Folland: *Real Analysis*

Danny Nygård Hansen

29th March 2022

0 • Introduction

- We use the letter d to denote vector spaces dimensions, freeing up n to be used as an index, e.g. in sequences. In particular we write $\mathbb{R}^d$ and $\mathbb{C}^d$.
- The symbol $\mathbb{K}$ denotes either the real or complex numbers.
- The unit sphere in $\mathbb{R}^{n+1}$ is denoted $\mathbb{S}^n$.
- We denote the power set of a set X by 2^X.
- The restriction of a function $f: X \to Y$ to a subset $A \subseteq X$ is denoted $f|_A$.
- Whenever we need to make the distinction, $\mathcal{L}^p(\mu)$ refers to the space of μ-p-integrable functions, while $L^p(\mu)$ denotes the quotient of $\mathcal{L}^p(\mu)$ with the subspace of functions that are zero μ-a.e.
- The space of bounded operators between normed spaces X and Y is denoted $\mathcal{B}(X, Y)$.
- The bounded and continuous complex-valued functions on a topological space X is denoted $C_b(X)$.
- A vector space equipped with an inner product is called an *inner product space*.

1 • Measures

1.2. σ-algebras

EXERCISE 1.1

Let $\mathcal{M}$ be an infinite σ-algebra.

(a) $\mathcal{M}$ contains an infinite sequence of disjoint sets.

(b) $|\mathcal{M}| \geq \mathfrak{c}$.

Of course part (a) is trivial unless we require the sets to be nonempty.

SOLUTION. (a) We show by contraposition that there exists a nonempty set $A \in \mathcal{M}$ such that the restriction of $\mathcal{M}$ to A^c is infinite. That is, assuming that no such set exists, we show that $\mathcal{M}$ is finite. Pick any nonempty $A \in \mathcal{M}$. Then the restriction of $\mathcal{M}$ to A and A^c respectively are both finite. For any $B \in \mathcal{M}$ we can write

$$B = (B \cap A) \cup (B \cap A^c).$$

But each set in the union lies in one of the restrictions, so there are finitely many decompositions like the one above, so there are finitely many sets $B \in \mathcal{M}$.

Now construct the sequence: Pick $A \in \mathcal{M}$ as above, restrict $\mathcal{M}$ to A^c, and continue recursively.

(b) Let (A_n) be the sequence constructed above. There is an injection $\varphi\colon 2^{\mathbb{N}} \to \mathcal{M}$ given by $\varphi(I) = \bigcup_{i\in I} A_i$ (injectivity follows since the sets in the sequence are disjoint). Hence $|\mathcal{M}| \geq |2^{\mathbb{N}}| = \mathfrak{c}$. □

1.3. Measures

EXERCISE 1.14

If μ is a semifinite measure and $\mu(E) = \infty$, for any $C > 0$ there exists $F \subseteq E$ with $C < \mu(F) < \infty$.

SOLUTION. Consider

$$S = \sup\{\mu(F) \mid F \subseteq E, \mu(F) < \infty\}.$$

If $S = \infty$, then the result is obvious. So assume towards a contradiction that $S < \infty$. For $n \in \mathbb{N}$ choose $F_n \subseteq E$ with $\mu(F_n) < \infty$ such that

$$S - \frac{1}{n} \leq \mu(F_n) \leq S.$$

Put $G_k = \bigcup_{n=1}^{k} F_n$. Then $G_k \subseteq E$ and $\mu(G_k) < \infty$, so the same inequality holds with F_n replaced by G_k. Now putting $G = \bigcup_{k\in\mathbb{N}} G_k$, continuity of μ gives

$$S - \frac{1}{n} \leq \mu(G) \leq S$$

for all $n \in \mathbb{N}$, so $\mu(G) = S$.

By assumption $\mu(E \setminus G) = \infty$, so $E \setminus G$ contains a set $G' \in \mathcal{M}$ such that $0 < \mu(G') < \infty$. But then

$$\mu(G \cup G') = \mu(G) + \mu(G') > S,$$

a contradiction. □

EXERCISE 1.16

Let $(X, \mathcal{M}, \mu)$ be a measure space. A set $E \subseteq X$ is called *locally measurable* if $E \cap A \in \mathcal{M}$ for all $A \in \mathcal{M}$ such that $\mu(A) < \infty$. Let $\widetilde{\mathcal{M}}$ be the collection of all locally measurable sets. Clearly $\mathcal{M} \subseteq \widetilde{\mathcal{M}}$; if $\mathcal{M} = \widetilde{\mathcal{M}}$, then μ is called *saturated*.

(a) If μ is σ-finite, then μ is saturated.

(b) $\widetilde{\mathcal{M}}$ is a σ-algebra.

(c) Define $\tilde{\mu}$ on $\widetilde{\mathcal{M}}$ by $\tilde{\mu}(E) = \mu(E)$ if $E \in \mathcal{M}$ and $\tilde{\mu}(E) = \infty$ otherwise. Then $\tilde{\mu}$ is a saturated measure on $\widetilde{\mathcal{M}}$, called the *saturation* of μ.

(d) If μ is complete, so is $\tilde{\mu}$.

(e) Suppose that μ is semifinite. For $E \in \widetilde{\mathcal{M}}$ define

$$\underline{\mu}(E) = \sup\{\mu(A) \mid A \in \mathcal{M} \text{ and } A \subseteq E\}.$$

Then $\underline{\mu}$ is a saturated measure on $\widetilde{\mathcal{M}}$ that extends μ.

(f) Let X_1, X_2 be disjoint uncountable sets, $X = X_1 \cup X_2$, and $\mathcal{M}$ the σ-algebra of countable or co-countable sets in X. Let μ_0 be counting measure on 2^{X_1}, and define μ on $\mathcal{M}$ by $\mu(E) = \mu_0(E \cap X_1)$. Then μ is a measure on $\mathcal{M}$, $\widetilde{\mathcal{M}} = 2^X$, and in the notation of parts (c) and (e), $\tilde{\mu} \neq \underline{\mu}$.

SOLUTION. (a) Assume that μ is σ-finite, and let $E \subseteq X$ be locally measurable. Let $(A_n) \subseteq \mathcal{M}$ be such that $X = \bigcup_{n \in \mathbb{N}} A_n$ and $\mu(A_n) < \infty$. Then $E \cap A_n \in \mathcal{M}$, and so $E = \bigcup_{n \in \mathbb{N}} (E \cap A_n) \in \mathcal{M}$.

(b) Clearly we have $X \in \widetilde{\mathcal{M}}$. Then let $(E_n) \subseteq \widetilde{\mathcal{M}}$, and let $A \in \mathcal{M}$ with $\mu(A) < \infty$. Then

$$A \cap \bigcup_{n \in \mathbb{N}} E_n = \bigcup_{n \in \mathbb{N}} (A \cap E_n) \in \mathcal{M},$$

so $\bigcup_{n \in \mathbb{N}} E_n \in \widetilde{\mathcal{M}}$. Finally let $E \in \widetilde{\mathcal{M}}$ and $A \in \mathcal{M}$ with $\mu(A) < \infty$. Then

$$E^c \cap A = A \setminus E = A \setminus (E \cap A) = (E \cap A)^c \cap A \in \mathcal{M}$$

since $E \cap A \in \mathcal{M}$, so $E^c \in \widetilde{\mathcal{M}}$.

(c) We first show that $\tilde{\mu}$ is a measure. Clearly $\tilde{\mu}(\emptyset) = 0$, so let (E_n) be a sequence of disjoint sets in $\widetilde{\mathcal{M}}$, and let $E = \bigcup_{n\in\mathbb{N}} E_n$. Say that E_m does not lie in $\mathcal{M}$ for some $m \in \mathbb{N}$. Then we must have $\tilde{\mu}(E) = \infty$, since otherwise $E \in \mathcal{M}$ with $\mu(E) < \infty$, and hence $E_m = E_m \cap E \in \mathcal{M}$. Thus we have

$$\sum_{n=1}^{\infty} \tilde{\mu}(E_n) \geq \tilde{\mu}(E_m) = \infty = \tilde{\mu}(E),$$

so $\sum_{n=1}^{\infty} \tilde{\mu}(E_n) = \tilde{\mu}(E)$. The same is obviously true if all E_n lie in $\mathcal{M}$.

Next we show that $\tilde{\mu}$ is saturated, i.e. that $\widetilde{\widetilde{\mathcal{M}}} \subseteq \widetilde{\mathcal{M}}$, so let $E \in \widetilde{\widetilde{\mathcal{M}}}$. For all $A \in \widetilde{\mathcal{M}}$ with $\tilde{\mu}(A) < \infty$ we then have $E \cap A \in \widetilde{\mathcal{M}}$. By definition of $\tilde{\mu}$ we must have $A \in \mathcal{M}$, so we also have

$$E \cap A = (E \cap A) \cap A \in \mathcal{M}.$$

And since this is true for all $A \in \mathcal{M}$ with $\mu(A) < \infty$, it follows that $E \in \widetilde{\mathcal{M}}$.

In some sense, the fact that $\tilde{\mu}$ is saturated is obvious: The more sets of finite measure, the harder it is to be saturated, and vice-versa. On the other hand, the sets of infinite measure are irrelevant, so since the only new sets in $\widetilde{\mathcal{M}}$ have infinite measure, they cannot affect whether the measure is saturated or not.

(d) Assume that μ is complete. Let $F \subseteq X$ be such that there is a set $E \in \widetilde{\mathcal{M}}$ with $F \subseteq E$ and $\tilde{\mu}(E) = 0$. Then also $E \in \mathcal{M}$, and since μ is complete we have $F \in \mathcal{M} \subseteq \widetilde{\mathcal{M}}$ as desired. Or more succinctly: Saturating a measure only introduces sets of infinite measure, so it does not introduce any null-sets.

(e) Assume that μ is semifinite. We first show that $\underline{\mu}$ is a measure. Clearly $\underline{\mu}(\emptyset) = 0$, so let $(E_n) \subseteq \widetilde{\mathcal{M}}$ be a sequence of disjoint sets. Clearly $\underline{\mu}$ is increasing, so sigma-additivity is obvious if any of the sets E_n have infinite measure. Assume then that $\underline{\mu}(E_n) < \infty$ for all $n \in \mathbb{N}$. Let $\varepsilon > 0$, and choose $A_n \in \mathcal{M}$ such that $A_n \subseteq E_n$ and $\underline{\mu}(E_n) \leq \mu(A_n) + \varepsilon/2^n$. Then

$$\underline{\mu}\Big(\bigcup_{n\in\mathbb{N}} E_n\Big) \geq \mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n) \geq \sum_{n=1}^{\infty} \mu(E_n) - \varepsilon.$$

Since this holds for all $\varepsilon > 0$, we obtain the first inequality. For the other inequality, let $E = \bigcup_{n\in\mathbb{N}} E_n$, and first assume that $\underline{\mu}(E) = \infty$. Pick $A \in \mathcal{M}$ with $A \subseteq E$. Since μ is semifinite, we can choose A such that $C < \mu(A) < \infty$ for any given $C > 0$. Letting $A_n = A \cap E_n \in \mathcal{M}$ we get

$$C < \mu(A) = \sum_{n=1}^{\infty} \mu(A_n) \leq \sum_{n=1}^{\infty} \underline{\mu}(E_n),$$

and since C is arbitrary, we get $\sum_{n=1}^{\infty}\underline{\mu}(E_n) = \infty$. If instead $\underline{\mu}(E) < \infty$, pick $A \subseteq E$ with $A \in \mathcal{M}$ and $\underline{\mu}(E) \le \mu(A) + \varepsilon$. Again letting $A_n = A \cap E_n$ we get

$$\underline{\mu}(E) - \varepsilon \le \mu(A) = \sum_{n=1}^{\infty} \mu(A_n) \le \sum_{n=1}^{\infty} \underline{\mu}(E_n).$$

And since ε is arbitrary, we obtain the other inequality.

Next we show that $\underline{\mu}$ is saturated. Letting E be locally $\underline{\mu}$-measurable, we must show that E is also locally μ-measurable. So let $A \in \mathcal{M}$ with $\mu(A) < \infty$. Then $\underline{\mu}(A) < \infty$, and so $E \cap A \in \widetilde{\mathcal{M}}$. But then

$$E \cap A = (E \cap A) \cap A \in \mathcal{M},$$

as desired.

(f) It is pretty obvious that μ is a measure on $\mathcal{M}$. Then let $E \subseteq X$ and $A \in \mathcal{M}$ with $\mu(A) < \infty$. Then $A \cap X_1$ must be finite, and so A is not co-countable. But then it is countable, and so is $E \cap A$, hence $E \cap A \in \mathcal{M}$. Thus every subset of X is locally measurable.

Notice that μ is semifinite. We have $\tilde{\mu}(X_2) = \infty$ since $X_2 \notin \mathcal{M}$, but $\underline{\mu}(X_2) = 0$ since every subset of X_2 is disjoint from X_1, and so is has measure zero. □

1.4. Outer Measures

EXERCISE 1.18

Let $\mathcal{A} \subseteq 2^X$ be an algebra, $\mathcal{A}_\sigma$ the collection of countable unions of sets in $\mathcal{A}$, and $\mathcal{A}_{\sigma\delta}$ the collection of countable intersections of sets in $\mathcal{A}_\sigma$. Let μ_0 be a premeasure on $\mathcal{A}$ and μ^* the induced outer measure.

(a) For any $E \subseteq X$ and $\varepsilon > 0$ there exists $A \in \mathcal{A}_\sigma$ with $E \subseteq A$ with $\mu^*(A) \le \mu^*(E) + \varepsilon$.

(b) If $\mu^*(E) < \infty$, then E is μ^*-measurable iff there exists $B \in \mathcal{A}_{\sigma\delta}$ with $E \subseteq B$ and $\mu^*(B \setminus E) = 0$.

(c) If μ_0 is σ-finite, the restriction $\mu^*(E) < \infty$ in (b) is superfluous.

SOLUTION. (a) Let $E \subseteq X$ and $\varepsilon > 0$. The definition of μ^* yields a sequence $(A_n) \subseteq \mathcal{A}$ such that $E \subseteq \bigcup_{n\in\mathbb{N}} A_n$ and $\sum_{n=1}^{\infty} \mu_0(A_n) \le \mu^*(E) + \varepsilon$. It follows that

$$\mu^*(E) + \varepsilon \ge \sum_{n=1}^{\infty} \mu_0(A_n) = \sum_{n=1}^{\infty} \mu^*(A_n) \ge \mu^*\Big(\bigcup_{n\in\mathbb{N}} A_n\Big).$$

(b) Let $E \subseteq X$. For $n \in \mathbb{N}$ there is a set $B_n \in \mathcal{A}_\sigma$ such that $E \subseteq B_n$ and $\mu^*(B_n) \leq \mu^*(E) + 1/n$. Letting $B = \bigcap_{n\in\mathbb{N}} B_n \in \mathcal{A}_{\sigma\delta}$ we get $\mu^*(B) \leq \mu^*(E)$, and since $E \subseteq B$ we also have the opposite inequality, so $\mu^*(B) = \mu^*(E)$.

Now assume that $\mu^*(E) < \infty$ and that E is μ^*-measurable. Then

$$\mu^*(B) = \mu^*(B \cap E) + \mu^*(B \cap E^c) = \mu^*(E) + \mu^*(B \setminus E),$$

from which it follows that $\mu^*(B \setminus E) = 0$.

Conversely, assume that there is a $B \in \mathcal{A}_{\sigma\delta}$ with $E \subseteq B$ and $\mu^*(B \setminus E) = 0$. Then B lies in the σ-algebra generated by $\mathcal{A}$, so it is μ^*-measurable. Let $A \subseteq X$. Then

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

and so

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) \leq \mu^*(A \cap B) + \mu^*(A \cap B^c) = \mu^*(A),$$

showing that E is μ^*-measurable. (Notice that we haven't used that $\mu^*(E) < \infty$ for the second implication.)

(c) We only need to prove the first implication above. By σ-finiteness of μ_0, let (E_n) be a sequence of subsets of X such that $\mu^*(E_n) < \infty$ and $E = \bigcup_{n\in\mathbb{N}} E_n$. Let $\varepsilon > 0$. Then there are sets $A_n \in \mathcal{A}_\sigma$ such that $\mu^*(A_n) \leq \mu^*(E_n) + \varepsilon/2^n$. Letting $B_\varepsilon = \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}_\sigma$ we get

$$\mu^*(B_\varepsilon \setminus E) = \mu^*\Big(\bigcup_{n\in\mathbb{N}} (A_n \cap E^c)\Big) \leq \mu^*\Big(\bigcup_{n\in\mathbb{N}} (A_n \cap E_n^c)\Big) \leq \sum_{n=1}^{\infty} \mu^*(A_n \setminus E_n) \leq \varepsilon.$$

Finally we let $B = \bigcap_{k\in\mathbb{N}} B_{1/k} \in \mathcal{A}_{\sigma\delta}$, and we get $\mu^*(B \setminus E) = 0$ as desired. □

Remark 1.1. Notice that (b) and (c) in particular show that any Lebesgue measurable set E, and therefore any Borel set, is the intersection of a G_δ set B and a Lebesgue null set $B \setminus E$. ⌟

Exercise 1.20

Let μ^* be an outer measure on X, $\mathcal{M}^*$ the σ-algebra of μ^*-measurable sets, $\overline{\mu} = \mu^*|_{\mathcal{M}^*}$, and μ^+ the outer measure induced by $\overline{\mu}$ as in (1.12) (with $\overline{\mu}$ and $\mathcal{M}^*$ replacing μ_0 and $\mathcal{A}$).

(a) If $E \subseteq X$, we have $\mu^*(E) \leq \mu^+(E)$, with equality iff there exists $A \in \mathcal{M}^*$ with $A \supseteq E$ and $\mu^*(A) = \mu^*(E)$.

(b) If μ^* is induced from a premeasure, then $\mu^* = \mu^+$.

(c) If $X = \{0,1\}$, there exists an outer measure μ^* on X such that $\mu^* \neq \mu^+$.

SOLUTION. (a) Recall that the definition of μ^+ means that

$$\mu^+(E) = \inf\left\{ \sum_{n=1}^{\infty} \overline{\mu}(A_n) \,\middle|\, A_n \in \mathcal{M}^*, E \subseteq \bigcup_{n\in\mathbb{N}} A_n \right\},$$

and that we by definition of $\overline{\mu}$ can replace $\overline{\mu}$ with μ^*. For any such sequence (A_n) we have

$$\mu^*(E) \leq \mu^*\left(\bigcup_{n\in\mathbb{N}} A_n \right) \leq \sum_{n=1}^{\infty} \mu^*(A_n) = \sum_{n=1}^{\infty} \overline{\mu}(A_n).$$

And since $\mu^+(E)$ is the infimum of all such sums, we have $\mu^*(E) \leq \mu^+(E)$.

Next assume that there is an $A \in \mathcal{M}^*$ with $E \subseteq A$ such that $\mu^*(A) = \mu^*(E)$. Using the sequence $A_1 = A$ and $A_n = \emptyset$ for $n > 1$ in the definition of μ^+ yields

$$\mu^+(E) \leq \overline{\mu}(A) = \mu^*(A) = \mu^*(E).$$

Hence $\mu^+(E) = \mu^*(E)$ as desired.

Conversely, assuming that $\mu^*(E) = \mu^+(E)$ we have

$$\mu^*(E) = \inf\left\{ \sum_{n=1}^{\infty} \mu^*(A_n) \,\middle|\, A_n \in \mathcal{M}^*, E \subseteq \bigcup_{n\in\mathbb{N}} A_n \right\}.$$

Given $\varepsilon > 0$, choose a sequence (A_n) such that

$$\mu^*\left(\bigcup_{n\in\mathbb{N}} A_n \right) \leq \sum_{n=1}^{\infty} \mu^*(A_n) \leq \mu^*(E) + \varepsilon,$$

and let $B_\varepsilon = \bigcup_{n\in\mathbb{N}} A_n$. Letting $A = \bigcap_{k\in\mathbb{N}} B_{1/k} \in \mathcal{M}^*$ we thus have $\mu^*(A) \leq \mu^*(E)$.

(b) Assume that μ^* is induced from a premeasure on an algebra $\mathcal{A}$, and let $E \subseteq X$. Recall that $\mathcal{A}$ consists of μ^*-measurable sets, so $\sigma(\mathcal{A}) \subseteq \mathcal{M}^*$. For $n \in \mathbb{N}$ choose, in accordance with Exercise 1.18(a), a set $A_n \in \mathcal{A}_\sigma$ with $E \subseteq A_n$ such that $\mu^*(A_n) \leq \mu^*(E) + 1/n$. Letting $A = \bigcap_{n\in\mathbb{N}} A_n$ we have $E \subseteq A$ and $\mu^*(A) \leq \mu^*(E)$. The other inequality is obvious, so $\mu^*(A) = \mu^*(E)$, and part (a) implies that $\mu^*(E) = \mu^+(E)$ as desired. □

EXERCISE 1.21

Let μ^* be an outer measure induced from a premeasure and $\overline{\mu}$ the restriction of μ^* to the μ^*-measurable sets. Then $\overline{\mu}$ is saturated.

SOLUTION. Let $\mathcal{A}$ denote the algebra on which the premeasure in question is defined, and denote by $\mathcal{M}^*$ the σ-algebra of μ^*-measurable sets. Recall that $\mathcal{A} \subseteq \mathcal{M}^*$.

Let $E \subseteq X$ be locally measurable. It suffices to show that

$$\mu^*(F) \geq \mu^*(F \cap E) + \mu^*(F \cap E^c)$$

for all $F \subseteq X$ with $\mu^*(F) < \infty$. Given $\varepsilon > 0$, Exercise 1.18(a) yields a set $A \in \mathcal{A}_\sigma$ such that $\mu^*(A) \leq \mu^*(F) + \varepsilon$. Then $\mu^*(A) < \infty$, and so $E \cap A \in \mathcal{M}^*$. It follows that

$$\begin{aligned} \mu^*(F) + \varepsilon \geq \mu^*(A) &= \mu^*(A \cap (E \cap A)) + \mu^*(A \cap (E \cap A)^c) \\ &= \mu^*(A \cap E) + \mu^*(A \cap E^c) \\ &\geq \mu^*(F \cap E) + \mu^*(F \cap E^c), \end{aligned}$$

and hence $E \in \mathcal{M}^*$. Thus $\overline{\mu}$ is saturated. □

EXERCISE 1.22

Let $(X, \mathcal{M}, \mu)$ be a measure space, μ^* the outer measure induced by μ according to (1.12), $\mathcal{M}^*$ the σ-algebra of μ^*-measurable sets, and $\overline{\mu} = \mu^*|_{\mathcal{M}^*}$.

(a) If μ is σ-finite, then $\overline{\mu}$ is the completion of μ.

(b) In general, $\overline{\mu}$ is the saturation of the completion of μ.

SOLUTION. (a) Let $\overline{\mathcal{M}}$ be the σ-algebra from Theorem 1.9 (namely, the σ-algebra generated by the sets in $\mathcal{M}$ along with all μ-null sets). This is clearly the smallest σ-algebra on which there can exist a complete extension of μ, so since $\overline{\mu}$ is also a complete extension of μ, we must have $\overline{\mathcal{M}} \subseteq \mathcal{M}^*$. Theorem 1.9 yields the uniqueness of a complete extension of μ on $\overline{\mathcal{M}}$, so it suffices to show that $\mathcal{M}^* \subseteq \overline{\mathcal{M}}$.

Now assume that μ is σ-finite, and let $E \in \mathcal{M}^*$. Then also $E^c \in \mathcal{M}^*$, and Exercise 1.18(c) ensures the existence of sets $B, D \in \mathcal{M}_{\sigma\delta} = \mathcal{M}$ with $E \subseteq B$ and $E^c \subseteq D$ such that

$$\mu^*(B \setminus E) = 0 \quad \text{and} \quad \mu^*(E \setminus D^c) = \mu^*(D \setminus E^c) = 0.$$

It follows that

$$\mu(B \setminus D^c) \leq \mu^*(B \setminus E) + \mu^*(E \setminus D^c) = 0,$$

so $E \setminus D^c$ is a μ-null set. Thus $E = D^c \cup (E \setminus D^c)$ is a union of a set in $\mathcal{M}$ and a μ-null set, and hence $E \in \overline{\mathcal{M}}$.

(b) Let $\hat{\mu}$ denote the completion of μ on $\overline{\mathcal{M}}$, and let $\widetilde{\mathcal{M}}$ denote the σ-algebra of locally $\hat{\mu}$-measurable sets. First we show that $\widetilde{\mathcal{M}} = \mathcal{M}^*$, so let $E \in \widetilde{\mathcal{M}}$. To show that E is μ^*-measurable it suffices to show that

$$\mu^*(F) \geq \mu^*(F \cap E) + \mu^*(F \cap E^c)$$

for all $F \subseteq X$ with $\mu^*(F) < \infty$. Calculations identical to the ones in the solution to Exercise 1.21 show this.

Conversely, let $E \in \mathcal{M}^*$ and consider $A \in \overline{\mathcal{M}}$ with $\hat{\mu}(A) < \infty$. Then also $A \in \mathcal{M}^*$, so $E \cap A \in \mathcal{M}^*$. The argument at the beginning of part (a) showed that $\overline{\mu}$ is an extension of $\hat{\mu}$, so $\mu^*(E \cap A) = \hat{\mu}(E \cap A) < \infty$. The same argument as in part (a), only now using Exercise 1.18(b) instead of (c), shows that $E \cap A \in \overline{\mathcal{M}}$, and so $E \in \widetilde{\mathcal{M}}$.

Finally, let $\tilde{\mu}$ denote the saturation of $\hat{\mu}$. We show that $\overline{\mu} = \tilde{\mu}$. Since the completion of μ on $\overline{\mathcal{M}}$ is unique, the two measures must agree here. Instead let $E \in \widetilde{\mathcal{M}} \setminus \overline{\mathcal{M}}$. By definition of $\tilde{\mu}$ we must then have $\tilde{\mu}(E) = \infty$. On the other hand, we just showed (for $E \cap A$ instead of E) that $\mu^*(E) < \infty$ implies $E \in \overline{\mathcal{M}}$. Since we have assumed that this is not the case, we must have $\overline{\mu}(E) = \mu^*(E) = \infty$. Thus $\overline{\mu} = \tilde{\mu}$. □

1.5. Borel Measures on the Real Line

EXERCISE 1.25

If $E \subseteq \mathbb{R}$, the following are equivalent.

(a) $E \in \mathcal{M}_\mu$.

(b) $E = V \setminus N_1$ where V is a G_δ set and $\mu(N_1) = 0$.

(c) $E = H \cup N_2$ where H is an F_σ set and $\mu(N_2) = 0$.

SOLUTION. Folland proves this claim when $\mu(E) < \infty$, so assume that $\mu(E) = \infty$. Since μ is σ-finite, there is a sequence $(E_n)_{n\in\mathbb{N}}$ in $\mathcal{M}_\mu$ with $\mu(E_n) < \infty$ for all $n \in \mathbb{N}$ and $E = \bigcup_{n\in\mathbb{N}} E_n$. Then there are sequences (H_n) of F_σ sets and (N_n) of null sets such that $E_n = H_n \cup N_n$. Then $H = \bigcup_{n\in\mathbb{N}} H_n$ is also an F_σ set and $N = \bigcup_{n\in\mathbb{N}} N_n$ a null set, and $E = H \cup N$.

Applying this to E^c yields a similar decomposition $E^c = H \cup N$. But then $E = H^c \setminus N$, and H^c is a G_δ set. □

2 • Integration

2.1. Measurable Functions

EXERCISE 2.10

The following implications are valid iff the measure μ is complete:

(a) If f is measurable and $f = g$ μ-a.e., then g is measurable.

(b) If f_n is measurable for $n \in \mathbb{N}$ and $f_n \to f$ μ-a.e., then f is measurable.

SOLUTION. (a) Let $f, g \colon (X, \mathcal{E}, \mu) \to (Y, \mathcal{F})$ be functions from a measure space to a measurable space where f is $(\mathcal{E}, \mathcal{F})$-measurable. Let $N = \{f \neq g\}$ and assume that $\mu(N) = 0$. Given $B \in \mathcal{F}$ we must show that $g^{-1}(B) \in \mathcal{E}$. But notice that

$$g^{-1}(B) = f^{-1}(B) \cup \{f \notin B, g \in B\} \setminus \{f \in B, g \notin B\},$$

and that the latter two sets are subsets of N, hence measurable. Thus $g^{-1}(B)$ is also measurable.

Conversely, let μ be a measure on a measurable space $(X, \mathcal{E})$ that is not complete, and let $N \subseteq X$ be a non-measurable μ-null set. Then $\mathbf{1}_N = 0$ μ-a.e., but $\mathbf{1}_N$ is not measurable.

(b) Consider the set A of points $x \in X$ such that $f_n(x)$ does not converge to $f(x)$. Then $f_n \mathbf{1}_{A^c} \to f \mathbf{1}_{A^c}$ pointwise everywhere, so Proposition 2.7 (or Corollary 2.9) implies that $f \mathbf{1}_{A^c}$ is measurable. By assumption $\mu(A) = 0$, so $f \mathbf{1}_{A^c} = f$ μ-a.e. and part (a) implies that f is measurable.

Conversely [TODO] □

2.7. Integration in Polar Coordinates

REMARK 2.1. We give a heuristic derivation of the radial measure ρ_d. Let $\mathrm{d}A$ be an area element in $\mathbb{R}^2$. In polar coordinates (r, θ) this has a radial size of $\mathrm{d}r$ and an angular size of $r\,\mathrm{d}\theta$. Notice that since θ is an angle, we multiply it by the distance r from the origin. Hence

$$\mathrm{d}A = r\,\mathrm{d}\theta\,\mathrm{d}r = (r\,\mathrm{d}r)\,\mathrm{d}\theta.$$

Going up one dimension we introduce another angular coordinate φ, which contributes a factor $f(\theta, \varphi) r\,\mathrm{d}\varphi$ to the volume element, where f is some function of the angular coordinates. Similarly when going up yet another dimension: this again introduces a factor r, and now f is a function of yet another angular coordinate. In d dimensions we have $d-1$ angular coordinates $\theta_1, \ldots, \theta_{d-1}$, so the volume element is on the form

$$\mathrm{d}V = f(\theta_1, \ldots, \theta_{d-1}) r^{d-1}\,\mathrm{d}r\,\mathrm{d}\theta_1 \cdots \mathrm{d}\theta_{d-1}.$$

The radial part is thus $r^{d-1}\,\mathrm{d}r$, so it makes sense to define the radial measure ρ_d on $(0,\infty)$ by
$$\rho_d(E) = \int_E r^{d-1}\,\mathrm{d}r.$$
⌟

3 • Signed Measures and Differentiation

3.1. Signed Measures

Exercise 3.2

If ν is a signed measure, E is ν-null iff $|\nu|(E) = 0$. Also, if ν and μ are signed measures, $\nu \perp \mu$ iff $|\nu| \perp \mu$ iff $\nu^+ \perp \mu$ and $\nu^- \perp \mu$.

Solution. Assume that E is ν-null, and let $P \cup N$ be a Hahn decomposition for ν. Then
$$\nu^+(E) = \nu(E \cap P) = 0,$$
since $E \cap P \subseteq E$. Similarly we get $\nu^-(E) = 0$, so $|\nu|(E) = 0$. Conversely, assume that $|\nu|(E) = 0$. Then $\nu^\pm(F) = 0$ for all measurable $F \subseteq E$, and so $\nu(F) = 0$.

The other claims follow directly from the above. □

Exercise 3.3

Let ν be a signed measure on $(X, \mathcal{M})$.

(a) $L^1(\nu) = L^1(|\nu|)$.

(b) If $f \in L^1(\nu)$,
$$\left| \int f\,\mathrm{d}\nu \right| \leq \int |f|\,\mathrm{d}|\nu|.$$

(c) If $E \in \mathcal{M}$,
$$|\nu|(E) = \sup\left\{ \left| \int_E f\,\mathrm{d}\nu \right| \,\middle|\, |f| \leq 1 \right\}.$$

Solution. (a) This follows directly from the definition of $L^1(\nu)$.

(b) For $f \in L^1(\nu)$ we have
$$\left| \int f\,\mathrm{d}\nu \right| = \left| \int f\,\mathrm{d}\nu^+ - \int f\,\mathrm{d}\nu^- \right| \leq \int |f|\,\mathrm{d}\nu^+ + \int |f|\,\mathrm{d}\nu^- = \int |f|\,\mathrm{d}|\nu|,$$
since $|\nu| = \nu^+ + \nu^-$.

(c) If $|f| \le 1$, then

$$\left|\int_E f \,\mathrm{d}\nu\right| \le \int_E |f| \,\mathrm{d}|\nu| \le |\nu|(E),$$

showing one inequality. For the other inequality, let $P \cup N$ be a Hahn decomposition for ν, and let $f = \mathbf{1}_P - \mathbf{1}_N$. Then

$$\begin{aligned}\int_E f \,\mathrm{d}\nu &= \int_E (\mathbf{1}_P - \mathbf{1}_N) \,\mathrm{d}\nu^+ - \int_E (\mathbf{1}_P - \mathbf{1}_N) \,\mathrm{d}\nu^- \\ &= \nu^+(E \cap P) - \nu^+(E \cap N) - \nu^-(E \cap P) + \nu^-(E \cap N) \\ &= \nu^+(E) + \nu^-(E) = |\nu|(E). \qquad \square\end{aligned}$$

Exercise 3.4

If ν is a signed measure and λ, μ are positive measures such that $\nu = \lambda - \mu$, then $\lambda \ge \nu^+$ and $\mu \ge \nu^-$.

Solution. Let $P \cup N$ be a Hahn decomposition for ν. Then

$$\nu^+(E) = \nu(E \cap P) = \lambda(E \cap P) - \mu(E \cap P) \le \lambda(E \cap P) \le \lambda(E),$$

and similarly for ν^-. □

Exercise 3.5

If ν_1, ν_2 are signed measures that both omit the value ∞ or $-\infty$, then $|\nu_1 + \nu_2| \le |\nu_1| + |\nu_2|$.

Solution. First notice that

$$\nu_1 + \nu_2 = (\nu_1^+ + \nu_2^+) - (\nu_1^- + \nu_2^-),$$

so by the previous exercise we have

$$|\nu_1 + \nu_2| = (\nu_1 + \nu_2)^+ + (\nu_1 + \nu_2)^+ \le (\nu_1^+ + \nu_2^+) + (\nu_1^- + \nu_2^-) = |\nu_1| + |\nu_2|. \qquad \square$$

Exercise 3.7

Suppose that ν is a signed measure on $(X, \mathcal{M})$ and $E \in \mathcal{M}$.

(a) $\nu^+(E) = \sup\{\nu(F) \mid F \in \mathcal{M}, F \subseteq E\}$ and $\nu^-(E) = -\inf\{\nu(F) \mid F \in \mathcal{M}, F \subseteq E\}$.

(b) We have

$$|\nu|(E) = \sup\left\{\sum_{i=1}^n |\nu(E_i)| \,\middle|\, n \in \mathbb{N}, E_1, \ldots, E_n \text{ disjoint}, \bigcup_{i=1}^n E_i = E\right\}.$$

SOLUTION. (a) We prove the first identity, the second is proved similarly. Denote the supremum on the right-hand side by $\mu(E)$, and let $P \cup N$ be a Hahn decomposition for ν. Since $E \cap P \subseteq E$ we have

$$\nu^+(E) = \nu(E \cap P) \leq \mu(E).$$

Furthermore, for $F \in \mathcal{M}$ with $F \subseteq E$ notice that

$$\nu(F) = \nu^+(F) - \nu^-(F) \leq \nu^+(F) \leq \nu^+(E),$$

showing that $\mu(E) \leq \nu^+(E)$.

(b) Denote the quantity on the right-hand side by $\rho(E)$, and let $P \cup N$ be a Hahn decomposition for ν. The disjoint union $E = (E \cap P) \cup (E \cap N)$ yields

$$\rho(E) \geq |\nu(E \cap P)| + |\nu(E \cap N)| = \nu^+(E) + \nu^-(E) = |\nu|(E).$$

Conversely, let $E_1, \dots, E_n$ be disjoint sets in $\mathcal{M}$ such that $\bigcup_{i=1}^n E_i = E$. For $i = 1, \dots, n$ we have

$$|\nu(E_i)| = |\nu^+(E_i) - \nu^-(E_i)| \leq \nu^+(E_i) + \nu^-(E_i) = |\nu|(E_i),$$

implying that

$$\sum_{i=1}^n |\nu(E_i)| \leq \sum_{i=1}^n |\nu|(E_i) = |\nu|(E).$$

It follows that $\rho(E) \leq |\nu|(E)$. □

4 • Point Set Topology

4.5. *Locally Compact Hausdorff Spaces*

EXERCISE 4.49

Let X be a compact Hausdorff space and $E \subseteq X$.

(a) If E is open, then E is locally compact in the relative topology.

(b) b

(c) c

SOLUTION. (a) Since every point in X has a compact neighbourhood (namely X itself), X is locally compact. So if $x \in E$ and E is open, then Proposition 4.30 yields a compact neighbourhood $K \subseteq E$ of x. But then K is also a compact neighbourhood of x in E, showing that E is locally compact.

(b) b

(c) c □

REMARK 4.1. Let X be a compact Hausdorff space, and let $x_0 \in X$ be any point in X. Exercise 4.49(a) then shows that $X \setminus \{x_0\}$ is locally compact, so we may consider the one-point compactification $(X \setminus \{x_0\})^*$. We claim that $(X \setminus \{x_0\})^* \cong X$.

Denote the adjoined point by ∞ and consider the inclusion map $i\colon X \setminus \{x_0\} \to X$ extended to $(X \setminus \{x_0\})^*$ by letting $i(\infty) = x_0$. This is a bijection, and restricted to $X \setminus \{x_0\}$ it is a homeomorphism onto its image. We claim that i is itself a homeomorphism, and since both its domain and codomain are compact Hausdorff it suffices to show that it is continuous. So let $U \subseteq X$ be open. If $x_0 \notin U$ then $i^{-1}(U) = (i|_{X\setminus\{x_0\}})^{-1}(U)$, which is open in $X \setminus \{x_0\}$, hence open in $(X \setminus \{x_0\})^*$. Otherwise of $x_0 \in U$ then $\infty \in i^{-1}(U)$, and we need to show that $i^{-1}(U)^c = i^{-1}(U^c)$ is compact. But U^c is a closed, hence compact, subset of $X \setminus \{0\}$, so its preimage under the homeomorphism $i|_{X\setminus\{x_0\}}$ is also compact. ⌟

EXERCISE 1 4.52

The one-point compactification of $\mathbb{R}^n$ is homeomorphic to the sphere $\mathbb{S}^n$.

SOLUTION. Let $x_0 \in \mathbb{S}^n$ be any point on the sphere. By stereographic projection, $\mathbb{S}^n \setminus \{x_0\}$ and $\mathbb{R}^n$ are homeomorphic. But then Remark 4.1 shows that

$$(\mathbb{R}^n)^* \cong (\mathbb{S}^n \setminus \{x_0\})^* \cong \mathbb{S}$$

as desired. □

4.7. The Stone–Weierstrass Theorem

REMARK 4.2. Notice that we never use the Hausdorff assumption in the proof of the Stone–Weierstrass theorem. However, if X is a topological space and there exists a family $\mathcal{F}$ of functions in $C(X)$ or $C(X,\mathbb{R})$ that separates points in X, then X is automatically Hausdorff: For let $x \neq y$ be points in X, and let $f \in \mathcal{F}$ be such that $f(x) \neq f(y)$. Choosing disjoint neighbourhoods U_x and U_y of x and y respectively, $f^{-1}(U_x)$ and $f^{-1}(U_y)$ are disjoint neighbourhoods of x and y in X. Hence X is Hausdorff.

In other words, Hausdorff is not a necessary condition in the statement of the theorem, but rather follows from the other hypotheses.

In contrast, the compactness hypothesis is used very explicitly in the proof of Lemma 4.49. ⌟

EXERCISE 4.66

Let $1 - \sum_{n=1}^{\infty} c_n t^n$ be the Maclaurin series for $(1-t)^{1/2}$.

(a) The series converges absolutely and uniformly on compact subsets of $(-1,1)$, as does the termwise differentiated series $-\sum_{n=1}^{\infty} nc_n t^{n-1}$. Thus, if $f(t) = 1 - \sum_{n=1}^{\infty} c_n t^n$, then $f'(t) = -\sum_{n=1}^{\infty} nc_n t^{n-1}$.

(b) By explicit calculation, $f(t) = -2(1-t)f'(t)$, from which it follows that $(1-t)^{-1/2} f(t)$ is constant. Since $f(0) = 1$, $f(t) = (1-t)^{1/2}$.

SOLUTION. (a) We first compute the coefficients c_n. If $g(t) = (1-t)^{1/2}$, then we claim that

$$g^{(n)}(t) = -\frac{(2n-3)(2n-5)\cdots(3)(1)}{2^n}(1-t)^{-(2n-1)/2}$$

for $n \in \mathbb{N}$ and $t \in (-1,1)$. Indeed, this follows easily by induction. Hence

$$c_n = \frac{1}{n!} g^{(n)}(0) = -\frac{1}{n!} \frac{(2n-3)(2n-5)\cdots(3)(1)}{2^n}.$$

Now let $\rho \in (0,1)$. Then

$$\left| \frac{c_{n+1}\rho^{n+1}}{c_n \rho^n} \right| = \frac{n!}{(n+1)!} \frac{2n-1}{2} \rho = \frac{2n-1}{2n} \rho \xrightarrow[n\to\infty]{} \rho < 1.$$

The ratio test then implies that the series $\sum_{n=1}^{\infty} c_n \rho^n$ converges, so it follows from the Weierstrass M-test that the series $1 - \sum_{n=1}^{\infty} c_n t^n$ converges absolutely and uniformly on the interval $[-\rho, \rho]$, and hence on all compact subsets of $(-1,1)$. We similarly find that

$$\left| \frac{(n+1)c_{n+1}\rho^n}{nc_n \rho^{n-1}} \right| = \frac{n!}{(n+1)!} \frac{n+1}{n} \frac{2n-1}{2} \rho = \frac{n+1}{n} \frac{2n-1}{2n} \rho \xrightarrow[n\to\infty]{} \rho < 1,$$

so the series $-\sum_{n=1}^{\infty} nc_n t^{n-1}$ also converges as claimed.

(b) Notice that

$$\begin{aligned}
-2(1-t)f'(t) &= 2(1-t)\sum_{n=1}^{\infty} nc_n t^{n-1} = 2\sum_{n=1}^{\infty} nc_n t^{n-1} - 2\sum_{n=1}^{\infty} nc_n t^n \\
&= 2\sum_{n=0}^{\infty} (n+1)c_{n+1} t^n - 2\sum_{n=1}^{\infty} nc_n t^n \\
&= 2\sum_{n=0}^{\infty} \big((n+1)c_{n+1} - nc_n\big) t^n.
\end{aligned}$$

A short calculation shows that $(n+1)c_{n+1} - nc_n = c_n/2$, so the above equals $f(t)$ as claimed. Thus we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(1-t)^{-1/2}f(t) = (1-t)^{-1/2}f'(t) + \frac{1}{2}(1-t)^{-3/2}f(t) = 0,$$

showing that $(1-t)^{-1/2}f(t)$ is constant. But $f(0) = 1$, so it follows that $f(t) = (1-t)^{1/2} = g(t)$. □

5 • Elements of Functional Analysis

5.1. Normed Vector Spaces

REMARK 5.1. We give a slightly different proof of Proposition 5.2.

Clearly if $T\colon X \to Y$ is continuous, then it is continuous at 0. And if this is so, then there is a $\delta > 0$ such that $\|h\| < \delta$ implies $\|Th\| \leq 1$, for $h \in X$. For all $x \in X$ we thus have

$$\|Tx\| = \frac{\|x\|}{\delta}\left\|T\left(\delta\frac{x}{\|x\|}\right)\right\| \leq \delta^{-1}\|x\|,$$

so T is bounded.

We let

$$\|T\| = \sup\left\{\|Tx\| \;\middle|\; x \in X, \|x\| \leq 1\right\}$$

If T is bounded, then clearly $\|T\| < \infty$. If conversely $\|T\| < \infty$, then

$$\|Tx\| = \left\|T\frac{x}{\|x\|}\right\|\|x\| \leq \|T\|\,\|x\|$$

for all $x \neq 0$, so T is bounded. Furthermore, if $K > 0$ is such that $\|Tx\| \leq K\|x\|$ for all $x \in X$, then $\|Tx\| \leq K$ whenever $\|x\| \leq 1$. But then $\|T\| \leq K$. ⌟

EXERCISE 5.3

If Y is complete, so is $\mathcal{B}(X,Y)$.

SOLUTION. We prove the following lemma:

> *Let X and Y be normed spaces, and let $(T_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{B}(X,Y)$. If (T_n) is Cauchy in the operator norm and converges to some $T\colon X \to Y$ in the strong operator topology, then $T \in \mathcal{B}(X,Y)$ and $T_n \to T$ in the operator norm.*

The map T is clearly linear. Choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies that $\|T_n - T_m\| \leq \varepsilon$. For $x \in X$ and $n \geq N$ we then have

$$\|(T_n - T)x\| = \lim_{m\to\infty} \|(T_n - T_m)x\| \leq \lim_{m\to\infty} \|T_n - T_m\| \|x\| \leq \varepsilon \|x\|.$$

Hence $T_n - T$ is bounded, and then so is T. Furthermore, $\|T_n - T\| \leq \varepsilon$, so $T_n \to T$ in the operator norm.

To prove the initial claim it thus suffices to produce, given a Cauchy sequence $(T_n)_{n\in\mathbb{N}}$, a map $T\colon X \to Y$ such that $\operatorname{s-lim}_{n\to\infty} T_n = T$. But notice that we for $x \in X$ we have

$$\|T_n x - T_m x\| \leq \|T_n - T_m\| \|x\|$$

for $m, n \in \mathbb{N}$, so $(T_n x)_{n\in\mathbb{N}}$ is a Cauchy sequence in Y. Defining $T\colon X \to Y$ by $Tx = \lim_{n\to\infty} T_n x$, T is the strong limit of T_n by construction. □

EXERCISE 5.4

If X and Y are normed spaces, the map $(T, x) \mapsto Tx$ is continuous from $\mathcal{B}(X, Y) \times X$ to Y.

SOLUTION. If $T, S \in \mathcal{B}(X, Y)$ and $x, y \in X$, then

$$\|Tx - Sy\| \leq \|Tx - Ty\| + \|Ty - Sy\| \leq \|T\| \|x - y\| + \|T - S\| \|y\|.$$

The claim follows.

Notice that this proof is identical to the proof that multiplication in a Banach algebra is continuous, but the Banach inequality is replaced with the inequality $\|Tx\| \leq \|T\| \|x\|$. The proof is also almost identical to the proof that multiplication on $\mathbb{R}$ or $\mathbb{C}$ is continuous, except here we have the *equality* $|xy| = |x||y|$. □

EXERCISE 5.6

Suppose that X is a finite-dimensional vector space. Let $(e_1, \ldots, e_d)$ be a basis for X, and define $\|\sum_{i=1}^d a_i e_i\|_1 = \sum_{i=1}^d |a_i|$.

(a) $\|\cdot\|_1$ is a norm on X.

(b) The map $T\colon (a_1, \ldots, a_d) \mapsto \sum_{i=1}^d a_i e_i$ is continuous from $\mathbb{K}^d$ with the usual Euclidean topology to X with the topology defined by $\|\cdot\|_1$.

(c) The set $S = \{x \in X \mid \|x\|_1 = 1\}$ is compact in the topology defined by $\|\cdot\|_1$.

(d) All norms on X are equivalent.

SOLUTION. (a) This is obvious.

(b) If we equip $\mathbb{K}^d$ with the 1-norm, then T is an isometry and thus continuous (in fact a homeomorphism since it is surjective).

(c) Since the unit sphere in $\mathbb{K}^d$ (with respect to the 1-norm) is compact and T is continuous, S is also compact.

(d) If $\|\cdot\|$ is any norm on X, we need to find $C_1, C_2 > 0$ such that

$$C_1\|x\|_1 \le \|x\| \le C_2\|x\|_1 \tag{5.1}$$

for all $x \in X$. This is obvious for $x = 0$, and if $x \neq 0$ we may divide through by $\|x\|_1$. The claim is then that

$$C_1 \le \|x\| \le C_2$$

for all $x \in X$ with $\|x\|_1 = 1$, i.e. all $x \in S$. We first show that $\|\cdot\|$ is continuous with respect to $\|\cdot\|_1$. For $x = \sum_{i=1}^d a_i e_i$ and $y = \sum_{i=1}^d b_i e_i$ in X we have

$$\|x-y\| = \left\|\sum_{i=1}^d (a_i - b_i)e_i\right\| \le \sum_{i=1}^d |a_i - b_i|\|e_i\| \le \|x-y\|_1 \max_{1\le i\le d}\|e_i\|.$$

Continuity of $\|\cdot\|$ now follows from the reverse triangle inequality. (In fact, this calculation also proves the second inequality of (5.1), but we give a second argument below.)

Since $\|\cdot\|$ is continuous and S is compact with respect to $\|\cdot\|_1$, there exist $x_0, x_1 \in S$ such that

$$\|x_0\|_1 \le \|x\| \le \|x_1\|_1$$

for all $x \in S$. And since both of x_0 and x_1 are nonzero then so are their norms, proving the claim. □

EXERCISE 5.9

Let $C^k([0,1])$ be space of functions on $[0,1]$ possessing continuous derivatives up to order k on $[0,1]$, including onesided derivatives at the endpoints.

(a) If $f \in C([0,1])$, then $f \in C^k([0,1])$ iff f is k times continuously differentiable on $(0,1)$ and $f^{(j)}(0+) = \lim_{x\downarrow 0} f^{(j)}(x)$ and $f^{(j)}(1-) = \lim_{x\uparrow 1} f^{(j)}(x)$ exist for $j \le k$.

(b) $\|f\| = \sum_{j=0}^k \|f^{(j)}\|_\infty$ is a norm on $C^k([0,1])$ that makes $C^k([0,1])$ into a Banach space.

Solution. (a) The 'only if' part is obvious. Conversely, we show by induction in j that $f \in C^j([0,1])$ for $j = 0,\dots,k$. This is true for $j = 0$ by assumption, so assume that it is true for some j. For $x \in (0,1)$ there is a $\xi \in (0,x)$ such that $f^{(j)}(x) - f^{(j)}(0) = f^{(j+1)}(\xi)(x-0)$. It follows that

$$\frac{f^{(j)}(x) - f^{(j)}(0)}{x-0} = f^{(j+1)}(\xi) \xrightarrow[x\downarrow 0]{} f^{(j+1)}(0+).$$

Thus $f^{(j)}$ has a one-sided derivative at 0, and since the derivative is precisely the limit $f^{(j+1)}(0+)$, this also shows that $f^{(j+1)}$ is continuous at 0. Similarly at 1, so $f \in C^{j+1}([0,1])$ as desired.

(b) Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $C^1([0,1])$ converging to a function f, such that the sequence (f_n') converges uniformly in $C([0,1])$ to a function g. Let $\varepsilon > 0$, and choose $N \in \mathbb{N}$ such that $n \geq N$ implies that $\|f_n' - g\|_\infty < \varepsilon$. For $n \geq N$ and fixed $x \in [0,1]$ we then have

$$\left| \int_0^x f_n'(t)\,\mathrm{d}t - \int_0^x g(t)\,\mathrm{d}t \right| \leq \int_0^x |f_n'(t) - g(t)|\,\mathrm{d}t \leq \varepsilon x.$$

It follows that

$$f(x) - f(0) = \lim_{n\to\infty}\bigl(f_n(x) - f_n(0)\bigr) = \lim_{n\to\infty} \int_0^x f'(t)\,\mathrm{d}t = \int_0^x g(t)\,\mathrm{d}t.$$

Thus we see that $f \in C^1([0,1])$ with $f' = g$.

Now let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $C^k([0,1])$. Then the sequences $(f_n^{(j)})$ are Cauchy sequences in $C([0,1])$ for $j = 0,\dots,k$, and so the sequences have uniform limits. But then we are in the situation above, so it follows by induction that $f_n^{(j)} \to f^{(j)}$ uniformly for all j. Hence $f_n \to f$ in $C^k([0,1])$, so this is a Banach space. □

Remark 5.2. As an application of the above we consider the following: Let $D\colon C^k([0,1]) \to C^{k-1}([0,1])$ be the differential operator $f \mapsto f'$. We claim that this is bounded with respect to the above norm. For $f \in C^k([0,1])$ we have

$$\|Df\| = \sum_{j=0}^{k-1} \|(Df)^{(j)}\|_\infty = \sum_{j=0}^{k-1} \|f^{(j+1)}\|_\infty = \sum_{j=1}^{k} \|f^{(j)}\|_\infty \leq \|f\|.$$

The usual counterexamples to the boundedness of D on e.g. $(C^1([0,1]), \|\cdot\|_\infty)$ do not work here. The norm $\|\cdot\|$ in effect takes into account the fact that functions that take on similar values may have derivatives that vary wildly. ⌟

REMARK 5.3: Riesz' lemma.
The statement of the lemma is as follows:

> *Let X be a normed vector space and M a proper closed subspace of X. For $\alpha \in (0,1)$ there exists an $x \in X$ with $\|x\| = 1$ such that*
> $$\inf_{m \in M} \|x - m\| \geq \alpha.$$

Since the quotient norm on X/M is given by $\|x + M\| = \inf_{m \in M} \|x - m\|$, this is precisely the statement of Exercise 5.12(b) [TODO: reference].

In Exercise 5.19(b) [TODO: reference] we use this to show that an infinite-dimensional normed vector space is not locally compact. It is easy to show that this is equivalent to the closed unit ball $\bar{B}_1(0)$ being compact.

Conversely, every normed space $(X, \|\cdot\|)$ of dimension $d < \infty$ is locally compact: Choose a linear isomorphism $T\colon \mathbb{C}^d \to X$ and let it induce a norm $\|\cdot\|'$ on X. With this norm T is an isometry, hence a homeomorphism, so the local compactness of $\mathbb{C}^d$ is transferred to $(X, \|\cdot\|')$. But all norms on finite-dimensional vector spaces are equivalent, so $(X, \|\cdot\|)$ is also locally compact.

This equivalence of local compactness and finite-dimensionality generalises to Hausdorff topological vector spaces. This is known as F. Riesz' theorem. ⌟

EXERCISE 5.12

Let X be a normed vector space and M a proper closed subspace of X.

(a) a

(b) For any $\varepsilon > 0$ there exists $x \in X$ such that $\|x\| = 1$ and $\|x + M\| \geq 1 - \varepsilon$.

(c) The projection map $\pi\colon X \to X/M$ has norm 1.

(d) d

(e) e

SOLUTION. (a) a

(b) Let $\varepsilon > 0$, and pick some $y \in X \setminus M$. By definition of the quotient norm there exists an $m \in M$ such that
$$\frac{\|y + M\|}{\|y - m\|} \geq 1 - \varepsilon.$$
Letting $x = (y - m)/\|y - m\|$ we have $\|x\| = 1$ and
$$\|x + M\| = \left\| \frac{y - m}{\|y - m\|} + M \right\| = \frac{\|y + M\|}{\|y - m\|} \geq 1 - \varepsilon$$

as desired.

(c) For any $x \in X$ we have $\|x + M\| \leq \|x + 0\|$, so $\|\pi\| \leq 1$. But given $\varepsilon > 0$, (b) shows that $\|x + M\| \geq 1 - \varepsilon$ for some $x \in X$ with $\|x\| = 1$, so $\|\pi\| \geq 1 - \varepsilon$. Since ε was arbitrary, $\|\pi\| \geq 1$.

(d) d

(e) e ☐

Exercise 5.15

Suppose that X and Y are normed vector spaces and $T \in \mathcal{B}(X, Y)$. Let $\mathcal{N}(T) = \{x \in X \mid Tx = 0\}$.

(a) $\mathcal{N}(T)$ is a closed subspace of X

(b) There is a unique bounded $\tilde{T}\colon X/\mathcal{N}(T) \to Y$ such that $T = \tilde{T} \circ \pi$, where $\pi\colon X \to X/\mathcal{N}(T)$ is the projection. Moreover, $\|\tilde{T}\| = \|T\|$.

Solution. (a) This is obvious since T is continuous.

(b) Basic linear algebra yields a unique (not necessarily bounded) linear map $\tilde{T}\colon X/\mathcal{N}(T) \to Y$ such that $T = \tilde{T} \circ \pi$. To compute its norm we begin with a lemma:

> *Let X be a normed vector space and M a closed subset of X. Define $B = \{x \in X \mid \|x\| < 1\}$ and $\tilde{B} = \{x + M \in X/M \mid \|x + M\| < 1\}$. Then $\pi(B) = \tilde{B}$.*

The inclusion $\pi(B) \subseteq \tilde{B}$ is obvious since $\|\pi\| = 1$. For the opposite inclusion, let $x + M \in \tilde{B}$. By definition of the quotient norm there exists an $m \in M$ such that $\|x - m\| < 1$, since $\|x + M\| < 1$. But then $x - m \in \tilde{B}$, and so

$$x + M = \pi(x) = \pi(x - m) \in \pi(U),$$

proving the second inclusion.

Returning to the solution of the exercise, notice the following:

$$\begin{aligned}
\|\tilde{T}\| &= \sup\Big\{\|\tilde{T}\xi\| \;\Big|\; \xi \in \tilde{B}\Big\} \\
&= \sup\Big\{\|\tilde{T}\xi\| \;\Big|\; \xi \in \pi(U)\Big\} \\
&= \sup\Big\{\|\tilde{T}(\pi(x))\| \;\Big|\; x \in B\Big\} \\
&= \sup\Big\{\|Tx\| \;\Big|\; x \in B\Big\} \\
&= \|T\|.
\end{aligned}$$

Here we use the fact that for an operator $T\colon X \to Y$ it suffices to consider $x \in X$ with $\|x\| < 1$ in computing its norm: For if $\|x\| = 1$, let $\varepsilon_n = 1 - 1/n$. Then $\|\varepsilon_n x\| < 1$, and

$$\|Tx\| = \frac{1}{\varepsilon_n}\|T(\varepsilon_n x)\| \leq \frac{1}{\varepsilon_n}\sup\Bigl\{\|Ty\| \Bigm| y \in B\Bigr\} \xrightarrow[n\to\infty]{} \sup\Bigl\{\|Ty\| \Bigm| y \in B\Bigr\}.$$

Hence $\|T\| \leq \sup\{\|Ty\| \mid y \in B\}$, and the opposite equality is obvious. ☐

5.2. Linear Functionals

EXERCISE 5.18

Let X be a normed vector space.

(a) If M is a closed subspace and $x \in X \setminus M$, then $M + \mathbb{C}x$ is closed.

(b) Every finite-dimensional subspace of X is closed.

SOLUTION. (a) Let $(y_n)_{n\in\mathbb{N}}$ and $(\lambda_n)_{n\in\mathbb{N}}$ be sequences in M and $\mathbb{C}$ respectively such that $y_n + \lambda_n x$ converges to some $z \in X$. By Theorem 5.8(b) there is a $\varphi \in X^*$ such that $\varphi(x) \neq 0$ and $\varphi|_M = 0$. Applying φ to the above sequence yields

$$\varphi(z) = \lim_{n\to\infty}\bigl(\varphi(y_n) + \lambda_n\varphi(x)\bigr) = \bigl(\lim_{n\to\infty}\lambda_n\bigr)\varphi(x),$$

which implies that λ_n converges to $\varphi(z)/\varphi(x)$. The sequence (y_n) is then also convergent with limit in M, and so

$$\lim_{n\to\infty}\bigl(y_n + \lambda_n x\bigr) = \lim_{n\to\infty}\left(y_n + \frac{\varphi(z)}{\varphi(x)}x\right) = \lim_{n\to\infty} y_n + \frac{\varphi(z)}{\varphi(x)}x,$$

which lies in $M + \mathbb{C}x$ as desired.

(b) We give two different arguments. If U is a finite-dimensional subspace of X and $(e_1,\dots,e_d)$ is a basis for U, then $U = \sum_{i=1}^d \mathbb{C}e_i$. Since $\{0\}$ is a closed subspace of X, the desired result follows from the above by induction.

We may also argue as follows: It suffices to show that U is complete. To this end, let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in U and write $x_n = \lambda_{n1}e_1 + \cdots + \lambda_{nd}e_d$. We claim that the sequence $(\lambda_{ni})_{n\in\mathbb{N}}$ is a Cauchy sequence for all i. For the norm $\|\cdot\|$ on U inherited from X is equivalent to the 1-norm $\|\cdot\|_1$, so

$$\|x_m - x_n\| \geq C\|x_m - x_n\|_1 \geq C|\lambda_{mi} - \lambda_{ni}|$$

for some $C > 0$. Since $\mathbb{C}$ is complete, the sequence $(\lambda_{ni})_{n\in\mathbb{N}}$ converges to some $\lambda_i \in \mathbb{C}$. Letting $x = \lambda_1 e_1 + \cdots + \lambda_d e_d$, we claim that $x_n \to x$ as $n \to \infty$. This follows since (choosing the e_i to be unit vectors)

$$\begin{aligned}\|x_n - x\| &= \|(\lambda_{n1} - \lambda_1)e_1 + \cdots + (\lambda_{nd} - \lambda_d)e_d\| \\ &\leq |\lambda_{n1} - \lambda_1| + \cdots + |\lambda_{nd} - \lambda_d|,\end{aligned}$$

and the right-hand side converges to zero. □

Exercise 5.19

Let X be an infinite-dimensional normed vector space.

(a) There is a sequence $(x_n)_{n\in\mathbb{N}}$ in X such that $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $\|x_n - x_m\| \geq 1/2$ for $m \neq n$.

(b) X is not locally compact.

Solution. (a) First pick any unit vector $x_1 \in X$. By Exercise 5.18 the subspace $M_1 = \mathbb{C}x_1$ is closed, so Exercise 5.12(b) yields a unit vector $x_2 \notin M_1$ such that $\|x_2 + M_1\| \geq 1/2$. Since $x_1 \in M_1$ we in particular have $\|x_2 - x_1\| \geq 1/2$. Similarly, letting $M_2 = M_1 + \mathbb{C}x_2$ we get a unit vector $x_3 \notin M_2$ with $\|x_3 + M_2\| \geq 1/2$. Single both x_1 and x_2 lie in M_2 we have $\|x_3 - x_1\| \geq 1/2$ and $\|x_3 - x_2\| \geq 1/2$. Continuing this process yields the desired sequence. [TODO: Exercise references]

(b) Assume towards a contradiction that X is locally compact. Then $0 \in X$ has a compact neighbourhood K, and by multiplying with an appropriate scalar we may assume that K contains the closed unit ball $\overline{B}_1(0)$. Thus K contains the sequence (x_n) constructed in part (a). Now Theorem 0.25 implies that K is sequentially compact, so (x_n) has a convergent subsequence. But this is impossible since $\|x_n - x_m\| \geq 1/2$ for $m \neq n$, so X is not locally compact. □

Remark 5.4. Let X and Y be normed spaces, and let $T \in \mathcal{B}(X, Y)$. If T is an isometry, then clearly $\|T\| = 1$. It is easy to think that the converse is also true, perhaps if T is also assumed to be boundedly invertible, but this is not the case: For instance, equip $\mathbb{R}^2$ with the supremum norm[1] and consider the operator $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by $S(x, y) = (x, y/2)$. Then $\|S\| = 1$, and $S^{-1}(x, y) = (x, 2x)$ is also bounded with $\|S^{-1}\| = 2$. But S is clearly not an isometry, since e.g.

$$\|S(0,2)\|_\infty = \|(0,1)\|_\infty = 1 \neq 2 = \|(0,2)\|_\infty.$$

The problem is already apparent, in that the norm of S^{-1} is *not* 1, so S^{-1} cannot be an isometry. This motivates the following result:

[1] In Remark 5.5 we will see that this makes $\mathbb{R}^2$ into the categorical product of $\mathbb{R}$ and $\mathbb{R}$. This has no relevance to the present discussion, as far as I know.

Let $T \in \mathcal{B}(X,Y)$ be a boundedly invertible map between normed spaces such that $\|T\| = \|T^{-1}\| = 1$. Then T is an isometry.

For if $x \in X$ and $y = Tx$, then

$$\|Tx\| \le \|T\|\,\|x\| = \|x\| = \|T^{-1}y\| \le \|T^{-1}\|\,\|y\| = \|y\| = \|Tx\|,$$

so $\|Tx\| = \|x\|$. ⌟

Remark 5.5: The categories **Nor** and $\mathbf{Nor}_1$ of normed spaces.
A map $f\colon (S,\rho) \to (T,\delta)$ between metric spaces having the property that

$$\delta(f(x), f(y)) \le \rho(x,y)$$

for all $x,y \in S$ is variously called a *short map*, a *metric map*, *nonexpansive* or *-expanding*, a *weak contraction*, or just a Lipschitz function with Lipschitz constant 1. We consider the category $\mathbf{Nor}_1$ whose objects are normed spaces and whose arrows are linear maps that are also short maps. Notice that a linear map $T\colon X \to Y$ between normed spaces is short just when $\|T\| \le 1$. Hence $\mathbf{Nor}_1$ is a subcategory of the category **Nor** of normed spaces and bounded linear maps.

A bounded linear map $T\colon X \to Y$ is an isomorphism in **Nor** just when it is boundedly invertible. In $\mathbf{Nor}_1$ the situation is slightly more complicated: The map S in Remark 5.4 is a short map but its inverse is not short. Hence the isomorphisms in $\mathbf{Nor}_1$ are the boundedly invertible maps with short inverses, and this latter assumption cannot be removed. Furthermore, we claim that in this case T is in fact an isometry. If T has a bounded inverse T^{-1}, then

$$1 = \|\mathrm{id}_X\| = \|T^{-1}T\| \le \|T^{-1}\|\,\|T\|.$$

Hence if both T and T^{-1} are short maps, then $\|T\| = \|T^{-1}\| = 1$. But then Remark 5.4 implies that T is an isometry. Conversely, surjective isometries[2] are clearly short maps whose inverses are also short, so any surjective isometry is an isomorphism in $\mathbf{Nor}_1$.

If X and Y are normed spaces we may equip the Cartesian product $X \times Y$ with different norms, two of which are of particular importance here, namely the supremum norm[3] $\|(x,y)\|_\infty = \max\{\|x\|, \|y\|\}$ and the 1-norm $\|(x,y)\|_1 = \|x\| + \|y\|$. We reserve the notation $X \times Y$ for the Cartesian product equipped with the supremum norm, and we use the notation $X \oplus Y$ when we equip the Cartesian product with the 1-norm.

[2] An isometry is in particular injective, so surjective isometries are bijective. The inverse is also clearly bounded.

[3] We denote any norm on a vector space other than $X \times Y$ by $\|\cdot\|$, relying on context to distinguish.

We claim that $X \times Y$ is a categorical product of X and Y. First notice that the projections $\pi_X \colon X \times Y \to X$ and $\pi_Y \colon X \times Y \to Y$ are indeed short maps. For instance,

$$\|\pi_X(x,y)\| = \|x\| \leq \max\Big\{\|x\|, \|y\|\Big\} = \|(x,y)\|_\infty.$$

Given short linear maps $T_X \colon Z \to X$ and $T_Y \colon Z \to Y$, the map $T \colon Z \to X \times Y$ given by $Tz = (T_X z, T_Y z)$ is certainly linear. It is also short, for

$$\|Tz\|_\infty = \|(T_X z, T_Y z)\|_\infty = \max\Big\{\|T_X z\|, \|T_Y z\|\Big\} \leq \|z\|.$$

Notice that the 1-norm would not in general make T into a short map, but that the supremum norm is in some sense natural: Bounding a pair (x,y) just means bounding *both* x and y separately. Furthermore, it clearly makes the diagram

$$\begin{tikzcd}
 & & X \\
Z \arrow[r, "T"] \arrow[urr, bend left, "T_X"] \arrow[drr, bend right, "T_Y"'] & X \times Y \arrow[ur, "\pi_X"'] \arrow[dr, "\pi_Y"] & \\
 & & Y
\end{tikzcd}$$

commute, and it is (even in **Set**) unique with this property, so $X \times Y$ is indeed a product of X and Y.

Next we claim that $X \oplus Y$ is a coproduct of X and Y. The inclusion maps $i_X \colon X \to X \oplus Y$ and $i_Y \colon Y \to X \oplus Y$ are given by $i_X(x) = (x,0)$ and $i_Y(0,y)$. Notice that e.g.

$$\|i_X(x)\|_1 = \|(x,0)\|_1 = \|x\| + \|0\| = \|x\|,$$

so the inclusion maps are isometries, in particular short maps. Furthermore, if $S_X \colon X \to W$ and $S_Y \colon Y \to W$ are short linear maps, we define a map $S \colon X \oplus Y \to W$ by $S(x,y) = S_X x + S_Y y$. This is then clearly linear, and it is also short since

$$\|S(x,y)\| = \|S_X x + S_Y y\| \leq \|S_X x\| + \|S_Y y\| \leq \|x\| + \|y\| = \|(x,y)\|_1.$$

Again notice that the supremum norm would not make S into a short map. But the 1-norm is natural in the sense that elements of $X \oplus Y$ are to be thought of, in some sense, *sums* of elements in X and Y. Hence the norm of such a sum is (naturally) the sum of the norms. Finally, it clearly makes the diagram

$$\begin{tikzcd}
X \arrow[dr, "i_X"] \arrow[drr, bend left, "S_X"] & & \\
& X \oplus Y \arrow[r, "S"] & W \\
Y \arrow[ur, "i_Y"'] \arrow[urr, bend right, "S_Y"'] & &
\end{tikzcd}$$

commute, and so $X \oplus Y$ is a coproduct of X and Y as claimed.

For completeness we note that the categories **Ban** and $\mathbf{Ban}_1$ of Banach spaces and, respectively, bounded and short linear maps are full subcategories of **Nor** and $\mathbf{Nor}_1$. If X and Y are Banach spaces, then so are $X \times Y$ and $X \oplus Y$: If $((x_n, y_n))_{n\in\mathbb{N}}$ is a Cauchy sequence in either, then (x_n) and (y_n) are Cauchy in X and Y respectively, converging to $x \in X$ and $y \in X$. We then have

$$\|(x_n, y_n) - (x, y)\|_\infty = \|(x_n - x, y_n - y)\|_\infty = \max\Big\{\|x_n - x\|, \|y_n - y\|\Big\},$$

which goes to zero as $n \to \infty$. We similarly have

$$\|(x_n, y_n) - (x, y)\|_1 = \|x_n - x\| + \|y_n - y\|,$$

which similarly goes to zero. In either case (x_n, y_n) converges to (x, y). Thus $X \times Y$ and $X \oplus Y$ are also a product and coproduct in **Ban** and $\mathbf{Ban}_1$.

Furthermore, if X and Y are Banach spaces and $T \in \mathcal{B}(X, Y)$ is bijective, then the Open Mapping Theorem implies that T^{-1} is bounded. The isomorphisms in **Ban** are thus simply the bijections. However, the example in Remark 5.4 shows that an isomorphism in **Ban** with norm 1 might have an inverse with norm greater than 1. Thus there does not seem to be a simpler characterisation of the isomorphisms of $\mathbf{Ban}_1$ than the bijections T such that both T and T^{-1} have norm 1. $\lrcorner$

Exercise 5.21

If X and Y are normed vector spaces, define $\alpha \colon X^* \oplus Y^* \to (X \times Y)^*$ by

$$\alpha(\varphi, \psi)(x, y) = \varphi(x) + \psi(y).$$

Then α is an isometric isomorphism.

This says that the dual functor $(-)^* \colon \mathbf{Nor} \to \mathbf{Nor}$ sends products to coproducts. [TODO: Is this more properly a functor on $\mathbf{Nor}_1$? And what about the dual space, can it contain functionals with norm > 1?]

SOLUTION. We first show that α is surjective, so let $\chi \in (X \times Y)^*$ and define $\varphi(x) = \chi(x,0)$ and $\psi(y) = \chi(0,y)$. These are then bounded linear functionals: e.g.,

$$|\varphi(x)| = |\chi(x,0)| \leq \|\chi\| \|(x,0)\| = \|\chi\| \|x\|,$$

and $\alpha(\varphi,\psi) = \varphi(x) + \psi(y) = \chi(x,y)$, so α is surjective.

Next we show that α is an isometry. We have

$$\begin{aligned} |\alpha(\varphi,\psi)(x,y)| &= |\varphi(x) + \psi(y)| \\ &\leq |\varphi(x)| + |\psi(y)| \\ &\leq \|\varphi\| \|x\| + \|\psi\| \|y\| \\ &\leq \bigl(\|\varphi\| + \|\psi\|\bigr) \max\bigl\{\|x\|, \|y\|\bigr\} \\ &= \|(\varphi,\psi)\| \|(x,y)\|, \end{aligned}$$

so $\|\alpha(\varphi,\psi)\| \leq \|(\varphi,\psi)\|$. Next, let $x \in X$ and $y \in Y$ be unit vectors. Theorem 5.8(b) then furnishes $\varphi \in X^*$ and $\psi \in Y^*$ with $\|\varphi\| = \|\psi\| = 1$, $\varphi(x) = \|x\| = 1$ and $\psi(y) = \|y\| = 1$. We thus have

$$\begin{aligned} |\alpha(\varphi,\psi)(x,y)| &= |\varphi(x) + \psi(y)| \\ &= \|x\| + \|y\| \\ &= 2 \cdot 1 \\ &= \bigl(\|\varphi\| + \|\psi\|\bigr) \max\bigl\{\|x\|, \|y\|\bigr\} \\ &= \|(\varphi,\psi)\| \|(x,y)\|, \end{aligned}$$

showing that $\|\alpha(\varphi,\psi)\| \geq \|(\varphi,\psi)\|$. In total, α is an isometry. Hence it is also injective and thus an isomorphism. □

REMARK 5.6. Let X be a vector space over a field k, and let X^* be the algebraic dual of X. If U is a subspace of X, then the *annihilator* of U is the subspace U^0 of X^* consisting of those functionals φ such that $\varphi(u)$ for all $u \in U$. We use U^0 to describe the algebraic dual U^* of U.

Let $i_U \colon U \to X$ be the inclusion map, and consider its pullback

$$\beta = i_U^* \colon X^* \to U^*$$

given by precomposition with i_U. This is surjective, since if $\psi \in U^*$ then we may extend this to a linear functional on X by letting $\psi(v) = 0$ for all $v \in V$, where V is any complement of U in X. Furthermore, a functional $\varphi \in X^*$ lies in the kernel of β just if φ vanishes on U, i.e. if $\varphi \in U^0$. The first isomorphism theorem then yields a linear isomorphism

$$\tilde{\beta} \colon X^*/U^0 \to U^*. \qquad ⌟$$

EXERCISE 5.23

Suppose that X is a Banach space. If M is a closed subspace of X and N is a closed subspace of X^*, let $M^0 = \{\varphi \in X^* \mid \varphi|_M = 0\}$ and $N^\perp = \{x \in X \mid \varphi(x) = 0 \text{ for all } \varphi \in N\}$.

(a) M^0 and $N^\perp$ are closed subspaces of X^* and X, respectively.

(b) $(M^0)^\perp = M$ and $(N^\perp)^0 \supseteq N$. If X is reflexive, $(N^\perp)^0 = N$.

(c) c

(d) Define $\beta\colon X^* \to M^*$ by $\beta(\varphi) = \varphi_M$; then β induces a map $\tilde{\beta}\colon X^*/M^0 \to M^*$, and $\tilde{\beta}$ is an isometric isomorphism.

SOLUTION. (a) First assume that X is a normed vector space over $\mathbb{K}$, and assume that M and N are merely (not necessarily closed) *subsets* of X and X^*. Then M^0 and $N^\perp$ are clearly subspaces. Consider the inclusion map $i_M\colon M \to X$ and its pullback $\beta = i_M^*\colon X^* \to M^*$. The former clearly has norm 1, so for $\varphi \in X^*$ the composition $\varphi \circ i_M$ is bounded. It follows that

$$\|\beta(\varphi)\| = \|\varphi \circ i_M\| \le \|\varphi\|\|i_M\| = \|\varphi\|$$

so β is bounded. But notice that $M^0 = \ker\beta$, so M^0 is closed. Furthermore, notice that

$$N^\perp = \bigcap_{\varphi\in N} \ker\varphi,$$

so $N^\perp$ is an intersection of closed sets, hence is closed.

(b) Let $x \in M$. Then $\varphi(x) = 0$ for all $\varphi \in M^0$, and so $x \in (M^0)^\perp$. Conversely, assume that M is now a closed subspace of M, and assume that $x \notin M$. Theorem 5.8(b) then yields a functional $\varphi \in X^*$ such that $\varphi_M = 0$ and $\varphi(x) \neq 0$. But this means that $\varphi \in M^0$ and that $x \notin (M^0)^\perp$.

Furthermore, if $\varphi \in N$ then clearly $\varphi(x)$ for all $x \in N^\perp$, so $\varphi \in (N^\perp)^0$ (even if N is neither closed or a subspace).

TODO: X reflexive.

(c) c

(d) Since $\ker\beta = M^0$ and β is surjective, $\tilde{\beta}$ is a linear isomorphism, and it is bounded since β is. It remains to be shown that it is an isometry. First let $\varphi \in X^*$ and notice that

$$\|\tilde{\beta}(\varphi + M^0)\| = \|\beta(\varphi)\| = \|\varphi \circ i_M\| \le \|\varphi\|,$$

since $\|i_M\| = 1$ (unless $M = 0$, but in this case the claim is trivial). Since $\|\varphi + M^0\|$ is the infimum of $\|\psi\|$ over all $\psi \in X^*$ such that $\psi + M^0 = \varphi + M^0$, it follows that $\|\tilde{\beta}(\varphi + M^0)\| \leq \|\varphi + M^0\|$. For the opposite inequality, consider the seminorm

$$p(x) = \|\tilde{\beta}(\varphi + M^0)\| \, \|x\|$$

on X. For $x \in M$ we have

$$|\varphi(x)| = |\beta(\varphi)(x)| = |\tilde{\beta}(\varphi + M^0)(x)| \leq p(x),$$

so the Hahn–Banach theorem furnishes a $\psi \in X^*$ that extends $\varphi|_M$ and satisfies[4] $|\psi| \leq p$, i.e. $\|\psi\| \leq \|\tilde{\beta}(\varphi + M^0)\|$. In other words, $\psi|_M = \varphi|_M$ or equivalently $\psi + M^0 = \varphi + M^0$. It follows that

$$\|\varphi + M^0\| = \|\psi + M^0\| \leq \|\psi\| \leq \|\tilde{\beta}(\varphi + M^0)\|.$$

In total, $\tilde{\beta}$ is an isometry. □

5.5. *Hilbert Spaces*

Remark 5.7. We give a different proof of the Cauchy–Schwarz inequality using (very) basic properties of orthogonal projections:

> *If X is a inner product space over $\mathbb{K}$, then*
>
> $$|\langle x, y \rangle| \leq \|x\| \, \|y\| \tag{5.2}$$
>
> *for all $x, y \in X$, with equality if and only if x and y are linearly dependent.*

This is obvious if $y = 0$, so assume not. The *projection* of x on y is the unique vector $p \in \operatorname{span}(y)$ such that $y \perp x - p$. This exists and is unique, for notice that for $\alpha \in \mathbb{K}$ we have

$$0 = \langle x - \alpha y, y \rangle = \langle x, y \rangle - \alpha \langle y, y \rangle$$

if and only if

$$\alpha = \frac{\langle x, y \rangle}{\langle y, y \rangle},$$

so $p = \alpha y$. Notice that p has the property that $x = p$ if and only if x and y are linearly dependent. The 'only if' part is obvious, and the converse follows since if $x = \beta y$ for some $\beta \in \mathbb{K}$ then, plugging in above, we find that $\alpha = \beta$.

Also notice that $p \perp x - p$. Writing $x = p + (x - p)$, Pythagoras' theorem thus implies that

$$\|x\|^2 = \|p\|^2 + \|x - p\|^2 \geq \|p\|^2, \tag{5.3}$$

[4] In the real case this follows since p is a seminorm.

with equality just when $x = p$, i.e. when x and y are linearly dependent. Inserting the formula above for p, the inequality (5.3) is equivalent to

$$\|x\| \geq \|p\| = \frac{|\langle x, y\rangle|}{\|y\|} = \frac{|\langle x, y\rangle|}{\|y\|^2}\|y\|$$

which in turn is equivalent to (5.2). ⌟

8 • Elements of Fourier Analysis

8.1. Preliminaries

REMARK 8.1: Completeness of the Schwartz space.
We elaborate on the proof of Proposition 8.2. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathcal{S}$. Then the sequence $(\partial^\alpha f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in the uniform norm for all multi-indices α, so $\partial^\alpha f_n$ converges uniformly to g_α. Folland then shows that $g_\alpha = \partial^\alpha g_0$. It remains to be shown that $\|f_n - g_0\|_{(N,\alpha)} \to 0$ for all α.

To this end, let $\varepsilon > 0$ and choose $M \in \mathbb{N}$ such that $m, n \geq M$ implies that $\|f_n - f_m\|_{(N,\alpha)} < \varepsilon$. For every $x \in \mathbb{R}^d$ we thus have

$$(1 + \|x\|)^N |\partial^\alpha f_n(x) - \partial^\alpha f_m(x)| < \varepsilon.$$

Letting $m \to n$ we get

$$(1 + \|x\|)^N |\partial^\alpha f_n(x) - \partial^\alpha g_0(x)| \leq \varepsilon.$$

Taking the supremum we find that $n \geq M$ implies that $\|f_n - g_0\|_{(N,\alpha)} \leq \varepsilon$, showing that $f_n \to g$ in $\mathcal{S}$. ⌟

8.2. Convolutions

REMARK 8.2: Associativity of convolution.
If $f, g, h \in \mathcal{M}(\mathcal{B}(\mathbb{R}^d))$, then we define the function $k \colon \mathbb{R}^{3d} \to \mathbb{C}$ by

$$k(x, y, z) = f(y)g(x - y - z)h(z).$$

This is clearly measurable, so we may consider the function $K \colon \mathbb{R}^d \to [0, \infty]$ given by

$$K(x) = \int_{\mathbb{R}^{2d}} |k(x, \cdot, \cdot)| \,\mathrm{d}\lambda_{2d}.$$

By Tonelli's theorem K is also measurable, so the set

$$\Delta(f, g, h) = \{x \in \mathbb{R}^d \mid k(x, \cdot, \cdot) \in \mathcal{L}^1(\lambda_d)\} = \{x \in \mathbb{R}^d \mid K(x) < \infty\}$$

is measurable. For $x \in \Delta(f, g, h)$, Fubini's theorem thus implies that

$$\begin{aligned}
(f * g) * h(x) &= (g * f) * h(x) \\
&= \int_{\mathbb{R}^d} g * f(x-z) h(z) \, \mathrm{d}z \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(y) g(x-z-y) \, \mathrm{d}y \right) h(z) \, \mathrm{d}z \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(y) g(x-z-y) h(z) \, \mathrm{d}y \right) \mathrm{d}z \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(y) g(x-y-z) h(z) \, \mathrm{d}z \right) \mathrm{d}y \\
&= \int_{\mathbb{R}^d} f(y) \left(\int_{\mathbb{R}^d} g(x-y-z) h(z) \, \mathrm{d}z \right) \mathrm{d}y \\
&= \int_{\mathbb{R}^d} f(y) h * g(x-y) \, \mathrm{d}y \\
&= f * (h * g)(x) \\
&= f * (g * h)(x).
\end{aligned}$$

Thus convolution is associative on $\Delta(f, g, h)$. If $f, g, h \in \mathcal{L}^1(\lambda_d)$, then it is easy to show that $\Delta(f, g, h)^c$ is a Lebesgue null-set. However, it is not clear whether (and I don't see why it should be true that) $\Delta(f, g, h)$ is the same as $\Delta(f * g, h)$ or $\Delta(f, g * h)$. ⌟