Mathematical Methods of Quantum Mechanics

Chapter 1

Preliminaries

1.1 Metric spaces

Definition 1.1.1. *Let M be a set. A* **metric** *on M is a mapping $d : M \times M \to \mathbb{R}$ with the following properties. For all $x, y, z \in M$,*

(i) $d(x, y) \geq 0$.

(ii) $d(x, y) = 0 \Leftrightarrow x = y$.

(ii) $d(x, y) = d(y, x)$

(iii) $d(x, y) \leq d(x, z) + d(z, y)$.

A **metric space** *is a pair (M, d) where M is a set and d is a metric on M.*

Remark 1.1.1. *Let (M, d) be a metric space and $A \subset M$. Then $d_A := d \restriction A \times A$ is obviously a metric on A. It is called the* **induced metric** *on A from (M, d).*

Example 1. *Let $M = \mathbb{R}^n$. Then the following expressions are metrics on M. Let $x, y \in M$.*

(a) $d_1(x, y) := \sum_{j=1}^{n} |y_j - x_j|$ *(straight forward to verify).*

(b) $d_\infty(x, y) := \max_{j=1,\ldots,n} |y_j - x_j|$ *(straight forward to verify).*

(c) $d_2(x, y) := (\sum_{j=1}^{n} (y_j - x_j)^2)^{1/2}$. *This is called the* **Euclidean** *metric. (Follows using the properties of inner products or normes, see later.)*

Example 2. *Let $M = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. $d(x, y) = d_2(x, y)$ for $x, y \in M$ defines a metric.*

Example 3. *Let $M = C(I)$ for some compact I interval in $\mathbb{R}^d$. Then $d_\infty(f, g) := \sup_{x \in I} |f(x) - g(x)|$ and $d_1(f, g) := \int_I |f(x) - g(x)| dx$. Are metrics on M as one easily verifies.*

Definition 1.1.2. *Let (M,d) be a metric space. A sequence $(x_n)_{n=0}^{\infty}$ in M is said to* **converge** *to an element $x \in M$, written*

$$x = \lim_{n\to\infty} x_n \quad \text{or} \quad x_n \overset{(n\to\infty)}{\longrightarrow} x \quad \text{or} \quad x_n \overset{d}{\longrightarrow} x,$$

iff for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $d(x, x_n) < \epsilon$ for all $n \geq N$. In that case x is called the **limit** *of (x_n).*

Lemma 1.1.1. *Let (M,d) be a metric space, (x_n) a sequence in M, and $x \in M$. Then the following holds.*

(a) We have $x_n \to x$ iff $d(x_n, x) \to 0$.

(b) If $y \in M$, $x_n \to y$, and $x_n \to x$, then $x = y$.

Proof. (a) Immediate consequence of the definition. (b) $d(x,y) \leq d(x,x_n) + d(x_n,y) \to 0$. □

Definition 1.1.3. *A sequence $(x_n)_{n\in\mathbb{N}}$ is called a* **Cauchy sequence** *if for all $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that $d(x_n, x_m) < \epsilon$ for all $m, n \geq N$.*

Proposition 1.1.2. *Any convergent sequence is a Cauchy sequence.*

Proof. Given $x_n \to x$ and $\epsilon > 0$, find N so $n \geq N$ impies $d(x_n, x) < \epsilon/2$. Then $n, m \geq N$ implies $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{1}{2}\epsilon + \frac{1}{2}\epsilon$. □

Remark 1.1.2. *Note that the metric d_1 restricted to $\mathbb{Q}$ is not dense.*

Definition 1.1.4. *A metric space in which all Cauchy sequences converge is called* **complete**.

Definition 1.1.5. *A set D in a metric space M is called* **dense** *if every $m \in M$ is a limit of elements in D.*

Remark 1.1.3. *$\mathbb{Q}$ is dense in $\mathbb{R}$.*

Definition 1.1.6. *Suppose X and Y are metric spaces and $f : X \to Y$.*

(a) f is called **continuous at** *$p \in M$ if for all $\epsilon > 0$ there exists a $\delta > 0$ such that $d_Y(f(x), f(p)) < \epsilon$ for all $x \in X$ for which $d_X(x, p) < \delta$.*

(b) f is called **continuous** *if it is continuous at every point of X.*

Proposition 1.1.3. *(Sequence Criterion) A function f from a metric space (X, d_X) to a metric space (Y, d_Y) is continuous at $p \in X$ if $f(x_n) \xrightarrow{d_Y} f(p)$ whenever $x_n \xrightarrow{d_X} p$.*

Proof. Analogous to the proof for functions in $\mathbb{R}$. □

Definition 1.1.7. *A bijection ϕ from (X, d_X) to (Y, d_Y) which preserves the metric, that is,*

$$d_Y(\phi(x_1), \phi(x_2)) = d_X(x_1, x_2) \quad \forall x_1, x_2 \in X$$

is called an **isometry** *(it is automatically continuous). (X, d_X) and (Y, d_Y) are said to be* **isometric** *if such an isometry exists.*

Theorem 1.1.4. *Let (M, d) be a metric space, which is not complete. Then there exists a metric space $(\tilde{M}, \tilde{d})$ so that M is isometric to a dense subset of $\tilde{M}$. This space $\tilde{M}$ is unique except for isometries.*

Proof. Define the following equivalence relation on the set of Cauchy sequences in M. Let (x_n) and (y_n) be two Cauchy sequences. They are called equivalent, $(x_n) \sim (y_n)$, if

$$\lim_{n\to\infty} d(x_n, y_n) = 0.$$

Define

$$\begin{aligned}
&\tilde{M} := \{[(x_n)]_\sim : (x_n) \text{ Cauchy in } M\} \\
&\tilde{d}([(x_n)]_\sim, [(y_n)]_\sim) = \lim_{n\to\infty} d(x_n, y_n) \\
&i(x) = (x)_n.
\end{aligned}$$

It is straightforward to verify that this definition satisfies the stated properties. □

To complete our discussion of metric spaces, we want to introduce the notations of open and closed sets. The reader should keep the examples of open and closed sets on the real line in mind.

Definition 1.1.8. *Let (M, d) be a metric space.*

(a) Let $x_0 \in M$ and $r > 0$. Then

$$U_r(x) := \{y \in M : d(x, y) < r\} \quad \text{and} \quad B_r(x) := \{y \in M : d(x, y) \leq r\}$$

are called **open ball** *and* **closed ball centered** *x with radius r, respectively.*

(b) A set $O \subset M$ is called **open** *if $\forall x \in O, \exists r > 0, U_r(x) \subset O$.*

(c) A set $V \subset M$ is called a **neighborhood** *of y if $\exists r > 0$, $U_r(y) \subset V$.*

(d) A point p is called **interior point** *of E if $\exists r > 0$, $U_r(x) \subset E$.*

Remark 1.1.4. We note the following about the above definitions.

(a) An open ball is indeed open (an application of the triangle inequality) A closed ball is indeed closed (will follow since the complement is open (see below) by an application of the triangle inequality).

(b) A set is open if and only if it is a neighborhood of each of its points. (Obvious)

(c) Every set is a neighborhood of its interior points. (Obvious)

(d) A set is open if and only if all its points are interior points. (Obvious)

Definition 1.1.9. *Let (M, d) be a metric space.*

(a) Let $E \subset M$. A point $x \in M$ is called a **limit point** *of E if for all $r > 0$ the open ball $U_r(x)$ contains an element of E different from x (i.e. $\forall r > 0, U_r(x) \cap E \setminus \{x\} \neq \emptyset$). We denote by $\mathrm{limpts}(E)$ the set of limit points of E.*

(b) A set $C \subset M$ is called **closed** *if $\mathrm{limpts}(C) \subset C$.*

(c) *For $E \subset M$ the set $\overline{E} := E \cup \text{limpts}(E)$ is called the* **closure** *of E.*

A limit point can be characterized using the notion of sequences. This allows equivalent characterizations of the definitions given in 1.1.9, which are often more useful in applications.

Lemma 1.1.5. *Let (M, d) be a metric space.*

(a) *Let $E \subset M$. Then $x \in M$ is a limit point of E if and only if there exists a sequence (x_n) in $E \setminus \{x\}$ such that $x_n \to x$.*

(b) *C is closed if and only if for any sequence (x_n) in C with $x_n \to x$ it follows that $x \in C$.*

(c) *We have $\overline{A} = \{x \in M : \exists (x_n), x_n \in A, x_n \to x\}$.*

Proof. (a) if: trivial. only if: For every $n \in \mathbb{N}$ we can choose an element $x_n \in U_{1/n}(x) \cap E \setminus \{x\}$. Then $x_n \to x$.
(b) if: Let $x \in \text{limpts}(C)$. Then there exists by (a) a sequence in C which converges to x. But then by assumption $x \in C$. Hence C is closed.
only if: Let (x_n) be a sequence in C with $x_n \to x$. If x were not an element of C, then it would be a limit point of C. By closedness of C this would imply $x \in C$, a contradition.
(c) Follows from (a).
(d) Follows from (c). □

Lemma 1.1.6. *Let (X, d) be a metric space. A closed subspace A of a complete metric space X is complete.*

Remark 1.1.5. Let (M, d) be a metric space. A set D in a metric space M is dense if and only if $\overline{D} = M$.
Proof: This follows immediately from the previous lemma part (c).

The reader can prove for himself the following collection of elementary statements.

Theorem 1.1.7. *Let (M, d) be a metric space.*

(a) A set is open if and only if its complement is closed.

(b) A set is closed if and only if its complement is open.

(c) M and $\emptyset$ are open. The intersection of two open sets is open. Arbitrary unions of open sets are open.

(d) M and $\emptyset$ are closed. The union of two closed sets is closed. Arbitrary intersections of closed sets are closed.

Proof. (a) First, suppose E^c is closed. Let $x \in E$. Then $x \notin E^c$, and x is not a limit point of E^c. Hence there exists an $r > 0$ such that $U_r(x) \cap E^c$ is empty, that is, $U_r(x) \subset E$. Thus E is open.
Next, suppose E is open. Let x be a limit point of E^c. Then for every $r > 0$ we have $U_r(x) \cap E^c \neq \emptyset$. Thus $x \neq E$ and $x \in E^c$. Hence E^c is closed.
(b) follows from (a).
(c) It is trivial that M and $\emptyset$ are open. Let O_1 and O_2 be open. If $x \in O_1 \cap O_2$, then there exist positive r_1, r_2 such that $U_{r_1}(x) \subset O_1$ and $U_{r_2}(x) \subset O_2$. It follows that $x \in U_{\min\{r_1,r_2\}} \subset O_1 \cap O_2$. Hence $O_1 \cap O_2$ are open. Let $(O_j)_{j\in J}$ be a family of open sets. Let $x \in \bigcup_{j\in J} O_j =: U$. Then $x \in O_l$ for some $l \in J$. Hence ther exists an $r > 0$ such that $U_r(x) \subset O_l \subset U$. Hence U is open.
(d) This follows from (c) and de Morgan's laws. □

Theorem 1.1.8. *Let (M, d) be a metric space and $E \subset M$. Then the following holds.*

(a) $\overline{E}$ is closed. In particular, $\overline{E}$ is a closed set containing E.

(b) $E = \overline{E}$ if and only if E is closed.

(c) $\overline{E} \subset F$ for every closed set F containing E we have

(d) Let $\mathcal{C} = \{C \subset M : C$ closed $, E \subset C\}$. Then $\overline{E} = \bigcap_{C\in\mathcal{C}} C$.

By (a) and (c), $\overline{E}$ is the **smallest** *closed subset of X that contains E.*

Proof. (a) If $p \in X$ and $p \notin \overline{E}$ then p is neither a point of E nor a limit point of E. Hence p has a neighborhood which does not intersect E. The complement of $\overline{E}$ is therefore open. Hence $\overline{E}$ is closed.
(b) If $E = \overline{E}$. Then (a) implies that E is closed. On the other hand, if E is closed, then $\text{limpts}(E) \subset E$, and hence $\overline{E} = E \cup \text{limpts}(E) = E$.
(c) If F is closed and $F \supset E$, then $\text{limpts}(F) \subset F$ and $\text{limpts}(E) \subset \text{limpts}(F)$, hence $\overline{E} = E \cup \text{limpts}(E) \subset F$.
(d) $\subset$: By (d) of the above theorem, it follows that $\bigcap_{C \in \mathcal{C}} C$ is a closed set which by definition of the intersection contains E. Thus the inclusion follows from (c). $\supset$ By (a) we know that $\overline{E} \in \mathcal{C}$. Thus the inclusion now follows from the definition of the intersection. $\square$

Theorem 1.1.9. *Let $f : X \to Y$ be a function from a metric space X to a metric space Y. Then the following holds.*

(a) f is continuous at x if and only if for every neighborhood V of $f(x)$ the set $f^{-1}(V)$ is a neighborhood of x.

(b) f is continuous if and only if for all open sets $O \subset Y$ the set $f^{-1}(O)$ is open.

Proof. (a). $\Rightarrow$ Let O be a neighborhood of $f(x)$. Then for some $r > 0$ we have $U_r(f(x)) \subset V$. By continuity in x there exist a $\delta > 0$ such that $y \in U_\delta(x)$ implies $f(y) \in U_r(f(x))$. It follows that $U_\delta(x) \subset f^{-1}(U_r(f(x)) \subset f^{-1}(V)$. Thus $f^{-1}(V)$ is a neighborhood of x.
$\Leftarrow$. Let $\epsilon > 0$. Then $U_\epsilon(f(x))$ is a neighborhood of $f(x)$. By assumption $f^{-1}(U_\epsilon(f(x)))$ is a neighborhood of x. Thus there exists a $\delta > 0$ such that $U_\delta(x) \subset f^{-1}(U_\epsilon(f(x)))$. This implies $f(y) \in U_\epsilon(f(x))$ for all $y \in U_\delta(x)$. Hence f is continuous in x.
(b) $\Rightarrow$ Let $O \subset Y$ be open. Let $x \in f^{-1}(O)$. Then $f(x) \in O$, so O is a neighborhood of $f(x)$. It follows from (a) that $f^{-1}(O)$ is a neighborhood of x. So $f^{-1}(O)$ is open.
$\Leftarrow$. Let $x \in X$. Let V be a neighborhood of $f(x)$. Then for some $r > 0$ we have $U_r(f(x)) \subset V$. By assumption $f^{-1}(U_r(f(x)))$ is open. Since $x \in f^{-1}(U_r(f(x))) \subset f^{-1}(V)$,

it follows that $f^{-1}(V)$ is a neighborhood of x. It now follows from (a) that f is continuous in x. Since $x \in X$ was arbitrary, f is continuous. $\square$

Lemma 1.1.10. *Let* (X, d_X) *and* (Y, d_Y) *be metric spaces. Then*

$$\begin{aligned} d_{X\times Y} : (X \times Y) \times (X \times Y) &\to \mathbb{R} \\ ((x, y), (u, v)) &\mapsto d_X(x, u) + d_Y(y, v) \end{aligned}$$

is a metric on $X \times Y$*. Moreover, the following holds.*

(a) Let $((x_n, y_n))_{n\in\mathbb{N}}$ *be a sequence in* $X \times Y$*. Then* $(x_n, y_n) \to (x, y)$ *iff* $x_n \to x$ *and* $y_n \to y$*.*

(b) The projection maps $\pi_1 : X \times Y \to X, (x, y) \mapsto x$ *and* $\pi_2 : X \times Y \to Y, (x, y) \mapsto y$ *are continuous.*

Proof. The first statement is straight forward to verify.
(a) Follows from $d_{X\times Y}((x_n, y_n), (x, y)) = d_X(x_n, x) + d_Y(y_n, y)$. If: properties of limits. Only if: nonnegativity of the metric.
(b) Follos from (a). $\square$

Appendix: Topological Spaces

Definition 1.1.10. *Let X be a set. A* **topology** *on X is a subset $\mathcal{O}$ of $\mathcal{P}(X)$ with the following properties.*

(i) $\emptyset \in \mathcal{O}$ *and* $X \in \mathcal{O}$.

(ii) $\mathcal{C} \subset \mathcal{O} \Rightarrow \bigcup_{O \in \mathcal{C}} O \in \mathcal{O}$.

(iii) $O_1, O_2 \in \mathcal{O} \Rightarrow O_1 \cap O_2 \in \mathcal{O}$.

A **toplogical space** *is a pair $(X, \mathcal{O})$ where X is a set and $\mathcal{O}$ is a topology on X. For a topological space $(X, \mathcal{O})$ any element of $\mathcal{O}$ is called an* **open** *set of $(X, \mathcal{O})$.*

1.2 Normed Vector Spaces

Let $\mathbb{F}$ stand for the fields $\mathbb{C}$ or $\mathbb{R}$. We assume that the reader is familiar with vector spaces and linear transformations.

Definition 1.2.1. *Let X and Y be vectorspaces. A mapping $A : X \to Y$ is called a linear operator (mapping, transformation) if*

$$A(\lambda x + \mu y) = \lambda Ax + \mu Ay\ , \qquad \forall x, y \in X, \quad \forall \lambda, \mu \in \mathbb{F}\ .$$

Definition 1.2.2. *Let X be a vector space over $\mathbb{F}$. A norm is a mapping $\| \cdot \| : X \to \mathbb{R}$ satisfying*

(i) $\|x\| \geq 0$

(ii) $\|x\| = 0 \Rightarrow x = 0$.

(iii) $\|\lambda x\| = |\lambda| \|x\|$.

(iv) $\|x + y\| \leq \|x\| + \|y\|$.

In that case the pair $(X, \| \cdot \|)$ is called a **normed space**. *If (ii) does not hold, $\| \cdot \|$ is called a seminorm. (Observe that (iii) implies $x = 0 \Rightarrow \|x\| = 0$.)*

Remark 1.2.1. A normed space $(X, \| \cdot \|)$ is a metric space with metric $d(x, y) := \|x - y\|$. *Proof:* This is straight forward to verify.

Example 4. *Consider the vector space $\mathbb{F}^n$. Then the following expressions are norms on $\mathbb{F}^n$. Let $x \in \mathbb{F}^n$.*

(a) $\|x\|_1 := \sum_{j=1}^n |x_j|$. *Proof: Straight forward to verify.*

(b) $\|x\|_\infty := \max_{j=1,\ldots,n} |y_j - x_j|$. *Proof: Straight forward to verify.*

(c) $\|x\|_2 := (\sum_{j=1}^n |x_j|^2)^{1/2}$. *Proof: Follows using the properties of inner products see later. Or as a special case of the following example.*

(d) *For* $p \geq 1$, $\|x\|_p := (\sum_{j=1}^n |x_j|^p)^{1/p}$. *Proof: Minkowski, see below.*

Example 5. *Let* $I \subset \mathbb{R}^n$ *be a compact interval. Consider the vector space* $C(I) := \{f : I \to \mathbb{F} : f \text{ continuous}\}$. *Then the following expressions are norms on* $C(I)$. *Let* $f \in C(I)$.

(a) $\|f\|_1 := \int_I |f(x)|dx$. *Proof: Straight forward to verify.*

(b) $\|f\|_\infty := \max_{x\in I} |f(x)|$. *Proof: Straight forward to verify.*

(c) $\|f\|_2 := \left(\int_I |f(x)|^2 dx\right)^{1/2}$. *Proof: Follows using the properties of inner products see later. Or as a special case of the following example.*

(d) *For* $p \geq 1$, $\|f\|_p := \left(\int_I |f(x)|^p dx\right)^{1/p}$. *Proof: Minkowski, see below.*

Lemma 1.2.1. *Young's Inequality. Let* $p, q > 1$ *with* $\frac{1}{p} + \frac{1}{q}$. *Then for any* $a, b \geq 0$ *we have*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Proof. Follows from the convexity of $x \mapsto e^x$.

$$ab = e^{\ln a + \ln b} = e^{p^{-1}\ln a^p + q^{-1}\ln b^q} \leq \frac{e^{\ln a^p}}{p} + \frac{e^{\ln b^q}}{q}.$$

□

Theorem 1.2.2. *Hölder. Let* $\frac{1}{p} + \frac{1}{q} = 1$.

(a) *We have for* $x, y \in \mathbb{F}^n$

$$\sum_{j=1}^n |x_j y_j| \leq \|x\|_p \|y\|_q. \tag{1.2.1}$$

(b) *We have for* $f, g \in C(I)$

$$\int_I |fg|dx \leq \|f\|_p \|g\|_q. \tag{1.2.2}$$

Proof. (a) If $\|x\|_q$ or $\|y\|_p$ is zero the claim is trivial as both sides of the inquality are equal to zero. Thus suppose that both are nonzero.

$$\sum_{j=1}^{n}|x_jy_j| = \sum_{j=1}^{n}\frac{|x_j|}{\|x\|_q}\frac{|y_j|}{\|y\|_p}(\|x\|_q\|y\|_p) \le \sum_j(\frac{|x_j|^p}{p\|x\|_p^p}+\frac{|y_j|^q}{q\|y\|_q^q})\|x\|_q\|y\|_p = \|x\|_q\|y\|_p.$$

(b) is analogous to (a). If $\|f\|_q$ or $\|g\|_p$ is zero the claim is trivial as both sides of the inquality are equal to zero. Thus suppose that both are nonzero.

$$\begin{aligned}\int_I |f(x)g(x)|dx &= \int_I \frac{|f(x)|}{\|f\|_p}\frac{|g(x)|}{\|g\|_q}dx\|f\|_p\|g\|_q \\ &\le \int_I\left(\frac{|f(x)|^p}{p\|f\|_p^p}+\frac{|g(x)|^q}{q\|g\|_q^q}\right)dx\|f\|_p\|g\|_q = \|f\|_p\|g\|_q.\end{aligned}$$

□

Theorem 1.2.3. *Minkowswi. Let $p \ge 1$.*

(a) We have for $x, y \in \mathbb{F}^n$

$$\|x+y\|_p \le \|x\|_p + \|y\|_p. \tag{1.2.3}$$

(b) We have for $f, g \in C(I)$

$$\|f+g\|_p \le \|f\|_p + \|g\|_p. \tag{1.2.4}$$

Proof. (a) The result is obvious if $p = 1$ or if $x + y = 0$. Otherwise observe that

$$|x+y|^p \le (|x|+|y|)|x+y|^{p-1}$$

and apply Hölder's inequality, noting that $(p-1)q = p$ when $q^{-1} = 1 - p^{-1}$:

$$\begin{aligned}\sum_j |x_j+y_j|^p &\le \|x\|_p\||x+y|^{p-1}\|_q + \|y\|_p\||x+y|^{p-1}\|_q \\ &= (\|x\|_p+\|y\|_p)\left[\sum_j |x_j+y_j|^p\right]^{1/q}.\end{aligned}$$

(b) This is shown analogous to (a). Here are the details. The result is obvious if $p = 1$ or if $f + g = 0$. Otherwise observe that

$$|f + g|^p \leq (|f| + |g|)|f + g|^{p-1}$$

and apply Hölder's inequality, noting that $(p - 1)q = p$ when $q^{-1} = 1 - p^{-1}$:

$$\begin{aligned} \int_I |f(x) + g(x)|^p dx &\leq \|f\|_p \||f + g|^{p-1}\|_q + \|g\|_p \||f + g|^{p-1}\|_q \\ &= (\|f\|_p + \|g\|_p) \left[\int_I |f(x) + g(x)|^p dx \right]^{1/q} . \end{aligned}$$

□

Theorem 1.2.4. *For any norm $\|\cdot\| : X \to \mathbb{R}$, we have*

$$|\|x\| - \|y\|| \leq \|x - y\| \ .$$

In particular $\|\cdot\|$ is a continuous function.

Definition 1.2.3. *Two norms, $\|\cdot\|_1$ and $\|\cdot\|_2$, on a normed linear space X are called equivalent if there are finite positive constants C and C' such that for all $x \in X$,*

$$C\|x\|_1 \leq \|x\|_2 \leq C'\|x\|_1 \ .$$

Remark 1.2.2.

- On $\mathbb{F}^n$ all norms are equivalent. However on an infinite dimensional vector spaces there exist inequivalent norms.
- If X is a normed space with equivalent norms $\|\cdot\|_1$ and $\|\cdot\|_2$, then the identity operator $\mathrm{id}_X : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$, $x \mapsto x$ is a bijective continuous linear operator with continuous inverse.
- If there exists a continuous and bijective linear mapping from $(X, \|\cdot\|_X)$ to $(Y, \|\cdot\|_Y)$ with continuous inverse, then X and Y are as normed spaces indistinguishable.

Remark 1.2.3.

- Let X and Y be two normed spaces. Let $X \times Y$ be the Cartesian product, i.e., the set of all pairs (x, y) with $x \in X$ and $y \in Y$. We equip $X \times Y$ with the natural vector space structure

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2), \qquad \alpha(x, y) := (\alpha x, \alpha y),\ \forall \alpha \in \mathbb{F}.$$

The maps $\pi_1 : X \times Y \to X, (x, y) \to x$ and $\pi_2 : X \times Y \to Y, (x, y) \mapsto y$ are called projection onto the first and second factor, respectively.

- Let e be a norm on $\mathbb{R}^2$. Then

$$\|(x, y)\|_e \ := \ e((\|x\|, \|y\|))$$

is a norm on $X \times Y$. Moreover, e' is any other norm on $\mathbb{R}^2$, then $\| \cdot \|_e$ and $\| \cdot \|_{e'}$ are equivalent norms on $X \times Y$.

- To be explicit we shall henceforth assume that $X \times Y$ is equipped with

$$\|(x, y)\|_1 = \|x\| + \|y\|.$$

- With respect to any of the norm's $\| \cdot \|_e$ we have

$$(x_n, y_n) \to (x, y) \quad \text{if and only if} \quad x_n \to x \text{ and } y_n \to y.$$

(Proof: $\|(x_n, y_n) - (x, y)\| = \|(x_n - x, y_n - y)\| = \|x_n - x\| + \|y_n - y\|$, thus if the right side tends to zero so must the left and the other way around.)

- X and Y are Banach spaces if and only if $X \times Y$ is a Banach space. Moreover the projections π_1 and π_2 are continuous.

Proposition 1.2.5. *Let X be normed space. Then the mappings*

$$X \times X \to X, \quad (x, y) \mapsto x + y$$
$$\mathbb{F} \times X \to X, \quad (\lambda, x) \mapsto \lambda x$$

are continuous. We equipp $X \times X$ and $\mathbb{F} \times X$ with the product metric.

Proof. The continuity of the addition follows from

$$\|(x+y)-(x'+y') \leq \|x-x'\| + \|y-y'\|.$$

The continuity of the scalar multiplication at (λ, x) follows from

$$\begin{aligned}\|\lambda x - \lambda' x'\| &\leq |\lambda - \lambda'|\|x\| + |\lambda'|\|x - x'\| \\ &\leq |\lambda - \lambda'|(\|x\| + \|x - x'\|) + |\lambda|\|x - x'\|,\end{aligned}$$

where we insertet $|\lambda'| \leq |\lambda| + |\lambda - \lambda'|$ into the r.h.s of the first line. □

Definition 1.2.4. *A* **bounded linear transformation** *(or bounded operator) from a normed linear space $(X, \|\cdot\|_X)$ to a normed linear space $(Y, \|\cdot\|_Y)$ is a linear map $T : X \to Y$ such that for some $C \geq 0$ we have*

$$\|Tv\|_Y \leq C\|v\|_X, \quad \forall v \in X. \tag{1.2.5}$$

We denote by $\mathcal{B}(X, Y)$ the set of all bounded operators from X to Y.

Lemma 1.2.6. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Let $T : X \to Y$ be a linear operator. Then the followign holds.*

(a) T is bounded if and only if

$$\|T\| := \sup_{\substack{v \in X \\ \|v\|_X = 1}} \|Tv\|_Y := \sup\{\|Tv\|_Y : v \in X,\ \|v\|_X = 1\}. \tag{1.2.6}$$

is finite.

(b) We have

$$\begin{aligned}\|T\| = \sup_{\substack{v \in X \\ \|v\|_X \leq 1}} \|T\| &= \sup_{\substack{v \in X \\ v \neq 0}} \frac{\|Tv\|_Y}{\|v\|_X} \\ &= \inf\{C \geq 0 : \|Tv\|_Y \leq C\|v\|_X,\ \forall v \in X\}.\end{aligned}$$

(c) $\mathcal{B}(X,Y)$ *is a vector space and* $\|\cdot\|$ *defined by* (1.2.6) *defines a norm on* $\mathcal{B}(X,Y)$*, called the (operator)* **norm**. *We have*

$$\|Tv\| \le \|T\|\|v\| \quad \forall v \in X.$$

Proof. (a) (1.2.5) implies that for $\|v\|_X = 1$ we have $\|Tv\| \le C$. So

$$\sup_{\substack{v\in X\\ \|v\|_X=1}} \|Tv\|_Y \le C. \tag{1.2.7}$$

On the other hand for $v \neq 0$ we find

$$\|Tv\| = \|v\|\|T(v/\|v\|)\| \le \|v\| \sup_{\substack{v\in X\\ \|v\|_X=1}} \|Tv\|_Y. \tag{1.2.8}$$

(b) The first line follows from

$$\begin{aligned}\|T\| &= \sup_{\|v\|_X=1} \|Tv\|_Y \le \sup_{\|v\|_X\le 1} \|Tv\|_Y = \sup_{0<\|v\|_X\le 1} \|Tv\|_Y \\ &\le \sup_{0<\|v\|_X\le 1} \frac{\|Tv\|_Y}{\|v\|_X} = \sup_{0<\|v\|_X\le 1} \left\| T\left(\frac{v}{\|v\|_X}\right)\right\|_Y = \sup_{\|v\|_X=1} \|Tv\|_Y = \|T\|.\end{aligned}$$

The see the second line can be seen as follows. "$\le$" follows from (1.2.7). "$\ge$" follows from (1.2.8).

(c) Since any fintie linear combination of bounded operators is again a bounded operator, $\mathcal{B}(X,Y)$ is a vector space. It is easy to see that $\|\cdot\|$ is a norm; for example, the triangle inequality is proven by the computation

$$\|A+B\| = \sup_{\|x\|=1} \|(A+B)x\| \le \sup_{\|x\|=1} (\|Ax\| + \|Bx\|) \le \sup_{\|x\|=1} \|Ax\| + \sup_{\|x\|=1} \|Bx\| = \|A\| + \|B\|.$$

The inequality follows from (1.2.8). □

Theorem 1.2.7. *Let* T *be a linear transformation between two normed linear spaces. Then the following are equivalent:*

(a) T is continuous at a sinlge point.

(b) T is continuous.

(c) T is bounded.

Proof. (b) $\Rightarrow$ (a) trivial.
(a) $\Rightarrow$ (b) Follows from linearity. Details: Suppose that T is continous at p. Let $x \in V$ be arbitrary. Let $\epsilon > 0$. Then there exists a $\delta > 0$ such that $\|Tv - Tp\| < \epsilon$ whenvever $\|v - p\| < \delta$. Hence for $\|y - x\| < \delta$, we have $\|Ty - Tx\| = \|T(y - x + p) - Tp\| < \epsilon$, since $\|(y - x + p) - p\| = \|y - x\| < \delta$.
(c) $\Rightarrow$ (b) Details: Let $x \in V$ be arbitrary. Let $\epsilon > 0$. Set $\delta := \epsilon/\|T\|$. Then for $\|y - x\| < \delta$ we have

$$\|Tx - Ty\| \leq \|T\|\|x - y\| < \|T\|\delta = \epsilon.$$

(b) $\Rightarrow$ (c) Since T is continuous in 0, there exists a $\delta > 0$ such that $\|Ty\| \leq 1$ whenever $\|y\| \leq \delta$. Thus for all nonzero $x \in V$ we have

$$\|Tx\| = \|x\|\delta^{-1}\|T(\delta x/\|x\|) \leq \delta^{-1}\|x\|.$$

□

Definition 1.2.5. *A complete normed space is called* **Banach space***.*

Theorem 1.2.8. *Let X be a normed vector space, which is not complete. Then there exists a Banach space $\tilde{X}$ and an linear mapping $i : X \to \tilde{X}$, which preserves the norm ($\|i(v)\| = \|v\|$), such that $i(X)$ is dense in $\tilde{X}$. The space $\tilde{X}$ is unique except for linear isometries.*

Proof. The same as for metric spaces, Theorem 1.1.4, with the obvious generalizations. □

Remark 1.2.4. (a) Let I be a compact Interval of $\mathbb{R}^d$. Consider the norm $\|f\|_1 = \int_I |f(x)|dx$ on $C(I)$. The completion of $(C(I), \|\cdot\|_1)$ is called the space Lebesgue space $L^1(I)$.

(b) Consider the space

$$C_c(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{C} : f \text{ continuous and vanishing outside of some ball}\}.$$

Consider the norm $\|f\|_1 = \int_{\mathbb{R}^d} |f(x)|dx$. One denotes the completion of this space by $L^1(\mathbb{R}^d)$.

(c) Consider the space

$$C_c(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{C} : f \text{ continuous and vanishing outside of some ball}\}.$$

Consider the norm $\|f\|_p = \left(\int_{\mathbb{R}^d} |f(x)|^p dx\right)^{1/p}$ for $p \geq 1$. One denotes the completion of this space by $L^p(\mathbb{R}^d)$.

Note that the completions are by consruction given as equivalence classes of Cauchy sequences. One can show that these equivalence classes can indeed be indentified with equivalence classes of functions. (The proof is similar to th completeness proofs of the L^p-spaces defined in context of measure theory, see later.) Representing the L^p spaces as equivalence classes of functions is the usual approach in integration theory.

Theorem 1.2.9. *(B.L.T. theorem) Suppose T is a bounded linear transformation from a normed linear space $(V_1, \|\cdot\|_1)$ to a complete normed linear space $(V_2, \|\cdot\|_2)$. Then T can be uniquely extended to a bounded linear transformation (with the same bound), $\tilde{T}$, from the completion of V_1 to V_2.*

Proof. Let $\tilde{V}_1$ be the completion of V_1. For each $x \in \tilde{V}_1$, there is a sequence of elements (x_n) in V_1 with $x_n \to x$ as $n \to \infty$. Since (x_n) converges, it is Cauchy, so given $\epsilon > 0$,we can find N so that $n, m > N$ implies $\|x_n - x_m\|_1 \leq \epsilon/\|T\|$. Then $\|Tx_n - Tx_m\|_2 = \|T(x_n - x_m)\|_2 \leq \|T\|\|x_n - x_m\|_1 \leq \epsilon$ which proves that Tx_n is a Cauchy sequence in V_2. Since V_2 is complete, $Tx_n \to y$ for some $y \in V_2$. Set $\tilde{T}x = y$. We must first show that this definition is independent of the sequence chosen. Suppose (x'_n) is another sequence in V_1 with $x'_n \to x$. Then as before $Tx'_n \to y'$ for some y'. By convergence thre exists for any $\eta > 0$ an m such that $\|y - Tx_m\| < \eta/2$ and $\|y - Tx'_m\| < \eta/2$. Then

$\|y - y'\| \leq \|y - Tx_m\| + \|Tx'_m - Ty'\| < \eta$. Since $\eta > 0$ was arbitrary $y = y'$. Moreover, we can show $\tilde{T}$ so defined is bounded because

$$\begin{aligned} \|\tilde{T}x\|_2 &= \lim_{n\to\infty} \|Tx_n\|_2 \quad \text{(by continuity of the norm)} \\ &\leq \limsup_{n\to\infty} \|T\|\|x_n\|_1 \quad \text{(comparison for sequences)} \\ &= \|T\|\|x\|_1 \quad \text{(by continuity of the norm)} \end{aligned}$$

Thus $\tilde{T}$ is bounded. The proofs of linearity and uniqueness are left to the reader. □

Remark 1.2.5.

(a) Consider the space

$$C_c(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{C} : f \text{ continuous and vanishing outside of some ball}\}$$

with the norm $\|f\|_1 = \int_{\mathbb{R}^d} |f(x)|dx$. We denote the completion by $L^1(\mathbb{R}^d)$. Define the map $R : C_c(\mathbb{R}^d) \to \mathbb{C}, f \mapsto R(f) := \int_{\mathbb{R}^d} f(x)dx$. Let $\tilde{R}$ denote the extension of R to $L^1(\mathbb{R}^d)$. Then for $f \in L^1(\mathbb{R}^d)$ one calls $\tilde{R}(f)$ the Lebesgue integral of f and one writes

$$\int_{\mathbb{R}^d} f(x)dx := \tilde{R}(f).$$

Theorem 1.2.10. *A normed vector space X is complete iff every absolutely convergent series in X converges.*

Proof. $\Rightarrow$: Suppose X is complete and $\sum_{n=1}^{\infty} \|x_n\| < \infty$. Let $S_N = \sum_{n=1}^{N} x_n$. Then for $N > M$ we have

$$\|S_N - S_M\| \leq \sum_{n=M+1}^{N} \|x_n\| \to 0, \quad M, N \to \infty,$$

so the sequence (S_N) is Cauchy and hence convergent.

$\Leftarrow$: Suppose that every absolutely convergent series converges, and let (x_n) be a Cauchy sequence. We can choose $n_1 < n_2 < \cdots$ such that $\|x_n - x_m\| < 2^{-j}$ for $m, n \geq n_j$. Let

$y_1 = x_{n_1}$ and $y_j = x_{n_j} - x_{n_{j-1}}$ for $j > 1$. Then $\sum_{j=1}^{k} y_j = x_{n_k}$, and

$$\sum_{j=1}^{\infty} \|y_j\| \leq \|y_1\| + \sum_{j=1}^{\infty} 2^{-j} = \|y_1\| + 1 < \infty,$$

so limt $x_{n_k} = \sum_{j=1}^{\infty} y_j$ exists. But since (x_n) is Cauchy, it is easily verified that(x_n) converges to the same limt as (x_{n_k}). □

Chapter 2

Measure Theory

2.0.1 Measurable Spaces and Measure Spaces

Definition 2.0.1. *Let X be a nonempty set and let $\mathcal{A} \subset \mathcal{P}(X)$ be a collection of subsets of X. We consider the following properties.*

(i) *$\mathcal{A}$ is closed under complements:*
If $E \in \mathcal{A}$, then $E^c \in \mathcal{A}$.

(ii) *$\mathcal{A}$ is closed under finite unions:*
If $n \in \mathbb{N}$ and $E_1, ..., E_n \in \mathcal{A}$, then $\bigcup_{n=1}^{n} E_i \in \mathcal{A}$.

(ii') *$\mathcal{A}$ is closed under countable unions:*
For any sequence $(E_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ we have $\bigcup_n E_n \in \mathcal{A}$.

If $\mathcal{A}$ is nomepty and satisfies (i) and (ii) it is called an **algebra** *of sets. If $\mathcal{A}$ is nonempty and satisfies (i) and (ii') it is called a σ-***algebra***. If $\mathcal{A}$ is a σ–algebra on X, we call the tuple $(X, \mathcal{A})$ a* **measurable space** *and the sets in $\mathcal{A}$* **measurable***.*

Example 6. *Let X be a nonempty set. Then obviously $\{\emptyset, X\}$ and $\mathcal{P}(X)$ are σ-algebras on X.*

Lemma 2.0.1. *Let X be a nonempty set and let $\mathcal{A} \subset \mathcal{P}(X)$. Consider the following properties.*

(i) $\emptyset, X \in \mathcal{A}$

(ii) $\mathcal{A}$ *is closed under finite intersections:*
If $n \in \mathbb{N}$ *and* $E_1, ..., E_n \in \mathcal{A}$, *then* $\bigcap_{n=1}^{n} E_i \in \mathcal{A}$.

(ii') $\mathcal{A}$ *is closed under countable intersections:*
For any sequence $(E_n)_{n\in\mathbb{N}}$ *in* $\mathcal{A}$ *we have* $\bigcap_n E_n \in \mathcal{A}$.

If $\mathcal{A}$ *is an algebra on* X, *then (i) and (ii) hold. If* $\mathcal{A}$ *is a* σ*-algebra on* X, *then (i) and (ii') hold. Every* σ*-algebra is an algebra. An algebra is a* σ*-algebra, if it is closed under countable disjoint unions. For any collection of algebras or* σ*-algebras* $(\mathcal{A}_\alpha)_{\alpha\in A}$ *the set*

$$\bigcap_{\alpha\in A} \mathcal{A}_\alpha$$

is again an algebra or a σ*-algebra, respectively.*

Proof. Elementary: Some Details. (i) Let $\mathcal{A}$ be an algebra or a σ-algebra and $E \in \mathcal{A}$. Then $\emptyset = E \cap E^c \in \mathcal{A}$ and $X = \emptyset^c \in \mathcal{A}$.
(ii), (ii') Follow from de Morgan's laws.
Every σ-algebra is an algebra, which can be seen by extending a finite collection of sets to a countable collection using the emptyset. □

Definition 2.0.2. *Let X be a nonempty set. Let $\mathcal{E} \subset \mathcal{P}(X)$. Then*

$$\mathcal{M}(\mathcal{E}) := \bigcap_{\substack{\mathcal{A} \text{ is } \sigma-\text{algebra of } X, \\ \mathcal{E}\subset\mathcal{A}}} \mathcal{A}$$

is called the σ*-algebra* **generated by** $\mathcal{E}$ *(or the smallest* σ*-algebra containing* $\mathcal{E}$*).*

Remark 2.0.1. By the lemma above, $\mathcal{M}(\mathcal{E})$ is indeed a σ-algebra. We say that $\mathcal{N}$ is a smallest σ-algebra containing $\mathcal{E}$ (w.r.t. inclusion of sets), if

(i) $\mathcal{N}$ is a σ-algebra containing $\mathcal{E}$,

(ii) if $\mathcal{A}$ is a σ-algebra containing $\mathcal{E}$, then $\mathcal{N} \subset \mathcal{A}$.

We note that if a smallest σ-algebra exists it is unique as a consequence of (i) and (ii). It is easy to see by definition that $\mathcal{M}(\mathcal{E})$ is the smallest σ-algebra of X containing $\mathcal{E}$.

Definition 2.0.3. *Let X be a metric space or more generally a topological space and let $\mathcal{O}_X$ denote the colllection of its open sets. Then $\mathcal{B}_X := \mathcal{M}(\mathcal{O}_X)$ is called the* **Borel** *σ-algebra of X. By $\mathcal{B}_{\mathbb{R}^d}$ we denote the Borel σ-algebra of $\mathbb{R}^d$, when $\mathbb{R}^n$ is equipped with a metric which is given by a norm (as for example the Euclidean metric).*

Lemma 2.0.2. *If $\mathcal{E} \subset \mathcal{M}(\mathcal{F})$ then $\mathcal{M}(\mathcal{E}) \subset \mathcal{M}(F)$.*

Proof. $\mathcal{M}(\mathcal{F})$ is a σ-algebra containing $\mathcal{E}$; it therefore contains $\mathcal{M}(\mathcal{E})$. □

Example 7. *$\mathcal{B}_{\mathbb{R}}$ is generated by each of the following.*

(a) the open intervals: $\mathcal{E}_1 = \{(a,b) : a < b\}$,

(b) the closed intervals: $\mathcal{E}_2 = \{([a,b] : a < b\}$,

(c) the half-open intervals: $\mathcal{E}_3 = \{(a,b] : a < b\}$ or $\mathcal{E}_4 = \{[a,b) : a < b\}$,

(d) the open rays: $\mathcal{E}_5 = \{(a,\infty) : a \in \mathbb{R}\}$ or $\mathcal{E}_6 = \{(-\infty,b) : b \in \mathbb{R}\}$,

(e) the closed rays: $\mathcal{E}_7 = \{[a,\infty) : a \in \mathbb{R}\}$ or $\mathcal{E}_8 = \{(-\infty,b] : b \in \mathbb{R}\}$.

Proof. (a) Clearly, $\mathcal{E}_1 \subset \mathcal{B}_{\mathbb{R}}$. The opposite inclusion follows, since every open subset of $\mathbb{R}$ is a countable union of open intervals, hence $\mathcal{O}_{\mathbb{R}} \subset \mathcal{E}_1$ and thus by the lemma $\mathcal{B}_R = \mathcal{M}(\mathcal{E}_1)$.
(b) Now $\mathcal{E}_1 \subset \mathcal{M}(\mathcal{E}_2)$, since $(a,b) = \mathbb{R} \setminus (\cup_{n\in\mathbb{N}}([-n,a] \cup [b,n]))$. Thus $\mathcal{B}_{\mathbb{R}} \subset \mathcal{M}(\mathcal{E}_2)$ by (a) and the lemma. To show the oposite inclusion we use $\mathcal{E}_2 \subset \mathcal{B}_{\mathbb{R}}$, since $[a,b] = \mathbb{R} \setminus ((-\infty,a) \cup (b,\infty))$.
(d) Clearly $\mathcal{M}(\mathcal{E}_j) \subset \mathcal{B}_{\mathbb{R}}$ for $j = 5, 6$, since half open intervals are open. To show the oposite inclusion we note that $\mathcal{E}_1 \subset \mathcal{M}(\mathcal{E}_j)$ for $j = 5$, since $(a,b) = (a,\infty) \setminus (\bigcap_{n\in\mathbb{N}}(b - 1/n, \infty))$. In case $j = 6$ one argues similarly.
The remaing is left as an excercise. □

Example 8. *$\mathcal{B}_{\mathbb{R}^d}$ is generated by each of the following.*

(a) *the open intervals: $\mathcal{E}_1 = \{I_1 \times \cdots I_d : I_j$ open interval of $\mathbb{R}\}$,*

(b) *the closed intervals: $\mathcal{E}_2 = \{I_1 \times \cdots I_d : I_j$ closed interval of $\mathbb{R}\}$*

Proof. (a) Since every open interval is open, we have $\mathcal{M}(\mathcal{E}_1) \subset \mathcal{B}_{\mathbb{R}^d}$.
On the other hand since every open set of $\mathbb{R}^d$ can be written as a countable union of open intervals, we have $\mathcal{O}_{\mathbb{R}^d} \subset \mathcal{M}(\mathcal{E}_1)$. Thus $\mathcal{B}_{\mathbb{R}^d} \subset \mathcal{M}(\mathcal{E}_1)$.
(b) Excercise. □

Example 9. *Let $X_1, ..., X_n$ be metric spaces and $X = X_1 \times \cdots \times X_n$ be equipped with the product metric. Then the following holds.*

(a) *$\mathcal{M}(\{E_1 \times \cdots \times E_d : E_j$ open set of $X_j\} \subset \mathcal{B}_X$.*

(b) *$\mathcal{M}(\{E_1 \times \cdots \times E_d : E_j$ open set of $X_j\} = \mathcal{B}_X$, if all X_j are separable.*

(c) *If $\pi_j : X \to X_j$ is the canonical projection, we have*

$$\mathcal{M}(\{E_1 \times \cdots \times E_d : E_j \text{ open set of } X_j\} = \mathcal{M}(\{\pi_j^{-1}(E) : E \text{ open set of } X_j\}.$$

Proof. (a) Since every Cartesian product of open sets in X_j, $j = 1, ..., d$, is open in X (a) follows.
(b) Since each X_j is dense, every open set of X can be written as a countable union of cartesian products of open sets in X_j. Thus $\mathcal{O}_X \subset \mathcal{M}(\{E_1 \times \cdots \times E_d : E_j$ open set of $X_j\})$.
(c) $\supset$: this is obvious, since the gerating set on the left hand side contains that on the right hand side.
$\subset$: Since $E_1 \times \cdots \times E_d = \bigcap_{j=1}^d \pi_j^{-1}(E_j)$, we see that

$$\{E_1 \times \cdots \times E_d : E_j \text{ open set of } X_j\}) \subset \mathcal{M}\left(\{\pi_j^{-1}(E_j) : E_j \text{ open set of } X_j\}\right)$$

and the result therefore follows from Lemma 2.0.2. □

Next we consider the product σ-algebra.

Definition 2.0.4. *Let* $(X_\alpha)_{\alpha\in A}$ *be a indexed collection of nonempty sets. Consider their Cartesian product*

$$X := \prod_{\alpha\in A} X_\alpha$$

and the coordinate maps $\pi_\alpha : X \to X_\alpha$*. If* $\mathcal{A}_\alpha$ *is a* σ*-algebra on* X_α *for each* $\alpha \in A$*, then*

$$\bigotimes_{\alpha\in A} \mathcal{A}_\alpha := \mathcal{M}\left(\left\{\pi_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathcal{A}_\alpha, \alpha \in A\right\}\right)$$

is called the **product** σ**-algebra** *on* X*.*

Proposition 2.0.3. *Let* $\mathcal{A}_\alpha$ *be a* σ*-algebra on* X_α *for each* $\alpha \in A$*. If* $\mathcal{A}_\alpha$ *is generated by* $\mathcal{E}_\alpha$*, then*

$$\bigotimes_{\alpha\in A} \mathcal{A}_\alpha = \mathcal{M}\left(\left\{\pi_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathcal{E}_\alpha, \alpha \in A\right\}\right)$$

Proof. $\supset$ is trivial (since the generating set of the right hand side is a subset of the other.)
$\subset$ Write $\mathcal{F} = \{\pi_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathcal{E}_\alpha, \alpha \in A\}$. Fix $\alpha \in A$. Then $\{E \in X_\alpha : \pi_\alpha^{-1}(E) \in \mathcal{M}(\mathcal{F}_1)\}$ is easily seen to be a σ-algebra on X_α that contains $\mathcal{E}_\alpha$ and hence $\mathcal{A}_\alpha$. In other words $\pi_\alpha^{-1}(E)$ for all $E \in \mathcal{A}_\alpha$. Since α was aribitrary the inclusion now follows from Lemma 2.0.2. $\square$

Lemma 2.0.4. *Let* $X_1, ..., X_d$ *be metric spaces and let* $X = X_1 \cdots \times \cdots X_d$ *be equipped with the product metric. Then the following holds.*

(a) $\bigotimes_{j=1}^d \mathcal{B}_{X_j} \subset \mathcal{B}_X$*.*

(b) $\bigotimes_{j=1}^d \mathcal{B}_{X_j} = \mathcal{B}_X$*, if* $X_1, ..., X_d$ *are separable.*

Proof. (a) By (a) of Proposition 2.0.3 $\bigotimes_{j=1}^d \mathcal{B}_{X_j}$ is by sets of the form $\pi_j^{-1}(O_j)$, $j = 1, ..., d$, where O_j is open in X_j. Since these sets are open in X, the inclusion follows from Lemma 2.0.2.
(b) Follows from Example 9 (b,c) and Proposition 2.0.3. $\square$

Corollary 2.0.5. *We have* $\mathcal{B}_{\mathbb{R}^d} = \bigotimes_1^d \mathcal{B}_{\mathbb{R}}$.

Proposition 2.0.6. *Let* $\mathcal{E}$ *be a family of subsets of* X *such that*

(i) $\emptyset \in \mathcal{E}$,

(ii) if $E, F \in \mathcal{E}$, *then* $E \cap F \in \mathcal{E}$,

(iii) if $E \in \mathcal{E}$ *then* E^c *is a finite disjoint union of members of* $\mathcal{E}$.

(such a family is called **elementary family***). Then the collection of finite disjoint unions of* $\mathcal{E}$ *is an algebra. (Clearly, this algebra is equal to the colletion of finite unions of* $\mathcal{E}$*.)*

Proof. Let $\mathcal{A}$ be the set of finite disjoint unions of $\mathcal{E}$. The set is nonempty by (i).
Step 1: If $A \in \mathcal{E}$ and $B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$.
Since $B \setminus A = B \cap A^c$ is by (iii) and (ii) a finite disjoint uniont of elements of $\mathcal{E}$, the claim follows from the identity $A \cup B = A \cup (B \setminus A)$.
Step 2: If $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$.
This follows from a repeated application of Step 1.
Step 3: If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$.
Suppose $A = \bigcup_{j=m}^n E_m$ with $E_m \in \mathcal{E}$, $m = 1, ..., n$, mutually disjoint. By (iii) we have $E_m^c = \bigcup_{j=1}^{J_m} B_m^j$ with $B_m^1, \cdots, B_m^{J_m}$ disjoint memebers of $\mathcal{E}$. Then

$$A^c = \bigcap_{m=1}^n E_m^c = \bigcap_{m=1}^n (\bigcup_{j=1}^{J_m} B_m^j) = \bigcup \{B_1^{j_1} \cap \cdots B_n^{j_n} : 1 \le j_m \le J_m, 1 \le m \le n\} \in \mathcal{A}.$$

□

Example 10. As an application of the previous proposition the following sets are algebras.

- In $\mathbb{R}^d$ for any $d \ge 1$. Finite unions of disjoint intervals. Finite unions of disjoint open intervals. Finite unions of disjoint closed intervals. Finite unions of disjoint left open right closed intervals. Finite unions of disjoint left closed right open intervals.

- Let X and Y be two sets and $\mathcal{X} \subset \mathcal{P}(X)$ and $\mathcal{Y} \subset \mathcal{P}(Y)$ an algebra (or an elementary family). Then finite disjoint unions of sets of the form $A \times B$ wit h$A \in \mathcal{X}$ and $B \in \mathcal{Y}$ form an algebra.

Next we introduce measures.

Definition 2.0.5. *Let X be a set and $\mathcal{A}$ a σ-algebra on X. A* **measure** *on $\mathcal{A}$ is a function*

$$\mu : \mathcal{A} \to [0, \infty]$$

with the following properties

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

(ii) if $(E_j)_{j\in\mathbb{N}}$ is a sequence of disjoint sets in $\mathcal{A}$, then $\mu(\bigcup_{j\in\mathbb{N}} E_j) = \sum_{j=1}^{\infty} \mu(E_j)$.

The triple $(X, \mathcal{A}, \mu)$ is called a **measure space***. The measure is called* **finite***, if $\mu(X) < \infty$. The measure is called* σ**-finite***, if there is a sequence $(E_j)_{j\in\mathbb{N}}$ in $\mathcal{A}$ with $\bigcup_j E_j = X$ and $\mu(E_j) < \infty$ for all $j \in \mathbb{N}$.*

Remark 2.0.2. Property (ii) is called countable additivity. It implies finite additivity:

(ii') if $E_1, \cdots E_n$ are disjoint sets in $\mathcal{A}$, then $\mu(\bigcup_{j=1}^{n} E_j) = \sum_{j=1}^{n} \mu(E_j)$.

(Proof: One can extend the finite collection of sets to a sequence by setting $E_j = \emptyset$ for $f > n$.)

Example 11. Let X be a set with $\mathcal{A}$ a σ-algebra on X. Let $x \in X$. Then the measure $\delta_x : \mathcal{A} \to [0, \infty]$ defined by

$$\delta_x(E) = \begin{cases} 1 \ , & \text{if } \ x \in E \\ 0 \ , & \text{if } \ x \notin E \ , \end{cases}$$

for all $E \in \mathcal{A}$, is called **Dirac point measure** at x. (It is indeed a measure as one readily verifies.)

The basic properties of measures are summarized in the following theorem.

Theorem 2.0.7. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Then the following holds.*

(a) (Monotonicity) If $E, F \in \mathcal{A}$ and $E \subset F$, then $\mu(E) \leq \mu(F)$.

(b) (Subadditivity) If $(E_j)_{j\in\mathbb{N}}$ is a sequence in $\mathcal{A}$, then $\mu(\bigcup_j E_j) \leq \sum_j \mu(E_j)$.

(c) (Continuity from below) If $(E_j)_{j\in\mathbb{N}}$ is a sequence in $\mathcal{A}$ with $E_1 \subset E_2 \subset \cdots$, then

$$\mu(\bigcup_j E_j) = \lim_{j\to\infty} \mu(E_j)$$

(d) (Continuity from above) If $(E_j)_{j\in\mathbb{N}}$ is a sequence in $\mathcal{A}$ with $E_1 \supset E_2 \supset \cdots$ and $\mu(E_1) < \infty$, then

$$\mu(\bigcap_j E_j) = \lim_{j\to\infty} \mu(E_j)$$

aybe use a better
tation

Proof. (a) if $E \subset F$, then $\mu(F) = \mu(E) + \mu(F \setminus E) \geq \mu(E)$.
(b) Let $F_1 = E_1$ and $F_k = E_k \setminus (\bigcup_{j=1}^{k-1} E_j)$ for $k > 1$. Then the F_k's are disjoint and $\bigcup_{j=1}^{n} F_j = \bigcup_{j=1}^{n} E_j$ for all n. Therefore, by (a),

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \mu\left(\bigcup_{j=1}^{\infty} F_j\right) = \sum_{j=1}^{\infty} \mu(F_j) \leq \sum_{j=1}^{\infty} \mu(E_j).$$

(c) Setting $E_0 = \emptyset$, we have

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j \setminus E_{j-1}) = \lim_{j\to\infty} \sum_{j=1}^{n} \mu(E_j \setminus E_{j-1}) = \lim_{n\to\infty} \mu(E_n).$$

(d) Let $F_j = E_1 \setminus E_j$; then $F_1 \subset F_2 \subset \cdots$, $\mu(E_1) = \mu(F_j) + \mu(E_j)$, and $\bigcup_{j=1}^{\infty} F_j = E_1 \setminus \left(\bigcap_{j=1}^{\infty} E_j\right)$. By (c), then,

$$\mu(E1) = \mu\left(\bigcap_{j=1}^{\infty} E_j\right) + \lim_{j\to\infty} \mu(F_j) = \mu\left(\bigcup_{j=1}^{\infty} E_j\right) + \lim_{j\to\infty} [\mu(E_1) - \mu(E_j)].$$

Since $\mu(E_1) < \infty$, we may subtract it from both sides to yield the desired result. $\square$

Next we consider sets of measure zero and complete measure spaces.

Definition 2.0.6. *Let $(X, \mathcal{A}, \mu)$ be a measure space. A sete $E \in \mathcal{A}$ with $\mu(E) = 0$ is called a μ-**null** set (or a set of measure zero). We say that a statement holds for μ-**a.e.** x, if there exists a measurable set N with $\mu(N) = 0$ such that the statements hold for all $x \in X \setminus N$.*

Definition 2.0.7. *Let $(X, \mathcal{A}, \mu)$ be a measure space. The measure μ is called* **complete** *if its domain includes all subsets of null sets (i.e, if $F \in \mathcal{A}$, $\mu(F) = 0$ and $E \subset F$ implies $E \in \mathcal{A}$).*

Theorem 2.0.8. *Suppose that $(X, \mathcal{A}, \mu)$ is a measure space. Let $\mathcal{N} = \{N \in \mathcal{A} : \mu(N) = 0\}$ and $\overline{\mathcal{A}} = \{E \cup F : E \in \mathcal{A}$ and $F \subset N$ for some $N \in \mathcal{N}\}$. Then $\overline{\mathcal{A}}$ is a σ-algebra, and there is a unique extension $\overline{\mu}$ of μ to a complete measure on $\overline{\mathcal{A}}$. The extension satisfies*

$$\overline{\mu}(A \cup F) = \mu(A), \tag{2.0.1}$$

whenever $A \in \mathcal{A}$ and $F \subset N$ for some $N \in \mathcal{N}$.

Proof. First we show that $\overline{\mathcal{A}}$ is a σ-algebra.
(ii) Since $\mathcal{A}$ and $\mathcal{N}$ are closed under countable unions, so is $\mathcal{A}$.
(i) Let $E \cup F \in \overline{\mathcal{A}}$ where $E \in \mathcal{A}$ and $F \subset N$ for some $N \in \mathcal{N}$. We can assume without loss that $E \cap N = \emptyset$ (otherwise, replace F by $F \setminus E$ and N by $N \setminus E$). Then $E \cup F = (E \cup N) \setminus (N \setminus F)$, so $(E \cup F)^c = (E \cup N)^c \cup (N \setminus F) \in \overline{\mathcal{A}}$, since $N \setminus F \subset N$.
We want to define the extension using (2.0.1). For this we have to make sure that the definition is well defined, that it is independent of the specific representation of the set. Suppose $E_1 \cup F_1 = E_1 \cup F_2$ with $E_j \in \mathcal{A}$ and $F_1 \subset N_j$ for some $N_j \in \mathcal{N}$. Then $E_1 \subset E_2 \cup N_2$ and so $\mu(E_1) \leq \mu(E_2) + \mu(N_2) = \mu(E_2)$, and likewise $\mu(E_1) \leq \mu(E_2)$. It is now easily verified that $\overline{\mu}$ is a complete measure on $\overline{\mathcal{A}}$, and that $\overline{\mu}$ is the only measure on $\overline{\mathcal{A}}$ that extends μ; details are left to the reader as an excercise. $\square$

Definition 2.0.8. *The measure $\overline{\mu}$ in the above theorem is called the* **completion** *of μ and $\overline{\mathcal{A}}$ is called the* **completion** *of $\mathcal{A}$ with respect to μ.*

2.0.2 Lebesgue Measure

In measure theory one adopt the convention that $0 \cdot \infty = \infty \cdot 0 = 0$. For a bounded intervall I of $\mathbb{R}$ with left endpoint a and right endpoint b (clearly $a \leq b$) we define its length by

$$|I| = b - a$$

If the intervall I is unbounded we define its length $|I| = \infty$. An interval in $\mathbb{R}^d$ is a set of the form

$$I = I_1 \times \cdots \times I_d,$$

where I_j is an interval in $\mathbb{R}$. We say that an intervall in $\mathbb{R}^d$ is left open right closed (or left closed an right open) if each of its factors in the Cartesian product has this property.

For an interval in $\mathbb{R}^d$ we define the volume by

$$|R| := |I_1| \cdots |I_d|.$$

Let $\mathcal{I}$ denote the set intervalls and $\mathcal{I}_o$ the set of open intervals.

Definition 2.0.9. *For any set $E \subset \mathbb{R}^d$ we define*

$$\lambda^*(E) = \inf\{\sum_{j=1}^{\infty} |R_j| : (R_j)_{j\in\mathbb{N}} \text{ sequence in } \mathcal{I}_o \text{ with } E \subset \bigcup_{j=1}^{\infty} R_j\}.$$

The function $\lambda^* : \mathcal{P}(\mathbb{R}^d) \to [0, \infty]$ is called Lebesgue outer measure.

Remark 2.0.3. If I is an open interval, then $\lambda^*(I) = |I|$ and hence the Lebesgue outer measure agrees on open intervals with the volume. To see this property, we note that clearly $\lambda^*(I) \leq |I|$ (take (R_j) with $R_1 = I$ and $R_j = \emptyset$ for $j \geq 2$). To see the opposite inequality, let $\epsilon > 0$. For $I = \prod_{j=1}^{d}(a_j, b_j)$ consider $J_\epsilon = \prod_{j=1}^{d}[a_j + \epsilon, b_j - \epsilon]$. Now let R_j be any cover of I with open intervals. Then it is an open cover of the compact interval J_ϵ. Thus there exists a finite subcover, i.e., an $N \in \mathbb{N}$ such that $R_1, ...R_N$ covers J_ϵ. Then by Riemann integration theory, for example we find $|J_\epsilon| = \int 1_{J_\epsilon} \leq \int \sum_{j=1}^{N} 1_{R_j} = \sum_{j=1}^{N} |R_j| \leq \sum_{j=1}^{\infty} |R_j|$. It follows that $|J_\epsilon| \leq \lambda^*(I)$ for all $\epsilon > 0$. Since $|J_\epsilon| \to |I|$ as $\epsilon \downarrow 0$. We find $|I| \leq \lambda^*(I)$.

Remark 2.0.4. We note that in the definition of λ^* we could replace the set of open intervals with the with the set of intervals, the set of closed intervalls, the set of left open right closed, or left closed right open intervas. Clearly, if I is any interval we find $\lambda^*(I) = |I|$.

Theorem 2.0.9. *$\lambda^*|_{\mathcal{B}(\mathbb{R}^d)}$ is a measure on $\mathbb{R}^d$.*

Proof. We will not prove this result. $\square$

Definition 2.0.10. *The completion of $\lambda^*|_{\mathcal{B}(\mathbb{R}^d)}$ is called* **Lebesgue maesure** *on $\mathbb{R}^d$. It is denoted by λ, its σ-algebra by $\mathcal{L}^d$ and the elements of $\mathcal{L}^d$ are called* **Lebesgue measurable sets**. *(Sometimes one also calls $\lambda^*|_{\mathcal{B}(\mathbb{R}^d)}$ Lebesgue measure.)*

The following theorem provides properties about Lebesgue measurable sets.

Theorem 2.0.10. *Let $E \subset \mathbb{R}^d$ be measurable . Then the following holds.*

(a)

$$\begin{aligned}\lambda(E) &= \inf\{\lambda(U) : U \supset E,\, U \text{ open }\} \quad \textit{(inner regularity)}\\ &= \sup\{\lambda(K) : K \subset E,\, K \text{ compact }\} \quad \textit{(outer regularity).}\end{aligned}$$

(b) If $\lambda(E) < \infty$, for any $\epsilon > 0$ there is a finite collection of open intervals $(R_j)_{j=1}^N$ such that $\lambda(E \triangle \bigcup_{j=1}^N R_j) < \epsilon$. (In view of Proposition 2.0.6 the intervals can be chosen to be disjoint but then they are not necessarily open anymore.) ($A \triangle B := (A \setminus B) \cup (B \setminus A)$).

Proof. (a). We consider the first identity first. By monotonicity of the measure, we find $\lambda(E) \leq \inf\{\cdots\}$. To show the opposite inequality, let $\epsilon > 0$. Then there exists a sequence (R_j) of open intervals, such that their union covers E and $\sum_j |R_j| < \lambda(E) + \epsilon$. Since $\lambda(\bigcup R_j) \leq \sum_j \lambda(R_j) = \sum_j |R_j|$ by subbaditivity (see also Remark 2.0.3), we find $\inf\{\cdot\} < \lambda(E) + \epsilon$. Since $\epsilon > 0$ is arbitrary, the first identity follows. Now to the second identity.

Again by monotonicity of measures, we find $\lambda(E) \geq \sup\{\cdot\}$. To show the opposite inequality assume first that E is bounded. Then it is contained in some large Intervall I. Let $\epsilon > 0$. Then by the first identity there exists an open set such U such that $I \setminus E \subset U$ and $\lambda(U) < \lambda(I \setminus E) + \epsilon$. Define $K = I \setminus U$. Then K is compact and $K \subset E$, moreover

$$\lambda(K) = \lambda(I) - \lambda(I \cap U) \geq \lambda(E) + \lambda(I \setminus E) - \lambda(U) > \lambda(E) - \epsilon.$$

. Thus $\lambda(E) = \sup\{\cdots\}$ in case E is bounded. Now suppose that E is unbounded. Let $I_n = [-n, n]^d$ and $J_n = I_n \setminus I_{n-1}$. Then (J_n) is a disjoint family of bouned sets covering $\mathbb{R}^d$. Given $\epsilon > 0$, there exists for each $E_n = E \cap J_n$, by what we have shown, a compact set $K_n \subset E_n$ such that $\lambda(E_n) < \lambda(K_n) + \epsilon 2^{-n}$. It follows that $H_m = \bigcup_{n=1}^m K_n$ is compact, satisfies $H_m \subset E$ and $\lambda(H_m) + \epsilon > \lambda(\bigcup_{n=1}^m E_n)$. Now since $\lim_m \lambda(\bigcup_{n=1}^m E_n) = \lambda(E)$ by continunity of measures, the claim follows.

(b) Let E have finite measure. Let $\epsilon > 0$. Then by (a) there exists a compact set $K \subset E$ such that

$$\lambda(E) < \lambda(K) + \epsilon/2. \tag{2.0.2}$$

By definition of the outer measure and compactness, there exist finitely many open intervalls (R_j) such that their union covers K and

$$\sum_j \lambda(R_j) < \lambda(K) + \epsilon/2. \tag{2.0.3}$$

Using the property of symmetric differences, we find

$$E \triangle \bigcup_{j=1}^N R_j \subset (E \triangle K) \cup (K \triangle \bigcup_{j=1}^N R_j) = (E \setminus K) \cup (\bigcup_{j=1}^N R_j \setminus K).$$

Using subbaditivity we find with (2.0.2) and (2.0.3)

$$\lambda(E \triangle \bigcup_{j=1}^N R_j) \leq \lambda(E \setminus K) + \lambda(\bigcup_{j=1}^N \setminus K) \leq \epsilon/2 + \epsilon/2.$$

□

Theorem 2.0.11. *The following holds about Lebesgue measurability and the Lebesgue measure. Let $E \subset \mathbb{R}^d$ then the following properties are equivalent.*

(a) E is Lebegue measurable.

(b) $E = H \cup N$, where H is a countable union of closed sets and $\lambda(N) = 0$.

(c) $E = V \setminus N$, where V is a countable intersection of open sets and $\lambda(N) = 0$.

Proof. That (a) or (b) implies (a) follows directly from the propererty that Lebesgue measurable sets contain the Borel sets.
(a) $\Rightarrow$ (b,c). Suppose E is measurable. Suppose $\lambda(E) < \infty$. By Theorem 2.0.10 there exists for each $j \in \mathbb{N}$ an open set $U_j \supset E$ and a compact set $K_j \subset E$ such that

$$\lambda(U_j) - 2^{-j} \leq \lambda(E) \leq \lambda(K_j) + 2^{-j}.$$

Let $V = \bigcap U_j$ and $H = \bigcup K_j$. Then $H \subset E \subset V$ and $\lambda(V) = \lambda(H) = \lambda(E) < \infty$ (by continuity of measures). So $\lambda(V \setminus E) = \lambda(E \setminus H) = 0$. The result is proved in the case $\lambda(E) < \infty$. The general case $\lambda(E) = \infty$ is left to the reader as an excercise. $\square$

2.0.3 Construction of Measures from outer measures

To construct a measure on a set X we proceed as follows. Typically we can easily construct a preliminary measure (premeasure) on some algebra. To extend this premeasure to a measure, one uses the notation of an outer measure. That is, one first extends the premeasure to an outer measure, which itself is not measure but defined on the set of all subsets of X. We then find a suitable collection of subsets of X on which the outer measure is in fact a measure.

Definition 2.0.11. *Let X be a nonempty set. On outer measure on X is a function*

$$\mu^* : \mathcal{P}(X) \to [0, \infty]$$

such that the following holds.

(i) $\mu^*(\emptyset) = 0$.

(ii) $\mu^*(A) \leq \mu^*(B)$ *if* $A \subset B$

(iii) If $(A_j)_{j\in\mathbb{N}}$ *is a sequence in* $\mathcal{P}(X)$*, then* $\mu^*(\cup_j A_j) \leq \sum_j A_j$.

Lemma 2.0.12. *Let X be a non-empty set, $\mathcal{E} \subset \mathcal{P}(X)$, with $\emptyset, X \in \mathcal{E}$, and $\rho : \mathcal{E} \to [0, \infty]$ a function with $\rho(\emptyset) = 0$. For any $A \in \mathcal{P}(X)$ define*

$$\mu^*(A) = \inf\{\sum_j \rho(E_j) : E_j \in \mathcal{E},\ A \subset \bigcup_j E_j\}. \tag{2.0.4}$$

Then μ^* is an outer measure.

Proof. For any $A \subset X$ there exists a $(E_j)_{j=1}^\infty \in \mathcal{E}$ such that $A \subset \bigcup_{j=1}^\infty E_j$ (take $E_j = X$ for all j) so the definition of μ^* makes sense.
(i): $\mu^*(\emptyset) = 0$ (take $E_j = \emptyset$ for all j).
(ii): $\mu^*(A) \leq \mu^*(B)$ for $A \subset B$ because the set over which the infimum is taken in the definition of $\mu^*(A)$ includes the corresponding set in the definition of $\mu^*(B)$.
(iii): Suppose $(A_j)_{j\in\mathbb{N}}$ is a sequence in $\mathcal{P}(X)$ and $\epsilon > 0$. Then for each j there exists

a sequence $(E_{j,k})_{k\in\mathbb{N}}$ in $\mathcal{E}$ such that $A_j \subset \bigcup_{k=1}^{\infty} E_{j,k}$ and $\sum_{k=1}^{\infty} \rho(E_{j,k}) \leq \mu^*(A_j) + \epsilon 2^{-j}$. Clearly, $\bigcup_{j=1}^{\infty} A_j \subset \bigcup_{j,k=1}^{\infty} E_{j,k}$and $\sum_{j,k} \rho(E_{k,j}) \leq \sum_{j=1}^{\infty}(\mu^*(A_j)+2^{-j}\epsilon) = \sum_{j=1}^{\infty} \mu^*(A_j)+\epsilon$. Since $\epsilon > 0$ is arbitrary, the inequality follows. □

Example 12. *As a simple application of Lemma 2.0.12, we see that Lebesgue outer measure λ^* is indeed an outer measure.*

Definition 2.0.12. *Let X be a nonempty set and μ^* an outer measure on X. A set $A \in \mathcal{P}(X)$ is called μ^*-measurable if*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c), \quad \forall E \subset X.$$

Remark 2.0.5. We note that in the above definition $\leq$ follows from subadditivity. Thus $\geq$ is the nontrivial property a μ^*-measurable set must satisfy.

Theorem 2.0.13. *(Caratheodory; Existence of Measures) Let X be a nonempty set and μ^* an outer measure on X. The collection, $\mathcal{A}$, of μ^*-measurable sets is a σ-algebra, and the restriction of μ^* to $\mathcal{A}$ is a complete measure.*

Proof. (The following proof is inspired from G. Teschl, notes online). First we observe that $\mathcal{A}$ is closed under complements since the defintion of μ^* measurability of a set is the same as its complement.
Next we show finite additivity. Let $A, B \in \mathcal{A}$ and $E \subset \mathcal{P}(X)$. Applying Caratheodory's condiditon twice

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\ &= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c). \end{aligned}$$

But $A \cup B = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$, so by subadditivity

$$\mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) \geq \mu^*(E \cap (A \cup B)),$$

and hence

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c).$$

It follows that $A \cup B \in \mathcal{A}$.
Next we want to show that $\mathcal{A}$ is closed with respect to countable unions. To this end let (A_n) be a sequence in $\mathcal{A}$. Without restriction we can assume that they are disjoint (otherwise replace A_n by $A_n \setminus \bigcup_{j=1}^{n-1} A_j$). Abbreviate $\tilde{A}_n = \bigcup_{k\le n} A_n$, $A = \bigcup_n A_n$. Then for any set $E \in \mathcal{P}(X)$ we have

$$\begin{aligned}
\mu^*(\tilde{A}_n \cap E) &= \mu^*(A_n \cap \tilde{A}_n \cap E) + \mu^*(A_n^c \cap \tilde{A}_n \cap E) \\
&= \mu^*(A_n \cap E) + \mu^*(\tilde{A}_{n-1} \cap E) \\
&= \ldots = \sum_{k=1}^{n} \mu^*(A_k \cap E).
\end{aligned}$$

Using $\tilde{A}_n \in \mathcal{A}$ and monotonicity of μ^*, we infer

$$\begin{aligned}
\mu^*(E) &= \mu^*(\tilde{A}_n \cap E) + \mu^*(\tilde{A}_n^c \cap E) && (2.0.5)\\
&\geq \sum_{k=1}^{n} \mu^*(A_k \cap E) + \mu^*(A^c \cap E). && (2.0.6)
\end{aligned}$$

Letting $n \to \infty$ and using subadditivity finally gives

$$\begin{aligned}
\mu^*(E) &\geq \sum_{k=1}^{\infty} \mu^*(A_k \cap E) + \mu^*(A^c \cap E) \\
&\geq \mu^*(A \cap E) + \mu^*(A^c \cap E) \geq \mu^*(E). && (2.0.7)
\end{aligned}$$

thus $A \in \mathcal{A}$, and we infer that $\mathcal{A}$ is a σ-algebra. Finally, setting $E = A$ in (2.0.7), we obtain

$$\mu^*(A) = \sum_{k=1}^{\infty} \mu^*(A_k \cap A) + \mu^*(A^c \cap A) = \sum_{k=1}^{\infty} \mu^*(A_k),$$

and we are done □

To show Theorem 2.0.9 it suffices to show that a generating set of the Borel σ-algebra (see Example 8) is λ^*-measurable. This is the content of the following lemma.

Lemma 2.0.14. *Let I be an open interval of $\mathbb{R}^d$, then*

$$\lambda^*(E) = \lambda^*(I \cap E) + \lambda^*(I^c \cap E), \quad \forall E \subset \mathbb{R}^d.$$

Proof. $\leq$: follows from the subadditivity of the outer measure.
$\geq$: For notational simplicity we only show this in $\mathbb{R}$ (the generalization to $\mathbb{R}^d$ is obvious). Let $I = (a, b)$. Let $\epsilon > 0$. By definition of the outer measure there exists a sequence $(I_j)_{j\in\mathbb{N}}$ of open intervals convering E such that

$$\sum_j |I_j| \leq \lambda^*(E) + \epsilon.$$

Then $I'_j := I_j \cap I$, $j \in \mathbb{N}$, defines a sequence of disjoint open intervals such that $E \cap I \subset \bigcup I'_j$ and $I''_j := I''_{j,-} \cup I''_{j,+}$ with $I''_{j,-} = I_j \cap (-\infty, a + \epsilon 2^{-j})$ and $I''_{j,+} = I_j \cap (b - \epsilon 2^{-j}, \infty)$, $j \in \mathbb{N}$, establishes a family (which we can easily arrange in a sequence) of open intervals such that $E \cap I^c \subset \bigcup I''_j$. It follows from the definition of the outer measure that

$$\lambda^*(I \cap E) + \lambda^*(I^c \cap E) \leq \sum (|I'_j| + (|I''_{j,-}| + |I''_{j,+})) \leq \sum_i (|I_j| + 2\epsilon 2^{-j}) \leq \lambda^*(E) + 3\epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have shown $\geq$. $\square$

Proof of Theorem 2.0.9. First recall that Lebesgue outer measure λ^* is by Lemma 2.0.12 indeed an outer measure. Thus by Caratheodory's extension theorem (Theorem 2.0.13) there exists a measure on the λ^*-measurable sets. Since open rectangles are λ^*-measurable by Lemma 2.0.14, and the set of open rectangles generates the Borel measurable sets of $\mathbb{R}^d$ (see Example 8), the Borel measurable sets must be contained in the σ-algebra of λ^*-measurable sets. $\square$

Theorem 2.0.15. *The following holds about Lebesgue measurability and the Lebesgue measure. A set of outer measure zero is measurable (and has zero measure). Let $E \subset \mathbb{R}^d$ then the following properties are equivalent.*

(a) E is Lebesgue measurable.

(b) For every $A \subset \mathbb{R}^d$ we have $\lambda^*(A) = \lambda^*(E \cap A) + \lambda^*(E^c \cap A)$. (Caratheodory)

Proof. Let M be a set of outer measure zero. Then by definition of the outer measure there exists a sequence of open sets (U_j) such that $M \subset U_j$ and $\lambda(U_j) \to 0$. Thus

$N := \bigcap U_j$ is measurable, satisfies $\lambda(N) = 0$ by continuity of measures, and $M \subset N$. Since Lebesgue measure is complete the claim follows.
(a) $\Rightarrow$ (b) This follows from Caratheodory's theorem, and the fact that open intervals satisfy the condition, see Lemma 2.0.14.
(b) $\Rightarrow$ (a) Suppose first that $\lambda^*(E) < \infty$. By definition of the outer measure, there exists a sequence of open sets (U_j) such that $E \subset U_j$ and $\lambda^*(E) = \lim_j \lambda(U_j)$. Then $V = \bigcap U_j$ satisfies $\lambda^*(E) = \lambda(V)$ and $E \subset V$. Now by the condition we find $\lambda^*(V) = \lambda^*(E \cap V) + \lambda^*(E^c \cap V) = \lambda^*(V) + \lambda^*(E^c \cap V)$. Thus $N := V \setminus E$ is a set of measure zero. We conclude that $E = V \setminus N$ and the claim follows form Theorem 2.0.11. For the general case we refer the reader to [Wheeden & Zygmund Thm (3.30)] $\square$

2.0.4 Uniqueness of Measures and Monotone Classes

Monotone class theorem is useful to prove that a measure is uniquely determined an a generating algebra.

Definition 2.0.13. *We define a* **monotone class** *on a space X to be a subset $\mathcal{M}$ of $\mathcal{P}(X)$ that is closed under countable increasing unions and countable decreasing intersections, i.e.*

(i) if $A_i \in \mathcal{M}$ for $i = 1, 2, ...,$ and if $A_1 \subset A_2 \subset \cdots$, then $\bigcup_i A_i \in \mathcal{M}$.

(ii) if $B_i \in \mathcal{M}$ for $i = 1, 2, ...,,$ and if $B_1 \supset B_2 \supset \cdots$, then $\bigcap_i B_i \in \mathcal{M}$.

Remark 2.0.6. It is easy to see, that $\mathcal{P}(XI$ is monotone class, and that the intersection of any family of monotone classes is monotone class. So for any $\mathcal{E} \subset \mathcal{P}(X)$ there is a unique smallest monotone class containing $\mathcal{E}$, called the **monotone class generated by** $\mathcal{E}$.

Theorem 2.0.16. *(Monotone class theorem). If $\mathcal{A}$ is an algebra of subsets of X, then the monotone class generated by $\mathcal{A}$ coincides with the σ-algebra gerated by $\mathcal{A}$.*

Proof. Let $\mathcal{M} := \mathcal{M}(\mathcal{A})$ and $\mathcal{C} := \mathcal{C}(\mathcal{A})$ denote the σ-algebra and the monotone class generated by $\mathcal{E}$, respectively. The claim we have to show is $\mathcal{C} = \mathcal{M}$.
$\subset$: Since $\mathcal{M}$ is monotone class, we have $\mathcal{C} \subset \mathcal{M}$.
$\supset$: It suffices to show that $\mathcal{C}$ is a σ-algebra.
Clearly, $\mathcal{C}$ is nonemtpy, since $\mathcal{A}$ is. Next we show that $\mathcal{C}$ is closed under complements. For this define

$$\mathcal{C}^{(c)} = \{F \in \mathcal{C} : E^c \in \mathcal{C}\}.$$

Then $\mathcal{C}^{(c)}$ contains $\mathcal{A}$, since $\mathcal{A}$ is an algebra. For any increasing sequence (A_i) in $\mathcal{C}^{(c)}$ and decreasing sequence (B_i) in $\mathcal{C}^{(c)}$ we have

$$(\bigcup A_i)^c = \bigcap A_i^c \in \mathcal{C}, \quad (\bigcap B_i)^c = \bigcup B_i^c \in \mathcal{C},$$

since $\mathcal{C}$ is monotone class. Thus $\mathcal{C}^{(c)}$ is a monotone class containing $\mathcal{A}$, and hence $\mathcal{C}^{(c)} = \mathcal{C}$. Let us now show that $\mathcal{C}$ is closed under unions. For this define for any $E \in \mathcal{P}(X)$, the set

$$\mathcal{C}_E = \{F \in \mathcal{C} : E \cup F \in \mathcal{C}\}.$$

First assume that $E \in \mathcal{A}$. Then $\mathcal{A} \subset \mathcal{C}_E$ (since $\mathcal{A}$ is an algebra). For any increasing sequence (A_i) in $\mathcal{C}_E$ and decreasing sequence (B_i) in $\mathcal{C}_E$, the sequences $(E \cup A_i)$ and $(E \cup B_i)$ are increasing and decreasing sequences in $\mathcal{C}$, respectively, and hence

$$E \cup \bigcup A_i = \bigcup (E \cup A_i) \in \mathcal{C}, \quad E \cup \bigcap B_i = \bigcap (E \cup B_i) \in \mathcal{C},$$

since $\mathcal{C}$ is monotone class. Thus $\mathcal{C}_E$ is a monotone class containing $\mathcal{A}$, and hence $\mathcal{C}_E = \mathcal{C}$. Next, assume that $E \subset \mathcal{C}$. Then $\mathcal{A} \subset \mathcal{C}_E$ by what we have just shown. A verbatim repetition of the previous argment, shows that $\mathcal{C}_E = \mathcal{C}$. We conclude that $\mathcal{C}$ is closed under finite unions. To show that it is closed under countable unions, let (E_j) be a sequence in $\mathcal{C}$. Then $\bigcup_{j=1}^{n} E_j \in \mathcal{C}$ for all $n \in \mathbb{N}$ (by what we have shown), and since $\mathcal{C}$ is closed under countable increasing unions it follows that $\bigcup E_j \in \mathcal{C}$. In short $\mathcal{C}$ is a σ-algebra and we are done.

□

Theorem 2.0.17. *(Uniqueness of measures). Let Ω be a set, $\mathcal{A}$ an algebra of substs of Ω and Σ the smallest σ-algebra that contains $\mathcal{A}$. Let μ_1 be a measure such that there exists a sequence of sets $A_i \in \mathcal{A}$ such that $\mu_1(A_i) < \infty$ and $\Omega = \bigcup_i A_i$. If μ_2 is a measure that coincides with μ_1 on $\mathcal{A}$, then $\mu_1 = \mu_2$ on all of Σ.*

Proof. We first prove the theorem under the assumption that μ_1 is a finite measure on Ω. Consider the set

$$\mathcal{M} = \{A \in \Sigma : \mu_1(A) = \mu_2(A)\}.$$

Clearly, this collection of sets contains $\mathcal{A}$ and we shall show that $\mathcal{A}$ is a monotone class. By the monotone class theoren we will then conclude that $\mathcal{M} = \Sigma$. Let (A_j) and (B_j) be an increasing and decreasing sequence in $\mathcal{M}$, respectively. Then by the continuity from below and above

$$\mu_1(\bigcup A_j) = \lim_{j\to\infty} \mu_1(A_j) = \lim_{j\to\infty} \mu_2(A_j) = \mu_2(\bigcup A_j),$$
$$\mu_1(\bigcap B_j) = \lim_{j\to\infty} \mu_1(B_j) = \lim_{j\to\infty} \mu_2(B_j) = \mu_2(\bigcap B_j),$$

Thus $\mathcal{M}$ is monotone class.

Next we return to the σ-finite case. Thus there exists a sequence (A_j) in $\mathcal{A}$ such that $\bigcup A_j = X$ and $\mu(A_j) < \infty$. Define $C_n = \bigcup_{k=1}^n A_j$. Then (C_n) is an increasing sequence in $\mathcal{A}$ with $\mu(C_n) < \infty$ and $\bigcup C_n = X$. Thus by continuity from below, and what we have just shown, we find for any $E \in \Sigma$ that

$$\mu_1(E) = \lim_n \mu_1(E \cap C_n) = \lim_n \mu_2(E \cap C_n) = \mu_2(E).$$

$\square$

2.1 Integration

Definition 2.1.1. *Let $(X, \mathcal{A}_X)$ and $(Y, \mathcal{A}_Y)$ be two measurable spaces. A function*

$$f : X \to Y$$

is called $(\mathcal{A}_X, \mathcal{A}_Y)$-**measurable**, *if $f^{-1}(E) \in \mathcal{A}_X$ for all $E \in \mathcal{A}_Y$.*

Remark 2.1.1. If not otherwise specified we shall always equip $\mathbb{R}^d$ with the Borel sigma algebra.

Lemma 2.1.1.

(i) The composition of measurable maps is measurable.

(ii) If $\mathcal{A}_Y$ is generated by $\mathcal{E}$, then $f : X \to Y$ is measurable iff $f^{-1}(E) \in \mathcal{A}_X$ for all $E \in \mathcal{E}$.

(iii) If Y and Y are metric spaces, every continuous map $f : X \to Y$ is $(\mathcal{B}_X, \mathcal{B}_Y)$ measurable.

Proof. (i) Follows from the properties of preimages.
(ii) The only if is trivial. For the converse, observe that $\{E \subset Y : f^{-1}(E) \in \mathcal{A}_X\}$ is a σ-algebra that contains $\mathcal{E}$ (by elemtary properties of preimages); it therefore contains $\mathcal{A}_Y$. The second statement follows from the first.
(iii) This follows (ii).

□

Proposition 2.1.2. *Let $(X, \mathcal{A}_X)$ and $(Y_\alpha, \mathcal{A}_{Y_\alpha})$ for $\alpha \in A$ be measurable spaces. Let $Y = \prod_{\alpha \in A} Y_\alpha$, and let $\pi_\alpha : Y \to Y_\alpha$ be the coordinate projection. Then $f : X \to \prod_{\alpha \in A} Y_\alpha$ ist $(\mathcal{A}_X, \bigotimes_{\alpha \in A} \mathcal{A}_{Y_\alpha})$ measurable if $\pi_\alpha \circ f$ is measurable for all $\alpha \in A$.*

Proof. $\Rightarrow$: The composition of measurable maps is measurable.
$\Leftarrow$: Suppose each $\pi_\alpha \circ f$ is measurable. Then for each $E_\alpha \in \mathcal{A}_{Y_\alpha}$ we have $f^{-1}(\pi_\alpha^{-1}(E_\alpha)) = (f \circ \pi_\alpha)^{-1}(E_\alpha) \in \mathcal{A}_X$. It follows that f is measurable by (ii) of the above Lemma 2.1.1 (since $\bigotimes_{\alpha \in A} \mathcal{A}_{Y_\alpha}$ is generatged by sets of the form $\pi_\alpha^{-1}(E_\alpha)$). □

Corollary 2.1.3. *Let $(X, \mathcal{A}_X)$ be a measurable space.*

(i) A function $X \to \mathbb{C}$ is measurable iff $\mathrm{Re} f$ and $\mathrm{Im} f$ are measurable.

(ii) If $f, g : X \to \mathbb{C}$ are measurable, then so are $f + g$ and fg.

Proof. (i) Follows since $\mathcal{B}_{\mathbb{C}} = \mathcal{B}_{\mathbb{R}^2} = \mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$ (see Lemma 2.0.4).
(ii) Define the map $A : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$, $(x, y) \to x + y$ and the map $(f, g) : X \to \mathbb{C} \times \mathbb{C}$, $x \mapsto (f(x), g(x))$. Now A is continuous and hence $(\mathcal{B}_{\mathbb{C}^2}, \mathcal{B}_{\mathbb{C}})$ measurable, and (f, g) is $(\mathcal{A}_X, \mathcal{B}_{\mathbb{C}}\otimes, \mathcal{B}_{\mathbb{C}}\ )$ measurable by Lemma 2.1.2. Since $\mathcal{B}_{\mathbb{C}^2} = \mathcal{B}_{\mathbb{C}}\otimes, \mathcal{B}_{\mathbb{C}}$ (see Lemma 2.0.4) it follows that $f + g = A \circ (f, g)$ is measurable as a composition of measurable maps. Likewise we show the measurability of fg usint the continuous map $M : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ $(x, y) \mapsto xy$. □

Remark 2.1.2. It is sometimes convenient to consider functions with values in the extended real number system $\overline{\mathbb{R}} = [-\infty, \infty]$. We define Borel sets in $\overline{\mathbb{R}}$ by $\mathcal{B}_{\overline{\mathbb{R}}} = \{E \subset \overline{\mathbb{R}} : E \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\}$.

Theorem 2.1.4. *Let $(X, \mathcal{A}_X)$ be a measurable space.*

(a) Let $(f_j)_{j\in\mathbb{N}}$ be a sequence of $\mathbb{R}$ valued measurable functions on X. Then the following functions are measurable

(i) $g_+(x) := \sup_j f_j(x)$, $g_-(x) := \inf_j f_j(x)$

(ii) $h_+(x) := \limsup_j f_j(x)$, $h_-(x) := \liminf_j f_j(x)$

(iii) $f(x) := \lim_{j\to\infty} f(x)$ provided the limit exists for every $x \in X$.

(b) Let $(f_j)_{j\in\mathbb{N}}$ be a sequence of $\mathbb{C}$ valued measurable functions on X. Then $f(x) := \lim_{j\to\infty} f(x)$ is measurable provided the limit exists for every $x \in X$.

Proof. (a)(i) Note that $g_+(-\infty, a) = \bigcup_{j=1}^{\infty} f_j^{-1}(a, \infty)$. Thus g_+ is measurable by (ii) of Lemma 2.1.1 and Example 7. Similarly one shows the measurability of g_-.
(ii) The claim for h_+ follows from the definition of limsup, i.e., $\limsup_j f_j(x) = \inf_k \sup_{k\leq l} f_l(x)$,

and a repeated application of (ii). Analogously the claim follows for h_-.
(iii) If the limit exists, then we have $\lim_j f_j(x) = \limsup_j f_j(x)$ and the claim follows from (ii).
(b) Follos from (iii) if (a) and (i) of Corollary 2.1.3. $\square$

Definition 2.1.2. *A complex value function s on a measurable space X whose range consists of finitely manh points will be called a* **simple function***.*

Lemma 2.1.5.

(a) Let $s : X \to \mathbb{C}$ be a simple function. Then there exists an $N \in \mathbb{N}$ and distinct numbers $\alpha_j \in \mathbb{C}$, $j = 1, ..., n$ such that $\mathrm{Ran} s = \{\alpha_1, ..., \alpha_n\}$ *for some $n \in \mathbb{N}$. Moreover,*

$$s = \sum_{j=1}^{n} \alpha_j 1_{s^{-1}(\{\alpha_j\})} = \sum_{c \in \mathrm{Ran} s} c 1_{s^{-1}(\{c\})}.$$

(b) A simple function s is measurable if and only if $s^{-1}(\{c\})$ is measurable for all $c \in$ $\mathrm{Ran} s$*.*

(c) For finitely many complex numbers $\beta_1, ..\beta_m$ and subsets $B_1, ..., B_m$ of X (not necessarily disjoint), the function $\sum_{j=1}^{m} \beta_j 1_{B_j}$ is simple.

(d) A finite linear combination of simple functions is simple.

Proof. Trivial. $\square$

Theorem 2.1.6. *Let $(X, \mathcal{A})$ be a measurable space. If $f : X \to [0, \infty]$ is measurable, there exists a sequence $(s_n)_{n \in \mathbb{N}}$ of simple functions such that*

- $0 \le s_1 \le s_2 \le \cdots \le f$
- *$s_n \to f$ pointwise, and $s_n \to f$ uniformly on any set on which f is bounded.*

Picture

Proof. For $n \in \mathbb{N}$ and $0 \leq k \leq 2^{2n} - 1$, let

$$E_{n,k} := f^{-1}\left((k2^{-n}, (k+1)2^{-n}]\right) \quad \text{and} \quad F_n := f^{-1}\left((2^n, \infty]\right),$$

and define

$$s_n = \sum_{k=0}^{2^{2n}-1} k2^{-n} 1_{E_{n,k}} + 2^n 1_{F_n},$$

One easily checks that this function satisfies the required properties.

$\square$

Definition 2.1.3. *Let $(X, \mathcal{A}, \mu)$ be a measure space and $E \in \mathcal{A}$. Let $s : X \to [0, \infty]$ be a measruable and simple function, then we define*

$$\int_E s d\mu = \sum_{c \in \operatorname{Ran} s} c\mu(s^{-1}(\{c\}) \cap E) \tag{2.1.1}$$

with convention $0 \cdot \infty = 0$.

Proposition 2.1.7. *Let s, t be nonnegative measurable simple functions on X.*

(a) The map

$$\varphi : \mathcal{A} \to [0, \infty], \quad E \mapsto \varphi(E) = \int_E s d\mu,$$

is a measure on $\mathcal{A}$.

(b) We have

$$\int_X (s+t) d\mu = \int_X s d\mu + \int_X t d\mu.$$

(c) If $s \leq t$, then $\int_X s d\mu \leq \int_X t d\mu$.

(This Proposition is only preliminary, we shall generalize it later to a much larger class of functions.)

Proof. (a) We have for any sequence $(E_j)_{j\in\mathbb{N}}$ of disjoint members of $\mathcal{A}$, by the countable additivity of μ that

$$\begin{aligned}\varphi(\bigcup_j E_j) &= \sum_{c\in\operatorname{Ran} s} \mu(s^{-1}(\{c\})\cap\bigcup_j E_j)\\ &= \sum_{c\in\operatorname{Ran} s}\sum_{j=1}^{\infty}\mu(s^{-1}(\{c\})\cap E_j)\\ &= \sum_{j=1}^{\infty}\sum_{c\in\operatorname{Ran} s}\mu(s^{-1}(\{c\})\cap E_j)\\ &= \sum_{j=1}^{\infty}\varphi(E_j).\end{aligned}$$

Also $\varphi(\emptyset)=0$.
(b) For $c\in\operatorname{Ran} s$ and $d\in\operatorname{Ran} t$ we have for $E_{c,d}=s^{-1}(\{c\})\cap t^{-1}(\{d\})$ that

$$\int_{E_{c,d}}(s+t)d\mu=(c+d)\mu(E_{c,d})=c\mu(E_{c,d})+d\mu(E_{c,d})=\int_{E_{c,d}} sd\mu+\int_{E_{c,d}} td\mu.$$

Thus (b) holds for $E_{c,d}$ in place of X. Since X is a disjoint union of the sets $\{E_{c,d} : c\in\operatorname{Ran} s,\ d\in\operatorname{Ran} t\}$, (b) follows from (a).
(c) This follows from (b) and $\int_X td\mu=\int_X sd\mu+\int_X(t-s)d\mu\geq\int_X sd\mu$. $\square$

Definition 2.1.4. *Let $(X,\mathcal{A},\mu)$ be a measure space and $E\in\mathcal{A}$. Define*

$$L^+(X,\mu):=\{f:X\to[0,\infty]:\ f \text{ measurable }\}.$$

If $f\in L^+(X,d\mu)$ we define the **Lebesgue integral** *of f over E with respect to the measure μ by*

$$\int_E fd\mu=\sup\left\{\int_E sd\mu : 0\leq s\leq f,\quad s \text{ simple, measurable}\right\}. \tag{2.1.2}$$

Remark 2.1.3. If from the context it is clear, what the measure space μ or the set X is we shall write $\int f$, $\int_X f$, or $\int fd\mu$ for $\int_X fd\mu$.

Remark 2.1.4. Observe that we apparently have two definitions for $\int_E f d\mu$ if f is simple, namely, (2.1.1) and (2.1.2). However, these assign the same value of the integral, since f is, in this case, the largest of the functions s which occur on the right of (2.1.2).

Remark 2.1.5. The following properties are immediate consequences of the definitions. The functinos and sets occuring in them are assumed to be measurable.

(a) If $0 \leq f \leq g$, then $\int_E f d\mu \leq \int_E g d\mu$.

(b) If $A \subset B$ and $f \geq 0$, then $\int_A f d\mu \leq \int_B f d\mu$.

(c) If $f \geq 0$ and $c \in [0, \infty)$, then $\int_E c f d\mu = c \int_E f d\mu$.

(d) if $f(x) = 0$ for all $x \in E$, then $\int_E f d\mu = 0$ (even if $\mu(E) = \infty$).

(e) If $\mu(E) = 0$, then $\int_E f d\mu = 0$ (even if $f(x) = \infty$ for every $x \in E$).

(f) If $f \geq 0$, then $\int_E f d\mu = \int_X 1_E f d\mu$.

Theorem 2.1.8. *(Lebesgue's Monotone Convergence Theorem) Let $(f_j)_{j \in \mathbb{N}}$ be a sequence in $L^+(X, \mu)$, such that*

(i) $0 \leq f_1(x) \leq f_2(x) \cdots \leq \infty$ for every $x \in X$, ($f_j \leq f_{j+1}$ for all $j \in \mathbb{N}$)

(ii) $f_n(x) \to f(x)$ as $n \to \infty$ for every $x \in X$. ($f = \lim_{j \to \infty} f_j$ pointwise)

Then f is measurable and

$$\int f_X d\mu = \lim_{n \to \infty} \int_X f_n d\mu.$$

Proof. Since $\int f_n d\mu \leq \int f_{n+1} d\mu$, the limit

$$\alpha := \lim_n \int f_n d\mu \tag{2.1.3}$$

exists as a number in $[0, \infty]$. By Theorem 2.1.4, f is measurable. Since $f_n \leq f$, we have $\int f_n \leq \int f$ for every n, so (2.1.3) implies

$$\alpha \leq \int f d\mu \tag{2.1.4}$$

Let s be a simple function such that $0 \leq s \leq f$. Let c be a constant with $0 < c < 1$, and define

$$E_n = \{x : f_n(x) \geq cs(x)\}, \quad n \in \mathbb{N}.$$

Then each E_n is measurable, $E_1 \subset E_2 \subset E_3 \subset \cdots$ (by montonicity), and

$$X = \bigcup_n E_n \tag{2.1.5}$$

as we now show. Let $x \in X$. In case $f(x) = 0$, we have $x \in E_1$; in case $f(x) > 0$, we have $cs(x) < f(x)$, since $c < 1$, and hence $x \in E_n$ for some n. Hence we have shown 2.1.5. Also

$$\int_X f_n d\mu \geq \int_{E_n} f_n d\mu \geq c \int_{E_n} s d\mu, \quad n \in \mathbb{N}. \tag{2.1.6}$$

Now taking $n \to \infty$ and and applying (a) of and continuity of measures from below (Theorem 2.0.7) to the right hand side of (2.1.6), we find

$$\alpha \geq c \int_X s d\mu. \tag{2.1.7}$$

Since (2.1.7) holds for every $c < 1$, we have $\alpha \geq \int_X s d\mu$ for every simple measurable s satisying $0 \leq s \leq f$, so that

$$\alpha \geq \int_X f d\mu. \tag{2.1.8}$$

The theorem now follows from (2.1.3), (2.1.4), and (2.1.8). □

Remark 2.1.6. The monotone convergence theorem is an essential tool in many situations, but its immediate significance for us is asfollows. The definition of $\int f$ involves the supremum over a huge (usually uncountable) family of simple functions, so it may be difficult to evaluate $\int f$ directly from the definition. The monotone convergence theorem, however, assures us that to compute $\int f$ it is enough to compute $\lim s_n$ where (s_n) is any sequence of measurable simple functions that increase to f, and Theorem 2.1.6 guarantees that such sequences exist. Thus for any $f \in L^+$ we have

$$\int f d\mu = \lim_{n\to\infty} \int s_n d\mu$$

for the sequence (s_n) as in Theorem 2.1.6.

Corollary 2.1.9. *Let (X, μ) be a measure space.*

(a) For $f, g \in L^+(X, \mu)$ we have

$$\int_X (f+g)\mu = \int_X f d\mu + \int_X g d\mu.$$

(b) If $(f_j)_{j\in\mathbb{N}}$ is a sequence in $L^+(X, \mu)$, then

$$\int_X \sum_{j=1}^{\infty} f_j d\mu = \sum_{j=1}^{\infty} \int_X f_j d\mu.$$

(c) If $f \in L^+(X, \mu)$, then $\int_X f d\mu = 0$ iff $f = 0$ a.e.

Proof. (a) By Theorem 2.1.6 we can find sequences (s_n) and (t_n) of measurable nonnegative simple functions that increase to f and g. Then $(s_n + t_n)$ increases to $f + g$, so by the monotone convergence theorem and Proposition 2.1,

$$\int (f+g) = \lim \int (s_n + t_n) = \lim (\int s_n + \int t_n) = \lim \int s_n + \lim \int t_n = \int f + \int g.$$

(b) By induction we find from (a) that $\int \sum_{j=1}^N f_j = \sum_{j=1}^N \int f_j$ for any finite N. Letting $N \to \infty$ and applying the monotone convergence theorem the claim follows.
(c) Define $f_n = f 1_{f\geq \frac{1}{n}}$. Then $f_n \to f$ pointwise and $0 \leq f_n \leq f_{n+1}$. $\frac{1}{n}\mu(\{f \geq 1/n\}) = \int \frac{1}{n} 1_{f\geq 1/n} \leq \int f_n \leq 0$. Hence $\mu(f \geq 1/n) = 0$. Thus $\bigcup \{f \geq 1/n\} = \{f > 0\}$ has measure zero. This implies that $f = 0$ a.e.. □

Theorem 2.1.10. *(Fatou's Lemma) let $(f_n)_{n\in\mathbb{N}}$ is any sequence in L^+, then*

$$\int \liminf{}_n f_n \leq \liminf{}_n \int f_n.$$

Proof. For each $j \geq k \geq 1$ we have $\inf_{n\geq k} f_n \leq f_j$, hence $\int \inf_{n\geq k} f_n \leq \int f_j$. It follows that $\int \inf_{n\geq k} f_n \leq \inf_{j\geq k} \int f_j$. Now taking $k \to \infty$ and applying the monotone convergence theorem, we find

$$\int \liminf f_n = \lim_{k\to\infty} \int \inf_{n\geq k} f_n \leq \liminf_n \int f_n.$$

□

The integral defined above can be extended to real-valued measurable functions f in an obvious way. Define $f_+ := \max(f, 0)$ and $f_- := \max(-f, 0)$.

Definition 2.1.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space and $E \in \mathcal{A}$. Let $f : X \to \mathbb{R}$ be measurable. We say that f is* **integrable** *over E if $\int_E f_=$ and $\int_E f_-$ are finite. In that case we define*

$$\int_E f := \int_E f_+ - \int_E f_-$$

and call it the **Lebegue integral**.

Remark 2.1.7. For measurable $f : X \to \mathbb{R}$ measurable one can define more generally

$$\int f := \int f_+ - \int f_-$$

if at least one $\int f_+$ or $\int f_-$ is finite.

Remark 2.1.8. For measurable $f : X \to \mathbb{R}$ it is clear that f is integrable iff $\int |f| < \infty$.

Proposition 2.1.11. *The set of integrable functions real-valued functions on X is a real vector space and the integral is a linear functional on it.*

Proof. The first assertion follows from the fact that $|af + bg| \leq |a||f| + |b||g|$, and it is easy to see that $\int af = a \int f$ for any $a \in \mathbb{R}$. To show additivity, suppose that f and g are integrable. Then $(f+g)_+ - (f+g)_- = f_+ - f_- + g_+ - g_-$, so $(f+g)_+ + f_- + g_- = (f+g)_- + f_+ + g_+$, and by (a) of Corollary 2.1.9,

$$\int (f+g)_+ + \int f_- + \int g_- = \int (f+g)_- + \int f_+ + \int g_+.$$

Bringing the terms involving the negative part $(\cdot)_-$ to the other side yields additivity. □

Next we can extend the definition of the integral to complex-valued measurable functions f again in an obvious way.

Definition 2.1.6. *Let $(X, \mathcal{A}, \mu)$ be a measure space and $E \in \mathcal{A}$. Define*

$$L^1(E, \mathbb{C}, d\mu) := \left\{ f : X \to \mathbb{C} : \ f \text{ measurable}, \ \int_E |f| d\mu < \infty \right\}$$

If $f \in L^1(E, \mathbb{C}, d\mu)$, define

$$\int_E f d\mu := \int_E \mathrm{Re} f d\mu + i \int_E \mathrm{Im} f d\mu.$$

(Note that since $\mathrm{Re} f_\pm \leq |f|$ and $\mathrm{Im} f_\pm \leq |f|$, the expressions on the right hand side are well defined.)

Remove the E

Proposition 2.1.12. *The space of complex-valued integralbe functions is a complex vector space and the integral is a complex linerar functional on it. If $f \in L^1$, then $|\int f| \leq \int |f|$.*

Proof. The first part follows from the previous proposition. This is trivial if $\int f = 0$ and almost trivial if f is real, since

$$\left| \int f \right| = \left| \int f_+ - \int f_- \right| \leq \int f_+ + \int f_- = \int |f|.$$

If f is complex-valued and $\int f \neq 0$, let $\alpha = \overline{\int f} / |\int f|$. Then $|\int f| = \alpha \int f = \int \alpha f$. In particular, $\int \alpha f$ is real, so

$$\left| \int f \right| = \mathrm{Re} \int \alpha f = \int \mathrm{Re}(\alpha f) \leq \int |\mathrm{Re}(\alpha f)| \leq \int |\alpha f| = \int |f|.$$

□

Theorem 2.1.13. *(Dominated Convergence) Let $(f_n)_{n \in \mathbb{N}}$ be a sequence in $L^1(X, \mathbb{C})$ such that*

(i) $f_n(x) \to f(x)$ *for every* $x \in X$,

(ii) there exists a function $g \in L^1(X, \mathbb{C})$ *such that* $|f_n| \leq g$, $\forall n \in \mathbb{N}$.

Then $f \in L^2(X, \mathbb{C})$ and $\lim_{n \to \infty} \int f_n = \int f$.

Proof. f is measurable, and since $|f| \leq g$ a.e., we have $f \in L^1$. By taking real and imaginary parts it suffices to assume that f_n and f are real valued, in which case we have $g + f_n \geq 0$ a.e. and $g - f_n \geq 0$. Thus by Fatou's lemma,

$$\int g + \int f \leq \liminf \int (g + f_n) = \int g + \liminf \int f_n,$$
$$\int g - \int f \leq \liminf \int (g - f_n) = \int g - \limsup \int f_n.$$

Therefore, $\liminf \int f_n \geq \int f \geq \limsup \int f_n$, and the result follows. □

Remark 2.1.9. *The role played by sets of measure zero. See Rudin.*

Theorem 2.1.14. *(Parameter dependent integrals) Suppose that $f : X \times [a, b] \to \mathbb{C}$ $(-\infty < a < b < \infty)$ and that $f(\cdot, t) : X \to \mathbb{C}$ is integrable for each $t \in [a, b]$. Let*

$$F(t) = \int_X f(x, t) d\mu(x).$$

(a) *Suppose there exists $g \in L^1$ such that $|f(x, t)| \leq g(x)$ for all x, t. If $f(x, \cdot)$ is continuous at t_0 for each x, then F is continuous at t_0 and*

$$\lim_{t \to t_0} \int_X f(x, t) d\mu(x) = \int_X \lim_{t \to t_0} f(x, t) d\mu(x)$$

(b) *Suppose the partial derivative of f exists with respect to t and there exists $g \in L^1$ such that $|(\partial f / \partial t)(x, t)| \leq g(x)$ for all x, t. Then F is differentiable and*

$$\frac{d}{dt} \int_X f(x, t) d\mu(x) = \int_X \partial_t f(x, t) d\mu(x).$$

Proof. From [1].
For (a) apppy the dominated convergence theorem to $f_n(x) = f(x, t_n)$ wherer (t_n) is any sequence in $[a, b]$ convergint to t_0.
For (b) observe that

$$\partial_t f(x, t_0) = \lim h_n, \qquad \text{where } h_n(x) = \frac{f(x, t_n) - f(x, t_0)}{t_n - t_0},$$

(t_n) again being any sequence convergint to t_0. It follows that $\partial_t f$ is measurable, and by the mean value theorem

$$h_n(x) \leq \sup_{t \in [a,b]} |\partial_t f(x,t)| \leq g(x),$$

so the dominated convergence theorem can be invoked again to give

$$F'(t_0) = \lim \frac{F(t_n) - F(t_0)}{t - t_0} = \lim \int h_n(x) d\mu(x) = \int \partial_t f(x,t) d\mu(x).$$

□

2.2 Product Measures

Let X_1 and x_2 be two nonempty sets. If $E \subset X_1 \times X_2$, we define for $x \in X_1$ the **1st-coordinate-section**

$$E_1(x) = \{y \in X_2 : (x, y) \in E\}$$

we define for $y \in X_2$ the **2nd-coordinate-section**

$$E_2(y) = \{x \in X_1 : (x, y) \in E\}.$$

Also if f is a function on $X \times Y$ we define the **1st-coordinate-section** $f_{1,x}$ and **2nd-coordinate-section** $f_{2,y}$ by

$$f_{1,x}(y) = f_{2,y}(x) = f(x, y).$$

Proposition 2.2.1. *(Section Property.) Let $(X, \mathcal{A}_X)$ and $(Y, \mathcal{A}_Y)$ be measure spaces.*

(a) If $E \in \mathcal{A}_X \otimes \mathcal{A}_Y$, then $E_1(x) \in \mathcal{A}_Y$ for all $x \in X$ and $E_2(y) \in \mathcal{A}_X$ for all $y \in Y$.

(b) If f is $\mathcal{A}_X \otimes \mathcal{A}_Y$ measurable, then $f_{1,x}$ is $\mathcal{A}_Y$ measurable for all $x \in X$ and $f_{2,y}$ is $\mathcal{A}_X$ measurable for all $y \in Y$.

ljust Notation

Proof. Let $\mathcal{R}$ be the collection fo all subsets of E of $X \times Y$ such that $E_{1,x} \in \mathcal{A}_Y$ for all x and $E_{2,y} \in \mathcal{A}_X$ for all y. Then $\mathcal{R}$ obviously contains all rectangles. Since $(\bigcup E_j)_{1,x} = \bigcup (E_j)_{1,x}$ and $(E^c)_{1,x} = (E_{1,x})^c$, and likewise for y-sections, $\mathcal{R}$ is a σ-algebra. Therefore $\mathcal{R} \supset \mathcal{A}_X \otimes \mathcal{A}_Y$, which proves (a). (b) follows from (a) because $(f_{1,x})^{-1}(B) = (f^{-1}(B))_{1,x}$ and $(f_{2,y})^{-1}(B) = (f^{-1}(B))_{2,y}$. $\square$

Theorem 2.2.2. *(Product measure) Let $(\Omega_1, \Sigma_1, \mu_1)$, $(\Omega_2, \Sigma_2, \mu_2)$ be two sigma-finite measure spaces. For any $E \in \Sigma_1 \otimes \Sigma_2$ the functions $x \mapsto \mu_2(E_1(x))$ on Ω_1 and $y \mapsto \mu_1(E_2(y))$ on Ω_2 are measurable on Ω_1 and Ω_2, respectively, and*

$$(\mu_1 \times \mu_2)(E) := \int_{\Omega_1} \mu_2(E_1(x)) d\mu_1(x) = \int_{\Omega_2} \mu_1(E_2(y)) d\mu_2(y). \tag{2.2.1}$$

Moreover, $\mu_1 \times \mu_2$ defined by (2.2.1) is a σ-finite measure on $\Omega_1 \times \Omega_2$ with σ-algebra $\Sigma_1 \otimes \Sigma_2$, which is called the **product of the measures** *μ_1 and μ_2.*

Proof. Let $\mathcal{C}$ be the set of all $E \in \Sigma_1 \otimes \Sigma_2$ such that $x \mapsto \mu_2(E_1(x))$ and $y \mapsto \mu_1(E_2(y))$ are measurable. If $E = A_1 \times A_2$ with $A_j \in \Sigma_j$, then

$$x \mapsto \mu_2(E_1(x)) = 1_{A_1}(x)\mu_2(A_2), \quad y \mapsto \mu_1(E_2(y)) = 1_{A_2}(y)\mu_1(A_1) \tag{2.2.2}$$

are measurable and thus $E \in \mathcal{C}$. By additivity, it follows that finite disjoint unions of rectangles are in $\mathcal{C}$. So by the monotone class theorem it will suffice to show that $\mathcal{C}$ is monotone class (note that disjoint unions of rectangles forms an algebra, see Proposition 2.0.6). Suppose (E_n) is an increasing sequence in $\mathcal{C}$ and $U = \bigcup E_n$. Then the functions $y \mapsto \mu_1((E_n)_2(y))$ and $x \mapsto \mu_2((E_n)_1(x))$ increase pointwise to $y \mapsto \mu_1(U_2(y))$ and $x \mapsto \mu_2(U_1(x))$, which thus are also measurable. Suppose (E_n) is a decreasing sequence in $\mathcal{C}$ and $C = \bigcap E_n$. Then the functions $y \mapsto \mu_1((E_n)_2(y))$ and $x \mapsto \mu_2((E_n)_1(x))$ converge pointwise to $y \mapsto \mu_1(C_2(y))$ and $x \mapsto \mu_2(C_1(x))$, which thus are also measurable. Thus $\mathcal{C} = \Sigma_1 \otimes \Sigma_2$.
Next we show that (2.2.1) defines a measure.
Consider any collection of disjoint sets $(A_j) \in \Sigma_1 \otimes \Sigma_2$. Clearly their sections $(A_j)_1(x)$, are measurable (by the section property, see Proposition above) and also disjoint and hence

$$\mu_2((\bigcup A_j)_1(x)) = \sum_{j=1}^{\infty} \mu_2((A_j)_1(x)).$$

The montone convergence theorem then yields the countable additivity of $\mu_1 \times \mu_2$. Similarly, the second integral defines a countable additive measure.
Next we want to show the second idenity of (2.2.1). For this we can use as above the montone class theorem and σ-additivity or we can directly invoke the uniqueness theorem of measures. Les us do the latter. If $E = A_1 \times A_2$ with $A_j \in \Sigma_j$, then using (2.2.2) we see that

$$\int_{\Omega_1} \mu_2(E_1(x))d\mu_1(x) = \mu_1(A_1)\mu_2(A_2) = \int_{\Omega_2} \mu_1(E_2(y))d\mu_2(y).$$

By additivity, it follows that the identity holds for finite disjoint unions of rectangles. Since the set of such finite unions is an algebra, the claim follows from Uniqueness of measures, provided the finiteness condition holds. To this end, we know by σ-additivity there exists for $k = 1, 2$ a sequence $(A_j^{(k)})$ in Σ_k, such that $\mu_k(A_j^{(k)}) < \infty$ and $\bigcup A_j^{(k)} = \Omega_k$. Thus $\mu_1 \times \mu_2(A_j^{(1)} \times A_j^{(2)}) = \mu_1(A_j^{(1)})\mu_2(A_j^{(2)}) < \infty$ and $\bigcup_j A_j^{(1)} \times A_j^{(2)} = \Omega_1 \times \Omega_2$. This yields the claim. $\square$

Theorem 2.2.3. *(Fubini-Tonelli) Let $(X, \mathcal{A}_X, \mu_X)$ and $(Y, \mathcal{A}_Y, \mu_Y)$ be σ-finite measure spaces.*

(a) (Tonelli) If $f \in L^+(X \times Y)$, then the functions

$$y \mapsto \int_X f(x,y) d\mu_X(x), \quad x \mapsto \int_Y f(x,y) d\mu_Y(y) \tag{2.2.3}$$

are in $L^+(X)$ and $L^+(Y)$, respectively, and

$$\int_{X \times Y} f d(\mu_X \times \mu_Y) = \int_Y \left[\int_X f(x,y) d\mu_X(x) \right] d\mu_Y(y) \tag{2.2.4}$$

$$= \int_X \left[\int_Y f(x,y) d\mu_Y(y) \right] d\mu_X(x). \tag{2.2.5}$$

(b) (Fubini) If $f \in L^1(\mu_X \times \mu_Y)$, then $f_{1,x} \in L^1(\mu_Y)$ a.e. $x \in X$, and $f_{2,y} \in L^2(\mu_X)$ for a.e. $y \in Y$, the functions (2.2.3) *are in $L^1(\mu_X)$ and $L^1(\mu_Y)$, respectively, and* (2.2.4) *and* (2.2.5) *hold.*

Proof. Tonelli's theorem reduces to Theorem 2.2.2 in case f is a chracteristic function, and it therefore holds for nonnegative simple functions by linearity. If $f \in L^+(X \times Y)$, let (f_n) be a sequence of simple functions that increase pointwise to f as in Theorem ccc. The monotone convergence theorem implies, first, the measurability of the functions (2.2.3) and, second that

$$\begin{aligned}
\int_X \left[\int_Y f(x,y) d\mu_Y(y) \right] d\mu_X(x) &= \int_X \lim_n \left[\int_Y f_n(x,y) d\mu_Y(y) \right] d\mu_X(x) \\
&= \lim_n \int_X \left[\int_Y f_n(x,y) d\mu_Y(y) \right] d\mu_X(x) \\
&= \lim_n \int_{X \times Y} f_n d(\mu_Y \times \mu_X) \\
&= \int_{X \times Y} f d(\mu_Y \times \mu_X),
\end{aligned}$$

which is (2.2.4). The second line (2.2.5) is shown analgously. This shows (a).
(b) If $f \in L^+(X \times Y)$ and $\int f d(\mu \times \nu) < \infty$, then then it follows from (a) that $\int_X f(x,y)d\mu_X(x) < \infty$ a.e. y and $\int_Y f(x,y)d\mu_Y(y) < \infty$ a.e. x, that is $f_{1,x} \in L^1(\mu_Y)$ for a.e. x and $f_{2,y} \in L^1(\mu_X)$ for a.e. y. If $f \in L^1(\mu_X \times \mu_Y)$, then, the claimed identities follow from the application of (a) to the positive and negative parts of the real and imaginary parts of f. $\square$

2.3 L^p spaces

Definition 2.3.1. *Let* $(X, \mathcal{A}, \mu)$ *be a measure space and Let* $0 < p < \infty$*. We define for every measurable function* $f : X \to \mathbb{C}$ *the expression*

$$\|f\|_p := \left\{ \int |f|^p d\mu \right\}^{1/p}.$$

We define

$$L^p(X, \mathbb{C}, \mu) := \{f : X \to \mathbb{C} : \ f \text{ measurable}, \ \|f\|_p < \infty\}.$$

We will abbreviate $L^p(X, \mathbb{C}, \mu)$ *by* $L^p(X, \mu)$*,* $L^p(X)$ *or simply* L^p*.*

Remark 2.3.1. *We identify functions which agree up to a set of measure zero.*

Lemma 2.3.1. L^p *is a vector space for* $p \geq 1$*.*

Proof. This follows from $|f + g|^p \leq 2^p(|f|^p + |g|^p)$ (trivial). □

Theorem 2.3.2. *(Hölder's Inequality) Suppose* $1 < p < \infty$ *and* $\frac{1}{p} + \frac{1}{q} = 1$*. If* f, g *are measurable functions on* X*, then*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Proof. Similar to the proof for the Riemann Integral. □

Theorem 2.3.3. *(Minkowski Inequality) If* $1 \leq p < \infty$ *and* $f, ginL^p$*, then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

In particular, $\| \cdot \|_p$ *is a norm on* L^p*.*

Proof. Similar to the proof for the Riemann Integral. □

Theorem 2.3.4. *For* $1 \leq p < \infty$*, the space* L^p *is a Banach space with norm* $\| \cdot \|_p$*. If* $(f_n)_{n \in \mathbb{N}}$ *is a sequence with converges in* L^p *to* f*, then some subsequence of* (f_n) *converges pointwise a.e. to* f*.*

Proof. We note that the proof uses a standard idea, which is used to show the if part of Theorem 1.2.10. The fact that $\|\cdot\|_p$ defines a norm follows from results already shown. It remains to show completeness. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence. Then we can choose

$$n_1 < n_2 < \cdots$$

such that

$$\|f_n - f_m\|_p \leq 2^{-k}$$

for all $m, n \geq n_k$. Define $g_1 = f_{n_1}$ and $g_j := f_{n_j} - f_{n_{j-1}}$ for $j > 1$. Then $\sum_{j=1}^k g_j = f_{n_k}$. Define

$$G_k := \sum_{j=1}^k |g_{n_j}|$$

Then $G_k \leq G_{k+1}$ and by the triangle inequality

$$\left[\int G_k^p d\mu\right]^{1/p} \leq \sum_{j=1}^k \|g_{n_j}\|_p \leq \|g_1\| + \sum_{n=1}^\infty 2^{-n}.$$

Hence by monotone convergence, the pointwise limit $G := \lim_{n\to\infty} G_n$ is in L^p. This implies $G(x) < \infty$ a.e. x. Hence (f_{n_k}) converges a.e. x. Let f be the pointwise limit of $(f_{n_k})_{k\in\mathbb{N}}$ a.e.. Since $|f_{n_k}| \leq G$ a.e., we have $|f - f_{n_k}|^p \leq 2^p G^p$ a.e.. Since $|f - f_{n_k}|^p \to 0$ poinwise a.e., we obtain from dominated convergence that

$$\|f - f_{n_k}\|_p \to 0.$$

Now the claim follows, since every Cauchy sequence with convergent subsequence, converges to the limit of the subsequence. □

From the proof we obtain the following result.

Corollary 2.3.5. *If $1 \leq p <$ (below for $p = \infty$) and if (f_n) is a Cauchy sequence in L^p with limit f, then (f_n) has a subsequence which converges pointwise almost everywhere to $f(x)$.*

Theorem 2.3.6. *For $1 \le p < \infty$, the set of simple functions $f = \sum_{j=1}^{n} a_j 1_{E_j}$, where $\mu(E_j) < \infty$ for all j is dense in L^p.*

Proof. Clearly, such functions are in L^p. Let $f \in L^p$. First assume that $f \ge 0$. Choose a sequence (f_n) of simple functions such that $f_n \to f$ a.e. and $f_n \le f$, according to Theorem. Then $f_n \in L^p$ and $|f_n - f|^p \le 2^p |f|^p \in L^1$, so by the dominated convergence theorem, $\|f_n - f\|_p \to 0$. For general $f \in L^p$, apply the above to the positive and negative of real and imaginary part and use that finite sums of simple functions are simple. Moreover, if $f_n = \sum a_j 1_{E_j}$, where the E_j are disjoint and the a_j are nonzero, we must have $\mu(E_j) < \infty$ since $\sum |a_j|^p \mu(E_j) = \int |f_n|^p < \infty$. □

Corollary 2.3.7. *$L^p(\mathbb{R}^d, \lambda)$ has a countable dense subset. $C_c(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$.*

Proof. (Sketch) Approximate the E_j by union of disjoint intervals in $\mathbb{R}^d$, with rational endpoints. Approximate the charachteristic function of an interval by a continuous function (for example piecewise linear one). □

Definition 2.3.2. *Let $(X, \mathcal{A}, \mu)$ be a measure space. We define for every measurable function $f : X \to \mathbb{C}$ the expression*

$$\operatorname{ess\,sup}_{x\in X} |f(x)| := \|f\|_\infty := \inf\{a \ge 0 : \mu(\{x : |f(x)| \ge a\}) = 0\}$$

We define

$$L^\infty = L^\infty(X, \mu) := \{f : X \to \mathbb{C} :\ f \text{ measurable and } \|f\|_\infty < \infty\}.$$

Remark 2.3.2. Note the following.

(a) Observe that the infimum in the above definition is actually attained, since

$$\{x : |f(x)| > a\} = \bigcup_{n=1}^{\infty} \{x : |f(x)| > a + n^{-1}\}$$

and if the sets on the right are null, so is the one on the left.

(b) We have $\|f\|_\infty = \inf\{\sup_{x\in X\setminus N} |f(x)| : \mu(N) = 0\}$.

Theorem 2.3.8. *The following holds.*

(a) If f and g are measurable funtions on X, then $\|fg\|_1 \leq \|f\|_1\|g\|_\infty$.

(b) L^∞ is a complex vector space and $\|\cdot\|_\infty$ is a norm.

(c) L^∞ is a Banach space.

(d) $\|f_n - f\|_\infty \to 0$ iff there exists a measurable set $E \subset X$, such that $\mu(E^c) = 0$ and $f_n \to f$

(e) The simple functions are dense in L^∞. uniformly on E.

Proof. (a) Note that

$$|f(x)g(x)| \leq \|f\|_\infty|g(x)|$$

for almost all x. Integrating the above inequality, we obtain the inequality.
(b) This follows from the inequality $|f+g| \leq |f| + |g|$.
(c,d) Suppose (f_n) is a Cauchy sequence in L^∞. Let A_k and $B_{m,n}$ be the sets where $|f_k(x)| > \|f_k\|_\infty$ and where $|f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty$, and let E be the union of these sets, for $k, m,\ n = 1, 2, 3, ...$ Then $\mu(E) = 0$, and on the complement of E the seqeunce (f_n) converges uniformly to a bounded function f. Define $f(x) = 0$ for $x \in E$. Then $f \in Ł^\infty$, and $\|f_n - f\|_\infty \to 0$ as $n \to \infty$.
(f) see Folland. $\square$

Chapter 3

Hilbert Spaces

In this chapter $\mathbb{F}$ stands for $\mathbb{R}$ or $\mathbb{C}$. Vector spaces are always assumed to be over $\mathbb{F}$.

3.1 Sesquilinearforms

In this section we introduce the notion of a sesquilinear form.

Definition 3.1.1. *Let V be a vector space over $\mathbb{F}$. A mapping*

$$s : V \times V \to \mathbb{F},$$

such that for all $x, y, z \in V$ and $\alpha, \beta \in \mathbb{F}$ we have

$$s(\alpha x + \beta y, z) = \overline{\alpha} s(x, z) + \overline{\beta} s(y, z),$$
$$s(x, \alpha y + \beta z) = \alpha s(x, z) + \beta s(x, z),$$

is called a **sesquilinear form** *on V. A sesquilinear form is called*

- **symmetric**, *if*

$$s(x, y) = \overline{s(y, x)} \quad \forall x, y \in V,$$

- **non-negative**, *if*

$$s(x, x) \geq 0, \quad \forall x \in V,$$

- **positive definite**, *if*

$$s(x,x) > 0 \quad \forall x \in V \setminus \{0\}.$$

A sesquilinear form over a real vector space is also called a **bilinear form**.

Lemma 3.1.1. *For a sesquilinearform s on a vector space V we have*

$$s(0,x) = s(x,0) = 0, \quad \forall x \in V.$$

Proof. $s(0,x) = s(0 \cdot 0, x) = 0 \cdot s(x,0) = 0$ □

Proposition 3.1.2. *(Parallelogramm Identity) Let s be a sesquilinearform on a vector space V, and let $q(x) = s(x,x)$ for all $x \in V$. Then*

$$q(x+y) + q(x-y) = 2q(x) + 2q(y) \quad \forall x, y \in V. \tag{3.1.1}$$

Proof. A calculation shows

$$\begin{aligned} q(x+y) + q(x-y) &= s(x+y, x+y) + s(x-y, x-y) \\ &= (s(x,x) + s(x,y) + s(y,x) + s(y,y)) \\ &\quad + (s(x,x) - s(x,y) - s(y,x) + s(y,y)) \\ &= 2s(x,x) + 2s(y,y). \end{aligned}$$

□

Proposition 3.1.3. *(Polarization) Let V be a vector space over $\mathbb{F}$ and s a sesquilinear form on V. Suppose $q(x) = s(x,x)$ für alle $x \in V$.*

(a) If $\mathbb{F} = \mathbb{C}$, then

$$s(x,y) = \frac{1}{4}\sum_{n=0}^{3} i^n q(x + (-i)^n y), \quad \forall x, y \in V. \tag{3.1.2}$$

(b) If $\mathbb{F} = \mathbb{R}$, then

$$s(x,y) + s(y,x) = \frac{1}{2}(q(x+y) - q(x-y)), \quad \forall x, y \in V. \tag{3.1.3}$$

If s is symmetric we have $s(x,y) = \frac{1}{4}(q(x+y) - q(x-y))$.

Proof. (a)

$$i^n s(x+(-i)^n y, x+(-i)^n y) = i^n s(x,x) + s(x,y) + (-1)^n s(y,x) + i^n s(y,y).$$

Summing this expression over $n = 0, 1, 2, 3$ yields the polarization identity.
(b)

$$\begin{aligned} q(x+y) - q(x-y) &= s(x,x) + s(x,y) + s(y,x) + s(y,y) \\ &\quad - (s(x,x) - s(x,y) - s(y,x) + s(y,y)) = 2s(x,y) + 2s(y,x). \end{aligned}$$

□

Corollary 3.1.4. *If $\mathbb{F} = \mathbb{C}$ every positive sesquilinear form ist symmetric.*

Proof. Polarization. We have

$$\overline{s(x,y)} = \frac{1}{4}\sum_{n=0}^{3}(-i)^n q(x+(-i)^n y) = \frac{1}{4}\sum_{n=0}^{3}(-i)^n q((i)^n x + y) = s(y,x). \tag{3.1.4}$$

□

3.2 Vector Spaces with Semi-Inner Products

Definition 3.2.1. *Let V be a vector space over $\mathbb{F}$. A mapping $s : V \times V \to \mathbb{K}$ is called* **semi-inner product** *on V, if forall $x, y, z \in V$ and $\alpha, \beta \in \mathbb{F}$*

(i) $s(x, x) \geq 0$

(ii) $s(x, \alpha y + \beta z) = \alpha s(x, y) + \beta s(x, z)$

(iii) $s(x, y) = \overline{s(y, x)}$.

A semi-inner product s is called **inner product**, *if in addition*

(i+) $s(x, x) = 0 \Rightarrow x = 0$

For s an inner product one often writes $\langle \cdot, \cdot \rangle$ instead of $s(\cdot, \cdot)$ and one calls $(V, \langle \cdot, \cdot \rangle)$ an inner product space.

Lemma 3.2.1. *If s is a semi-inner product, then one has antilinearity in the first argument, i.e., $s(\alpha y + \beta z, x) = \bar{\alpha} s(y, x) + \bar{\beta} s(z, x)$.*

Proof.

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

□

Remark 3.2.1. We have
semi-inner product = a symmetric nonnegative sesquilinearform
inner product = symmetric strictly positive sesquilinearform.

Example 13. We have

(a) $\mathbb{F}^N$ with $\langle x, y \rangle = \sum_{j=1}^{N} \bar{x}_j y_j$ is an inner product space.

(b) $C[0, 1]$ with $\langle f, g \rangle = \int_0^1 \overline{f(x)} g(x) dx$ is an inner product space.

(c) On $\mathcal{R}[a,b]$ the expression $s(f,g) = \int_a^b \overline{f(t)}g(t)dt$ is a semi-inner product.

Proposition 3.2.2. *(Cauchy-Schwarz-Bunyakowski) Let $s(\cdot,\cdot)$ be a semiscalar product on V. Then for all $x, y \in V$*

$$|s(x,y)| \leq s(x,x)^{1/2} s(y,y)^{1/2}.$$

If s is an inner product, equality holds iff x and y are linearly dependent.

Proof. Case 1: $s(x,x) > 0$.
Then

$$\begin{aligned} 0 &\leq s(y - \frac{s(x,y)x}{s(x,x)}, y - \frac{s(x,y)x}{s(x,x)}) \\ &= s(y,y) - \frac{s(x,y)s(y,x)}{s(x,x)} - \frac{\overline{s(x,y)}s(x,y)}{s(x,x)} + \frac{|s(x,y)|^2}{s(x,x)} \\ &= s(y,y) - \frac{|s(x,y)|^2}{s(x,x)}. \end{aligned}$$

This yields $|s(x,y)|^2 \leq s(y,y)s(x,x)$.
Case 2: Suppose $s(y,y) > 0$. This is shown analogously as Case 1.
Case 3: $s(x,x) = 0$ and $s(y,y) = 0$. We can assume w.l.o.g. that $s(x,y) \geq 0$ (otherwise replace y by $y\overline{s(x,y)}$). Then

$$0 \leq s(x-y, -y) = s(x,x) - s(x,y) - s(y,x) + s(y,y) = -2s(x,y).$$

Thus $s(x,y) = 0$, and the inequality holds.
Now suppose s is an inner product.
If $x = 0$ or $y = 0$, then we have equality and x and y are linearly independent.
Suppose $x \neq 0$ and $y \neq 0$. Then from the proof Case 1, we see that equality holds iff $y - \frac{s(x,y)x}{s(x,x)} = 0$. But this is equivalent to x and y being linearly dependent. □

Remark 3.2.2. If one has an inner product, then one refers to above inequality as the Cauchy-Schwarz inequality.

Definition 3.2.2. *Let V be a vectorspace over $\mathbb{F}$. A mapping $\rho : V \to \mathbb{R}$ is called* **semi-norm** *on V, if for all $x, y \in V$ and $\alpha \in \mathbb{F}$*

(i) $\rho(\alpha x) = |\alpha|\rho(x)$ *(absolute homogeneity)*

(ii) $\rho(x+y) \leq \rho(x) + \rho(y)$ *(triangle inequality or subadditivity)*

Lemma 3.2.3. *The following holds.*

(a) For a seminorm ρ we have $\rho(0) = 0$ and $\rho(x) \geq 0$.

(b) Every norm is a seminorm.

(c) A seminorm ρ is a norm if and only if $\rho(x) = 0$ implies $x = 0$.

Proof. (a) Aus (i) folgt, dass $\rho(0) = \rho(0 \cdot 0) = |0|\rho(0) = 0$.
Aus (i) und (ii) folgt, dass ρ positiv ist, denn $0 = \rho(0) = \rho(v-v) \leq \rho(-v)+\rho(v) = 2\rho(v)$.
(b) This is trivial.
(c) This follows from (a) and the definition of a norm. □

Remark 3.2.3. For semi-norms one usually uses the symbol ρ and for norms one usually uses $\|\cdot\|$. However, we shall occasionally also use $\|\cdot\|$ for semi-norms.

Proposition 3.2.4. *Suppose s is a semi-scalar product on V. Then*

$$\rho(x) := s(x,x)^{1/2}$$

is a semi-norm on V. If s is a scalar product then ρ is a norm.

Proof. (i): $\rho(\lambda x)^2 = s(\lambda x, \lambda x) = \overline{\lambda}\lambda s(x,x) = |\lambda|^2\rho(x)^2$.
(ii): Using first Cauchy-Schwarz and then the squaring inequality, we find

$$\begin{aligned}
\rho(x+y)^2 &= s(x+y, x+y) \\
&= s(x,x) + s(x,y) + s(y,x) + s(y,y) \\
&\leq s(x,x) + 2s(x,x)^{1/2}s(y,y)^{1/2} + s(y,y) \\
&\leq (\rho(x)^2 + \rho(y)^2)^2 \ .
\end{aligned}$$

Now suppose s is a scalar product. Then (iii): $\rho(x) = 0$ iff $s(x,x) = 0$ iff $x = 0$. □

Definition 3.2.3. *Let V be a vector space with semi-inner product s. Then $x, y \in V$ are called* **orthogonal**, *written $x \perp y$, if*

$$s(x, y) = 0.$$

For $A \subset V$ we define the so called **orthogonal complement** $A^\perp := \{u \in V : s(v, u) = 0$ *für alle* $v \in A\}$.

Remark 3.2.4. Let V be a vector space and $M \subset V$. Then one defines the **linear span**

$$\mathrm{lin}M := \{\sum_{i=1}^{N} c_i v_i : N \in \mathbb{N}, c_i \in \mathbb{F}, v_i \in M\}$$

Lemma 3.2.5. *(Algebraic Properties) Let V be a vector space with semi-inner product s. Then the following holds*

(a) Ist $A \subset V$, so ist $A^\perp$ ein Vektorraum.

(b) Ist $A \subset B \subset V$, so gilt $A^\perp \supset B^\perp$.

(c) Ist $A \subset V$, so gilt $A \subset A^{\perp\perp}$.

(d) Ist $A \subset V$, so gilt $(\mathrm{lin}A)^\perp = A^\perp$

(e) Es gilt $\{0\}^\perp = V$.

Ist s zusätzlich ein Skalarprodukt, so gilt

(f) Ist $A \subset V$, so gilt $A \cap A^\perp = \{0\} \cap A$.

(g) Es gilt $V^\perp = \{0\}$. If $s(x, y) = s(x, z)$ for all $x \in V$, then $y = z$.

Proof. (a): Seien $x, y \in A^\perp$ und $\lambda, \mu \in \mathbb{F}$. Dann folgt für alle $z \in A$, dass $s(z, \lambda x + \mu y) = \lambda s(z, x) + \mu s(z, y) = 0$. Also gilt $\lambda x + \mu y \in A^\perp$. (b): $x \in B^\perp \Rightarrow \forall y \in B, s(x, y) = 0 \Rightarrow \forall y \in A, s(x, y) = 0 \Rightarrow x \in A^\perp$.
(c): $x \in A \Rightarrow \forall y \in A^\perp, s(x, y) = 0 \Rightarrow x \in A^{\perp\perp}$

(d): $(\text{lin}A)^{\perp} \subset A^{\perp}$ ist klar wegen (b). Die Umgekehrte Inklusion zeigt man wie folgt. Sei $x \in A^{\perp}$. Dann gilt für $y_i \in A$ und c_i mit $i = 1, ..., N$, dass $s(x, \sum_i c_i y_i) = \sum_i c_i s(x, y_i) = 0$. Damit folgt $x \in (\text{lin}A)^{\perp}$.
(e): Gleichung folgt aus $s(0, x) = 0$ für alle $x \in V$.
(f): Sei $x \in A \cap A^{\perp}$, so folgt $\langle x, x\rangle = 0$ und damit gilt $x = 0$.
(g): Dies folgt direkt aus (f). □

Definition 3.2.4. *Let V be a vector space with semi-scalar product s. Let I be an index set and e_j, $j \in I$, an element of V. Then $(e_j)_{j\in I}$ and also $\{e_j : j \in I\}$ are called* **orthonormal system** *(ONS), if*

$$s(e_i, e_j) = \delta_{i,j}$$

for all $i, j \in I$.

Proposition 3.2.6. *(Pythagoras) Let V be a vector space with semi-inner product s and let $\|\cdot\| = s(\cdot,\cdot)^{1/2}$. Then for all $x, y \in V$ with $x \perp y$ we have*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

Proof. We have

$$\|x + y\|^2 = s(x + y, x + y) = s(x, x) + s(x, y) + s(y, x) + s(y, y) = \|x\|^2 + \|y\|^2.$$

□

Definition 3.2.5. *Suppose $(a_j)_{j\in I}$ is a family of nonnegative numbers. Then one defines their* **sum** *by*

$$\sum_{j\in I} a_j := \sup\{\sum_{j\in A} a_j : A \subset I \textit{ finite}\}.$$

Lemma 3.2.7. *Suppose $(a_j)_{j\in I}$ is a family of nonnegative numbers. If $\sum_{j\in I} a_j < \infty$, then $\{j \in J : a_j \neq 0\}$ is at most countable.*

Proof. First observe that

$$\{j \in I : a_j \neq 0\} = \bigcup_{n \in \mathbb{N}} I_n, \quad I_n := \{j \in I : a_j \geq 1/n\}.$$

Since $\sum_{j \in I} a_j < \infty$, the set I_n must be finite. $\square$

Theorem 3.2.8. *(Bessel) Let V be a vector space with semi-inner product s and $\| \cdot \| = s(\cdot, \cdot)^{1/2}$.*

(a) (Bessel Inequality) If $\{e_j : j \in I\}$ is an ONS, then for all $x \in V$ we have

$$\|x\|^2 \geq \sum_{i \in I} |s(e_i, x)|^2. \tag{3.2.1}$$

(b) If e_j, $j = 1, ..., N$ is a finite ONS, then for all $x \in V$ we have

$$\|x\|^2 = \|x - \sum_{i=1}^{M} s(e_i, x) e_i\|^2 + \sum_{i=1}^{N} |s(e_i, x)|^2. \tag{3.2.2}$$

(c) If e_j, $j = 1, ..., N$ is a finite ONS, then for all $x \in V$ we have

$$x - \sum_{i=1}^{N} s(e_i, x) e_i \perp e_j, \quad j = 1, ..., N.$$

Proof. (Bessel) (c) folgt aus

$$\begin{aligned} s(x - \sum_{j=1}^{N} e_j s(e_j, x), e_i) &= s(x, e_i) - \sum_{j=1}^{N} \overline{s(e_j, x)} s(e_j, e_i) \\ &= s(x, e_i) - \overline{s(e_i, x)} = 0. \end{aligned}$$

(b): Aus (c) folgt $x - \sum_{i=1}^{N} e_i s(e_i, x) \perp \sum_{i=1}^{N} e_i s(e_i, x)$ und durch zweimalige Anwendung von Pythagoras

$$\|x\|^2 = \|x - \sum_{i=1}^{N} e_i s(e_i, x)\|^2 + \| \sum_{i=1}^{N} e_i s(e_i, x)\|^2 = \|x - \sum_{i=1}^{N} e_i s(e_i, x)\|^2 + \sum_{i=1}^{N} |s(e_i, x)|^2.$$

(a): Es folgt aus (b), dass

$$\|x\|^2 \geq \sum_{i=1}^{N} |s(e_i, x)|^2.$$

Damit folgt die Behauptung für beliebige Indexmengen. $\square$

Remark 3.2.5. Note that Bessels inquality for a single vector e_1 implies the Cauchy-Schwarz-Bunyakowski inequality.

Die vorangehenden Theoreme zeigen die Nützlichkeit von Orthonormalsystemen. Aus jedem abzählbaren System linear unabhängiger Vektoren kann man ein Orthonormalsystem gewinnen. Dies ist der Inhalt der nächsten Proposition.

Proposition 3.2.9. *(Gram/Schmidtsches Orthogonalisierungsverfahren) Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Seien die Vektoren $v_1, ..., v_N$ (bzw. v_j, $j \in \mathbb{N}$, d.h. $N = \infty$) linear unabhängig. Dann bilden die induktiv definierten Vektoren*

$$\begin{aligned} e_1 &:= \frac{v_1}{\|v_1\|} \\ e_{k+1} &:= \frac{v_{k+1} - \sum_{j=1}^{k} \langle e_j, v_{k+1} \rangle e_j}{\|v_{k+1} - \sum_{j=1}^{k} \langle e_j, v_{k+1} \rangle e_j\|} \end{aligned}$$

ein Orthonormalsystem mit

$$\mathrm{Lin}\{e_1, ..., e_k\} = \mathrm{Lin}\{v_1, ..., v_k\} \tag{3.2.3}$$

für alle $k = 1, ..., N$.

Proof. Der Beweis wird durch Induktion nach k geführt über die Aussage $e_1, ..., e_k$ sind Orthonormalsystem mit (3.2.3).
$k = 1$ ist klar.
$k \Rightarrow k + 1$: Es ist $w_{k+1} := v_{k+1} - \sum_{j=1}^{k} \langle e_j, v_{k+1} \rangle e_j$ orthogonal auf e_l, $l = 1, ..., k$ (nach obigem Theorem) und verschwindet nicht (aufgrund der linearen Unabhängigkeit der v_j

und der Bedingung an die linearen Hüllen.) Durch normieren erhalten wir e_{k+1} und $e_1, ..., e_{k+1}$ bilden ein Orthnormalsystem. Weiterhin gilt

$$\begin{aligned}
\mathrm{Lin}\{e_1, ..., e_k, e_{k+1}\} &= \mathrm{Lin}\{e_1, ..., e_k, w_{k+1}\} \\
&= \mathrm{Lin}\{e_1, ..., e_k, v_{k+1}\} \\
&= \mathrm{Lin}\{v_1, ..., v_k, v_{k+1}\},
\end{aligned}$$

wobei wir im zweiten Schritt die Definition der w's verwendeten und im letzten Schritt die Induktionsannahme. $\square$

Bemerkung 1. (Verallgemeinertes Gram/Schmidtsche Verfahren) Auch wenn die (v_j) nicht linear unabhängig sind, kann man das Gram/Schmidtsche Verfahren in einer einfachten Modifikation anwenden. Dazu streicht man alle dijenigen N bei denen $w_N = 0$ gilt. Die Details sind wie folgt: Definiere induktiv

$$w_1 := v_1 \ , \quad e_1 := \begin{cases} \frac{w_1}{\|w_1\|} \ , & w_1 \neq 0 \\ 0 \ , & \text{sonst.} \end{cases}$$

$$w_{k+1} := v_{k+1} - \sum_{j=1}^{k} \langle e_j, v_{k+1} \rangle e_j, \quad e_{k+1} := \begin{cases} \frac{w_{k+1}}{\|w_{k+1}\|} \ , & w_k \neq 0 \\ 0 \ , & \text{sonst.} \end{cases}$$

Dann bilden $\{e_k : k \in \mathbb{N}, e_k \neq 0\}$ ein Orthonormalsystem mit

$$\mathrm{lin}\{e_j : j \in \{1, ..., k\}, e_j \neq 0\} = \mathrm{lin}\{v_1, ..., v_k\}$$

3.3 Hilberträume

Lemma 3.3.1. *Ist $\|\cdot\|$ eine Norm auf einem Vektorraum, so ist $d : V \times V \to [0, \infty)$ definiert durch $d(x, y) := \|x-y\|$ eine Metrik auf V. Man nennt d die durch $\|\cdot\|$ induzierte Metrik.*

Proof. Nur hinschreiben. □

Das heisst ein Vektorraum mit Norm ist immer mit der induzierten Metrik ausgestattet. Vektorraum mit Skalarprodukt ist mit der induzierten Norm und der daraus induzierten Metrik ausgestattet.

Proposition 3.3.2. *Sei V ein Vektorraum mit Skalarprodukt mit induzierter Norm $\|\cdot\|$ und induzierter Metrik d.*

(a) $\|\cdot\| : V \to \mathbb{R}$ ist stetig.

(b) Die Abbildung $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{C}$ ist stetig, d.h., gilt $x_n \to x$ and $y_n \to y$ in V, so folgt $\langle x_n, y_n\rangle \to \langle x, y\rangle$.

Proof. (a) Die Stetigkeit folgt unmittelabar aus der umgekehrten Dreiecksungleichung

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Diese zeigt man wie folgt: $\|x\| \leq \|x - y\| + \|y\|$ impliziert $\|x\| - \|y\| \leq \|x - y\|$. Analog zeigt man $\|y\| - \|x\| \leq \|x - y\|$. Damit folgt die umgekehrte Dreiecksungleichung.
(b) Dies zeigt man wie in $\mathbb{R}$. Stetigkeit in $(a, b) \in V \times V$ folgt aus der Abschätzung

$$\begin{aligned} |\langle a, b\rangle - \langle x, y\rangle| &\leq |\langle a, b - y\rangle + \langle a - x, y\rangle| \leq \|a\|\|b - y\| + \|a - x\|\|y\| \\ &\leq \|a\|\|b - y\| + \|a - x\|(\|y - b\| + \|b\|) \end{aligned}$$

□

Bemerkung 2. Zu (b). Sind M_1 und M_2 metrische Räume mit Metriken d_1 und d_2, so ist $M_1 \times M_2$ eine metrsicher Raum mit Metrik $d_{M_1\times M_2}((x_1, x_2), (y_1, y_2)) = d_1(x_1, y_1) + d_2(x_2, y_2)$.

Definition 3.3.1. *Ein Vektorraum V mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und induzierter Metrik d heisst Hilbertraum, wenn er bezüglich d vollständig ist (d.h. wenn jede Cauchy Folge bzgl. d einen Grenzwert in V hat).*

Example 14. We have

(a) $\mathbb{F}^N$ with $\langle x, y \rangle = \sum_{j=1}^N \bar{x}_j y_j$ is a Hilbert space.

(b) $C[-1,1]$ with $\langle f, g \rangle = \int_{-1}^1 \overline{f(x)} g(x) dx$ is an inner product space, which is not complete. (Let $f_n(x) = 1$, if $x < 0$, $f_n(x) = 1 - nx$, if $0 \leq x \leq 1/n$, and $f_n(x) = 0$ otherwise. Then (f_n) is a Cauchy sequence. It converges to $f(x) = 1$ if $x \leq 0$ and $f(x) = 0$ otherwise. This function is not continuous.)

Lemma 3.3.3. *(Continuity properties) Let V be an inner product space and $A \subset V$ nonempty. Then*

$$\overline{\mathrm{lin}(A)}^\perp = A^\perp = \overline{A^\perp}.$$

Proof. Folgt aus der Stetigkeit der Norm. Erste Gleichheit: $\overline{\mathrm{lin} A}^\perp \subset A^\perp$ haben wir oben schon gezeigt. Es bleibt die umgekehrte Inklusion $\overline{\mathrm{lin} A}^\perp \supset A^\perp$ zu zeigen. Sei also $x \in A^\perp$. Dann gilt $\langle x, y \rangle = 0$ für alle $y \in A$. Damit folgt wegen der Stetigkeit des Skalarproduktes auch $\langle x, y \rangle = 0$ für alle $y \in \overline{\mathrm{lin} A}$. Also gilt $x \in \overline{\mathrm{lin} A}^\perp$.
Zweite Gleichheit: Die Inklusion $A^\perp \subset \overline{A^\perp}$ ist klar. Es bleibt die umgekehrte Inklusion $A^\perp \supset \overline{A^\perp}$ zu zeigen. Sei also $x \in \overline{A^\perp}$. Dann gibt es eine Folge (x_n) in $A^\perp$ mit $x_n \to x$. Für alle $y \in A$ gilt $\langle x, y \rangle = \lim_n \langle x_n, y \rangle = 0$. Also folgt $x \in A^\perp$. □

Lemma 3.3.4. *Let V be an inner product space. Then*

$$\|x\| = \sup_{y \in V, \, \|y\| \leq 1} |\langle y, x \rangle|$$

Let V be a normed space and W an inner product spaces and $A \in \mathcal{L}(V, W)$. Then

$$\|A\| = \sup_{\substack{z \in V, \, \|z\| \leq 1, \\ y \in W, \, \|y\| \leq 1}} |\langle y, Az \rangle|$$

Proof. First equality: $x = 0$ is trivial. By Cauchy-Schwarz we have for $x \neq 0$

$$\|x\| = \langle x/\|x\|, x\rangle \leq \sup_{y\in V,\, \|y\|\leq 1} |\langle y, x\rangle| \leq \sup_{y\in V,\, \|y\|\leq 1} \|y\|\|x\| = \|x\|.$$

Second equality:

$$\|A\| = \sup_{z\in V,\, \|z\|\leq 1} \|Az\| = \sup_{z\in V,\, \|z\|\leq 1} \sup_{y\in W,\, \|y\|\leq 1} |\langle y, Az\rangle| = \sup_{\substack{z\in V,\, \|z\|\leq 1,\\ y\in W,\, \|y\|\leq 1}} |\langle y, Az\rangle|$$

□

Proposition 3.3.5. *Let* $l^2(\mathbb{N}) := \{x : \mathbb{N} \to \mathbb{C} : \sum_{l=1}^{\infty} |x(n)|^2 < \infty\}$. *This is a vector space and* $\langle x, y\rangle = \sum_{i=1}^{\infty} \overline{x(i)}y(i)$ *is an inner product.* $l^2(\mathbb{N})$ *with this inner product is a Hilbert space.*

Proof. The vector space property follows from $|x(n) + y(n)|^2 \leq 2|x(n)|^2 + 2|y(n)|^2$. The expression $\langle\cdot,\cdot\rangle$ is well defined since $|\overline{x(n)}y(n)| \leq \frac{1}{2}|x(n)|^2 + \frac{1}{2}|y(n)|^2$. The properties of an inner product are straight forward to verify using the properties of limits.
It remains to show completeness. Sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $l^2(\mathbb{N})$. Für $l \in \mathbb{N}$ gilt

$$|x_n(l) - x_m(l)| \leq (\sum_{i=1}^{\infty} |x_n(i) - x_m(i)|^2)^{1/2} \leq \|x_n - x_m\| \to 0 \quad (n, m \to \infty).$$

Also konvergiert, da $\mathbb{C}$ vollständig ist, $(x_n(l))_{l\in\mathbb{N}}$ für jedes l. Definiere für $l \in \mathbb{N}$

$$x(l) := \lim_{n\to\infty} x_n(l)$$

Wir müssen noch zeigen, dass $x \in l^2(\mathbb{N})$ und dass $x_n \to x$ bezüglich der Norm in $l^2(\mathbb{N})$. Sei $\epsilon > 0$. Dann existierte ein $N_\epsilon \in \mathbb{N}$, so dass

$$\|x_n - x_m\|^2 < \epsilon, \quad \forall n, m \geq N_\epsilon.$$

Ist $n \geq N_\epsilon$, so gilt für jedes $N \in \mathbb{N}$

$$\sum_{j=1}^{N} |x(j) - x_n(j)|^2 = \lim_{k\to\infty} \sum_{j=1}^{N} |x_k(j) - x_n(j)|^2 \leq \lim_{k\to\infty} \|x_k - x_n\|^2 \leq \epsilon.$$

Da $N \in \mathbb{N}$ beliebig war, folgt dann für $n \geq N_\epsilon$

$$\sum_{j=1}^{\infty} |x(j) - x_n(j)|^2 \leq \epsilon$$

Damit folgt $x - x_n \in l^2(\mathbb{N})$ und somit $x = (x - x_n) + x_n \in l^2(\mathbb{N})$, und es folgt $x_n \to x$. $\square$

Satz 1. *(Approximationssatz) Sei $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ ein Hilbertraum. Ist C eine abgeschlossene konvexe Teilmenge von $\mathcal{H}$, so gibt es zu jedem $x \in \mathcal{H}$ genau ein $y \in C$ mit*

$$\|x - y\| = d(x, C) := \inf\{\|z - y\| : z \in C\}.$$

Damit existiert also die beste Approximation an x in C. (Zeichnung)

Bei dem Satz handelt es sich um eine fundamentale Eigenschaft eines Hilbertraumes. Entsprechend spielen die fundamentalen Eigenschaften des Hilbertraum nämlich Vollständigkeit und Parallelogrammidentität eine Rolle. Zur Erinnerung, die Parallelogrammidentität besagt, dass für eine Norm $\| \cdot \|$, welche von einem Skalarprodukten herrührr, gilt

$$\|u - v\|^2 + \|u + v\|^2 = 2(\|u\|^2 + \|v\|^2) \ .$$

Beweis von Satz 1. Sei $x \in \mathcal{H}$ fest. Setze

$$\delta := \inf_{y \in C} \|x - y\|.$$

Es folgt aus der Parallelogrammidentität und der Konvexheit von C, dass für alle $u, v \in C$

$$\begin{aligned} \|u - v\|^2 \quad &= \|(u - x) - (v - x)\|^2 \\ &= 2\|u - x\|^2 + 2\|v - x\|^2 - \|(u - x) + (v - x)\|^2 \\ &= 2\|u - x\|^2 + 2\|v - x\|^2 - 4\|x - \frac{1}{2}(u + v)\|^2 \\ &\leq 2\|u - x\|^2 + 2\|v - x\|^2 - 4\delta^2 \end{aligned} \tag{3.3.1}$$

Existenz: Nach Definition des Infimums existiert eine Folge $(y_n)_{n=1}^{\infty}$ in C so dass $\|x - y_n\| \to \delta$ für $n \to \infty$. Damit folgt aus (3.3.1), dass

$$\|y_n - y_m\|^2 \leq 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\delta^2 \to 2\delta^2 + 2\delta^2 - 4\delta^2 = 0 \ ,$$

für $n, m \to \infty$. Also ist $(y_n)_{n=1}^{\infty}$ eine Cauchy Folge in $\mathcal{H}$, welche gegen ein $y \in \mathcal{H}$ konvergiert. Da C abgeschlossen ist, gilt $y \in C$. Es folgt aus der Stetigkeit der Norm $\delta = \lim_n \|y_n - x\| = \|y - x\|$.

Eindeutigkeit: Gibt es y_1 und y_2 mit $\|y_1 - x\| = \|y_1 - x\| = \delta$, so folgt mit (3.3.1)

$$\|y_1 - y_2\|^2 \leq 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2 - 4\delta^2 = 0 .$$

Damit folgt $y_1 = y_2$. □

Satz 2. *(Projektionssatz) Sei $\mathcal{H}$ ein Hilbert-Raum und M ein abgeschlossener Unterraumvon $\mathcal{H}$. Dann kann jedes $x \in \mathcal{H}$ eindeutig geschrieben werden als*

$$x = z + w \qquad \textit{mit } z \in M \textit{ and } w \in M^{\perp}. \tag{3.3.2}$$

Das Eindeutige z in (3.3.2) *erfüllt* $\|x - z\| = \inf\{\|x - y\| : y \in M\}$.

Bemerkung 3. *Man beschreibt den Sachverhalt* (3.3.2) *oft mit der Notation*

$$\mathcal{H} = M \oplus M^{\perp} .$$

Ebenso schreibt man für $x = x^{\|} + x^{\perp}$ mit $x^{\|} \in M$ and $x^{\perp} in M^{\perp}$.

Proof. Eindeutigkeit: Sei $x = z + w = z' + w'$ mit $z, z' \in M$ und $w, w \in M^{\perp}$. Dann gilt $z - z' = w' - w$ und $\|z - z'\| = \langle z - z', z - z'\rangle = \langle z - z', w - w'\rangle = 0$. Damit folgt $z = z'$, und daraus $w = w'$.
Existenz: Nach dem Approximationssatz (Satz1) gibt es ein eindeutiges $z \in M$ mit

$$\|x - z\| = \inf\{\|x - y\| : y \in M\} =: \delta. \tag{3.3.3}$$

Setze $w = x - z$. Dann gilt $x = z + w$ mit $z \in M$. Es bleibt also zu zeigen, dass $w = x - z \in M^{\perp}$. Es gilt für alle $v \in M$ und $t \in \mathbb{R}$, dass

$$\delta^2 \leq \|(x - (z + tv)\|^2 = \langle w, w\rangle + \langle w, tv\rangle + \langle tv, w\rangle + t^2\langle v, v\rangle ,$$

und somit $0 \leq t(\langle w, v\rangle + \langle v, w\rangle) + t^2\|v\|^2$. Es folgt

$$\langle w, v\rangle + \langle v, w\rangle = 0 \ , \forall v \in M. \tag{3.3.4}$$

(To see insert $t = -(\langle w, v\rangle + \langle v, w\rangle)/(2\|v\|^2)$) Im Fall $\mathbb{K} = \mathbb{R}$, folgt damit $w \in M^\perp$. Im Fall $\mathbb{K} = \mathbb{C}$ argumentieren wir wie folgt. For any $u \in M$, insert into (3.3) first $v = u$ and then $v = iu \in M$. This gives

$$\langle u, z\rangle + \langle u, z\rangle = 0 \quad , \quad -i\langle u, z\rangle + i\langle u, z\rangle = 0 \ .$$

Multiplying the first equation by i and then adding both equations together, shows that $\langle u, z\rangle = 0$ for all $u \in M$. □

Lemma 3.3.6. *(Completness properties) Let $\mathcal{H}$ be a Hilbert space. then*

(a) Ist V eine abgeschlossener Unterraum von $\mathcal{H}$, so gilt $V^{\perp\perp} = V$.

(b) Ist $A \subset \mathcal{H}$, so gilt $\overline{\mathrm{lin} A} = A^{\perp\perp}$.

(c) Ist $A \subset \mathcal{H}$, so gilt

$$\overline{\mathrm{lin} A} = \mathcal{H} \quad \Leftrightarrow \quad A^\perp = \{0\}.$$

Proof. (a) Wir haben $V^{\perp\perp} \supset V$ schon gezeigt. Es bleibt die Umgekehrte Inklusion $V^{\perp\perp} \subset V$ zu zeigen. Sei $x \in V^{\perp\perp}$. Dann gilt nach dem Projektionssatz $x = z + w$ mit $z \in V$ und $w \in V^\perp$. Damit folgt $w = x - z \in V^{\perp\perp} \cap V^\perp = \{0\}$. Damit folgt $w = 0$ und $x = z \in V$.
(b) Mit (a) folgt $\overline{\mathrm{lin} A} = \overline{\mathrm{lin} A}^{\perp\perp} = A^{\perp\perp}$.
(c) $\Rightarrow$ folgt aus schon gezeigtem: $A^\perp = \overline{\mathrm{lin} A}^\perp = \mathcal{H}$.
$\Leftarrow$ Mit (b) folgt $\overline{\mathrm{lin} A} = A^{\perp\perp} = \{0\}^\perp = \mathcal{H}$. □

Definition 3.3.2. *Let V be a vector space. A projection in V is linear map $P : V \to V$ such that $P^2 = P$.*

Lemma 3.3.7. *Let P be a projection. Then also $(1 - P)$ is a projection.* $\mathrm{Ran}(1 - P) = \ker P$ *and* $\ker(1 - P) = \mathrm{Ran} P$. $\mathrm{Ran} P + \ker P = V$. .

Proof. $(1-P)^2 = 1-2P+P^2 = 1-P$. $P(1-P) = P-P^2 = P-P = 0$. Hence $\mathrm{Ran}(1-P) \subset \ker P$. If $Px = 0$, then $x = (1-P)x$ and $x \in \mathrm{Ran}(1-P)$.
Now $\mathrm{Ran}P + \ker P = \mathrm{Ran}P + \mathrm{Ran}(1-P) = V$. □

Definition 3.3.3. *If V is an inner product space, a projection P in V is called orthogonal projection if* $\mathrm{Ran}P \perp \ker P$.

Lemma 3.3.8. *If P is an orthogonal projection. Then P is bounded with $\|P\| \leq 1$ and* $\mathrm{Ran}P$ *is closed.*

Proof. P is bounded: We have $x = Px+(1-P)x$. Then $(1-P)x \in \ker P$ and $Px \in \mathrm{Ran}P$. Thus we find by Pythogoras $\|x\|^2 = \|(1-P)x\|^2 + \|Px\|^2$. It follows that $\|Px\| \leq \|x\|$. Hence P is continuous.
$\mathrm{Ran}P$ is closed: Let $y \in \mathcal{H}$. Suppose there exists a sequence $(y_n) \in \mathrm{Ran}P$ such that $y_n \to y$. It follows that $y_n = Py_n$ and $y \leftarrow y_n = Py_n \to Py$. Thus $y \in \mathrm{Ran}P$ and $\mathrm{Ran}P$ is closed. □

Proposition 3.3.9. *Let M be a closed subspace of a Hilbert space $\mathcal{H}$. Define Then*

$$P_M : \mathcal{H} \to \mathcal{H}, \quad x \mapsto P_M x := x^{\|}$$

where $x^{\|}$ is the unique element in M such that $x = x^{\|} + z$ for some $z \in M^{\perp}$. Then P_M is a projection with $\mathrm{Ran}P_M = M$ *and* $\ker P_M = M^{\perp}$, *and hence an orthonormal projection. P_M is the unique orthogonal projection in $\mathcal{H}$ with* $\mathrm{Ran}P_M = M$. *We call P_M the* **orthogonal projection onto** *M.*

Proof. By definition $P_M^2 x = P_M(x^{\|}) = x^{\|}$. Thus P_M is a projection. We have $\mathrm{Ran}P_M = M$, since $x \in M$ iff $P_M x = x$. We have $\ker P_M = M^{\perp}$ since $P_M x = 0$ iff $x \in M^{\perp}$. Let P'_M be a second orthogonal projection with $\mathrm{Ran}P'_M = M$. Let $x \in \mathcal{H}$. Since P'_M is an orthogonal projection we have $(1-P'_M)x \in M^{\perp}$. On the other hand $x = P'_M x + (1-P'_M)x$, so by uniqueness $P'_M x = P_M x$. □

Wir wenden uns nun Entwicklungen nach ONS zu.

Definition 3.3.4. *(Orthonormalbasis) Ein ONS* $\{e_j : j \in I\}$ *in einem Hilbertraum* $\mathcal{H}$ *heisst* **Orthonormalbasis (ONB)**, *wenn gilt* $\overline{\text{lin}\{e_j : j \in I\}} = \mathcal{H}$.

Bemerkung 4. An orthonormal basis and an algebraic basis are in general not the same. For finite dimensional spaces they agree.

Beispiel 1.

(a) The vector space $\mathbb{F}^n$ has $e_j \in \mathbb{F}^n$ with $j = 1, ..., n$ as an orthonormal basis, where $(e_j)_i := \delta_{i,j}$.

Proposition 3.3.10. *In* $l^2(\mathbb{N})$ *the set of vectors* $(e_j)_{j\in\mathbb{N}}$ *where* $e_j(n) = \delta_{j,n}$ *is an orthonormal basis.*

Proof. Offensichtlich gilt $\langle e_j, e_k\rangle = \delta_{j,k}$. Ist $x \in l^2(\mathbb{N})$, so gilt

$$\|x - \sum_{n=1}^{N} x_n e_n\|^2 = \sum_{n=N+1} |x_n|^2 \to 0 \quad (N \to \infty).$$

□

Definition 3.3.5. *Let* X *be a normed space and* J *a set. For a mapping* $x : J \to X$ *and an element* $y \in X$ *we write*

$$\sum_{j\in J} x_j = y$$

and say that the sum converges to the limit y, *if for every* $\epsilon > 0$, *there exists a finite set* $J_0 \subset J$ *such that*

$$\|\sum_{j\in A} x_j - y\| < \epsilon$$

for all finite set A *with* $J_0 \subset A \subset J$.

Theorem 3.3.11. *(Coefficient Representations) Sei* $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ *ein Hilbertraum und sei* $\{e_j : j \in I\}$ *ein ONS. Seien* $c_j \in \mathbb{K}$ *gegeben mit* $\sum_{j\in I} |c_j|^2 < \infty$. *Dann gibt es ein eindeutiges* $x \in H$ *mit*

$$x = \sum_{j\in I} c_j e_j$$

und es gilt

$$\|x\|^2 = \sum_{j\in I} |c_j|^2.$$

Proof. Wegen $\sum_{j\in I} |c_j|^2 < \infty$ können nur abzählbar viele c_j nicht verschwinden. Daher können wir ohne Einschränkung annehmen, dass die Indexmenge abzählbar ist und $I = \mathbb{N}$. Wir zeigen zuerst, dass $S_N := \sum_{j=1}^N c_j e_j$ eine Cauchy-Folge ist. Es gilt

$$\|S_N - S_M\|^2 = \sum_{j=N+1}^{M} |c_j|^2 \to 0, \quad N, M \to \infty.$$

Da V als Hilbertraum vollständig ist, konvergiert $(S_N)_{N\in\mathbb{N}}$ gegen ein $x \in V$.
Aufgrund des Satz des Pythagoras folgt daraus

$$\|x\|^2 = \lim_{N\to\infty} \|S_N\|^2 = \lim_{N\to\infty} \sum_{j=1}^{N} |c_j|^2.$$

□

Wir haben folgende Charaktersierung einer Orthonormalbasis

Lemma 3.3.12. *(Charakterisierung Basis) Sei $(V, \langle\cdot,\cdot\rangle)$ ein Hilbertraum und $(e_j)_{j\in I}$ ein ONS. Dann sind äquivalent:*

(i) $\{e_j : j \in I\}$ ist maximal (d.h. für jedes ONS S mit $\{e_j : j \in I\} \subset S$ gilt $S = \{e_j : j \in I\}$.

(ii) Es gilt $\{e_j : j \in I\}^\perp = \{0\}$.

(iii) Es gilt $\overline{\mathrm{lin}\{e_j : j \in I\}} = V$. (d.h. $\mathrm{lin}\{e_j : j \in I\}$ ist eine Basis.)

(iv) Für jedes $x \in V$ gilt $x = \sum \langle e_j, x\rangle e_j$.

(v) Für jedes $x \in V$ gilt $\|x\|^2 = \sum_j |\langle e_j, x\rangle|^2$.

Proof. (i) $\Longrightarrow$ (ii): Wäre die Aussage $\{e_j : j \in I\}^\perp = \{0\}$ falsch, so gäbe es ein $x \in \{e_j : j \in I\}^\perp$ mit $\|x\| \neq 0$ und das ONS $\{e_j : j \in I\} \cup \{\frac{x}{\|x\|}\}$ würde der Maximalität von $\{e_j : j \in I\}$ widersprechen.
(ii) $\Longrightarrow$ (i): Let $\{e'_\alpha : \alpha \in A\}$ be an ONB containing $\{e_j : j \in I\}$. Suppose there exists an $x \in \{e'_\alpha : \alpha \in A\} \setminus \{e_j : j \in I\}$. Then $\langle x, e_j\rangle$ for all $j \in I$ und hence by assumption $x = 0$. This is a contradiction.
(ii) $\Longleftrightarrow$ (iii): This follows from (c) of Lemma 3.3.6.
(ii) $\Longrightarrow$ (iv): Let $x \in V$. By Bessels Inequality $\sum_j |\langle e_j, x\rangle|^2 < \infty$. Thus there exists by the theorem about coefficient representations $y = \sum_j \langle e_j, x\rangle e_j$. By construction $\langle x - y, e_j\rangle = 0$. From (ii) it now follows that $x - y = 0$ and hence $x = \sum_j \langle e_j, x\rangle e_j$.
(iv) $\Longrightarrow$ (v): Follows from the theorem about coefficient representations (Theorem 3.3.11).
(v) $\Longrightarrow$ (iv): Let $x \in V$. By the theorem about coefficient representations and Bessels inequality the sum $y = \sum_j \langle e_j, x\rangle e_j$ exists and $\|y\|^2 = \sum_j |\langle e_j, x\rangle|^2$. By the usual calculation we find $y \perp x - y$. Thus it follows from (v) and the Theorem of Phythagoras that

$$\|y\|^2 = \sum_j |\langle e_j, x\rangle|^2 \overset{(v)}{=} \|x\|^2 \overset{\text{Pythagoras}}{=} \|x - y\|^2 + \|y\|^2.$$

Thus $\|y - x\| = 0$ und hence $y = x$.
(iv) $\Longrightarrow$ (ii): Let $x \perp e_j$ for all $j \in I$. Then by (iv) we can write

$$x = \sum_j \langle e_j, x\rangle e_j = \sum_j 0 e_j = 0.$$

This implies (ii). □

Corollary 3.3.13. *Let $\mathcal{H}$ be a Hilbert space. Then e_j, $j \in I$, is an ONB if and only if for all $x, y \in \mathcal{H}$ we have $\langle x, y\rangle = \sum_j \langle x, e_j\rangle\langle e_j, y\rangle$.*

Remark 3.3.1. The relation in the above corollary is often written in physics literature as $I = \sum_j |e_j\rangle\langle e_j|$.

Proof. $\Rightarrow$: The convergence of the right hand side follows since

$$|\langle x, e_j\rangle\langle e_j, y\rangle| \leq \frac{1}{2}(|\langle x, e_j\rangle|^2 + |\langle e_j, y\rangle|^2.$$

Now the equality follows from the continuity of the norm and (iv) of the previous lemma.
$\Leftarrow$: Setting $x = y$. We obtain (v) of the above Lemma. $\square$

Theorem 3.3.14. *Every Hilbert space has an orthonormal basis.*

Proof of Theorem 3.3.14. Let $\mathcal{E}$ denote the set of orthonormal subsets of $\mathcal{H}$. This set is ordered by inclusion of sets. Every subset of $\mathcal{E}$ which is totally ordered has an upper bound in $\mathcal{E}$, which is given by the union. Thus by Lemma of Zorn $\mathcal{E}$ has a maximal element, say S. By (i) of the previous Lemma we see that S is an orthonormal basis. $\square$

Proposition 3.3.15. *Let $\mathcal{H}$ be a Hilbert space and V a closed subspace. Let (e_j) be an orthonormal basis of V. Then*

$$P_V x := \sum_{j \in I} \langle e_j, x \rangle e_j$$

is an orthonormal projection with range V. (I.e. it is the unique orthogonal projection with this property.)

Proof. Clearly $\mathrm{Ran} P_V \subset V$. If $x \in V$, then $P_V x = x$ by (iv) of the above Lemma. So $\mathrm{Ran} P_V = V$ and $P_V^2 = P_V$. Now $P_V y = 0$ iff $\langle e_j, y \rangle = 0$ (by the theorem about coeficient represenations) iff $y \perp V$. Thus $\mathrm{Ran} P_V \perp \ker P_V$ and P_V is an orthogonal projection. The stemement in brackets has been shown above. $\square$

Separable Hilbert Spaces

Definition 3.3.6. *Let (M, d) be a metric space. We say that a subset D of M is* **dense** *in M, if $\overline{D} = M$. We call M* **separable** *if it contains a countable dense subset.*

Definition 3.3.7. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. A surjective linear mapping $U : \mathcal{H}_1 \to \mathcal{H}_2$ is called* **unitary** *if $\langle Ux, Uy\rangle_{\mathcal{H}_2} = \langle x, y\rangle_{\mathcal{H}_1}$ for all $x, y \in \mathcal{H}_1$. If such a mapping exists $\mathcal{H}_1$ and $\mathcal{H}_2$ are called* **unitarily equivalent**.

Lemma 3.3.16. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. A bijective linear mapping $U : \mathcal{H}_1 \to \mathcal{H}_2$ is unitary iff and only if $\|Ux\|_{\mathcal{H}_2} = \|x\|_{\mathcal{H}_1}$ for all $x \in \mathcal{H}_1$.*

Proof. Follows from polarization, Proposition 3.1.3. □

Proposition 3.3.17. *Sei $\mathcal{H}$ ein Hilbertraum. Dann sind folgende Aussagen q̈uivalent.*

(a) $\mathcal{H}$ ist separabel.

(b) $\mathcal{H}$ besitzt eine höchstens abzählbare ONB.

(c) Es gibt eine Menge $I \subset \mathbb{N}$ und eine unitäre Abbildung $U : l^2(I) \to \mathcal{H}$.

Proof. (a) $\Rightarrow$ (b): Sei $\mathcal{H}$ separabel und D eine dichte höchstens abzählbare Teilmenge von $\mathcal{H}$. Wende das verallgemeinerte Gram-Schmidt-Verfahren auf D an. Dann erhalten wir eine ONB, welche höchstens abzählbar ist.
(b) $\Rightarrow$ (a): Sei $\{e_j : j \in I\}$ eine ONB und $I = \{1, ..., N\}$ oder $I = \mathbb{N}$. Sei $x \in \mathcal{H}$. Dann gibt es zu jedem $\epsilon > 0$, ein $N \in \mathbb{N}$, Zahlen $c_1, ..., c_N \in \mathbb{C}$, so dass $\|x - \sum_{j=1}^N c_j e_j\| < \epsilon/2$. Da $\mathbb{Q}$ dicht in $\mathbb{R}$ gibt es $d_1, ..., d_N \in \mathbb{Q} + i\mathbb{Q}$ so dass $|c_j - d_j| < \epsilon/(2N)$. Damit folgt $\|x - \sum_{j=1}^N d_j e_j\| < \epsilon$. Wir folgern, dass die Menge

$$D := \left\{ \sum_{j=1}^{N} x_j c_j : N \in I,\ c_j \in \mathbb{Q} + i\mathbb{Q} \right\}$$

dicht in $\mathcal{H}$ liegt. Die Menge D ist höchstens abzählbar, da $\mathbb{Q}$ und I höchstens abzählbar sind.

(b) $\Rightarrow$ (c): Sei $(e_j)_{j\in I}$ one ONB. Dann ist die Abbildung $U : l^2(I) \to \mathcal{H}$ mit $(x_j)_{j\in I} \mapsto \sum_{j\in I} x_j e_j$ unitär.
(c) $\Rightarrow$ (b): $e_{j\in I}$ mit $e_j(n) = \delta_{j,n}$ ist eine Basis von $l^2(I)$. Es folgt, dass $(Ue_j)_{j\in I}$ eine Basis von $\mathcal{H}$ ist. $\square$

Bemerkung 5. Die Zustände eines quantenmechanischen Systems werden mittels eines Hilbertraumes beschrieben. Man fordert, dass der Zustand eines quantenmechanischen Systems durch höchstens abzählbare viele Messungen bestimmt werden kann. Mathematisch ausgedrückt bedeutet diese Eigenschaft, dass der Hilbertraum eines quantenmechanischen Systems separabel sein soll.

Completion

Definition 3.3.8. *Let $(V, \langle\cdot,\cdot\rangle)$ be a inner product space. A Hilbert space space $(\hat{V}, \langle\cdot,\cdot\rangle_{\hat{V}})$ together with a mapping $j : V \to \hat{V}$ is called a* **completion** *of V if*

- *j preserves the inner product,*

- *$j(V)$ is dense in $\hat{V}$.*

Theorem 3.3.18. *(Existence and uniqueness of completions) Let $(V, \langle\cdot,\cdot\rangle)$ be a inner product space which is not complete. Then there exists a completion. The completion is unique up to unitary mappings.*

Proof. This is an immediate consequence of Theorem 1.1.4. And the fact that norm preserving linear maps between inner product spaces, preserve the inner product (Lemma 3.3.16). $\square$

3.4 Riesz's Theorem

Lemma 3.4.1. *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Let $v \in V$. Then*

$$l_v : V \to \mathbb{K}, \quad x \mapsto \langle v, x \rangle$$

is a continuous linear mapping with $\|l_v\| = \|v\|$.

Proof. Observe that by definition. $\|l_v\| = \sup\{|l_v(x)| : x \in V, \|x\| \leq 1\}$. We have for $x \in V$ that

$$|l_v(x)| = |\langle v, x \rangle| \leq \|v\|\|x\|,$$

which shows $\|l_v\| \leq \|v\|$, and

$$|l_v(v)| = |\langle v, v \rangle| = \|v\|^2,$$

which showns $\|l_v\| \geq \|v\|$. □

The follwing theorem gives a converse of the above lemma, that is that every linear functional on a Hilbert space is given by the inner product with some vector.

Theorem 3.4.2. *(Riesz's Theorem) For each $l \in \mathcal{L}(\mathcal{H}, \mathbb{C})$, there exists a unique $y_l \in \mathcal{H}$ such that $l(x) = \langle y_l, x \rangle$ for all $x \in \mathcal{H}$. In addition $\|l\| = \|y_l\|$.*

Proof. Existence: If $l(x) = 0$ for all l, then choose $y_l = 0$ and we are done. Otherwise, if l is not the zero functional, then $M := \{x \in \mathcal{H} | l(x) = 0\}$ is not the whole Hilbert space. Since, l is continuous M is a closed subspace of $\mathcal{H}$. By the projection theorem, there exists a normalized vector $z \in M^\perp$ (pick a nonzero $w \in \mathcal{H} \setminus M$ and write $w = w^\| + w^\perp$ with $w^\| \in M$ and $w^\perp \in M^\perp$. Then $w^\perp \neq 0$ and $z = w^\perp / \|w^\perp\|$ is a normalized vector in $M^\perp$). We will show that $y_l = \overline{l(z)} z$ satisfies, $l(x) = \langle y_l, x \rangle$. To show this, oberve that for all $x \in \mathcal{H}$, $l(x)z - l(z)x \in M$. This implies,

$$0 = \langle z, l(x)z - l(z)x \rangle = l(x)\langle z, z \rangle - l(z)\langle z, x \rangle \ .$$

Therefore, $l(x) = l(z)\langle z, x \rangle = \langle \overline{l(z)} z, x \rangle = \langle y_l, x \rangle$.

Uniqueness: Suppose $\langle y, x\rangle = \langle y', x\rangle$ for all $x \in \mathcal{H}$. Then $\langle y - y', x\rangle = 0$ for all $x \in \mathcal{H}$, and $y' - y = 0$ (choose $x = y - y'$).
The last statement follows from the above lemma. $\square$

Corollary 3.4.3. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. Let $h : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathbb{F}$ be a function which satisfies*

(i) $h(u, \alpha x + \beta y) = \alpha h(u, x) + \beta h(v, y)$

(ii) $h(\alpha u + \beta v, x) = \overline{\alpha} h(u, x) + \overline{\beta} h(v, x)$

(iii) $|h(u, x)| \leq C\|u\|\|x\|$

for all $u, v \in \mathcal{H}_1$, $x, y \in \mathcal{H}_2$, $\alpha, \beta \in \mathbb{F}$. Then there exists a unique bounded linear transformation $A : \mathcal{H}_1 \to \mathcal{H}_2$ so that

$$h(x, y) = \langle Ax, y\rangle$$

for all $x, y \in \mathcal{H}_1$. The norm of A is the smallest constant C such that (iii) holds.

Proof. Uniqueness follows from (g) of Lemma 3.2.5.
Fix $u \in \mathcal{H}_1$. Then

$$l_u : \mathcal{H} \to \mathbb{F}, \quad y \mapsto h(u, y)$$

is by (i) and (iii) a continuous linear functional on $\mathcal{H}_2$. Thus by Riesz's theorem, there is a $z_u \in \mathcal{H}_2$ such that

$$\langle z_u, y\rangle = h(u, y) \quad \text{for all } y \in \mathcal{H}_2.$$

Define $Au = z_u$. It is not difficult to show that A is a continuous linear operator with the right properties: For this let $u, v \in \mathcal{H}_1$ and $\alpha, \beta \in \mathbb{F}$. Then for all $y \in \mathcal{H}_2$

$$\begin{aligned} \langle A(\alpha u + \beta v), y\rangle &= h(\alpha u + \beta v, y) \\ &= \overline{\alpha} h(u, y) + \overline{\beta} h(v, y) \\ &= \overline{\alpha}\langle Au, y\rangle + \overline{\beta}\langle v, y\rangle \\ &= \langle \alpha Au + \beta Av, y\rangle. \end{aligned}$$

This shows linearity. We have $\|A\| \leq C$ by Lemma 3.3.4. On the other hand we have $|h(u,v)| = |\langle Au, v\rangle| \leq \|A\|\|u\|\|v\|$. Thus $\|A\|$ is equal to the smallest constant C such that (iii) holds. $\square$

Remark 3.4.1. We note that a mapping h satisfying (i) and (ii) of Corollary 3.4.3 is called a sesquilinear form. The assertion of Corollary 3.4.3 extends naturally to the case where $\mathcal{H}_1$ is any normed vector space.

3.5 Hilbert Adjoint for bounded Operators

In this section we consider complex Hilbert spaces. We recall the following lemma, which follows directly from (g) of Lemma 3.2.5.

Lemma 3.5.1. *let H be a inner product space and for all $x \in H$*

$$\langle v, x\rangle = \langle w, x\rangle.$$

Then $v = w$.

For convenience we repeat the proof.

Proof. If the assumption hold we have for all $x \in H$ $\langle v - w, x\rangle = 0$. In particular $\langle v - w, v - w\rangle = 0$. This implies the claim. □

Lemma 3.5.2. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces. For $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ there exists a unique $T^* \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ such that*

$$\langle x, Ty\rangle_{\mathcal{H}_2} = \langle T^*x, y\rangle_{\mathcal{H}_1}$$

for all $x \in \mathcal{H}_2$ and $y \in \mathcal{H}_1$. We call T^* the **Hilbert adjoint** *of T.*

The lemma follows directly from Corollary 3.4.3. For convenience we give an independent proof.

Proof. For each $x \in \mathcal{H}_2$ the mapping

$$l_{x,T} : \mathcal{H}_1 \to \mathbb{F}, \quad y \mapsto \langle x, Ty\rangle_{\mathcal{H}_2}$$

is linear and continuous ($\|l_{x,T}(y)\| \leq \|x\|\|T\|\|y\|$). Thus by the Riesz Lemma there exists a unique element $y_{x,T} \in \mathcal{H}_1$ such that

$$\langle y_{x,T}, y\rangle_{\mathcal{H}_1} = l_{x,T}(y) = \langle x, Ty\rangle_{\mathcal{H}_2}$$

for all $y \in \mathcal{H}_1$. This shows the uniqueness part. Define

$$T^*x := y_{x,T}.$$

Then T^* is linear, since for all $y \in \mathcal{H}_1$ we have

$$\begin{aligned}
\langle y_{\alpha x+\alpha' x',T}, y\rangle_{\mathcal{H}_1} &= \langle \alpha x + \alpha' x', Ty\rangle_{\mathcal{H}_2} \\
&= \alpha\langle x, Ty\rangle_{\mathcal{H}_2} + \alpha'\langle x', Ty\rangle_{\mathcal{H}_2} \\
&= \langle \alpha y_{x,T} + \alpha' y_{x',T}, y\rangle_{\mathcal{H}_1}.
\end{aligned}$$

□

Remark 3.5.1. By taking complex conjugates we see

$$\langle Ty, x\rangle_{\mathcal{H}_2} = \langle y, T^*x\rangle_{\mathcal{H}_1}.$$

Lemma 3.5.3. *Let $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ be Hilbert spaces, $S, T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, $R \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$. Then*

(a) $(\alpha S + \beta T)^* = \overline{\alpha} S^* + \overline{\beta} T^*$ *for all* $\alpha, \beta \in \mathbb{F}$.

(b) $(RS)^* = S^*R^*$.

(c) $\|T^*\| = \|T\|$.

(d) $(T^*)^* = T$.

(e) $\|TT^*\| = \|T^*T\| = \|T\|^2$.

(f) $\mathrm{ker} S = (\mathrm{ran} S^*)^\perp$, $\mathrm{ker} S^* = (\mathrm{Ran} S)^\perp$.

Proof. (a) This follows, since

$$\langle (\alpha S + \beta T)^* x, y\rangle = \langle x, (\alpha S + \beta T)y\rangle = \alpha\langle x, Sy\rangle + \beta\langle x, Ty\rangle = \langle (\overline{\alpha} S^* + \overline{\beta} T^*)x, y\rangle$$

(b) This follows, since

$$\langle (RS)^* x, y\rangle = \langle x, RSy\rangle = \langle R^*x, Sy\rangle = \langle S^*R^*x, y\rangle.$$

(c) $\|T^*\| = \sup_{x,y} |\langle x, T^*y\rangle| = \sup_{x,y} |\langle T^*y, x\rangle| = \sup_{x,y} |\langle y, Tx\rangle| = \|T\|$.
(d) $\langle x, Ty\rangle = \langle T^*x, y\rangle = \langle x, (T^*)^*y\rangle$.
(e) We have

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle \leq \|x\|\|T^*Tx\|.$$

Thus

$$\|T\|^2 = \sup_{\|x\|\leq 1} \|Tx\|^2 \leq \sup_{\|x\|\leq 1} \|x\|\|T^*Tx\| \leq \|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2.$$

Hence

$$\|T^*T\| = \|T\|^2.$$

This implies $\|T\|^2 = \|T^*\|^2 = \|(T^*)^*T\| = \|TT^*\|$.
(f) We have $\ker T = (\mathrm{ran} T^*)^\perp$, since

$$\begin{aligned} Sx = 0 \ &\Leftrightarrow\ \langle Sx, y\rangle = 0 \quad \forall y \in \mathcal{H}_2 \\ &\Leftrightarrow\ \langle x, S^*y\rangle = 0 \quad \forall y \in \mathcal{H}_2 \\ &\Leftrightarrow\ x \in (\mathrm{ran} T^*)^\perp, \end{aligned}$$

and therefore also $\ker T^* = (\mathrm{ran} T^{**})^\perp = (\mathrm{ran} T)^\perp$ $\square$

Definition 3.5.1. *Let $\mathcal{H}$ be a Hilbertspace and $T \in \mathcal{L}(\mathcal{H})$. Then*

- *T is called* **selfadjoint** *if $T^* = T$.*

- *T is called* **normal** *if $T^*T = T^*T$.*

Lemma 3.5.4. *A selfajdoint operator is normal.*

Lemma 3.5.5. *(Chracterization Normality) Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{L}(\mathcal{H})$. Then the following statements are equivalent.*

(a) T is normal.

(b) $\|Tx\| = \|T^*x\|$ for all $x \in \mathcal{H}$.

In any of teh above cases it follows that $\ker T = \ker T^*$.

Proof. (a) $\Rightarrow$ (b): $\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle = \langle TT^*x, x\rangle = \langle T^*x, T^*x\rangle = \|T^*x\|^2$.
(b) $\Rightarrow$ (a) Follows from polarization. We have

$$\langle T^*Tx, y\rangle = \frac{1}{4}\sum_{n=0}^{3} i^{-n}\|T(x+i^n y)\|^2 = \frac{1}{4}\sum_{n=0}^{3} i^{-n}\|T^*(x+i^n y)\|^2 = \langle TT^*x, y\rangle.$$

□

Lemma 3.5.6. *(Chracterization Selfadjointness) Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{L}(\mathcal{H})$. Then the following statements are equivalent.*

(a) T is selfadjoint.

(b) $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in \mathcal{H}$.

(c) $\langle Tx, x\rangle \subset \mathbb{R}$ for all $x \in \mathcal{H}$.

In any of the above cases we have

$$\|T\| = \sup_{\|x\|\le 1} |\langle Tx, x\rangle|.$$

Proof. (a) $\Rightarrow$ (b): Follows from the definition.
(b) $\Rightarrow$ (a): It follows for all $x, y \in \mathcal{H}$ that $\langle T^*x, y\rangle = \langle x, Ty\rangle = \langle Tx, y\rangle$. Hence $T^*x = Tx$ for all $x \in \mathcal{H}$.
(b) $\Rightarrow$ (c): We have $\overline{\langle Tx, x\rangle} = \langle x, Tx\rangle = \langle Tx, x\rangle$.
(c) $\Rightarrow$ (b): First note that by (c) we have $\langle Tx, x\rangle = \overline{\langle Tx, x\rangle} = \langle x, Tx\rangle$. This it follows from polarization, that

$$\langle Tx, y\rangle = \frac{1}{4}\sum_{n=0}^{3} i^n\langle T(x+i^{-n}y), (x+i^{-n}y)\rangle = \frac{1}{4}\sum_{n=0}^{3} i^n\langle (x+i^{-n}y), T(x+i^{-n}y)\rangle = \langle x, Ty\rangle.$$

Sei $M := \sup_{\|x\|\le 1} |\langle Tx, x\rangle|$. Offensichtlich gilt $M \le \|T\|$. Die Umgekehrte Gleichung folgt aus der Polarisierungsidentität: Mit $\|x\|, \|y\| \le 1$ und (c) gilt

$$\mathrm{Re}\langle Tx, y\rangle = \frac{1}{4}\langle T(x+y), (x+y)\rangle \le \frac{1}{4}M\|x+y\|^2 \le M.$$

Mulitplying with a suitable $e^{i\alpha}$ we find for $\|x\|, \|y\| \le 1$ that

$$|\langle Tx, y\rangle| = \langle Tx, e^{i\alpha} y\rangle \le M.$$

This implies $\|T\| \le M$. $\square$

Corollary 3.5.7. *Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{L}(\mathcal{H})$ selfadjoint. If $\langle Tx, x\rangle =$ for all $x \in \mathcal{H}$, then $T = 0$.*

Definition 3.5.2. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbertspaces. Then $U \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is called* **unitary** *if $U^*U = 1_{\mathcal{H}_1}$ and $UU^* = 1_{\mathcal{H}_2}$.*

Lemma 3.5.8. *(Chracterization Unitarity) Let $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbertspaces and $U \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Then the following statements are equivalent.*

(a) U is unitary

(b) U is bijective and $U^* = U^{-1}$.

(c) U is bijective and $\langle Ux, Uy\rangle = \langle x, y\rangle$.

(d) U is a bijective and preserves the norm.

Proof. (a) $\Rightarrow$ (b): $U^*U = 1_{\mathcal{H}_1}$ and $UU^* = 1_{\mathcal{H}_2}$ implies that U is bijecitve and $U^* = U^{-1}$. (b) $\Rightarrow$ (a): This is trivial. (b) $\Rightarrow$ (c): It follows that $\langle Ux, Uy\rangle = \langle x, U^*Uy\rangle = \langle x, y\rangle$. (c) $\Rightarrow$ (b): We have $\langle x, y\rangle = \langle Ux, Uy\rangle = \langle x, U^*Uy\rangle$, which implies $U^*U = 1$. Bijectivity implis that $U^* = U^{-1}$. (c) $\Rightarrow$ (d): This is trivial and follows if we set $x = y$. (d) $\Rightarrow$ (c): This follows from polarization. That is $\langle Ux, Uy\rangle = \frac{1}{4}\sum_{n=0}^{3} i^{-n}\|U(x+i^n y)\|^2 = \frac{1}{4}\sum_{n=0}^{3} i^{-n}\|x + i^n y\|^2 = \langle x, y\rangle$. $\square$

Lemma 3.5.9. *(Chacterization Orthogonal Projection) Let $\mathcal{H}$ be a Hilbert space and $P \neq 0$ a projection in $\mathcal{H}$. The following satements are equivalent.*

(a) P is an orthogonal projection.

(b) P is selfadjoint.

(c) P is normal.

(d) $\|P\| = 1$.

Proof. (a) $\Rightarrow$ (b): We have $\langle Px, y\rangle = \langle Px, Py + (1-P)y\rangle = \langle Px, Py\rangle = ... = \langle x, Py\rangle$.
(b) $\Rightarrow$ (c): trivial.
(c) $\Rightarrow$ (a): $\ker P = \ker P^* = (\mathrm{Im} P)^\perp$.
(a) $\Rightarrow$ (d): $\|P\| \leq 1$ has already been shown. Equality follows from $P(Px) = Px$ and $Px \neq 0$ for some $x \in \mathcal{H}$.
(d) $\Rightarrow$ (b): Let $x \in \ker P$ and $y \in \mathrm{Im} P$. Then for all $\lambda \in \mathbb{R}$ we find

$$\|\lambda y\|^2 = \|P(x + \lambda y)\|^2 \leq \|x + \lambda y\|^2 = \|x\|^2 + \lambda(\langle x, y\rangle + \langle y, x\rangle) + \|\lambda y\|^2.$$

Thus $0 \leq \|x\|^2 + \lambda(\langle x, y\rangle + \langle y, x\rangle)$, which implies $\langle x, y\rangle + \langle y, x\rangle = 0$. For $u \in \ker P$ we find with $x = iu$ and $x = u$ that $-\langle u, y\rangle + \langle y, u\rangle = 0$ and $\langle u, y\rangle + \langle y, u\rangle = 0$. Thus $\langle u, y\rangle$ for all $u \in \ker P$ and $y \in \mathrm{Im} P$. $\square$

3.6 Constructions of Hilbert spaces

Direct Sums of Hilbert spaces

Proposition 3.6.1. *(Direct sum of Hilbert spaces.)*

(a) *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. Then $\mathcal{H}_1 \times \mathcal{H}_2$ is a Hilbert space with inner product*

$$\langle (x_1, y_1), (x_2, y_2) \rangle = \langle x_1, x_2 \rangle_{\mathcal{H}_1} + \langle y_1, y_2 \rangle_{\mathcal{H}_2} \, .$$

This space is called the **direct sum** *of the spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ and is denoted by $\mathcal{H}_1 \oplus \mathcal{H}_2$.*

(b) *Suppose $(\mathcal{H}_n)_{n=1}^N$, with $N \in \mathbb{N}$ or $N = \infty$, is a collection of Hilbert spaces. The* **direct sum** *is defined by*

$$\bigoplus_{n=1}^{N} \mathcal{H}_n := \{(x_n)_{n=1}^N : \; x_n \in \mathcal{H}_n, \; \sum_{n=1}^{N} \|x_n\|^2_{\mathcal{H}_n} < \infty\},$$

and it is a Hilbert space with inner product

$$\langle (x_n)_{n=1}^N, (y_n)_{n=1}^N \rangle := \sum_{n=1}^{N} \langle x_n, y_n \rangle_{\mathcal{H}_n} \, .$$

Proof. (a) Is straight forward to verify or follows directly from (b).
(b) The proof is analogous to the proof of 3.3.5 □

Remark 3.6.1. By unitary equivalence we do not need distinguish between inner and and outer direct sum.

Hilbert space of vector valued functions

Proposition 3.6.2. *(Hilbert space of vector valued functions) Suppose (M, μ) is a measure space and $\mathcal{H}'$ a separable Hilbert space. Let $L^2(M, d\mu; \mathcal{H}')$ be the set of measurable functions on M with values in $\mathcal{H}'$, which satisfy*

$$\int_M \|f(x)\|^2_{\mathcal{H}'} d\mu(x) < \infty.$$

This set is a Hilbert space with the inner product

$$\langle f, g\rangle = \int_M \langle f(x), g(x)\rangle_{\mathcal{H}'} d\mu(x).$$

(A function $f : M \to \mathcal{H}'$ is called measurable if for every $y \in \mathcal{H}'$ the function $x \mapsto \langle y, f(x)\rangle_{\mathcal{H}'}$ is measurable. Note there are different notions of measurability for vector valued functions, which agree for separable spaces.)
If $(\varphi_j)_{j\in I}$ is an ONB of $\mathcal{H}'$, then

$$U : L^2(M, d\mu; \mathcal{H}') \to \bigoplus_{j\in I} L^2(M, d\mu), \quad f \mapsto (\langle \varphi_j, f\rangle_{\mathcal{H}'})_{j\in I}. \tag{3.6.1}$$

is a unitary map.

For a proof see Reed & Simon.

Tensor product of Hilbert spaces

Definition 3.6.1. *(***Tensor product of Hilbert spaces***) Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces. For each $\varphi_1 \in \mathcal{H}_1$ and $\varphi_2 \in \mathcal{H}_2$ define $\varphi_1 \otimes \varphi_2$ as the conjugate bilinear form which acts on $\mathcal{H}_1 \times \mathcal{H}_2$ by*

$$(\varphi_1 \otimes \varphi_2)(x_1, x_2) := \langle x_1, \varphi_1\rangle_{\mathcal{H}_1} \langle x_2, \varphi_2\rangle_{\mathcal{H}_2}, \quad \forall (x_1, x_2) \in \mathcal{H}_1 \times \mathcal{H}_2.$$

Define

$$\mathcal{E} := \mathrm{lin}\{\varphi_1 \otimes \varphi_2 : \varphi_i \in \mathcal{H}_i\}$$

and an inner product by extending

$$\langle \varphi_1 \otimes \varphi_2, \psi_1 \otimes \psi_2\rangle_{\mathcal{E}} := \langle \varphi_1, \psi_1\rangle_{\mathcal{H}_1} \langle \varphi_2, \psi_2\rangle_{\mathcal{H}_2}.$$

by linearity to $\mathcal{E}$. (It is straight forward to verify, see the lemma below, that $\langle\cdot,\cdot\rangle_{\mathcal{E}}$ is well defined and an inner product). Define $\mathcal{H}_1 \otimes \mathcal{H}_2$ as the completion of $\mathcal{E}$.

Lemma 3.6.3. *$\langle\cdot,\cdot\rangle_{\mathcal{E}}$ is well defined and an inner product.*

Proof. To show that $\langle \cdot, \cdot \rangle$ is well defined, we must show that $\langle \lambda, \lambda' \rangle$ does not depend on which finite linear combinations are used to express λ and λ'. To do this it is sufficientto show that if $\mu = \sum_{j=1}^{M} a_j(\eta_j \otimes \tau_j)$ is the zero form, then $\langle \eta, \mu \rangle = 0$ for all $\eta \in \mathcal{E}$. To see that this is true, let $\eta = \sum_{i=1}^{N} c_i(\varphi_i \otimes \psi_i)$, then

$$\begin{aligned}
\langle \eta, \mu \rangle &= \langle \sum_{i=1}^{N} c_i(\varphi_i \otimes \psi_i), \mu \rangle \\
&= \sum_{i=1}^{N} c_i \langle \varphi \otimes \psi_i, \mu \rangle \\
&= \sum_{i=1}^{N} c_i \sum_{j=1}^{M} a_j \langle \varphi_i \otimes \psi_i, \eta_j \otimes \tau_j \rangle \\
&= \sum_{i=1}^{N} c_i \sum_{j=1}^{M} a_j (\eta_j \otimes \tau_j)(\varphi_i, \psi_i) \\
&= 0,
\end{aligned}$$

since μ is the zero form. Thus, $\langle \cdot, \cdot \rangle_{\mathcal{E}}$ is well defined.

It remains to show positive definiteness. Suppose $\lambda = \sum_{k=1}^{M} d_k(\eta_k \otimes \mu_k)$. Then $(\eta_k)_{k=1}^{m}$ and $(\mu_k)_{k=1}^{M}$ span subspaces $M_1 \subset \mathcal{H}_1$ and $M_2 \subset \mathcal{H}_2$, respectively. If we let $(\varphi_j)_{j=1}^{N_1}$ and $(\varphi_l)_{l=1}^{N_2}$ be ONB of M_1 and M_2 respectively, we can express each η_k in terms of the φ_j's and each μ_k in terms of the ψ_l's obtaining

$$\lambda = \sum_{j=1, l=1}^{N_1, N_2} c_{j,l}(\varphi_j \otimes \psi_l).$$

But

$$\begin{aligned}
\langle \lambda, \lambda \rangle_{\mathcal{E}} &= \langle \sum c_{j,l}(\varphi_j \otimes \psi_l), \sum c_{i,m}(\varphi_i \otimes \psi_m) \rangle_{\mathcal{E}} \\
&= \sum \overline{c_{j,l}} c_{i,m} \langle \varphi_j, \varphi_i \rangle \langle \varphi_l, \varphi_m \rangle \\
&= \sum_{j,l} |c_{j,l}|^2.
\end{aligned}$$

so if $\langle \lambda, \lambda \rangle_{\mathcal{E}} = 0$, then all the $c_{j,l} = 0$ and λ is the zero form. Thus $\langle \lambda, \lambda \rangle_{\mathcal{E}}$ is positive definite. $\square$

Remark 3.6.2. Let $\mathcal{H}_1, ..., \mathcal{H}_n$ one can define the n-fold tensor product inductively by

$$\mathcal{H}_1 \otimes \cdots \mathcal{H}_n := ((\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \cdots) \otimes \mathcal{H}_n,$$

or alternatively using multilinear forms.

Proposition 3.6.4. *For $i = 1, 2$ let $(\varphi_{i,k})_{k \in N_i}$ be an ONB of $\mathcal{H}_i$. Then $(\varphi_{1,l} \otimes \varphi_{2,k})_{(l,k) \in N_1 \times N_2}$ is an orthonormal basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$.*

Proof. Clearly $(\varphi_{1,l} \otimes \varphi_{2,k})_{(l,k) \in N_1 \times N_2}$ is an orthonormal set. Let $\eta_1 \otimes \eta_2$ with $\eta_i \in \mathcal{H}_i$ and

$$0 = \langle \eta_1 \otimes \eta_2, \varphi_{1,l} \otimes \varphi_{2,k} \rangle, \quad \forall (l, k) \in N_1 \times N_2.$$

It follows, that if $\langle \eta_1, \varphi_{1,l} \rangle \neq 0$ for some l, then $\langle \eta_2, \varphi_{2,k} \rangle = 0$ for all k, and $\eta_2 = 0$. Thus we conclude that $\eta_1 \otimes \eta_2 = 0$. Since the space $\mathcal{E}$ is dense, the claim follows. □

Example 15. (General Fock Spaces). Let $\mathcal{H}$ be a Hilbert space and denote by $\mathcal{H}^{\otimes n}$ the n-fold tensor product $\mathcal{H}^{\otimes n} = \mathcal{H} \otimes \mathcal{H} \otimes \cdots \otimes \mathcal{H}$. Set $\mathcal{H}^{\otimes 0} := \mathbb{C}$ and define

$$\mathcal{F}(\mathcal{H}) = \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes n}.$$

$\mathcal{F}(\mathcal{H})$ is called the Fock space over $\mathcal{H}$; it will be separable if $\mathcal{H}$ is. Actually, if is not $\mathcal{F}(\mathcal{H})$ itself, but two of its subspaces which are used most frequently in quantum field theory. The two subspacesare constructed as follows. Let $\mathcal{P}_n$ be the permutation group of n elements and let (φ_k) be a basis for $\mathcal{H}$. For each $\sigma \in \mathcal{P}_n$, we define the operator (which we also denote by σ) on basis elements of $\mathcal{H}^{\otimes n}$ by

$$\sigma(\varphi_{k_1} \otimes \varphi_{k_2} \otimes \cdots \otimes \varphi_{k_n}) = \varphi_{k_{\sigma(1)}} \otimes \varphi_{k_{\sigma(2)}} \otimes \cdots \otimes \varphi_{k_{\sigma(1n)}}.$$

σ extends by linearity to a bouned operator (of norm one) on $\mathcal{H}^{\otimes n}$ so we can define

$$S_n = \frac{1}{n!} \sum_{\sigma \in \mathcal{P}_n} \sigma.$$

It is an easy excercise to show that $S_n^2 = S_n$ and $S_n^* = S_n$, so S_n is an orthgonal projection. The range of S_n, denoted by $\mathcal{H}^{\otimes_s n}$, is called the n-fold **symmetric tensor product** of $\mathcal{H}$. We now define

$$\mathcal{F}_s(\mathcal{H}) = \bigoplus_{n=0}^{\infty} S_n(\mathcal{H}^{\otimes n}).$$

$\mathcal{F}_s(\mathcal{H})$ is called the **symmetric Fock space over** $\mathcal{H}$ or the **Boson Fock space over** $\mathcal{H}$. Let $\epsilon(\cdot)$ be the function from $\mathcal{P}_n$ to $\{-1, 1\}$ which is one on even permutations and minus one on odd permutations. Define

$$A_n = \frac{1}{n!} \sum_{\sigma \in \mathcal{P}_n} \epsilon(\sigma)\sigma.$$

Again it is an easy excercise to show that $A_n^2 = A_n$ and $A_n^* = A_n$, so A_n is an orthogonal projection. The range of A_n, denoted by $\mathcal{H}^{\otimes_a n}$, is called the n-fold **antisymmetric tensor product** of $\mathcal{H}$. We now define

$$\mathcal{F}_a(\mathcal{H}) = \bigoplus_{n=0}^{\infty} A_n(\mathcal{H}^{\otimes n}).$$

$\mathcal{F}_a(\mathcal{H})$ is called the **antisymmetric Fock space over** $\mathcal{H}$ or the **Fermion Fock space over** $\mathcal{H}$.

Tensor product of L^2-spaces

Theorem 3.6.5. *Let (M_i, μ_i), $i = 1, 2$ be measure spaces, such that $L^2(M_i, d\mu_i)$, $i = 1, 2$, are separable. Then*

(a) There is a unique isomorphism from $L^2(M_1, d\mu_1) \otimes L^2(M_2, d\mu_2)$ to $L^2(M_1 \times M_2, d\mu_1 \otimes d\mu_2)$ so that $f \otimes g \mapsto fg$, where $fg : (x_1, x_2) \mapsto f(x_1)g(x_2)$.

(b) If $\mathcal{H}'$ is a separable Hilbert space, then there is a unique isomorphism from $L^2(M_1, d\mu_1) \otimes \mathcal{H}'$ to $L^2(M_1, d\mu_1; \mathcal{H}')$ so that $f \otimes g \mapsto fg$.

(c) There is a unique isomorphism from $L^2(M_1 \times M_2, d\mu_1 \otimes d\mu_2)$ to $L^2(M_1, d\mu_1; L^2(M_2, d\mu_2))$ such that f is taken to $x_1 \mapsto (x_2 \mapsto f(x_1, x_2))$.

For a proof See Reed & Simon.

Example 16. As an example of (a) we have $L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) = L^2(\mathbb{R}^2)$. Or more generally we have $L^2(\mathbb{R}^{d_1}) \otimes \cdots \otimes L^2(R^{d_n}) = L^2(\mathbb{R}^{d_1+\cdots+d_n})$.

Example 17. The Hilbert space in the quantum-mechanical description of a single Schrödinger particle of spin one-half is $L^2(\mathbb{R}^3, dx; \mathbb{C}^2)$, that is, the set of pairs $(\psi_1(x), \psi_2(x))$ of square integrable functions. By what we have shown above, $L^2(\mathbb{R}^3, dx; \mathbb{C}^2)$ is naturally isomorphic to $L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$.

Example 18. (Fock spaces over L^2-spaces). We refer to notation introduced in Example 15. We consider Fock spaces in the special case where $\mathcal{H} = L^2(\mathbb{R}^3)$. In that case we have for $n \in \mathbb{N}$

$$\mathcal{H}^{\otimes n} = L^2(\mathbb{R}^{3n})$$

and $\mathcal{F}(\mathcal{H})$ consists of sequences $\psi = (\psi_n)_{n\in\mathbb{N}}$ such that

$$\psi_0 \in \mathbb{C}, \quad \psi_n \in L^2(\mathbb{R}^{3n}), \quad n \in \mathbb{N},$$

and

$$\|\psi\|^2 := |\psi_0|^2 + \sum_{n=1}^{\infty} \int_{\mathbb{R}^{3n}} |\psi_n(x_1, ..., x_n)|^2 dx_1 \cdots dx_n < \infty.$$

The symmetric and antisymmetric tensor products, respectively are given by

$$\begin{aligned}
\mathcal{H}^{\otimes_s n} &= \{\psi \in L^2(\mathbb{R}^{3n}) : \psi(x_1, ..., x_n) = \psi(x_{\sigma(1)}, ..., x_{\sigma(n)}), \quad \forall \sigma \in \mathcal{P}_n\} \\
\mathcal{H}^{\otimes_a n} &= \{\psi \in L^2(\mathbb{R}^{3n}) : \psi(x_1, ..., x_n) = \epsilon(\sigma)\psi(x_{\sigma(1)}, ..., x_{\sigma(n)}), \quad \forall \sigma \in \mathcal{P}_n\}.
\end{aligned}$$