefly write R_λ for $R_\lambda(A)$ and leave implicit that R_λ depends on A.

First, we prove some topological properties of the spectrum and resolvent. Recall the following lemma, which generalizes the geometric series.

3.8.6 Lemma (Carl Neumann) *Let $A \in \mathfrak{B}(\mathcal{H})$. If $\|A\| < 1$, then $I - A$ is invertible,*

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n,$$

and

$$\left\|(I - A)^{-1}\right\| \leqslant \frac{1}{1 - \|A\|}.$$

Proof. Since $\|A\| < 1$ and $\|A^n\| \leqslant \|A\|^n$ by submultiplicativity of the operator norm, we know $\sum_{n=0}^{\infty} \|A^n\| < \infty$. This implies that the family $(A^n)_{n\in\mathbb{N}}$ is absolutely summable, so $\sum_{n=0}^{\infty} A^n$ exists. Furthermore, for every $N \in \mathbb{N}$ we have

$$(I - A) \sum_{n=0}^{N} A^n = \left(\sum_{n=0}^{N} A^n\right)(I - A) = \sum_{n=0}^{N} A^n - \sum_{n=1}^{N+1} A^n = I - A^{N+1},$$

which implies that

$$\lim_{N\to\infty} (I - A) \sum_{n=0}^{N} A^n = \lim_{N\to\infty} \left(\sum_{n=0}^{N} A^n\right)(I - A) = I.$$

By continuity of multiplication in $\mathfrak{B}(\mathcal{H})$ one gets

$$(I - A) \sum_{n=0}^{\infty} A^n = \left(\sum_{n=0}^{\infty} A^n\right)(I - A) = I,$$

which proves that $I - A$ is invertible and $(I - A)^{-1} = \sum_{n=0}^{\infty} A^n$.

Finally, one concludes by the triangle inequality and submultiplicativity of the operator norm

$$\left\|(I - A)^{-1}\right\| \leqslant \sum_{n=0}^{\infty} \|A^n\| \leqslant \sum_{n=0}^{\infty} \|A\|^n = \frac{1}{1 - \|A\|}.$$

□

3.8.7 Proposition *Let $A \in \mathfrak{B}(\mathcal{H})$.*

(i) *For any $\lambda \in \varrho(A)$, one has*

$$B_{\|R_\lambda\|^{-1}}(\lambda) \subset \varrho(A) \ .$$

Hence, $\varrho(A) \subset \mathbb{C}$ is open.

(ii) *The spectrum $\sigma(A)$ is compact and*

$$\sigma(A) \subset \bar{B}_{\|A\|}(0) \ .$$

(iii) *If the complex number λ satisfies $|\lambda| > \|A\|$, then $\lambda \in \varrho(A)$ and*

$$R_\lambda = -\frac{1}{\lambda} - \sum_{n=1}^{\infty} \lambda^{-n-1} A^n \ ,$$

where convergence is with respect to the operator norm.

Proof. *ad (i)*. Fix $\lambda \in \varrho(A)$ and set $r = \|R_\lambda\|^{-1}$. Let $\mu \in B_r(\lambda)$. Then

$$\|(\mu - \lambda)R_\lambda\| = |\mu - \lambda|\,\|R_\lambda\| < 1.$$

Thus, by Lemma 3.8.6, one knows that $I - (\mu - \lambda)R_\lambda$ is invertible. Since $A - \lambda$ is invertible, the composition

$$(A - \lambda)\big(I - (\mu - \lambda)R_\lambda\big) = A - \mu$$

is invertible, which proves that $\mu \in \varrho(A)$. Hence $\varrho(A)$ is open.

ad (ii). Since $\varrho(A)$ is open, the complement $\sigma(A) = \mathbb{C}\backslash\varrho(A)$ is closed. Furthermore, if $|\lambda| > \|A\|$, then $\|\lambda^{-1}A\| < 1$, so $I - \lambda^{-1}A$ and hence $A - \lambda$ are invertible by Lemma 3.8.6. This implies that $\lambda \in \varrho(A)$, so $\sigma(A) \subset \bar{B}_{\|A\|}(0)$. Since $\sigma(A)$ is closed and bounded, it is compact.

ad (iii). If $|\lambda| > \|A\|$, then $I - \lambda^{-1}A$ is invertible by Lemma 3.8.6 and

$$(I - \lambda^{-1}A)^{-1} = \sum_{n=0}^{\infty} \lambda^{-n}A^n.$$

Since $-\lambda(A - \lambda)^{-1} = (I - \lambda^{-1}A)^{-1}$, one obtains

$$R_\lambda = -\frac{1}{\lambda}\sum_{n=0}^{\infty}\lambda^{-n}A^n = -\frac{1}{\lambda} - \sum_{n=1}^{\infty}\lambda^{-n-1}A^n,$$

as desired. □

Next, we prove some algebraic properties of the resolvent. Hereby, $[A, B] = AB - BA$ denotes the commutator of two operators, as usual.

3.8.8 Proposition *Let $A, B \in \mathfrak{B}(\mathcal{H})$. Then the following holds true.*

(i) *The resolvent commutes with the operator which means that*

$$[A, R_\lambda(A)] = 0 \quad \textit{for all } \lambda \in \varrho(A)\ .$$

(ii) *The values of the resolvent commute with each other that is*

$$[R_\lambda(A), R_\mu(A)] = 0 \quad \textit{for all } \lambda, \mu \in \varrho(A)\ .$$

(iii) (First resolvent identity) *For all $\lambda, \mu \in \varrho(A)$*

$$R_\lambda(A) - R_\mu(A) = (\lambda - \mu)R_\lambda(A)R_\mu(A)\ .$$

(iv) (Second resolvent identity) *For all $\lambda \in \varrho(A) \cap \varrho(B)$*

$$R_\lambda(A) - R_\lambda(B) = R_\lambda(A)\,(B - A)\,R_\lambda(B)\ .$$

Proof. ad (i). Obviously $[A, A - \lambda] = 0$, so

$$0 = R_\lambda [A, A - \lambda] R_\lambda = R_\lambda A - A R_\lambda,$$

as desired.

ad (iii). We compute

$$\begin{aligned}(R_\lambda - R_\mu)(A - \mu)(A - \lambda) &= (R_\lambda A - \mu R_\lambda)(A - \lambda) - (A - \lambda) \\ &= (A - \mu) R_\lambda (A - \lambda) - (A - \lambda) \\ &= \lambda - \mu,\end{aligned}$$

where we used part (i) to commute R_λ past A in the second step. Now multiplying both sides with $R_\lambda R_\mu$ from the right yields the desired equality.

ad (ii). For $\lambda = \mu$, one obviously has $[A_\lambda, A_\mu] = 0$. For $\lambda \neq \mu$, one concludes from (ii)

$$R_\mu R_\lambda = \frac{R_\mu - R_\lambda}{\mu - \lambda} = \frac{R_\lambda - R_\mu}{\lambda - \mu} = R_\lambda R_\mu,$$

so $[R_\lambda, R_\mu] = 0$ for $\lambda \neq \mu$ as well.

ad (iv). The last equality follows by

$$R_\lambda(A)\,(B - A)\,R_\lambda(B) = R_\lambda(A)\,\big((B - \lambda) - (A - \lambda)\big)\,R_\lambda(B) = R_\lambda(A) - R_\lambda(B)\ . \qquad \square$$

The resolvent $R_\bullet(A)$ also has some nice analytic properties which we are going to prove next.

3.8.9 Proposition *The resolvent $R_\bullet(A) : \varrho(A) \to \mathfrak{B}(\mathcal{H})$, $\lambda \mapsto R_\lambda$ is continuous and complex differentiable with derivative given by*

$$R_\bullet(A)' : \varrho(A) \to \mathfrak{B}(\mathcal{H}),\ \lambda \mapsto \lim_{\mu \to \lambda} \frac{R_\mu - R_\lambda}{\mu - \lambda} = R_\lambda^2$$

Proof. Fix $\lambda \in \varrho(A)$ and $\varepsilon > 0$. Let $0 < |\mu - \lambda| < \delta$, where

$$\delta = \min\left(\frac{\varepsilon}{2\,\|R_\lambda\|^2}, \frac{1}{2\,\|R_\lambda\|}\right)\ .$$

Note that $\mu \in \varrho(A)$ by Proposition 3.8.7. Moreover, $\|(\mu - \lambda)R_\lambda\| < 1$, so $I - (\mu - \lambda)R_\lambda$ is invertible with norm less than $(1 - \|(\mu - \lambda)R_\lambda\|)^{-1}$ by Lemma 3.8.6. Now observe that the first resolvent identity can be rearranged to

$$R_\mu = R_\lambda [I - (\mu - \lambda) R_\lambda]^{-1}\ .$$

Hence

$$\begin{aligned}\|R_\mu - R_\lambda\| &\leqslant |\mu - \lambda|\,\|R_\mu\|\,\|R_\lambda\| \\ &\leqslant |\mu - \lambda|\,\|R_\lambda\|^2\,\big\|(I - (\mu - \lambda)R_\lambda)^{-1}\big\| \\ &\leqslant \frac{|\mu - \lambda|\,\|R_\lambda\|^2}{1 - \|(\mu - \lambda)R_\lambda\|} \\ &< \frac{\varepsilon/2}{1 - 1/2} = \varepsilon\ .\end{aligned}$$

This proves that $\lambda \mapsto R_\lambda$ is continuous.

As for complex differentiability, we simply use the first resolvent identity and continuity to conclude

$$\lim_{\mu\to\lambda} \frac{R_\mu - R_\lambda}{\mu - \lambda} = \lim_{\mu\to\lambda} R_\mu R_\lambda = R_\lambda^2. \qquad \square$$

3.8.10 Proposition *Let $A \in \mathfrak{B}(\mathcal{H})$. Then $\lambda R_\lambda \to -I$ as $|\lambda| \to \infty$. In particular, $R_\lambda \to 0$ as $|\lambda| \to \infty$.*

Proof. Fix $\varepsilon > 0$. For $|\lambda| > \|A\|$, we have by Proposition 3.8.7 (iii)

$$\lambda R_\lambda = -I - \sum_{n=1}^{\infty} \lambda^{-n} A^n.$$

Since

$$\left\| \sum_{n=1}^{\infty} \lambda^{-n} A^n \right\| \leqslant \frac{\|A\|}{|\lambda| - \|A\|},$$

one sees that $\lambda R_\lambda \to -I$ as $|\lambda| \to \infty$. Similarly, for $|\lambda| > \|A\|$ one has

$$\|R_\lambda\| \leqslant \frac{1}{|\lambda|} + \frac{1}{|\lambda|} \sum_{n=1}^{\infty} \left\|\lambda^{-n} A^n\right\| \leqslant \frac{1}{|\lambda|} + \frac{1}{|\lambda|} \frac{\|A\|}{|\lambda| - \|A\|},$$

which shows that $R_\lambda \to 0$ as $|\lambda| \to \infty$. $\square$

3.8.11 Proposition *For all $v, w \in \mathcal{H}$, the map*

$$\langle R_\bullet(A)v, w\rangle : \ \varrho(A) \to \mathbb{C}, \ \lambda \mapsto \langle R_\lambda v, w\rangle$$

is holomorphic with derivative

$$\langle R_\bullet(A)v, w\rangle' : \ \varrho(A) \to \mathbb{C}, \ \lambda \mapsto \langle R_\lambda^2 v, w\rangle.$$

Proof. Given $\lambda \in \varrho(A)$, we compute

$$\lim_{\mu\to\lambda} \frac{\langle R_\mu v, w\rangle - \langle R_\lambda v, w\rangle}{\mu - \lambda} = \lim_{\mu\to\lambda} \frac{\langle (\mu - \lambda) R_\mu R_\lambda v, w\rangle}{\mu - \lambda} = \lim_{\mu\to\lambda} \langle R_\mu R_\lambda v, w\rangle = \langle R_\lambda^2 v, w\rangle,$$

where we have used the first resolvent identity in the first step and continuity of the inner product in the last. $\square$

3.8.12 Proposition *The spectrum of an operator $A \in \mathfrak{B}(\mathcal{H})$ is nonempty.*

Proof. Suppose $\sigma(A) = \varnothing$, hence $\varrho(A) = \mathbb{C}$. The map

$$\mathbb{C} \to \mathbb{C}, \ \lambda \mapsto \langle R_\lambda v, w\rangle$$

then is entire for every $v, w \in \mathcal{H}$. Furthermore, one has for $\|v\|, \|w\| \leqslant 1$

$$|\langle R_\lambda v, w\rangle| \leqslant \|R_\lambda\| \, \|v\| \, \|w\| \leqslant \|R_\lambda\| \ .$$

Since $\lambda \mapsto \|R_\lambda\|$ is continuous and $\|R_\lambda\| \to 0$ as $|\lambda| \to \infty$, one sees that $\|R_\lambda\|$ is bounded. Hence $\langle R_\bullet v, w\rangle$ is a bounded entire function, which by Liouville's theorem implies that it is zero for every pair $v, w \in \mathcal{H}$ with $\|v\| = \|w\| = 1$. This entails that $R_\lambda = 0$ for every $\lambda \in \mathbb{C}$, which is a contradiction to R_λ being invertible. Hence $\sigma(A) \neq \varnothing$. $\square$

3.9. Unbounded linear operators

3.9.1 In this section let V, W always denote Banach spaces over the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. The symbols $\mathcal{H}$, $\mathcal{H}_1$, ... will always stand for Hilbert spaces over $\mathbb{K}$.

3.9.2 Definition By an *unbounded* $\mathbb{K}$*-linear operator* or shortly by an *unbounded operator* from V to W we understand a linear map $A : \mathrm{Dom}(A) \to \mathrm{W}$ defined on a $\mathbb{K}$-linear subspace $\mathrm{Dom}(A) \subset \mathrm{V}$. As usual, $\mathrm{Dom}(A)$ is called the *domain* of the operator A. The space of unbounded $\mathbb{K}$-linear operators from V to W will be denoted $\mathfrak{L}_{\mathbb{K}}(V, W)$ or just $\mathfrak{L}(V, W)$.

3.9.3 Remark In this work, the term "unbounded" is meant in the sense of "not necessarily bounded". Sometimes we just say *linear operator* or even only *operator* instead of "unbounded linear operator".

3.9.4 Observe that besides the domain $\mathrm{Dom}(A)$ of an unbounded operator $A \in \mathfrak{L}(\mathrm{V}, \mathrm{W})$ the *kernel*

$$\mathrm{Ker}(A) = \big\{ v \in \mathrm{V} \bigm| Av = 0 \big\} \subset \mathrm{V} \ ,$$

the *image*

$$\mathrm{Im}(A) = \big\{ w \in \mathrm{W} \bigm| \exists v \in \mathrm{Dom}(A) : w = Av \big\} \subset \mathrm{W} \ ,$$

and the *graph*

$$\mathrm{Gr}(A) = \big\{ (v, w) \in \mathrm{Dom}(A) \times \mathrm{W} \bigm| w = Av \big\} \subset \mathrm{V} \times \mathrm{W}$$

of A are all linear subspaces. We will frequently make use of this.

3.9.5 Definition An unbounded operator $A \in \mathfrak{L}(\mathrm{V}, \mathrm{W})$ is called *densely defined* if $\mathrm{Dom}(A)$ is dense in V, and *closed* if the graph $\mathrm{Gr}(A)$ is closed in $\mathrm{V} \times \mathrm{W}$. The operator $A \in \mathfrak{L}(V, W)$ is called *closable* if the closure $\overline{\mathrm{Gr}(A)}$ is the graph of an unbounded operator from V to W.

An operator $A \in \mathfrak{L}(V, W)$ is called an *extension* of $B \in \mathfrak{L}(V, W)$ if $\mathrm{Gr}(B) \subset \mathrm{Gr}(A)$. One writes in this situation $B \subset A$.

II.4. C*-Algebras

4.1. Elementary Definitions and Properties

A C^*-algebra is a Banach algebra with additional structure. We recall the definition of a Banach algebra below.

4.1.1 Definition A *Banach algebra* is a Banach space A with an associative, bilinear multiplication operation $A \times A \to A$, $(a,b) \mapsto ab$ which is *submultiplicative* with respect to the norm:

$$\|ab\| \leqslant \|a\| \, \|b\| \,, \qquad \forall\, a, b \in A. \tag{4.1.1}$$

We say A is *unital* if there exists a *unit* $1 \in A$ satisfying $\|1\| = 1$ and

$$1a = a1 = a, \qquad \forall\, a \in A. \tag{4.1.2}$$

In a unital Banach algebra we can speak of inverses of elements, but not every element of a Banach algebra is invertible. By an inverse we mean a two-sided inverse unless otherwise specified. We write $A^\times$ for the set of invertible elements in A.

We state without proof some obvious facts about Banach algebras.

4.1.2 Proposition *Let A be a Banach algebra.*

1. *For all $a \in A$, we have $0a = a0 = 0$.*

If A is unital, then the following hold.

2. *The unit is unique.*
3. *Inverses are unique.*
4. *The additive identity is not invertible.*
5. *The multiplicative identity is its own inverse.*
6. *If $a, b \in A^\times$, then $ab \in A^\times$ and $(ab)^{-1} = b^{-1}a^{-1}$.*
7. *If $a \in A$ has a left inverse and a right inverse, then these inverses are equal, so $a \in A^\times$.*

4.1.3 Proposition *The multiplication operation on a Banach algebra is continuous.*

Proof. Take sequences (a_n) and (b_n) in A such that $a_n \to a$ and $b_n \to b$. Using the triangle inequality and (4.1.1),

$$\begin{aligned} \|ab - a_n b_n\| &\leqslant \|ab - ab_n\| + \|ab_n - a_n b_n\| \\ &\leqslant \|a\| \, \|b - b_n\| + \|a - a_n\| \, \|b_n\| \end{aligned} \tag{4.1.3}$$

This manifestly approaches zero, so we conclude that $a_n b_n \to ab$. □

4.1.4 Proposition *If A is a Banach algebra and $a \in A^\times$, then*

$$\|a\|^{-1} \leqslant \left\|a^{-1}\right\|. \tag{4.1.4}$$

Proof. By (4.1.1), we have

$$1 = \|1\| = \left\|a^{-1}a\right\| \leqslant \left\|a^{-1}\right\| \|a\|. \tag{4.1.5}$$

The result follows by dividing by $\|a\|$. □

4.1.5 Proposition *Inversion $A^\times \to A^\times$, $a \mapsto a^{-1}$ is continuous.*

Proof. Take a sequence (a_n) in $A^\times$ such that $a_n \to a \in A^\times$. We compute

$$\begin{aligned} \left\|a^{-1} - a_n^{-1}\right\| &= \left\|a^{-1}(a_n - a)a_n^{-1}\right\| \\ &\leqslant \left\|a^{-1}\right\| \|a_n - a\| \left\|a_n^{-1}\right\| \\ &\leqslant \left\|a^{-1}\right\| \|a_n - a\| \left\|a^{-1}\right\| + \left\|a^{-1}\right\| \|a_n - a\| \left\|a^{-1} - a_n^{-1}\right\|. \end{aligned} \tag{4.1.6}$$

Moving the rightmost term to the other side yields

$$\left(1 - \left\|a^{-1}\right\| \|a_n - a\|\right) \left\|a^{-1} - a_n^{-1}\right\| \leqslant \|a_n - a\| \left\|a^{-1}\right\|^2 \tag{4.1.7}$$

For large enough n, the term in parentheses is nonzero, and we may divide by it, yielding

$$\left\|a^{-1} - a_n^{-1}\right\| \leqslant \frac{\|a_n - a\| \left\|a^{-1}\right\|^2}{1 - \left\|a^{-1}\right\| \|a_n - a\|}. \tag{4.1.8}$$

The left hand side manifestly approaches zero, so we conclude $a_n^{-1} \to a^{-1}$. □

4.1.6 Definition A ***C****-*algebra* is a Banach algebra A with an antilinear *star operation* $A \to A$, $a \mapsto a^*$ satisfying

(i) *involutivity:* $a^{**} = a$ for all $a \in A$,

(ii) *contravariance:* $(ab)^* = b^* a^*$ for all $a, b \in A$,

(iii) the ***C****-*property:* $\|a^* a\| = \|a\|^2$ for all $a \in A$.

An element $a \in A$ is *self-adjoint* if $a^* = a$.

A subset $B \subset A$ is a ***C****-*subalgebra* of A if B is a C^*-algebra under the restrictions to B of all operations defined on A. Equivalently, B must be a topologically closed subset which is closed under all the operations on A. Topological closedness is equivalent to completeness. We say B

is a *unital* C^*-subalgebra of A if B is a unital C^*-algebra and the unit in B is the same as the unit in A.[1]

If S is a subset of a C^*-algebra A, then the intersection of all C^*-subalgebras of A containing S is a C^*-subalgebra, called the ***C^*-subalgebra generated by S***.

If A and B are C^*-algebras, a *$*$-homomorphism* from A to B is a map $\pi : A \to B$ respecting the algebraic operations on A and B. More precisely, π is a linear map satisfying

$$\begin{aligned} \pi(ab) &= \pi(a)\pi(b) \\ \pi(a^*) &= \pi(a)^* \end{aligned} \tag{4.1.9}$$

for all $a, b \in A$. Notice that we do not require π to be continuous; we will show later that this is automatically so. If A and B are unital, we say π is a *unital $*$-homomorphism* if π is a $*$-homomorphism and $\pi(1) = 1$. The terms $*$-isomorphism and $*$-automorphism will be used in the natural sense.

4.1.7 Proposition *If A and B are C^*-algebras and $\pi : A \to B$ is a $*$-isomorphism, then $\pi^{-1} : B \to A$ is a $*$-isomorphism. If π is unital, then so is π^{-1}.*

Proof. We know from linear algebra that π^{-1} is a linear map. Given $a, b \in B$, let $a', b' \in A$ such that $\pi(a') = a$ and $\pi(b') = b$. Then

$$\begin{aligned} \pi^{-1}(ab) &= \pi^{-1}(\pi(a')\pi(b')) = \pi^{-1}(\pi(a'b')) = a'b' = \pi^{-1}(a)\pi^{-1}(b) \\ \pi^{-1}(a^*) &= \pi^{-1}(\pi(a')^*) = \pi^{-1}(\pi(a'^*)) = a'^* = \pi^{-1}(a)^*. \end{aligned} \tag{4.1.10}$$

If π is unital, then $\pi(1) = 1$, so $\pi^{-1}(1) = 1$ as well. $\square$

We will study $*$-homomorphisms more in a later section, for now focusing on properties of elements of C^*-algebras.

4.1.8 Definition If A is a C^*-algebra, a subset $B \subset A$ is a *Banach subalgebra* of A if it is a Banach algebra under the restrictions of all operations defined on A. Equivalently, B must be a topologically closed subset which is closed under all the operations on A. Topological closedness is equivalent to completeness. We say B is a *unital* Banach subalgebra of A if B is a unital C^*-algebra and the unit in B is the same as the unit in A. I pray I never have to consider a case where B is a Banach subalgebra of A and has a different unit from A.

If S is a subset of a Banach algebra or C^*-algebra A, then the intersection of all subalgebras of A containing S is a (Banach or C^*) subalgebra, called the *subalgebra generated by **S***.

A (unital) C^*-subalgebra is a (unital) Banach subalgebra $B \subset A$ which is closed under the star operation.

[1]The unitization of a unital C^*-algebra (discussed in a later section) provides an instance where we have a unital C^*-algebra and a C^*-subalgebra which has a different unit.

4.1.9 Definition Let A and B be C^*-algebras. A *$*$-homomorphism* is a linear map $\pi : A \to B$ such that

$$\begin{aligned} \pi(ab) &= \pi(a)\pi(b) \\ \pi(a^*) &= \pi(a)^* \end{aligned} \tag{4.1.11}$$

for all $a, b \in A$. In other words, π respects all algebraic operations on A and B. Note that we do not require π to be continuous. We will use the terms $*$-isomorphism and $*$-automorphism in the natural way.

We will study $*$-homomorphisms more in a later section. For now, let us establish a few more basic properties of C^*-algebras.

4.1.10 Proposition *If A is a C^*-algebra, then 0 is self-adjoint. If A is unital, then 1 is self-adjoint as well.*

Proof. For all $a \in A$, we have

$$0^* + a = 0^* + a^{**} = (0 + a^*)^* = a^{**} = a. \tag{4.1.12}$$

Hence, $0^* = 0$ by uniqueness of the additive identity. Furthermore,

$$1^*a = 1^*a^{**} = (a^*1)^* = a^{**} = a, \tag{4.1.13}$$

from which it follows that 1 is self-adjoint by uniqueness of the multiplicative identity. □

4.1.11 Proposition *Every $a \in A$ has a unique expression in the form $a = a_1 + ia_2$, where a_1 and a_2 are self-adjoint.*

Proof. This is evident upon setting $a_1 = (a + a^*)/2$ and $a_2 = (a - a^*)/2i$. If $a = a_1' + ia_2'$ for some self-adjoint a_1', a_2', then $a_1 - a_1' = i(a_2' - a_2)$, which can be self-adjoint only if it is zero. □

4.1.12 Proposition *Let A be a nontrivial C^*-algebra. If there exists $1 \in A$ satisfying*

$$1a = a1 = a, \qquad \forall\, a \in A, \tag{4.1.14}$$

then $\|1\| = 1$, i.e. A is unital.

Proof. Setting $a = 1$ in the C^*-property yields

$$\|1\| = \|1^*1\| = \|1\|^2. \tag{4.1.15}$$

Hence, $\|1\| = 0$ or $\|1\| = 1$. If $\|1\| = 0$, then $1 = 0$, so that $A = \{0\}$. Since A is nontrivial by hypothesis, this cannot be the case, so $\|1\| = 1$. □

4.1.13 Proposition *Let A be a unital C^*-algebra and let $a \in A^\times$. Then $a^* \in A^\times$ and*

$$(a^*)^{-1} = (a^{-1})^*. \tag{4.1.16}$$

Proof. We compute

$$a^*(a^{-1})^* = (a^{-1}a)^* = 1 = (aa^{-1})^* = (a^{-1})^*a^*, \tag{4.1.17}$$

which proves the result. □

4.1.14 Proposition *If A is a C^*-algebra and $a \in A$, then*

$$\|a^*\| = \|a\| . \tag{4.1.18}$$

Proof. The conclusion is trivial if $a = 0$, so suppose $a \neq 0$. The C^*-property and submultiplicativity yield

$$\|a\|^2 = \|a^* a\| \leqslant \|a\| \, \|a^*\| . \tag{4.1.19}$$

Dividing by $\|a\|$ yields $\|a\| \leqslant \|a^*\|$. Applying this result to a^* yields $\|a^*\| \leqslant \|a^{**}\| = \|a\|$. □

4.1.15 Corollary *The star operation $A \to A$, $a \mapsto a^*$ is continuous.*

Proof. This is immediate from Proposition 4.1.14. □

We conclude this section with several examples.

4.1.16 Example The complex numbers $\mathbb{C}$ give a fairly trivial unital C^*-algebra.

4.1.17 Example The bounded linear operators $\mathfrak{B}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$ are the prototypical example of a C^*-algebra. The star operation is given by the adjoint. Note that this is, of course, a unital C^*-algebra.

4.1.18 Example In a similar vein to the previous example, the set $M_n(\mathbb{C})$ of $n \times n$ matrices with complex entries is a C^*-algebra, where the star operation is given by Hermitian conjugation.

4.1.19 Example Let X be a compact Hausdorff space and let $C(X)$ be the space of continuous functions $X \to \mathbb{C}$. This is a unital C^*-algebra with the norm given by the supremum norm and the star operation given by complex conjugation. We note that

$$\|fg\| = \sup_{x \in X} |fg| \leqslant \sup_{x \in X} |f| \cdot \sup_{x \in X} |g| = \|f\| \, \|g\| \tag{4.1.20}$$

and

$$\|f^* f\| = \sup_{x \in X} |f|^2 = \left(\sup_{x \in X} |f| \right)^2 = \|f\|^2 , \tag{4.1.21}$$

so that this satisfies the nontrivial properties of a C^*-algebra.

4.1.20 Example Let X be a locally compact Hausdorff space and let $C_0(X)$ be the space of continuous functions $f : X \to \mathbb{C}$ which *vanish at infinity*, meaning for every $\varepsilon > 0$ there exists a compact $K \subset X$ such that $|f(x)| < \varepsilon$ for $x \notin K$. This space is a C^*-algebra with the supremum norm and the star operation given by complex conjugation. If X is not compact, then this is a non-unital C^*-algebra.

Finite Direct Sums

Let $A_1, \ldots, A_n$ be a finite collection of Banach algebras. We define the *direct sum*

$$\bigoplus_{i=1}^{n} A_i := \{(a_1, \ldots, a_n) : a_i \in A_i\} \tag{4.1.22}$$

with addition and multiplication defined componentwise. If $A_1, \ldots, A_n$ are C^*-algebras, define the star operation on A componentwise as well. Finally, set

$$\|(a_1, \ldots, a_n)\| = \max(\|a_1\|, \ldots, \|a_n\|). \tag{4.1.23}$$

We want to show that the direct sum so defined is a Banach algebra. It is easy to check that the norm above is indeed a norm using the properties of the max and the definition of the norms on A_i. Submultiplicativity also follows easily from submultiplicativity of the norms on the A_i. If $(a_{1,k}, \ldots, a_{n,k})_{k\in\mathbb{N}}$ is a Cauchy sequence in the direct sum, then each sequence $(a_{i,k})_{k\in\mathbb{N}}$ is Cauchy, hence convergent, in A_i for $i = 1, \ldots, n$. If $a_{i,k} \to a_i$ for each i, then $(a_{1,k}, \ldots, a_{n,k}) \to (a_1, \ldots, a_n)$ by definition of the norm on the direct sum. Thus, $\bigoplus_{i=1}^{n} A_i$ is complete, and is therefore a Banach algebra. If each A_i is a C^*-algebra, it is again easy to check that $\bigoplus_{i=1}^{n} A_i$ is a C^*-algebra using the properties of the C^*-algebras A_i.

We may also define the direct sum of a sequence of Banach algebras or C^*-algebras $\{A_n\}_{n\in\mathbb{N}}$. We define

$$\bigoplus_{n=1}^{\infty} A_n := \left\{(a_n)_{n\in\mathbb{N}} : a_n \in A_n \text{ and } \lim_{n\to\infty} \|a_n\| = 0\right\} \tag{4.1.24}$$

Again, the algebraic operations are defined componentwise and the norm is defined by

$$\|(a_n)_{n\in\mathbb{N}}\| = \max_{n\in\mathbb{N}} \|a_n\|. \tag{4.1.25}$$

The definition of $\bigoplus_{n=1}^{\infty} A_n$ ensures that the max exists.

One easily checks that this satisfies all algebraic properties of a Banach or C^*-algebra, but completeness is more subtle. If $\mathbf{a}_k = (a_{n,k})_{n\in\mathbb{N}} \in \bigoplus_{n=1}^{\infty} A_n$ and $(\mathbf{a}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence in the direct sum, then it follows in the same way as before that the sequence $(a_{n,k})_{k\in\mathbb{N}}$ is Cauchy in A_n, hence convergent with limit $a_{n,k} \to a_n \in A_n$. We must show that $\lim_{n\to\infty} \|a_n\| = 0$. For any $n, k, K \in \mathbb{N}$, we have

$$\|a_n\| \leqslant \|a_n - a_{n,k}\| + \|a_{n,k} - a_{n,K}\| + \|a_{n,K}\| \leqslant \|a_n - a_{n,k}\| + \|\mathbf{a}_k - \mathbf{a}_K\| + \|a_{n,K}\| \tag{4.1.26}$$

Fix $\varepsilon > 0$. Since $(\mathbf{a}_k)_{k\in\mathbb{N}}$ is Cauchy, we may choose $K \in \mathbb{N}$ such that $k, \ell \geqslant K$ implies $\|\mathbf{a}_k - \mathbf{a}_\ell\| < \varepsilon/3$. We may choose $N \in \mathbb{N}$ such that $n \geqslant N$ implies $\|a_{n,K}\| < \varepsilon/3$. Finally, for any $n \geqslant N$, we may choose $k \geqslant K$ such that $\|a_n - a_{n,k}\| < \varepsilon/3$. Thus, for $n \geqslant N$, we have $\|a_n\| < \varepsilon$, so $\lim_{n\to\infty} \|a_n\| = 0$.

Finally, we show that $\mathbf{a}_k \to \mathbf{a} := (a_n)_{n\in\mathbb{N}}$. Fix $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $\|a_n\| < \varepsilon/2$ if $n \geqslant N$. Choose $K \in \mathbb{N}$ such that $k, \ell \geqslant K$ and $n < N$ implies

$$\|\mathbf{a}_k - \mathbf{a}_\ell\| < \frac{\varepsilon}{2} \quad \text{and} \quad \|a_n - a_{n,k}\| < \varepsilon. \tag{4.1.27}$$

Then for $k \geqslant K$, we have

$$\|\mathbf{a} - \mathbf{a}_k\| < \max\left(\varepsilon, \max_{n \geqslant N} \|a_n - a_{n,k}\|\right). \tag{4.1.28}$$

But for $n \geqslant N$ we may choose ℓ large enough such that

$$\|a_n - a_{n,k}\| \leqslant \|a_n - a_{n,\ell}\| + \|\mathbf{a}_\ell - \mathbf{a}_k\| < \varepsilon. \tag{4.1.29}$$

Thus, $\|\mathbf{a} - \mathbf{a}_k\| < \varepsilon$, so $\mathbf{a}_k \to \mathbf{a}$. This proves that the direct sum is complete.

Both of these constructions come equipped with natural algebra homomorphisms $\iota_i : A_i \to \bigoplus A_n$ satisfying the following universal property. If B is another Banach or C^*-algebra with algebra homomorphisms $f_i : A_i \to B$, then there exists a unique algebra homomorphism $f : \bigoplus A_n \to B$ such that the diagram

$$\begin{tikzcd} \bigoplus A_i \arrow[r, "f"] & B \\ A_i \arrow[u, "\iota_i"] \arrow[ur, "f_i"'] & \end{tikzcd} \tag{4.1.30}$$

commutes. We define

$$f(a) = \sum f_i(\pi_i(a)) \tag{4.1.31}$$

4.2. Spectral Theory

4.2.1. Spectral Theory in Banach Algebras

Throughout this section, let A be a unital Banach algebra. We actually do not need to require A to be a C^*-algebra for the foundations of spectral theory, but the existence of a unit is essential. We will adopt the shorthand of writing λ for $\lambda 1$ for all $\lambda \in \mathbb{C}$.

4.2.2 Definition Given $a \in A$, we define the *resolvent set* of a as

$$\rho(a) = \left\{\lambda \in \mathbb{C} : \lambda - a \in A^\times\right\} \tag{4.2.1}$$

An element of $\rho(a)$ is called a *regular value* of a. If $\rho(a) \neq \varnothing$, the map $r_a : \rho(a) \to A$ defined as

$$r_a(\lambda) = (\lambda - a)^{-1} \tag{4.2.2}$$

is called the *resolvent* of a.

Likewise, we define the *spectrum* of a as

$$\sigma(a) = \left\{\lambda \in \mathbb{C} : \lambda - a \notin A^\times\right\} = \mathbb{C} \backslash \rho(a). \tag{4.2.3}$$

An element of $\sigma(a)$ is called a *spectral value* of a. If $\sigma(a) \neq \varnothing$, we define the *spectral radius* of a as

$$r(a) = \sup_{\lambda \in \sigma(a)} |\lambda|. \tag{4.2.4}$$

A priori we do not know that $\rho(a)$ or $\sigma(a)$ is nonempty. The following exposition will establish that in fact both of them are nonempty, so the resolvent $r_a(\lambda)$ and the spectral radius $r(a)$ are always defined.

4.2.3 Theorem *Let $a \in A$. If $\lambda \in \mathbb{C}$ such that $\|a\| < |\lambda|$, then the* Neumann series $\sum_{n=0}^{\infty}(a/\lambda)^n$ *converges, $\lambda \in \rho(a)$, and*

$$r_a(\lambda) = \frac{1}{\lambda}\sum_{n=0}^{\infty}\left(\frac{a}{\lambda}\right)^n. \tag{4.2.5}$$

Furthermore,

$$\|r_a(\lambda)\| \leqslant \frac{1}{|\lambda| - \|a\|}. \tag{4.2.6}$$

Proof. We begin by showing that the sequence of partial sums is a Cauchy sequence. Given $M, N \in \mathbb{N}$ with $M < N$, we have

$$\left\|\sum_{n=0}^{N}\left(\frac{a}{\lambda}\right)^n - \sum_{n=0}^{M}\left(\frac{a}{\lambda}\right)^n\right\| = \left\|\sum_{n=M+1}^{N}\left(\frac{a}{\lambda}\right)^n\right\| \leqslant \sum_{n=M+1}^{N}\left\|\left(\frac{a}{\lambda}\right)^n\right\| \leqslant \sum_{n=M+1}^{N}\left(\frac{\|a\|}{|\lambda|}\right)^n, \tag{4.2.7}$$

where we have used submultiplicativity in the last step. Since $\|a\|/|\lambda| < 1$, the rightmost expression can be made arbitrarily small by taking M and N to be large. Thus, the sequence of partial sums is Cauchy, hence convergent, since A is complete.

Now, for any $N \in \mathbb{N}$, we have

$$(\lambda - a)\left[\frac{1}{\lambda}\sum_{n=0}^{N}\left(\frac{a}{\lambda}\right)^n\right] = \left[\frac{1}{\lambda}\sum_{n=0}^{N}\left(\frac{a}{\lambda}\right)^n\right](\lambda - a) = 1 - \left(\frac{a}{\lambda}\right)^{N+1} \tag{4.2.8}$$

Submultiplicativity and $\|a/\lambda\| < 1$ imply $(a/\lambda)^n \to 0$. Thus, taking the limit as $N \to \infty$ of the above line yields

$$(\lambda - a)\left[\frac{1}{\lambda}\sum_{n=0}^{\infty}\left(\frac{a}{\lambda}\right)^n\right] = \left[\frac{1}{\lambda}\sum_{n=0}^{\infty}\left(\frac{a}{\lambda}\right)^n\right](\lambda - a) = 1, \tag{4.2.9}$$

as desired.

Finally, we note that for $N \in \mathbb{N}$, using the formula for a geometric series yields

$$\left\|\frac{1}{\lambda}\sum_{n=0}^{N}\left(\frac{a}{\lambda}\right)^n\right\| \leqslant \frac{1}{|\lambda|}\sum_{n=0}^{N}\left(\frac{\|a\|}{|\lambda|}\right)^n \leqslant \frac{1}{|\lambda|}\frac{1}{1 - \|a\|/|\lambda|} = \frac{1}{|\lambda| - \|a\|}. \tag{4.2.10}$$

Taking the limit as $N \to \infty$ yields (4.2.6). □

The following corollary rephrases some of the key points of the above theorem.

4.2.4 Corollary *Given $a \in A$, the resolvent set $\rho(a)$ is nonempty and*

$$r(a) \leqslant \|a\|. \tag{4.2.11}$$

4.2.5 Corollary *Given $a \in A$, the resolvent $r_a : \rho(a) \to A$ is continuous.*

Proof. Since $\rho(a) \neq \varnothing$, the resolvent is defined. Continuity then follows from continuity of addition, scalar multiplication, and inversion. □

4.2.6 Corollary *Let $a \in A$ and $\lambda_0 \in \rho(a)$. If $|\lambda - \lambda_0| < \|r_a(\lambda_0)\|^{-1}$, then $\lambda \in \rho(a)$ and*

$$r_a(\lambda) = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n r_a(\lambda_0)^{n+1}. \tag{4.2.12}$$

In particular, it follows that

$$B_{\|r_a(\lambda_0)\|^{-1}}(\lambda_0) \subset \rho(a) \tag{4.2.13}$$

for all $\lambda_0 \in \rho(a)$, so $\rho(a)$ is open in $\mathbb{C}$.

Proof. Let $|\lambda - \lambda_0| < \|r_a(\lambda_0)\|^{-1}$. Then $\|(\lambda_0 - \lambda)r_a(\lambda_0)\| < 1$, so $1 - (\lambda_0 - \lambda)r_a(\lambda_0)$ is invertible by Theorem 4.2.3. Since $\lambda_0 - a$ is invertible, the product

$$(\lambda_0 - a)[1 - (\lambda_0 - \lambda)r_a(\lambda_0)] = \lambda_0 - a - (\lambda_0 - \lambda) = \lambda - a \tag{4.2.14}$$

is also invertible, so $\lambda \in \rho(a)$. Furthermore, using the Neumann series for the inverse of $1 - (\lambda_0 - \lambda)r_a(\lambda_0)$, we obtain

$$\begin{aligned} r_a(\lambda) &= [1 - (\lambda_0 - \lambda)r_a(\lambda_0)]^{-1} r_a(\lambda_0) \\ &= \left[\sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n r_a(\lambda_0)^n\right] r_a(\lambda_0) \\ &= \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n r_a(\lambda_0)^{n+1}, \end{aligned} \tag{4.2.15}$$

as desired. □

In a similar vein, we have the following Corollary.

4.2.7 Corollary *The set $A^\times$ is open in A.*

Proof. Let $a \in A^\times$ and let $b \in A$ such that $\|b - a\| < \|a^{-1}\|^{-1}$. This implies that

$$\|1 - ba^{-1}\| \leqslant \|a - b\| \, \|a^{-1}\| < 1, \tag{4.2.16}$$

so $1 - (1 - ba^{-1}) = ba^{-1}$ is invertible, which implies b is invertible. □

Having studied a few properties of the resolvent set, we now turn to the spectrum. In particular, we want to show that the spectrum is nonempty. We will be aided by a few algebraic properties of the resolvent. We denote the commutator of two elements $a, b \in A$ by $[a, b] = ab - ba$.

4.2.8 Lemma *Let $a \in A$. For all $\lambda, \mu \in \rho(a)$, we have:*

(i) $[a, r_a(\lambda)] = 0$,

(ii) $r_a(\mu) - r_a(\lambda) = (\lambda - \mu) r_a(\mu) r_a(\lambda)$

(iii) $[r_a(\lambda), r_a(\mu)] = 0$.

Proof. (i). It is clear that $[\lambda - a, a] = 0$. So, since $r_a(\lambda) = (\lambda - a)^{-1}$, we have

$$0 = r_a(\lambda)[\lambda - a, a]r_a(\lambda) = r_a(\lambda)a - ar_a(\lambda) = [r_a(\lambda), a]. \tag{4.2.17}$$

(ii). We compute

$$\begin{aligned} [r_a(\mu) - r_a(\lambda)]\,(\lambda - a)(\mu - a) &= r_a(\mu)(\lambda - a)(\mu - a) - (\mu - a) \\ &= (\lambda - a) - (\mu - a) \\ &= \lambda - \mu. \end{aligned} \tag{4.2.18}$$

We used the fact that $[r_a(\mu), \lambda - a] = 0$ in the second step. Multiplying by $r_a(\mu)r_a(\lambda)$ on the left now yields the desired result.

(iii). If $\lambda = \mu$, the result is trivial. If $\lambda \neq \mu$, we may divide both sides of (ii) by $\lambda - \mu$ to obtain

$$r_a(\mu)r_a(\lambda) = \frac{r_a(\mu) - r_a(\lambda)}{\lambda - \mu}. \tag{4.2.19}$$

The right hand side is invariant under exchange of μ and λ, so the result follows. □

4.2.9 Lemma *Let $a \in A$. If f is in the continuous dual A^*, then $f \circ r_a : \rho(a) \to \mathbb{C}$ is holomorphic*

Proof. Let $\lambda \in \rho(a)$. For any $\mu \in \rho(a)$, $\mu \neq \lambda$, we have

$$\frac{f(r_a(\mu)) - f(r_a(\lambda))}{\mu - \lambda} = f\left(\frac{r_a(\mu) - r_a(\lambda)}{\mu - \lambda}\right) = -f(r_a(\mu)r_a(\lambda)). \tag{4.2.20}$$

Since f and r_a are continuous, the limit of the above as $\mu \to \lambda$ exists and is

$$(f \circ r_a)'(\lambda) = -f\big(r_a(\lambda)^2\big). \tag{4.2.21}$$

This proves that $f \circ r_a$ is holomorphic. □

4.2.10 Corollary *Given $a \in A$, the spectrum $\sigma(a)$ is nonempty and compact.*

Proof. We know $\sigma(a)$ is closed and bounded since $\rho(a)$ is open and $r(a) \leqslant \|a\|$, so it just remains to show that $\sigma(a)$ is nonempty. If $\sigma(a) = \varnothing$, then $\rho(a) = \mathbb{C}$, so $f \circ r_a$ is entire for all $f \in A^*$. Since f is bounded, we have

$$|(f \circ r_a)(\lambda)| \leqslant \|f\|\,\|r_a(\lambda)\|. \tag{4.2.22}$$

Furthermore, since $\lambda \mapsto \|r_a(\lambda)\|$ is continuous, it is bounded by a constant for $|\lambda| \leqslant 1 + \|a\|$ by the extreme value theorem. For $|\lambda| > 1 + \|a\|$, we have a bound from the Neumann series:

$$\|r_a(\lambda)\| \leqslant \frac{1}{|\lambda| - \|a\|} \leqslant 1. \tag{4.2.23}$$

Thus, $f \circ r_a$ is bounded and entire, so by Liouville's theorem it is constant. But if $\mu \neq \lambda$, then $r_a(\mu) \neq r_a(\lambda)$, for otherwise we would have

$$0 = r_a(\mu) - r_a(\lambda) = (\lambda - \mu)r_a(\mu)r_a(\lambda), \tag{4.2.24}$$

which would imply that $r_a(\mu) = r_a(\lambda) = 0$, but 0 is not invertible. By the Hahn-Banach theorem, there must be some $f \in A^*$ such that $f \circ r_a$ is not constant, which is a contradiction. Therefore $\rho(a) \neq \mathbb{C}$. □

In fact, using Lemma 4.2.9, we can say exactly what the spectral radius $\rho(a)$ is.

4.2.11 Theorem *Let $a \in A$. The spectral radius is given by*

$$r(a) = \lim_{n\to\infty} \|a^n\|^{1/n}\,, \tag{4.2.25}$$

where the limit on the right hand side is guaranteed to exist.

Proof. The result is trivial if $a = 0$, so assume $a \neq 0$. We will show that

$$\limsup \|a^n\|^{1/n} \leqslant r(a) \leqslant \liminf \|a^n\|^{1/n}\,, \tag{4.2.26}$$

which immediately yields the result.

Suppose $\lambda \in \sigma(a)$. If $\lambda^n \in \rho(a^n)$ for some $n \in \mathbb{N}$, then

$$\lambda^n - a^n = (\lambda - a)\big(\lambda^{n-1} + \lambda^{n-2}a + \cdots + \lambda a^{n-2} + a^{n-1}\big) \tag{4.2.27}$$

is invertible. Let b be the rightmost term in parentheses and note that b commutes with $\lambda - a$. But then

$$(\lambda - a)b(\lambda^n - a^n)^{-1} = 1 = (\lambda^n - a^n)^{-1}b(\lambda - a), \tag{4.2.28}$$

so $\lambda - a$ has a left inverse and a right inverse. Hence $\lambda - a$ has a two-sided inverse, so $\lambda \in \rho(a)$, which is a contradiction. Therefore $\lambda^n \in \sigma(a^n)$. Since $\rho(a^n) \leqslant \|a^n\|$, we see that

$$|\lambda| = (|\lambda|^n)^{1/n} \leqslant \|a^n\|^{1/n}\,. \tag{4.2.29}$$

This is true for all $\lambda \in \rho(a)$ and $n \in \mathbb{N}$, so

$$\rho(a) \leqslant \inf_{n\in\mathbb{N}} \|a^n\|^{1/n} \leqslant \liminf \|a^n\|^{1/n}\,. \tag{4.2.30}$$

To prove the other half of (4.2.26), suppose $\rho(a) > 0$, let $f \in A^*$, and consider the function $g : B_{\rho(a)^{-1}}(0) \to \mathbb{C}$ defined as

$$g(\lambda) = \begin{cases} f(r_a(\lambda^{-1})) & : \lambda \neq 0 \\ 0 & : \lambda = 0 \end{cases} \tag{4.2.31}$$

If $\rho(a) = 0$, we may define g in this way on all of $\mathbb{C}$. This function is holomorphic on the deleted disk $B_{\rho(a)^{-1}}(0)\backslash\{0\}$ by Lemma 4.2.9 and the fact that $\lambda \mapsto \lambda^{-1}$ is holomorphic on $\mathbb{C}\backslash\{0\}$. Furthermore, for $|\lambda| < \|a\|^{-1}/2$, we have

$$|g(\lambda)| \leqslant \|f\|\,\|r_a(\lambda^{-1})\| \leqslant \frac{\|f\|}{|\lambda|^{-1} - \|a\|} = \frac{\|f\|\,|\lambda|}{1 - \|a\|\,|\lambda|} \leqslant 2\,\|f\|\,|\lambda|, \tag{4.2.32}$$

so g is continuous at zero. It is then a consequence of Morera's theorem that g is holomorphic on the whole disk $B_{\rho(a)^{-1}}(0)$.

Furthermore, for $0 < |\lambda| < \|a\|^{-1}$, we can use the Neumann series to obtain a power series expansion:

$$g(\lambda) = f\left(\lambda \sum_{n=0}^{\infty} (\lambda a)^n\right) = \lambda \sum_{n=0}^{\infty} f(a^n)\lambda^n. \tag{4.2.33}$$

Since g is holomorphic on $B_{\rho(a)^{-1}}(0)$ or on $\mathbb{C}$ if $\rho(a) = 0$, the above gives its power series expansion on its entire domain by the unique representability of g by a power series. The radius of convergence of this power series is therefore at least $\rho(a)^{-1}$, so the series converges absolutely for every λ in the domain of g. Hence, for any $f \in A^*$ and $\lambda \in \mathbb{C}$ with $|\lambda| < \rho(a)^{-1}$, the sequence $|\lambda^n f(a^n)|$ is bounded as n varies across the natural numbers.

Recall that the map $\Psi : A \to A^{**}$ defined as $\Psi(a)(f) = f(a)$ is a linear isometry. Then the boundedness of $|\lambda^n f(a^n)|$ for all $f \in A^*$ indicates that the family $\Psi(\lambda^n a^n)$ is pointwise bounded. By the uniform boundedness principle, the family $\Psi(\lambda^n a^n)$ is uniformly bounded, i.e. for each λ there exists $M_\lambda > 0$ such that

$$|\lambda^n f(a^n)| < M_\lambda, \tag{4.2.34}$$

for all $f \in A^*$ with $\|f\| \leqslant 1$. By the Hahn-Banach theorem, there exists $f \in A^*$ with $\|f\| \leqslant 1$ such that $|f(a^n)| = \|a^n\|$. Thus, we have $|\lambda|^n \|a^n\| < M_\lambda$ for all $n \in \mathbb{N}$, or

$$\|a^n\|^{1/n} < M_\lambda^{1/n} |\lambda|^{-1}, \tag{4.2.35}$$

assuming $|\lambda| > 0$. Taking the limit supremum of both sides yields

$$\limsup \|a^n\|^{1/n} \leqslant \limsup M_\lambda^{1/n} |\lambda|^{-1} = \lim_{n\to\infty} M_\lambda^{1/n} |\lambda|^{-1} = |\lambda|^{-1}. \tag{4.2.36}$$

If $\rho(a) = 0$, this is valid for all $|\lambda| > 0$, which implies $\limsup \|a^n\|^{1/n} = 0 = \rho(a)$. If $\rho(a) > 0$, this is valid for all $|\lambda|^{-1} > \rho(a)$, which implies that

$$\limsup \|a^n\|^{1/n} \leqslant \rho(a), \tag{4.2.37}$$

as desired. □

The proof that $r(a) \leqslant \liminf \|a^n\|^{1/n}$ in Theorem 4.2.11 contains some interesting observations worth highlighting.

4.2.12 Lemma *If $a_1, \ldots, a_n, b \in A$ such that*

$$b = a_1 a_2 \cdots a_n, \tag{4.2.38}$$

and $[a_i, a_j] = 0$ for all i, j, then b is invertible if and only if each a_i is invertible.

Proof. It is obvious that b is invertible if each a_i is invertible. If b is invertible, then for any $i \leqslant n$, we have

$$\left(b^{-1} \prod_{j \neq i} a_j\right) a_i = b^{-1} b = 1 = b b^{-1} = a_i \left(\prod_{j \neq i} a_j\right) b^{-1}. \tag{4.2.39}$$

Thus, a_i has a left inverse and a right inverse, which must be equal. □

4.2.13 Theorem *Let $p = \sum_{i=0}^n \alpha_i z^i$ be a complex polynomial and let $a \in A$. Then*

$$\sigma(p(a)) = p(\sigma(a)). \tag{4.2.40}$$

Proof. Fix $\lambda \in \mathbb{C}$ and factorize $\lambda - p(a)$:

$$\lambda - p(z) = \beta_0 \prod_{i=1}^{n} (\beta_i - z) \tag{4.2.41}$$

for some $\beta_0, \ldots, \beta_n \in \mathbb{C}$. Then

$$\lambda - p(a) = \beta_0 \prod_{i=1}^{n} (\beta_i - a). \tag{4.2.42}$$

By Lemma 4.2.12, $\lambda \in \sigma(p(a))$ if and only if $\beta_i \in \sigma(a)$ for some $i \geqslant 1$. But there exists $i \geqslant 1$ such that $\beta_i \in \sigma(a)$ if and only if $\lambda \in p(\sigma(a))$, so we're done. □

Theorem 4.2.13 gives one example of how algebraic manipulations in A affect the spectra of the elements being manipulated. Let us give more results in this vein.

4.2.14 Theorem *If $a, b \in A$, then*

$$\sigma(ab) \cup \{0\} = \sigma(ba) \cup \{0\}. \tag{4.2.43}$$

If $a \in A^{\times}$, then

$$\sigma(a^{-1}) = \sigma(a)^{-1}. \tag{4.2.44}$$

Proof. Suppose $\lambda \in \rho(ab)$. Then using $(\lambda - ba)b = b(\lambda - ab)$ we compute

$$(\lambda - ba)\big[1 + b(\lambda - ab)^{-1}a\big] = (\lambda - ba) + ba = \lambda. \tag{4.2.45}$$

and likewise

$$\big[1 + b(\lambda - ab)^{-1}a\big](\lambda - ba) = (\lambda - ba) + ba = \lambda. \tag{4.2.46}$$

Therefore $\lambda \in \rho(ba)$ if $\lambda \neq 0$. Of course, the same result holds with a and b switched. In other words,

$$\rho(ab)\backslash\{0\} = \rho(ba)\backslash\{0\} \tag{4.2.47}$$

Taking complements yields (4.2.43).

If $a \in A^{\times}$, then clearly $0 \notin \sigma(a)$ and $0 \notin \sigma(a^{-1})$. If $\lambda \neq 0$, then

$$\lambda^{-1} - a = \lambda^{-1}a(a^{-1} - \lambda), \tag{4.2.48}$$

which implies $\lambda^{-1} \in \sigma(a)$ if and only if $\lambda \in \sigma(a^{-1})$ by Lemma 4.2.12. Since $\lambda^{-1} \in \sigma(a)$ if and only if $\lambda \in \sigma(a)^{-1}$, this is the desired result. □

4.2.15. [

Spectral Theory in C*-Algebras]Spectral Theory in C*-Algebras

We continue where we left off in the previous section by showing how the spectrum behaves with respect to the star operation. We now let A be a unital C^*-algebra for the rest of this section.

4.2.16 Proposition *Let $a \in A$. Then*

$$\sigma(a^*) = \sigma(a)^*. \tag{4.2.49}$$

Proof. We note that $\lambda - a^*$ is invertible if and only if $\lambda - a^*$ is invertible by Proposition 4.1.13. Hence, $\lambda \in \sigma(a^*)$ if and only if $\lambda^* \in \sigma(a)$ if and only if $\lambda \in \sigma(a)^*$. □

We now investigate the spectra of several special classes of elements of A.

4.2.17 Definition An element $a \in A$ is

(i) *normal* if $[a, a^*] = 0$,

(ii) an *isometry* if $a^*a = 1$, and

(iii) *unitary* if $a^*a = aa^* = 1$, i.e. $a \in A^\times$ and $a^{-1} = a^*$.

Note that both unitary and self-adjoint elements are normal. Furthermore, if a is an isometry or a unitary, then $\|a\| = 1$ by the C^*-property.

4.2.18 Corollary *If $a \in A$ is normal, then $\rho(a) = \|a\|$.*

Proof. We claim that

$$\left\|a^{2^n}\right\|^2 = \|a\|^{2^{n+1}}. \tag{4.2.50}$$

for all normal $a \in A$. For $n = 0$ this is trivial. Suppose it is true for some $n = k-1$ where $k \in \mathbb{N}$. Using normality of a, the C^*-property, and the fact that a^*a is self-adjoint, we compute

$$\begin{aligned} \left\|a^{2^k}\right\|^2 &= \left\|\left(a^{2^k}\right)^* a^{2^k}\right\| = \left\|(a^*)^{2^k} a^{2^k}\right\| = \left\|(a^*a)^{2^k}\right\| \\ &= \left\|(a^*a)^{2^{k-1}}\right\|^2 = \|a^*a\|^{2^k} = \|a\|^{2^{k+1}}. \end{aligned} \tag{4.2.51}$$

This proves (4.2.50). Now, using the formula for the spectral radius, we obtain

$$r(a) = \lim_{n\to\infty} \left\|a^{2^n}\right\|^{1/2^n} = \lim_{n\to\infty} \|a\|^{2^n/2^n} = \|a\|, \tag{4.2.52}$$

as desired. □

4.2.19 Corollary *If $a \in A$ is isometric, then $r(a) = 1$.*

Proof. We note that

$$\|a^n\|^2 = \left\|(a^n)^*(a^n)\right\| = \|(a^*)^n(a^n)\| = \|1\| = 1. \tag{4.2.53}$$

Thus,

$$r(a) = \lim_{n\to\infty} \|a^n\|^{1/n} = \lim_{n\to\infty} 1 = 1, \tag{4.2.54}$$

as desired. □

4.2.20 Corollary *If $a \in A$ is unitary, then $\sigma(a) \subset S^1$.*

Proof. We have

$$\sigma(a)^{-1} = \sigma(a^{-1}) = \sigma(a^*) = \sigma(a)^*. \tag{4.2.55}$$

Thus, if $\lambda \in \sigma(a)$, then $\lambda^{-1} = \mu^*$ for some $\mu \in \sigma(a)$. Since $r(a) \leqslant 1$, we know $1 \leqslant |\lambda|^{-1} = |\mu| \leqslant 1$, so $|\lambda| = 1$. □

4.2.21 Theorem *If $a \in A$ is self-adjoint, then*

$$\sigma(a) \subset [-\|a\|, \|a\|]. \tag{4.2.56}$$

In particular, $\sigma(a^2) = \sigma(a)^2 \subset [0, \|a\|^2]$.

Proof. Let $\lambda \in \mathbb{R}$ such that $\lambda^{-1} = |i\lambda^{-1}| > \|a\|$. Then $\lambda^{-1} \in \rho(a)$, so $1 + i\lambda a = -i\lambda(i\lambda^{-1} - a)$ is invertible. Note that $(1 + i\lambda a)^* = 1 - i\lambda a$ is invertible as well. Define

$$u = (1 - i\lambda a)(1 + i\lambda a)^{-1}. \tag{4.2.57}$$

Observe that $u^* = (1 - i\lambda a)^{-1}(1 + i\lambda a)$. Using the fact that $1 - i\lambda a$ commutes with $1 + i\lambda a$, we see that

$$u^*u = (1 - i\lambda a)^{-1}(1 + i\lambda a)(1 - i\lambda a)(1 + i\lambda a)^{-1} = 1. \tag{4.2.58}$$

Since u is the product of two invertible elements, we know u is invertible, and the above shows that $u^{-1} = u^*$, so u is unitary.

Given $\mu \in \mathbb{C}$, $\mu \neq i\lambda^{-1}$, observe that

$$\left|\frac{1 - \mathrm{i}\,\lambda\mu}{1 + \mathrm{i}\,\lambda\mu}\right| = \sqrt{\frac{(1 + \lambda\,\mathfrak{Im}\mu)^2 + (\lambda\,\mathfrak{Re}\mu)^2}{(1 - \lambda\,\mathfrak{Im}\mu)^2 + (\lambda\,\mathfrak{Re}\mu)^2}}, \tag{4.2.59}$$

which equals 1 if and only if $\mathfrak{Im}\mu = 0$, i.e. $\mu \in \mathbb{R}$. Since $\sigma(u) \subset S^1$, we see that if $\mu \in \mathbb{C}\backslash\mathbb{R}$, then $(1 - i\lambda\mu)(1 + i\lambda\mu)^{-1} \in \rho(u)$. In particular, if $\mu \neq i\lambda^{-1}$, then

$$\begin{aligned}(1 - i\lambda\mu)(1 + i\lambda\mu)^{-1} - u &= (1 + i\lambda\mu)^{-1}[(1 - i\lambda\mu)(1 + i\lambda a) - (1 + i\lambda\mu)(1 - i\lambda a)]\,(1 + i\lambda a)^{-1} \\ &= 2i\lambda(1 + i\lambda\mu)^{-1}(a - \mu)(1 + i\lambda a)^{-1}.\end{aligned} \tag{4.2.60}$$

If $\mu \in \mathbb{C}\backslash\mathbb{R}$, then the left hand side is invertible, so $\mu - a$ is invertible, so $\mu \in \rho(a)$. Furthermore, $i\lambda^{-1} \notin \sigma(a)$ since $|i\lambda^{-1}| > \|a\|$. Thus, $\sigma(a) \subset \mathbb{R}$. □

4.2.22 Corollary *If $a \in A$, then*

$$\|a\| = \sqrt{r(a^*a)}. \tag{4.2.61}$$

Hence, the norm is completely determined by the algebraic operations, i.e. the C^*-norm is unique.

Proof. Note that a^*a is normal, so

$$\|a\|^2 = \|a^*a\| = r(a^*a). \tag{4.2.62}$$

The result follows by taking a square root. □

4.2.23 Theorem (Spectral Permanence) *Let B be a unital C^*-subalgebra of A. For any $a \in B$, the spectrum of a in B is the same as the spectrum of a in A.*

Proof. Let us temporarily use the notation $\sigma_A(a)$ and $\sigma_B(a)$ to distinguish the spectrum of a in A and B respectively. We shall use the notation $\rho_A(a)$, $\rho_B(a)$, $r_A(a)$, and $r_B(a)$ similarly. If $\lambda \in \sigma_A(a)$, then $\lambda - a$ is not invertible in A, so $\lambda - a$ is certainly not invertible in B. Thus, $\sigma_A(a) \subset \sigma_B(a)$.

The reverse inclusion $\sigma_B(a) \subset \sigma_A(a)$ is trivial if $a = 0$, so suppose $a \neq 0$. It is easy to check that $\sigma_B(a) \subset \sigma_A(a)$ follows from the inclusion $B \cap A^\times \subset B^\times$, and this is what we will show. First suppose $a \in B \cap A^\times$ and a is self-adjoint. Let $\lambda_0 = 2i\,\|a\|$ and note that $\lambda_0 \in \rho_B(a)$ since $|\lambda_0| > \|a\|$. We know from Corollary 4.2.6 that $B_{\|r_a(\lambda_0)\|^{-1}}(\lambda_0) \subset \rho_B(a)$; thus we want to show that $0 \in B_{\|r_a(\lambda_0)\|^{-1}}(\lambda_0)$ to show that $a \in B^\times$.

Observe that self-adjointness of a implies that $r_a(\lambda_0)^* = r_a(\lambda_0^*)$, from which it follows that $r_a(\lambda_0)$ is normal since $[r_a(\lambda), r_a(\mu)] = 0$ for all $\lambda, \mu \in \rho_B(a)$. Normality then yields

$$\|r_a(\lambda_0)\| = r_A(r_a(\lambda_0)). \tag{4.2.63}$$

But we also know that

$$\sigma_A(r_a(\lambda_0)) = \sigma_A(\lambda_0 - a)^{-1} = [\lambda_0 - \sigma_A(a)]^{-1} \tag{4.2.64}$$

Hence,

$$r_A(r_a(\lambda_0)) = \operatorname{dist}(\lambda_0, \sigma_A(a))^{-1}. \tag{4.2.65}$$

But since λ_0 is purely imaginary and $\sigma_A(a) \subset \mathbb{R}$, we know that $\operatorname{dist}(\lambda_0, \sigma_A(a)) \geqslant |\lambda_0|$! In fact, we know $0 \notin \sigma_A(a)$ since $a \in A^\times$, so this is a strict inequality and taking inverses yields

$$\|r_a(\lambda_0)\| = \operatorname{dist}(\lambda_0, \sigma_A(a))^{-1} < |\lambda_0|^{-1}. \tag{4.2.66}$$

This implies $0 \in B_{\|r_a(\lambda_0)\|^{-1}}(\lambda_0)$, proving the theorem for self-adjoint a.

Finally, consider $a \in B \cap A^\times$, not necessarily self-adjoint. However, we see that $a^*a \in B \cap A^\times$ and a^*a is self-adjoint, so $a^*a \in B^\times$. Defining

$$b = (a^*a)^{-1}a^*, \tag{4.2.67}$$

we see that $ba = 1$, so $b = a^{-1}$ since a was assumed to be invertible in A. Since $b \in B$ manifestly, we conclude that $a \in B^\times$, as desired. □

Let us now consider how the spectrum of an element behaves under $*$-homomorphisms.

4.2.24 Proposition *Let A and B be unital C^*-algebras and let $\pi : A \to B$ be a unital $*$-homomorphism. Then $\pi(A^\times) \subset B^\times$ and*

$$\pi(a^{-1}) = \pi(a)^{-1} \tag{4.2.68}$$

for all $a \in A^\times$. If π is bijective, then $\pi(A^\times) = B^\times$.

Proof. Let $a \in A^\times$ and observe

$$\pi(a^{-1})\pi(a) = \pi(a^{-1}a) = \pi(1) = 1 = \pi(1) = \pi(aa^{-1}) = \pi(a)\pi(a^{-1}). \tag{4.2.69}$$

This proves that $\pi(a) \in B^\times$ and $\pi(a)^{-1} = \pi(a^{-1})$. If π is bijective, then $\pi^{-1}(B^\times) \subset A^\times$ by the same argument, so $\pi\pi^{-1}(B^\times) = B^\times \subset \pi(A^\times)$. □

4.2.25 Proposition *Let A and B be unital C^*-algebras and let $\pi : A \to B$ be a unital $*$-homomorphism. Then*

$$\sigma(\pi(a)) \subset \sigma(a) \tag{4.2.70}$$

and

$$\pi(r_a(\lambda)) = r_{\pi(a)}(\lambda) \tag{4.2.71}$$

for all $\lambda \in \rho(a)$. If π is bijective, then $\sigma(\pi(a)) = \sigma(a)$.

Proof. Let $\lambda \in \rho(a)$. Then by Proposition 4.2.24, we know $\pi(\lambda - a) = \lambda - \pi(a) \in B^\times$ and

$$r_{\pi(a)}(\lambda) = (\lambda - \pi(a))^{-1} = \pi(\lambda - a)^{-1} = \pi(r_a(\lambda)). \tag{4.2.72}$$

We see that $\rho(a) \subset \rho(\pi(a))$, so $\sigma(\pi(a)) \subset \sigma(a)$ by taking complements. If π is bijective, then $\sigma(a) = \sigma(\pi^{-1}(\pi(a))) \subset \sigma(\pi(a))$ as well. □

4.2.26 Proposition *Let A and B be unital C^*-algebras and let $\pi : A \to B$ be a unital $*$-homomorphism. Then*

$$\|\pi(a)\| \leqslant \|a\| \tag{4.2.73}$$

for all $a \in A$. In particular, π is continuous and $\|\pi\| = 1$.

Proof. By Proposition 4.2.25, we know that $r(\pi(a)) \leqslant r(a)$ for all $a \in A$. Then by Corollary 4.2.22,

$$\|\pi(a)\| = \sqrt{r(\pi(a)^*\pi(a))} = \sqrt{r(\pi(a^*a))} \leqslant \sqrt{r(a^*a)} = \|a\|. \tag{4.2.74}$$

This shows that π is continuous, and $\|\pi\| = 1$ since $\|\pi(1)\| = \|1\| = 1$. □

With these simple propositions in hand, we can prove a powerful theorem, known as the *continuous functional calculus* for self-adjoint elements.

4.2.27 Theorem *Let A be a unital C^*-algebra and let $a \in A$ be self-adjoint. There exists a unique unital $*$-homomorphism $C(\sigma(a)) \to A$, $f \mapsto f(a)$ such that $p(a) = \sum_{i=0}^n \alpha_i a^i$ for all complex polynomials $p(z) = \sum_{i=0}^n \alpha_i z^i$. Furthermore, for all $f \in C(\sigma(a))$, we have*

(i) $\|f(a)\| = \|f\|$,

(ii) *$f(a)$ is in the C^*-algebra generated by 1 and a. In particular, $[f(a), a] = 0$,*

(iii) *$\pi(f(a)) = f(\pi(a))$ for any unital $*$-homomorphism $\pi : A \to B$,*

(iv) *$\sigma(f(a)) = f(\sigma(a))$.*

Finally, if $g \in C(\sigma(f(a)))$, then

(v) *$(g \circ f)(a) = g(f(a))$.*

Note that $f(\pi(a))$ is well-defined in (iii) since $\sigma(\pi(a)) \subset \sigma(a)$, and $(g \circ f)(a)$ is well-defined by (iv).

Proof. It clear that the map $p \mapsto p(a)$ defined on polynomials $p(z) = \sum_{i,j=0}^{n} \alpha_i z^i$ is linear. Furthermore, since $p(a)$ is self-adjoint for any polynomial p, we have

$$\begin{aligned} \|p(a)\| &= r(p(a)) \\ &= \sup\{|\lambda| : \lambda \in \sigma(p(a))\} \\ &= \sup\{|\lambda| : \lambda \in p(\sigma(a))\} \\ &= \sup\{|p(\lambda)| : \lambda \in \sigma(a)\} \\ &= \|p\|, \end{aligned} \tag{4.2.75}$$

so the map $p \mapsto p(a)$ is continuous. Since $\sigma(a)$ is a compact subset of $\mathbb{R}$, the Weierstrass approximation theorem (and the Tietze extension theorem) imply that the set of polynomials is dense in $C(\sigma(a))$. Therefore the map $p \mapsto p(a)$ extends uniquely to a linear map $f \mapsto f(a)$ on $C(\sigma(a))$. It follows by standard continuity arguments that $\|f(a)\| = \|f\|$ for all $f \in C(\sigma(a))$ since $\|p(a)\| = \|p\|$ for all polynomials.

It is clear that for polynomials p and q, we have $(pq)(a) = p(a)q(a)$ and $(p^*)(a) = (p(a))^*$, the latter relying on self-adjointness of a. That $(fg)(a) = f(a)g(a)$ and $(f^*)(a) = (f(a))^*$ for arbitrary $f, g \in C(\sigma(a))$ again follows by standard continuity arguments using the fact that the polynomials are dense in $C(\sigma(a))$. Thus, $f \mapsto f(a)$ is a unital $*$-homomorphism.

Once again, (ii) and (iii) clearly hold for polynomials. Thus, (ii) holds for f by a standard argument using density of polynomials and completeness of the C^*-algebra generated by 1 and a. Likewise, (iii) holds by density of polynomials and by continuity of $*$-homomorphisms as shown in Proposition 4.2.26.

To prove (iv), let (p_n) be a sequence of polynomials such that $p_n \to f$. Given $\lambda \in \sigma(a)$, we know

$$p_n(\lambda) \in p_n(\sigma(a)) = \sigma(p_n(a)), \tag{4.2.76}$$

so $p_n(\lambda) - p_n(a)$ is not invertible. Since the complement of $A^\times$ is closed, taking the limit as $n \to \infty$ yields $f(\lambda) - f(a) \notin A^\times$, so $f(\lambda) \in \sigma(f(a))$. Hence $f(\sigma(a)) \subset \sigma(f(a))$. On the other hand, if $\lambda \notin f(\sigma(a))$, then $\lambda - f$ is invertible in $C(\sigma(a))$ with inverse $g \in C(\sigma(a))$. Then

$$(\lambda - f(a))g(a) = g(a)(\lambda - f(a)) = 1, \tag{4.2.77}$$

so $\lambda - f(a)$ is invertible, i.e. $\lambda \notin \sigma(f(a))$. This proves $\sigma(f(a)) \subset f(\sigma(a))$, as desired.

To prove (v), we note that $g \mapsto g \circ f$ is a unital $*$-homomorphism $C(\sigma(f(a))) \to C(\sigma(a))$, so $g \mapsto g \circ f \mapsto (g \circ f)(a)$ is a unital $*$-homomorphism $C(\sigma(f(a))) \to A$. Furthermore, if p is a polynomial, then $(p \circ f)(a) = p(f(a))$ since the map $C(\sigma(a)) \to A$ respects addition and multiplication. Therefore $(g \circ f)(a) = g(f(a))$ by uniqueness of the $*$-homomorphism $C(\sigma(f(a))) \to A$. □

4.2.28 Corollary *Let A and B be unital C^*-algebras. If $\pi : A \to B$ is an injective unital $*$-homomorphism, then π is an isometry:*

$$\|\pi(a)\| = \|a\| \tag{4.2.78}$$

for all $a \in A$.

Proof. Since $\|\pi(a)\| \leqslant \|a\|$ by Proposition 4.2.26, we need only show the reverse inequality. First we show $\|a\| \leqslant \|\pi(a)\|$ for all self-adjoint elements $a \in A$. Suppose a is self-adjoint and $\|\pi(a)\| < \|a\|$. Recall that $\rho(a) = \|a\|$ and $\sigma(a) \subset \mathbb{R}$, so $\|a\| \in \sigma(a)$ or $-\|a\| \in \sigma(a)$, and likewise for $\pi(a)$. Choose $f : [-\|a\|, \|a\|] \to \mathbb{R}$ such that f vanishes on $[-\|\pi(a)\|, \|\pi(a)\|]$ and $f(\|a\|) = f(-\|a\|) = 1$. Then $f(\pi(a)) = \pi(f(a)) = 0$ but $\|f(a)\| = \|f\| > 1$, contradicting injectivity of π. Therefore $\|a\| = \|\pi(a)\|$ for self-adjoint a.

For arbitrary $a \in A$, we have

$$\|\pi(a)\|^2 = \|\pi(a)^*\pi(a)\| = \|\pi(a^*a)\| = \|a^*a\| = \|a\|^2, \tag{4.2.79}$$

which concludes the proof. □

4.2.29. Positive Elements

We shall continue to let A be a unital C^*-algebra.

4.2.30 Definition An element $a \in A$ is *positive* if a is self-adjoint and $\sigma(a) \subset [0, \infty)$. We let A_+ denote the set of all positive elements of A.

4.2.31 Proposition *Let $a \in A$ be self-adjoint and let $\lambda \in \mathbb{R}$ such that $\lambda \geqslant \|a\|$. Then a is positive if and only if $\|\lambda - a\| \leqslant \lambda$.*

Proof. If $a \in A_+$, then

$$\|\lambda - a\| = r(\lambda - a) = \sup\{|\mu| : \mu \in \sigma(\lambda - a) = \lambda - \sigma(a)\} \leqslant \lambda \tag{4.2.80}$$

since $\sigma(a) \subset [0, \|a\|] \subset [0, \lambda]$.

Suppose $\|\lambda - a\| \leqslant \lambda$. If $\mu \in \sigma(a)$, then

$$|\lambda - \mu| \leqslant r(\lambda - a) = \|\lambda - a\| < \lambda, \tag{4.2.81}$$

which implies that $\mu \geqslant 0$, hence $a \in A_+$. □

4.2.32 Proposition *The set of positive elements A_+ is closed.*

Proof. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in A_+ converging to $a \in A$. Since the star operation is continuous, we have $a_n^* \to a^*$, but since $a_n^* = a_n$ for all $n \in \mathbb{N}$, we see that a is self-adjoint. Since the norm is also continuous, we see that $\|a_n\| \to \|a\|$. In particular, there exists $M > 0$ such that $\|a_n\| < M$ for all $n \in \mathbb{N}$. Then $\|a\| \leqslant M$ and $\|M - a_n\| \leqslant M$ for all $n \in \mathbb{N}$ by Proposition 4.2.31, so

$$\|M - a\| = \lim_{n\to\infty} \|M - a_n\| \leqslant M. \tag{4.2.82}$$

Since $\|a\| \leqslant M$ and $\|M - a\| \leqslant M$, Proposition 4.2.31 implies that a is positive. □

4.2.33 Proposition *The sum of two positive elements is positive.*

Proof. If $a, b \in A_+$, then $a+b$ is self-adjoint and Proposition 4.2.31 implies

$$\big\| \|a\| + \|b\| - (a+b) \big\| \leqslant \big\| \|a\| - a \big\| + \big\| \|b\| - b \big\| \leqslant \|a\| + \|b\| . \tag{4.2.83}$$

A second application of Proposition 4.2.31 yields $a+b \in A_+$. □

4.2.34 Proposition *Let $a \in A$ be self-adjoint. The following are equivalent.*

(i) *The element a is positive.*

(ii) *There exists a unique positive $b \in A$ such that $a = b^2$.*

(iii) *There exists a self-adjoint $b \in A$ such that $a = b^2$.*

(iv) *There exists $c \in A$ such that $a = c^* c$.*

Proof. The implications (ii) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (iv) are trivial.

(i) $\Rightarrow$ (ii). By the continuous functional calculus, we can take $\sqrt{a}$ and by the composition property and the fact that $(\sqrt{x})^2 = x$ on $\sigma(a)$, we have that $a = (\sqrt{a})^2$. The fact that the square root is real-valued and that the continuous functional calculus is a $*$-homomorphism imply that $\sqrt{a}$ is self-adjoint, and $\sqrt{\sigma(a)} = \sigma(\sqrt{a})$ implies that $\sqrt{a}$ is positive. If $a = b^2$ for any other positive $b \in A$, then the fact that $\sqrt{x^2} = x$ for $x \in \sigma(a)$ and the composition property imply that $\sqrt{a} = \sqrt{b^2} = b$.

(iv) $\Rightarrow$ (i). First we prove a lemma. Given $d \in A$, we claim that $\sigma(-d^* d) \subset [0, \infty)$ implies $d = 0$. Write $d = d_1 + i d_2$, where d_1 and d_2 are self-adjoint. Then

$$d^* d + d d^* = 2 d_1^2 + 2 d_2^2. \tag{4.2.84}$$

If $\sigma(-d^* d) \subset [0, \infty)$, then $\sigma(-d d^*) \subset \sigma(-d^* d) \cup \{0\} \subset [0, \infty)$. Thus $d^* d = 2 d_1^2 + 2 d_2^2 - d d^*$ is positive, since it is the sum of positive elements. But then $\sigma(d^* d) = \{0\}$, which implies that $d^* d = 0$ and therefore $d = 0$ by the C^*-property of the norm.

Continuing, we suppose $a = c^* c$ for some $c \in A$. by the continuous functional calculus, the elements $a_+ = (|a| + a)/2$ and $a_- = (|a| - a)/2$ are positive and $a = a_+ - a_-$. Furthermore, observe that

$$a_+ a_- = \frac{1}{4}\Big(|a|^2 - a^2\Big) = 0. \tag{4.2.85}$$

Defining $d = c a_-$, we compute

$$- d^* d = -a_- c^* c a_- = -a_- (a_+ - a_-) a_- = (a_-)^3, \tag{4.2.86}$$

which implies that $-d^* d$ is positive. By the previous paragraph, we know $d = 0$. Thus,

$$0 = c^* d = c^* c a_- = a a_- = -(a_-)^2 \tag{4.2.87}$$

which implies that $a_- = 0$, for example by self-adjointness of a_- and the C^*-property of the norm. Thus, $a = a_+$ is positive. □

4.2.35 Definition We define a partial ordering on A_+ by setting $a \leqslant b$ for $a, b \in A_+$ if and only if $b - a \in A_+$. Reflexivity and antisymmetry are easy to check and transitivity follows since $(c-b)+(b-a) = c-a \in A_+$ given $c-b, b-a \in A_+$. In fact, A_+ is a directed set since given $a, b \in A_+$, we have $a+b \in A_+$ and $a, b \leqslant a+b$.

4.2.36 Proposition *Let $a, b \in A_+$ and let $c \in A$. If $a \leqslant b$, then $c^*ac \leqslant c^*bc$.*

Proof. Since $a \in A_+$, there exists $d \in A$ such that $a = d^*d$. Then $c^*ac = c^*d^*dc = (dc)^*dc$, so $c^*ac \in A_+$. Likewise, $c^*bc \in A_+$ since $b \in B_+$. Likewise, $c^*(b-a)c \in A_+$ since $b - a \in A_+$, so $c^*ac \leqslant c^*bc$. □

4.2.37 Proposition *If $a \in A_+$ and $\lambda \geqslant 0$, then $a \leqslant \lambda$ if and only if $\|a\| \leqslant \lambda$.*

Proof. We have the following equivalences:

$$a \leqslant \lambda \quad \Longleftrightarrow \quad \sigma(\lambda - a) = \lambda - \sigma(a) \subset [0, \infty) \quad \Longleftrightarrow \quad r(a) = \|a\| \leqslant \lambda, \tag{4.2.88}$$

as desired. □

If we replace λ in the above proposition by an arbitrary element, then we only have an implication in one direction.

4.2.38 Proposition *If $a, b \in A_+$ and $a \leqslant b$, then $\|a\| \leqslant \|b\|$.*

Proof. We know $b \leqslant \|b\|$ by Proposition 4.2.37, so $a \leqslant \|b\|$ by transitivity. But this implies $\|a\| \leqslant \|b\|$ by another application of Proposition 4.2.37. □

4.2.39 Proposition *If $a, b \in A_+ \cap A^\times$ and $a \leqslant b$, then $b^{-1} \leqslant a^{-1}$.*

Proof. Note that $a^{-1}, b^{-1} \in A_+$ by the continuous functional calculus. Since $\sqrt{a^{-1}}$ is self-adjoint, we have

$$1 = \sqrt{a^{-1}}a\sqrt{a^{-1}} \leqslant \sqrt{a^{-1}}b\sqrt{a^{-1}}. \tag{4.2.89}$$

Thus, $\sigma(\sqrt{a^{-1}}b\sqrt{a^{-1}}) \subset [1, \infty)$ and by the continuous functional calculus,

$$1 \geqslant \left(\sqrt{a^{-1}}b\sqrt{a^{-1}}\right)^{-1} = \sqrt{a^{-1}}^{-1}b^{-1}\sqrt{a^{-1}}^{-1}. \tag{4.2.90}$$

Multiplying by $\sqrt{a^{-1}}$ to the left and right as in the first step now yields $b^{-1} \leqslant a^{-1}$. □

4.3. Infinite tensor products

4.3.1 Infinite tensor products of Hilbert spaces were introduced by von Neumann (1939). They were motivated by mathematical physics where one needs to describe quantum systems with infinitely many degrees of freedom, see e.g. Emch (2009); Bratteli and Robinson (1997). The original construction of infinite tensor products was generalized to von Neumann and C^*-algebras by Guichardet (1966), Blackadar (1977), and others. Meanwhile, the topic has been studied in quite some detail in the operator algebra literature, see e.g. Nakagami (1970a,b); Størmer

(1971). A purely algebraic or better categorical approach allowing the construction of infinite tensor products of modules over a given commutative ring has been given in (Chevalley, 1956, Sec. III.10). The work Ng (2013) is also in that spirit. We will essentially follow Chevalley (1956) and construct the infinite tensor product as a module universal with respect to multilinear maps. First we present the main algebraic construction, then we explain some of the subtleties which distinguish infinite from finite tensor products, and finally we construct infinite Hilbert tensor products and infinite tensor products of C^*-algebras.

4.3.2 Let R be a commutative ring and $(M_i)_{i\in I}$ a possibly infinite family of R-modules. Consider $\prod_{i\in I} M_i$, the product of the family $(M_i)_{i\in I}$ within the category of R-modules. For each $j \in I$ let $\pi_j : \prod_{i\in I} M_i \to M_j$ denote the natural projection onto the j-th factor and $\iota_j : M_j \hookrightarrow \prod_{i\in I} M_i$ the uniquely determined natural embedding such that

$$\pi_j \circ \iota_i = \begin{cases} \mathrm{id}_{M_i} & \text{for } i = j \text{ and} \\ 0 & \text{else.} \end{cases}$$

Given an R-module N one then understands by a *multilinear map* from $\prod_{i\in I} M_i$ to N a map $f : \prod_{i\in I} M_i \to N$ such that for each $j \in I$ and $x \in \prod_{i\in I} M_i$ with $\pi_j(x) = 0$ the map $M_j \to N$, $m \mapsto f(\iota_j(m) + x)$ is linear. The set of multilinear maps from $\prod_{i\in I} M_i$ to N will be denoted by $\mathfrak{Mlin}\big(\prod_{i\in I} M_i, N\big)$. It carries a natural structure of an R-module given by pointwise addition of multilinear maps and pointwise action of a scalar on a multilinear map that is by

$$f + g = \left(\prod_{i\in I} M_i \ni x \mapsto f(x) + g(x) \in N\right) \quad \text{and} \quad rf = \left(\prod_{i\in I} M_i \ni x \mapsto rf(x) \in N\right)$$

for all $f, g \in \mathfrak{Mlin}\big(\prod_{i\in I} M_i, N\big)$ and $r \in R$. Since for $j \in I$ and $x \in \prod_{i\in I} M_i$ with $\pi_j(x) = 0$ the maps $M_j \to N$, $m \mapsto (f + g)(\iota_j(m) + x) = f(\iota_j(m) + x) + g(\iota_j(m) + x)$ and $M_j \to N$, $m \mapsto rf(\iota_j(m) + x)$ are linear by assumption on f and g, the maps $f + g$ and rf are multilinear again, so $\mathfrak{Mlin}\big(\prod_{i\in I} M_i, N\big)$ is an R-module indeed with zero element the constant function mapping to $0 \in N$.

4.3.3 Remarks Before proceeding further let us make several explanations concerning the notation used.

(a) The space of multilinear maps $\mathfrak{Mlin}\big(\prod_{i\in I} M_i, N\big)$ actually depends on the family $(M_i)_{i\in I}$ and the R-module N, so in principle one should write $\mathfrak{Mlin}\big((M_i)_{i\in I}, N\big)$ instead of $\mathfrak{Mlin}\big(\prod_{i\in I} M_i, N\big)$. Nevertheless we stick to the latter notation since it is closer to standard notation for linear maps and since it will not lead to any confusion.

(b) In case the index set I has just two elements i_1, i_2, one calls a multilinear map $\prod_{i\in I} M_i = M_{i_1} \times M_{i_2} \to N$ a *bilinear map*. If the cardinality of I is 3, one sometimes calls a multilinear map $\prod_{i\in I} M_i \to N$ a *trilinear map*.

(c) In the following, when saying that $(I_a)_{a\in A}$ is a partition of the set I we mean that each I_a is a non-empty subset of I, that $I_a \cap I_b = \varnothing$ for $a \neq b$ and that $\bigcup_{a\in A} I_a = I$. The empty family is regarded as a partition of the empty set.

(d) We will frequently use in this section the same symbol for maps with the same "universal" properties despite those maps might be strictly speaking different. For example, π_k will stand for the canonical projections $\prod_{i\in I} M_i \to M_k$ and $\prod_{j\in J} M_j \to M_k$ whenever $k \in J \subset I$. Likewise we use the same notation for the two canonical embeddings $M_k \hookrightarrow \prod_{i\in I} M_i$ and $M_k \hookrightarrow \prod_{j\in J} M_j$ defined in 4.3.2 and denote them both by ι_k.

4.3.4 Lemma (cf. (Chevalley, 1956, Sec. III.10, Lemma 1 & 2)) *Assume that $(M_i)_{i\in I}$ is a family of R-modules, N an R-module, and $f : \prod_{i\in I} M_i \to N$ a mutilinear map.*

(i) *If $g : N \to N'$ is an R-module map, then $g \circ f : \prod_{i\in I} M_i \to N'$ is multilinear.*

(ii) *Let $J \subset I$ be non-empty, $y = (y_i)_{i\in I\setminus J}$ an element of the product $\prod_{i\in I\setminus J} M_i$, and $\iota_{J,y} : \prod_{j\in J} M_j \to \prod_{i\in I} M_i$ the unique map such that for all $x = (x_j)_{j\in J} \in (M_j)_{j\in J}$ and $k \in I$*

$$\pi_k \circ \iota_{J,y}(x) = \begin{cases} x_k & \text{for } k \in J, \\ y_k & \text{for } k \in I\setminus J. \end{cases}$$

Then the composition $f \circ \iota_{J,x} : \prod_{j\in J} M_j \to N$ is multilinear.

(iii) *Let $(I_a)_{a\in A}$ be a partition of the index set I which is assumed to be non-empty. Let $(N_a)_{a\in A}$ be a family of R-modules, $(g_a)_{a\in A}$ a family of multilinear maps $g_a : \prod_{i\in I_a} M_i \to N_a$, and $h : \prod_{a\in A} N_a \to N$ multilinear. Define $g : \prod_{i\in I} M_i \to \prod_{a\in A} N_a$ as the unique map such that*

$$\pi_b \circ g = g_b \circ \pi_{I_b} \quad \text{for } b \in A,$$

where π_J for $J \subset I$ as on the right side stands for the projection $\pi_J : \prod_{i\in I} M_i \to \prod_{j\in J} M_j$ uniquely determined by $\pi_j \circ \pi_J = \pi_j$ for all $j \in J$. Then the composition $h\circ g : \prod_{i\in I} M_i \to N$ is multilinear.

Proof. ad (i). Let $j \in I$ and $x \in \prod_{i\in I} M_i$ with $\pi_j(x) = 0$. By multilinearity of f and linearity of g, the map $M_j \to N'$, $m \mapsto gf(\iota_j(m) + x)$ then has to be linear, hence $g \circ f$ is multilinear.

ad (ii). Let $j \in J$ and $x \in \prod_{i\in J} M_i$ with $\pi_j(x) = 0$. Then $\pi_j(\iota_{J,y}(x)) = 0$ and $f\iota_{J,y}(\iota_j(m) + x) = f(\iota_j(m) + \iota_{J,y}(x)$ for all $m \in M_j$ by construction of $\iota_{J,y}$. Hence the map $M_j \to N$, $m \mapsto f\iota_{J,y}(\iota_j(m) + x)$ is linear by multilinearity of f. This proves that $f \circ \iota_{J,y}$ is multilinear.

ad (iii). Given $j \in I$ let b be the unique element of A such that $j \in I_b$. Assume that $x \in \prod_{i\in I} M_i$ with $\pi_j(x) = 0$. By construction one has $\pi_j(\pi_{I_b}(x)) = 0$. Now let $y \in \prod_{a\in A} N_a$ such that

$$\pi_a(y) = \begin{cases} 0 & \text{for } a = b, \\ g_a\pi_{I_a}(x) & \text{for } a \neq b. \end{cases}$$

One then obtains for $m \in M_j$

$$\pi_a g(\iota_j(m) + x) = \begin{cases} g_b\pi_{I_b}(\iota_j(m) + x) = g_b(\iota_j(m) + \pi_{I_b}(x)) & \text{for } a = b, \\ g_a\pi_{I_a}(x) = \pi_a(y) & \text{for } a \neq b. \end{cases}$$

Hence

$$hg(\iota_j(m) + x) = h\big(\iota_b\big(g_b(\iota_j(m) + \pi_{I_b}(x)\big) + y\big) \ ,$$

and the map $M_j \to N$, $m \mapsto hg(\iota_j(m) + x)$ is linear as the composition of two linear maps. □

4.3.5 Lemma *Assume to be given a non-empty family of R-modules $(M_i)_{i\in I}$ and a partition $(I_a)_{a\in A}$ of the index set I. Then there exists a natural ismorphism*

$$\kappa_{I,A} : \prod_{i\in I} M_i \to \prod_{a\in A} \prod_{i\in I_a} M_i$$

uniquely determined by the condition that $\pi_a \circ \kappa_{I,A} = \pi_{I_a}$ for all $a \in A$.

Proof. By the universal property of the product the R-module map $\kappa = \kappa_{I,A} : \prod_{i\in I} M_i \to \prod_{a\in A} \prod_{i\in I_a} M_i$ exists and is uniquely determined by the requirement that $\pi_a \circ \kappa_{I,A} = \pi_{I_a}$ for all $a \in A$. Naturality also follows from the universal property of the product. It remains to show that κ is an isomorphism. By construction, $\pi_i(x) = \pi_i\pi_a\kappa(x) = 0$ for all $i \in I$ and $a(i) \in A$ such that $i \in I_{a(i)}$, hence $x = 0$. So κ is injective. It is also surjective. To see this pick $x_a \in \prod_{i\in I_a} M_i$ for each $a \in A$. With $a(i)$ for $i \in I$ defined as before put $x = \big(\pi_i(x_{a(i)})\big)_{i\in I}$. Then, by construction, $\pi_i\pi_a\kappa(x) = \pi_i\pi_a(x) = \pi_i(x) = \pi_i(x_a)$ for all $a \in A$ and $i \in I_a$, hence $\big(\pi_a\kappa(x)\big)_{a\in A} = (x_a)_{a\in A}$ and κ is surjective. □

4.3.6 Proposition (Exponential law for multilinear maps) *Let $(M_i)_{i\in I}$ be a family of R-modules over a commutative ring R, N an R-module, and assume that $J \subset I$ is a non-empty subset such that the complement $K = I\backslash J$ is also non-empty. Then the map*

$$\eta_{I,J} : \mathfrak{Mlin}\left(\prod_{j\in J} M_j, \mathfrak{Mlin}\left(\prod_{k\in K} M_k, N\right)\right) \to \mathfrak{Mlin}\left(\prod_{i\in I} M_i, N\right),$$

$$f \mapsto \left(\prod_{i\in I} M_i \ni (x_i)_{i\in I} \mapsto f\big((x_j)_{j\in J}\big)\,((x_k)_{k\in K}) \in N\right)$$

is an isomorphism which is natural in $(M_i)_{i\in I}$ and N.

Proof. We first show that $\eta = \eta_{I,J}$ is linear. To this end let $f, g \in \mathfrak{Mlin}\Big(\prod_{j\in J} M_j, \mathfrak{Mlin}\left(\prod_{k\in K} M_k, N\right)\Big)$ and $r \in R$. Then, for all $x = (x_i)_{i\in I} \in \prod_{i\in I} M_i$,

$$\begin{aligned}\big(\eta(f+g)\big)(x) &= \big(f+g\big)\big((x_j)_{j\in J}\big)\,((x_k)_{k\in K}) = \big(f((x_j)_{j\in J}) + g((x_j)_{j\in J})\big)\,((x_k)_{k\in K}) = \\ &= f\big((x_j)_{j\in J}\big)\,((x_k)_{k\in K}) + g\big((x_j)_{j\in J}\big)\,((x_k)_{k\in K}) = \big(\eta f\big)(x) + \big(\eta g\big)(x) = \big(\eta f + \eta g\big)(x)\end{aligned}$$

and

$$\begin{aligned}\big(\eta(rf)\big)(x) &= (rf)((x_j)_{j\in J})\,((x_k)_{k\in K}) = \big(rf((x_j)_{j\in J})\big)\,((x_k)_{k\in K}) = r\big(f((x_j)_{j\in J})\,((x_k)_{k\in K})\big) = \\ &= r\big(\eta f(x)\big) = \big(r(\eta f)\big)(x)\ .\end{aligned}$$

Hence η is an R-module map.

Next we show that η is an isomorphism by constructing an inverse. Given $f \in \mathfrak{Mlin}\big(\prod_{i\in I} M_i, N\big)$ we define $f^\sharp : \mathfrak{Mlin}\big(\prod_{j\in J} M_j\big) \to \mathfrak{Mlin}\big(\prod_{k\in K} M_k, N\big)$ by the requirement that

$$f^\sharp(y)(z) = f(x_{y,z}) \quad \text{for all } \ y = (y_j)_{j\in J} \ \text{ and } \ z = (z_k)_{k\in K}\ ,$$

where $x_{y,z}$ is the element of $\prod_{i\in I} M_i$ uniquely determined by

$$\pi_i(x_{y,z}) = \begin{cases} y_i & \text{for } \ i \in J, \\ z_i & \text{for } \ i \in K. \end{cases}$$

One thus obtains an R-module map

$$(-)^\sharp_{I,J} : \mathfrak{Mlin}\left(\prod_{i\in I} M_i, N\right) \to \mathfrak{Mlin}\left(\prod_{j\in J} M_j, \mathfrak{Mlin}\left(\prod_{k\in K} M_k, N\right)\right), \quad f \mapsto f^\sharp$$

which by construction is inverse to $\eta_{I,J}$.

Naturality of $\eta_{I,J}$ in $(M_j)_{j\in J}$ and N is clear by definition. □

4.3.7 Definition Let $(M_i)_{i\in I}$ be a family of R-modules over a commutative ring R. By a *tensor product* of $(M_i)_{i\in I}$ one understands an R-module $\bigotimes_{i\in I} M_i$ together with a multilinear map $\tau : \prod_{i\in I} M_i \to \bigotimes_{i\in I} M_i$ such that the following universal property is fulfilled:

(ITensor) For every R-module N and every multilinear map $f : \prod_{i\in I} M_i \to N$ there exists a unique R-module map $\overline{f} : \bigotimes_{i\in I} M_i \to N$ such that the diagram

$$\begin{tikzcd} \prod\limits_{i\in I} M_i \arrow[r, "f"] \arrow[d, "\tau"'] & N \\ \bigotimes\limits_{i\in I} M_i \arrow[ur, "\overline{f}"'] & \end{tikzcd}$$

commutes.

The linear map $\overline{f}$ making the diagram comute will sometimes be called the *linearization* of the multilinear map f.

Given a tensor product $\left(\bigotimes_{i\in I} M_i, \tau\right)$, we will usually denote the image of an element $(x_i)_{i\in I} \in \prod_{i\in I} M_i$ under the map τ by $\otimes_{i\in I} x_i$.

4.3.8 Remarks (a) Strictly speaking, a tensor product of a family $(M_i)_{i\in I}$ of R-modules is a pair $\left(\bigotimes_{i\in I} M_i, \tau\right)$ having the above properties. By slight abuse of language, one usually denotes a tensor product just by its first component, the R-module $\bigotimes_{i\in I} M_i$. When helpful for clarity, the associated map $\tau : \prod_{i\in I} M_i \to \bigotimes_{i\in I} M_i$ will be denoted by $\tau_{(M_i)_{i\in I}}$ or by τ_I.

(b) In the case where the index set I of the family $(M_i)_{i\in I}$ is infinite, one sometimes calls $\bigotimes_{i\in I} M_i$ an *infinite tensor product*.

4.3.9 Theorem *Let $(M_i)_{i\in I}$ be a family of R-modules over a commutative ring R. Then the following holds true.*

(i) *A tensor product $\bigotimes_{i\in I} M_i$ of the family $(M_i)_{i\in I}$ exists and is unique up to isomorphism. If I is the empty set, then $\bigotimes_{i\in I} M_i = R$, if I contains a single element $i_\circ$, then $\bigotimes_{i\in I} M_i = M_{i_\circ}$.*

(ii) *If $(N_i)_{i\in I}$ is a second family of R-modules and $(f_i)_{i\in I}$ a family R-module maps $f_i : M_i \to N_i$, then there exists a unique linear map $\bigotimes_{i\in I} f_i : \bigotimes_{i\in I} M_i \to \bigotimes_{i\in I} N_i$ making the diagram*

$$\begin{tikzcd} \prod\limits_{i\in I} M_i \arrow[r, "f"] \arrow[d, "\tau"'] & \bigotimes\limits_{i\in I} N_i \\ \bigotimes\limits_{i\in I} M_i \arrow[ur, "\bigotimes\limits_{i\in I} f_i"'] & \end{tikzcd}$$

commute, where $f : \prod_{i\in I} M_i \to \bigotimes_{i\in I} N_i$ *is the multilinear map* $(x_i)_{i\in I} \mapsto \otimes_{i\in I} f_i(x_i)$.

(iii) *Let* $J \subset I$ *be a finite non-empty subset set such that* M_j *is isomorphic to* R *for all* $j \in J$. *Denote for each* $j \in J$ *by* 1_j *the image of the unit* $1 \in R$ *under the isomorphism* $R \cong M_j$ *and by* 1_J *the family* $(1_j)_{j\in J}$. *Moreover, for every family* $y = (y_j)_{j\in J}$ *let* $\iota_{J,y} : \prod_{i\in I\setminus J} M_i \to \prod_{i\in I} M_i$ *be the map which associates to* $x \in \prod_{i\in I\setminus J} M_i$ *the family* $(x_i)_{i\in I}$ *such that* $x_i = \pi_i(x)$ *for* $i \in I\setminus J$ *and* $x_i = y_i$ *for* $i \in J$. *Then the linearization* $\bar\iota_{J,1_J} : \bigotimes_{i\in I\setminus J} M_i \to \bigotimes_{i\in I} M_i$ *of the multilinear map* $\tau_I \circ \iota_{J,1_J} : \prod_{i\in I\setminus J} M_i \to \bigotimes_{i\in I} M_i$ *is an isomorphism.*

Proof. ad (i). By its universal property, the tensor product of the family $(M_i)_{i\in I}$ is uniquely determined up to isomorphism. Hence it remains to show the existence of the tensor product. To this end consider the free R-module over the set $\prod_{i\in I} M_i$ and denote it by F. Let $\delta : \prod_{i\in I} M_i \hookrightarrow F$ be the canonical injection and U be the submodule of F spanned by the elements

$$\delta\big(\iota_j(ry_j + z_j) + (x_i)_{i\in I}\big) - r\delta\big(\iota_j(y_j) + (x_i)_{i\in I}\big) - \delta\big(\iota_j(z_j) + (x_i)_{i\in I}\big) ,$$

where $j \in I$, $y_j, z_j \in M_j$, $r \in R$, and $(x_i)_{i\in I} \in \pi_j^{-1}(0)$. Then put $\bigotimes_{i\in I} M_i = F/U$ and define τ as the composition of the canonical projection $\pi : F \to \bigotimes_{i\in I} M_i$ with $\delta : \prod_{i\in I} M_i \to F$. By construction, τ is multilinear. Assume that N is an R-module and $f : \prod_{i\in I} M_i \to N$ is a multilinear map. By the universal property of free R-modules, f lifts to a unique R-linear map $f' : F \to N$ such that $f = f' \circ \delta$. By multilinearity of f, the map f' vanishes on the submodule U, hence descends to an R-linear $\overline{f} : \bigotimes_{i\in I} M_i \to N$ such that $f' = \overline{f} \circ \pi$. Hence $f = f' \circ \delta = \overline{f} \circ \pi \circ \delta = \overline{f} \circ \tau$. By surjectivity of δ and uniqueness of f', $\overline{f}$ is the unique R-linear map satisfying $f = \overline{f} \circ \tau$. Hence $\big(\bigotimes_{i\in I} M_i, \tau\big)$ is a tensor product of the family $(M_i)_{i\in I}$.

In case $I = \varnothing$, the cartesian product $\prod_{i\in I} M_i$ is final in the category of sets, hence consists of only one element $\star$ only. This means in particular that for an R-module N any map $f : \prod_{i\in I} M_i = \{\star\} \to N$ is multilinear. Put $\bigotimes_{i\in I} M_i = R$ and let $\tau : \{\star\} \to R$ be the map $\star \mapsto 1$. Now let $\overline{f} : R \to N$ be the unique linear map such that $\overline{f}(1) = f(\star)$. Then $f = \overline{f} \circ \tau$ and the pair (R, τ) fulfills the universal property of the tensor product.

If I is a singleton with unique element i_0, then $\prod_{i\in I} M_i = M_{i_0}$ and a map $f : \prod_{i\in I} M_i \to N$ is multilinear if and only if f as a map from $M_{i_\circ}$ to N is linear. This implies that the pair $(M_{i_0}, \mathrm{id}_{M_{i_\circ}})$ then is a tensor product for the family $(M_i)_{i\in I}$.

ad (ii). This is an immediate consequence of the universal property of the tensor product.

ad (iii). We construct an inverse to $\bar\iota_{J,1_J} : \bigotimes_{i\in I\setminus J} M_i \to \bigotimes_{i\in I} M_i$. Let $x = (x_i)_{i\in I}$ be an element of $\prod_{i\in I} M_i$ and put

$$\lambda(x) = \left(\prod_{j\in J} x_j\right) \cdot \otimes_{i\in I\setminus J} x_i \left(\prod_{j\in J} x_j\right) \cdot \tau_{I\setminus J}((x_i)_{i\in I\setminus J}) .$$

Then $\lambda : \prod_{i\in I} M_i \to \bigotimes_{i\in\setminus J} M_i$ is multilinear by construction, hence factors through a linear map $\overline{\lambda} : \bigotimes_{i\in I} M_i \to \bigotimes_{i\in I\setminus J} M_i$. By definition, $\overline{\lambda}$ is a left inverse of $\bar\iota_{J,1_J}$. It is also a right inverse since for all $(x_i)_{i\in I} \in \prod_{i\in I} M_i$ by multilinearity of τ_I

$$\bar\iota_{J,1_J} \circ \overline{\lambda} \circ \tau_I\left((x_i)_{i\in I}\right) = \bar\iota_{J,1_J}\left(\left(\prod_{j\in J} x_j\right) \cdot \otimes_{i\in I\setminus J} x_i\right) = \left(\prod_{j\in J} x_j\right) \cdot \left(\bar\iota_{J,1_J} \circ \tau_{I\setminus J}\left((x_i)_{i\in I\setminus J}\right)\right) =$$
$$= \left(\prod_{j\in J} x_j\right) \cdot \left(\tau_I \circ \iota_{J,1_J}\left((x_i)_{i\in I\setminus J}\right)\right) = \tau_I \circ \iota_{J,(x_j)_{j\in J}}\left((x_i)_{i\in I\setminus J}\right) = \tau_I\left((x_i)_{i\in I}\right)$$

and since by construction of the tensor product the image of τ_I is a generating system for the R-module $\bigotimes_{i\in I} M_i$. □

4.3.10 Lemma *Assume that $(M_i)_{i\in I}$ is a finite family of R-modules such that for every $i \in I$ a generating set S_i of the R-module M_i has been given. Then the set $S = \tau\left(\prod_{i\in I} S_i\right)$ is a generating set of the tensor product $\bigotimes_{i\in I} M_i$.*

Proof. By construction of the tensor product in the proof of Theorem 4.3.9 it is clear that a generating set of $\bigotimes_{i\in I} M_i$ is given by the set of elements of the form $\otimes_{i\in I} x_i$ where $(x_i)_{i\in I} \in \prod_{i\in I} M_i$. Each of the x_i can now be represented in the form

$$x_i = \sum_{k=1}^{n_i} r_{i,k} s_{i,k} \quad \text{with } r_{i,1}, \ldots, r_{i,n_i} \in R,\ \ s_{i,1}, \ldots, s_{i,n_i} \in S_i\ .$$

Hence, by multilinearity of τ and with $I = \{i_1, \ldots, i_d\}$,

$$\otimes_{i\in I} x_i = \tau\left((x_i)_{i\in I}\right) = \sum_{k_{i_1}=1}^{n_{i_1}} \cdots \sum_{k_{i_d}=1}^{n_{i_d}} r_{i_1,k_{i_1}} \cdot \ldots \cdot r_{i_d,k_{i_d}} \cdot \tau\left((s_{i,k_i})_{i\in I}\right)\ ,$$

so $\otimes_{i\in I} x_i$ is a linear combination of elements of S and the claim is proved. □

4.3.11 Lemma *Let $(M_i)_{i\in I}$ be a family of R-modules, $(I_a)_{a\in A}$ a finite partition of the index set I, and N an R-module. For $a \in A$ put $N_a = \bigotimes_{i\in I_a} M_i$ and let $\tau_a : \prod_{i\in I_a} M_i \to N_a$ denote the canonical map. Assume that $f : \prod_{a\in A}\prod_{i\in I_a} M_i \to N$ is a map which is* componentwise multilinear *in the following sense.*

(CM) *Let $b \in A$ and $y = (y_a)_{a\in A} \in \prod_{a\in A}\prod_{i\in I_a} M_i$ a family with $y_b = 0$. If for all $j \in I_b$ and families $x = (x_i)_{i\in I_b} \in \prod_{i\in I_b} M_i$ with $x_j = 0$ the map*

$$M_j \to N,\ \ m \mapsto f(\iota_b(\iota_j(m) + x) + y)$$

is linear, then f factors through $(\tau_a)_{a\in A} : \prod_{a\in A}\prod_{i\in I_a} M_i \to \prod_{a\in A} N_a$. More precisely, there exists a unique multilinear map $\overline{f} : \prod_{a\in A} N_a \to N$ such that

$$f = \overline{f} \circ (\tau_a)_{a\in A}\ .$$

Proof. We prove the claim by induction on the cardinality of A. If A is a singleton, then $\prod_{a\in A}\prod_{i\in I_a} M_i$ canonically coincides with $\prod_{i\in I} M_i$ and $f : \prod_{i\in I_a} M_i \to N$ is multilinear, hence by the universal property of the tensor product there exists a unique linear map $\overline{f} : N_a \to N$ such that $f = \overline{f} \circ \tau_a$.

Now assume that the claim holds whenever the cardinality of the index set A is $\leqslant n$ for some $n \in \mathbb{N}^*$. Assume to be given initial data $(M_i)_{i\in I}$ and N, a partition $(I_a)_{a\in A}$ of A with $|A| = n+1$ and componentwise multilinear map $f : \prod_{a\in A}\prod_{i\in I_a} M_i \to N$. Fix $a \in A$ and put $B = A\backslash\{a\}$. Let $x = (x_i)_{i\in I_a} \in \prod_{i\in I_a} M_i$ and $\widetilde{x}$ be the element of $\prod_{d\in A}\prod_{i\in I_d} M_i$ such that

$$\pi_d(\widetilde{x}) = \begin{cases} x & \text{for } d = a \ , \\ 0 & \text{else} \ . \end{cases}$$

The map

$$f_x : \prod_{b\in B}\prod_{i\in I_b} M_i \to N, \ \ y \mapsto f(\iota_B(y) + \widetilde{x})$$

then is componentwise multilinear. Hence by inductive assumption there exists a unique multilinear map $\overline{f_x} : \prod_{b\in B} N_b \to N$ such that $f_x = \overline{f_x} \circ (\tau_b)_{b\in B}$. By assumption on f the map $\prod_{i\in I_a} M_i \to \mathfrak{Map}\left(\prod_{b\in B}\prod_{i\in I_b} M_i, N\right)$, $x \mapsto f_x$ is multilinear which implies multilinearity of

$$\overline{f_\bullet} : \prod_{i\in I_a} M_i \to \mathfrak{Mlin}\left(\prod_{b\in B} N_b, N\right), \ \ x \mapsto \overline{f_x} \ .$$

Let $F : N_a \to \mathfrak{Mlin}\left(\prod_{b\in B} N_b, N\right)$ be its linearization. Application of the exponential law for multilinear maps, Proposition 4.3.6, now gives a multilinear map $\eta(F) : \prod_{d\in A} N_d \to N$ which we denote by $\overline{f}$. Given a family $(x_d)_{d\in A}$ of families $x_d = (x_i)_{i\in I_d}$ one checks

$$\overline{f}\left(\left(\tau_d(x_d)\right)_{d\in A}\right) = F\big(\tau_a(x_a)\big)\left(\left(\tau_b(x_b)\right)_{b\in B}\right) = \overline{f}_{x_a}\left(\left(\tau_b(x_b)\right)_{b\in B}\right) = f_{x_a}\left((x_b)_{b\in B}\right) = f\left((x_d)_{d\in A}\right) \ .$$

Hence $\overline{f} \circ (\tau_d)_{d\in A} = f$. To finish the induction step it remains to prove uniqueness. So let $\overline{g} : \prod_{d\in A} N_d \to N$ be another multilinear map such that $\overline{g} \circ (\tau_d)_{d\in A} = f$ and consider the induced linear map $\overline{g}^\sharp = \eta^{-1}(\overline{g}) : N_a \mapsto \mathfrak{Mlin}(\prod_{b\in B} N_b, N)$. Then for every $x \in \prod_{i\in I_a} M_i$ the relation

$$\overline{g}^\sharp(\tau_a(x)) \circ (\tau_b)_{b\in B} = f_x = \overline{f}_x \circ (\tau_b)_{b\in B}$$

is satisfied. Hence $\overline{g}^\sharp(\tau(x)) = \overline{f}_x$ for all $x \in \prod_{i\in I_a} M_i$ which entails that $\overline{g}^\sharp$ coincides with F. By Proposition 4.3.6 one obtains $\overline{g} = \overline{f}$. This finishes the induction step and the lemma is proved.□

4.3.12 Proposition *Let $(M_i)_{i\in I}$ be a family of R-modules and $(I_a)_{a\in A}$ a finite partition of the index set I. Then there exists a natural isomorphism*

$$\alpha_{I,A} : \bigotimes_{i\in I} M_i \to \bigotimes_{a\in A}\bigotimes_{i\in I_a} M_i.$$

Proof. Put $N_a = \bigotimes_{i\in I_a} M_i$ for $a \in A$ and let $\tau_a : \prod_{i\in I_a} M_i \to N_a$ be the canonical map to the tensor product. Let $\tau_A : \prod_{a\in A} N_a \to \bigotimes_{a\in A} N_a$ be the canonical map to the tensor product of the modules N_a. Define $\tau_{I,A} : \prod_{i\in I} M_i \to \prod_{a\in A} N_a$ as the unique map so that $\pi_a \circ \tau_{I,A} = \tau_a \circ \pi_{I_a}$ for all $a \in A$. By construction $\tau_{I,A} = (\tau_a)_{a\in A} \circ \kappa_{I,A}$, where $\kappa_{I,A} : \prod_{i\in I} M_i \to \prod_{a\in A}\prod_{i\in I_a} M_i$ is the natural isomorphism from Lemma 4.3.5. The composition $\tau_A \circ \tau_{I,A}$ then is multilinear by Lemma 4.3.4 (iii), hence factors through a linear map $\alpha_{I,A} : \bigotimes_{i\in I} M_i \to \bigotimes_{a\in A} N_a$ that is

$$\tau_A \circ (\tau_a)_{a\in A} \circ \kappa_{I,A} = \alpha_{I,A} \circ \tau_I \ . \tag{4.3.1}$$

Naturality of $\alpha_{I,A}$ in $(M_i)_{i\in I}$ is clear by definition so it remains to construct an inverse to $\alpha_{I,A}$. Consider the composition $\tau_I \circ \kappa^{-1} : \prod_{a\in A}\prod_{i\in I_a} M_i \to \bigotimes_{i\in I} M_i$. Assume that $a \in A$ and $(y_b)_{b\in A\setminus\{a\}} \in \prod_{b\in A\setminus\{a\}}\prod_{i\in I_b} M_i$ have been chosen. Let $y_a \in \prod_{i\in I_a} M_i$ be 0, put $\widetilde{y} = (y_d)_{d\in A} \in \prod_{d\in A}\prod_{i\in I_d} M_i$, and let $y \in \prod_{i\in I} M_i$ be the family such that $\pi_i(y) = \pi_i(y_{a(i)})$ for all $i \in I$, where $a(i)$ denotes the unique element of A such that $i \in I_{a(i)}$. In other words let $y = \kappa^{-1}(\widetilde{y})$. For every $j \in I_a$ and $x = (x_i)_{i\in I_a} \in \prod_{i\in I_a} M_i$ with $\pi_j(x) = 0$ the map

$$M_j \to \bigotimes_{i\in I} M_i,\ \ m \mapsto \tau_I \circ \kappa^{-1}\left(\iota_a(\iota_j(m) + x) + \widetilde{y}\right) = \tau_I\left(\iota_j(m) + \iota_{I_a}(x) + y\right)$$

then is multilinear since τ_I is multilinear and $\pi_j(\iota_{I_a}(x) + y) = \pi_j(x) + \pi_j(y_a) = 0$. Hence $\tau_I \circ \kappa^{-1}$ is componentwise multilinear and therefore, by Lemma 4.3.11, factors through the map $(\tau_a)_{a\in A} : \prod_{a\in A}\prod_{i\in I_a} M_i \to \prod_{a\in A} N_a$ which means that

$$\tau_I \circ \kappa^{-1} = \lambda_{I,A} \circ (\tau_a)_{a\in A} \tag{4.3.2}$$

for some uniquely defined multilinear map $\lambda_{I,A} : \prod_{a\in A} N_a \to \bigotimes_{i\in I} M_i$. Let

$$\overline{\lambda}_{I,A} : \bigotimes_{a\in A} N_a \to \bigotimes_{i\in I} M_i$$

be the linearization of $\lambda_{I,A}$. We claim that $\overline{\lambda_{I,A}}$ is inverse to $\alpha_{I,A}$. By definition of $\overline{\lambda_{I,A}}$ and Eqs. (4.3.1) and (4.3.2) one concludes

$$\overline{\lambda_{I,A}} \circ \alpha_{I,A} \circ \tau_I = \overline{\lambda_{I,A}} \circ \tau_A \circ (\tau_a)_{a\in A} \circ \kappa_{I,A} = \lambda_{I,A} \circ (\tau_a)_{a\in A} \circ \kappa_{I,A} = \tau_I\ .$$

Since the image of τ_I generates $\bigotimes_{i\in I} M_i$ as an R-module, $\overline{\lambda_{I,A}}$ has to be left inverse to $\alpha_{I,A}$. Using Eqs. (4.3.1) and (4.3.2) again compute

$$\alpha_{I,A} \circ \overline{\lambda_{I,A}} \circ \tau_A \circ (\tau_a)_{a\in A} = \alpha_{I,A} \circ \lambda_{I,A} \circ (\tau_a)_{a\in A} = \alpha_{I,A} \circ \tau_A \circ \kappa_{I,A}^{-1} = \tau_A \circ (\tau_a)_{a\in A}\ .$$

Since by Lemma 4.3.10 the image of $\tau_A \circ (\tau_a)_{a\in A}$ generates $\bigotimes_{a\in A}\bigotimes_{i\in I_a} M_i$, the equality

$$\alpha_{I,A} \circ \overline{\lambda_{I,A}} = \mathrm{id}_{\bigotimes_{a\in A}\bigotimes_{i\in I_a} M_i}$$

follows and the proposition is proved. □

4.3.13 Proposition and Definition *Let $(A_i)_{i\in I}$ be a family of R-algebras. Then the tensor product $A = \bigotimes_{i\in I} A_i$ carries in a natural way the structure of an R-algebra where the product map is defined by*

$$\cdot : A \times A \to A,\ \ (\otimes_{i\in I} a_i, \otimes_{i\in I} b_i) \mapsto \otimes_{i\in I}(a_i \cdot b_i)\ .$$

In case each of the algebras A_i is commutative, then A is commutative as well. Likewise, if each A_i is unital and 1_i denotes the unit element of A_i, then A is unital with unit given by $1 = \otimes_{i\in I} 1_i$. One calls A the tensor product algebra *of the family of algebras $(A_i)_{i\in I}$.*

Proof. The map

$$\prod_{(i,k)\in I\times\{1,2\}} A_i \to A,\ \ (a_{i,k})_{(i,k)\in I\times\{1,2\}} \mapsto \otimes_{i\in I}(a_{i,1} \cdot a_{i,2})$$

is multilinear by bilinearity of the product maps on the A_i and multilinearity of τ_I, so factors through a linear map $\mu : A \otimes A \cong \bigotimes_{(i,k)\in I\times\{1,2\}} A_i \to A$. Composition of μ with the canonical bilinear map $A \times A \to A \otimes A$ gives the product map $\cdot : A \times A \to A$ and shows that the product on A is well-defined. By construction, the product map $\cdot$ is bilinear. Given $\otimes_{i\in I} a_i, \otimes_{i\in I} b_i, \otimes_{i\in I} c_i \in A$ one computes

$$\big(\otimes_{i\in I}\, a_i \cdot \otimes_{i\in I} b_i \big) \cdot \otimes_{i\in I} c_i = \otimes_{i\in I} ((a_i \cdot b_i) \cdot c_i) = \otimes_{i\in I} (a_i \cdot (b_i \cdot c_i)) = \otimes_{i\in I} a_i \cdot \big(\otimes_{i\in I}\, b_i \cdot \otimes_{i\in I} c_i \big) \ .$$

This entails that the product on A is associative. In the same way one shows that A is commutive respectively unital if each of the A_i is. □

4.3.14 As we have seen, the infinite tensor product construction works well for objects of algebraic categories like R-modules, vector spaces or R-algebras. As soon as a topologies compatible with the algebraic structure come in it becomes difficult and sometimes even impossible to construct or even define

Bibliography

Alexandroff, P. and Hopf, H. (1965). *Topologie.* Erster Band. Grundbegriffe der mengentheoretischen Topologie, Topologie der Komplexe, topologische Invarianzsätze und anschliessende Begriffsbildungen, Verschlingungen im n-dimensionalen euklidischen Raum, stetige Abbildungen von Polyedern. Chelsea Publishing Co., New York.

Baez, J. (1997). Higher-Dimensional Algebra II. 2-Hilbert Spaces. *Adv. Math.*, 127:125–189.

Birkhoff, G. and von Neumann, J. (1936). The logic of quantum mechanics. *Ann. of Math. (2)*, 37(4):823–843.

Blackadar, B. E. (1977). Infinite tensor products of C^*-algebras. *Pacific J. Math.*, 72(2):313–334.

Bourbaki, N. (1998). *General topology. Chapters 1–4.* Elements of Mathematics (Berlin). Springer-Verlag, Berlin. Translated from the French, Reprint of the 1989 English translation.

Bratteli, O. and Robinson, D. W. (1997). *Operator algebras and quantum statistical mechanics. 2.* Texts and Monographs in Physics. Springer-Verlag, Berlin, second edition. Equilibrium states. Models in quantum statistical mechanics.

Chevalley, C. (1956). *Fundamental concepts of algebra.* Academic Press Inc., New York.

Emch, G. G. (2009). *Algebraic Methods in Statistical Mechanics and Quantum Field Theory.* Dover Publications, Inc., New York.

Frobenius, F. G. (1878). Über lineare Substitutionen und bilineare Formen. *Journal für die reine und angewandte Mathematik*, 84:1–63.

Gouvêa, F. Q. (1997). *p-adic numbers.* Universitext. Springer-Verlag, Berlin, second edition. An introduction.

Grothendieck, A. (1955). Produits tensoriels topologiques et espaces nucléaires. *Mem. Amer. Math. Soc.*, No. 16:140.

Guichardet, A. (1966). Produits tensoriels infinis et représentations des relations d'anticommutation. *Ann. Sci. École Norm. Sup. (3)*, 83:1–52.

Hirzebruch, F. and Scharlau, W. (1991). *Einführung in die Funktionalanalysis*, volume 296 of *B.I.-Hochschultaschenbücher [B.I. University Paperbacks].* Bibliographisches Institut, Mannheim. Reprint of the 1971 original.

Lang, S. (2002). *Algebra*, volume 211 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 3rd. edition.

Nakagami, Y. (1970a). Infinite tensor products of von Neumann algebras. I. *Kōdai Math. Sem. Rep.*, 22:341–354.

Nakagami, Y. (1970b). Infinite tensor products of von Neumann algebras. II. *Publ. Res. Inst. Math. Sci.*, 6:257–292.

Ng, C.-K. (2013). On genuine infinite algebraic tensor products. *Rev. Mat. Iberoam.*, 29(1):329–356.

Ostrowski, A. (1916). Über einige Lösungen der Funktionalgleichung $\psi(x) \cdot \psi(x) = \psi(xy)$. *Acta Math.*, 41(1):271–284.

Pietsch, A. (1972). *Nuclear locally convex spaces.* Springer-Verlag, New York-Heidelberg. Translated from the second German edition by William H. Ruckle, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 66.

Steen, L. A. and Seebach, Jr., J. A. (1995). *Counterexamples in topology.* Dover Publications, Inc., Mineola, NY. Reprint of the second (1978) edition.

Størmer, E. (1971). On infinite tensor products of von Neumann algebras. *Amer. J. Math.*, 93:810–818.

von Neumann, J. (1939). On infinite direct products. *Compositio Math.*, 6:1–77.

Licensing

GNU Free Documentation License
Version 1.3, 3 November 2008

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "publisher" means any person or entity that distributes copies of the Document to the public.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission. B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement. C. State on the Title page the name of the publisher of the Modified Version, as the publisher. D. Preserve all the copyright notices of the Document. E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices. F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below. G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice. H. Include an unaltered copy of this License. I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence. J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission. K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein. L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles. M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version. N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section. O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties–for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled “History” in the various original documents, forming one section Entitled “History”; likewise combine any sections Entitled “Acknowledgements”, and any sections Entitled “Dedications”. You must delete all sections Entitled “Endorsements”.

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an “aggregate” if the copyright resulting from the compilation is not used to limit the legal rights of the compilation’s users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

11. RELICENSING

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

```
Copyright (c)  YEAR  YOUR NAME.
Permission is granted to copy, distribute and/or modify this document
under the terms of the GNU Free Documentation License, Version 1.3
or any later version published by the Free Software Foundation;
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
A copy of the license is included in the section entitled "GNU
Free Documentation License".
```

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with… Texts." line with this:

```
with the Invariant Sections being LIST THEIR TITLES, with the
Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.
```

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.