

# Integration and Modern Analysis

John Benedetto and

A Functional analysis

Wojtek Czaja

## A.1 Definitions of spaces

This appendix lists results from functional analysis that are used in this book. There are many excellent texts and expositions including [134], [152], [235], [194], [302], [379], [393], [451].

### Definition A.1.1. Topological space

A *topological space*  $X$  is a pair  $(X, \mathcal{T})$ , where  $X$  is a non-empty set,  $\mathcal{T} \subseteq \mathcal{P}(X)$ , and  $\mathcal{T}$  satisfies the conditions:

- i.  $\emptyset \in \mathcal{T}$ ,  $X \in \mathcal{T}$ ,
- ii.  $\{U_\alpha : \alpha \in I, \text{ an index set}\} \subseteq \mathcal{T} \implies \bigcup_{\alpha \in I} U_\alpha \in \mathcal{T}$ ,
- iii.  $\{U_j : j = 1, \dots, n\} \subseteq \mathcal{T} \implies \bigcap_{j=1}^n U_j \in \mathcal{T}$ .

The elements of  $\mathcal{T}$  are called *open sets* and  $\mathcal{T}$  is a *topology* for the set  $X$ . The *interior* of  $S \subseteq X$ , denoted by  $\text{int } S$ , is the largest open set contained in  $S$ . The complement of an open set is a *closed set*. A set  $S$  in a topological space  $(X, \mathcal{T})$  is a *neighborhood* of  $x \in X$  if  $x \in U \subseteq S$  for some  $U \in \mathcal{T}$ .  $\mathcal{B} \subseteq \mathcal{T}$  is a *basis* for the topological space  $(X, \mathcal{T})$  if for each  $x \in X$  and each neighborhood  $S$  of  $x$ , we have  $x \in V \subseteq S$  for some  $V \in \mathcal{B}$ .  $\mathcal{B}_x$  is a *basis at*  $x \in X$  if each element of  $\mathcal{B}_x$  is a neighborhood of  $x$  and, for every neighborhood  $S$  of  $x$ , we have  $x \in B \subseteq S$  for some  $B \in \mathcal{B}_x$ .

### Theorem A.1.2. Characterization of a basis

A family  $\mathcal{B}$  is a basis for some topology  $\mathcal{T}$  for  $X = \bigcup\{B : B \in \mathcal{B}\}$  if and only if

$$\forall U, V \in \mathcal{B} \text{ and } \forall x \in U \cap V, \exists W \in \mathcal{B} \text{ such that } x \in W \subseteq U \cap V.$$

In this case,  $\mathcal{T}$  is the family of all unions of members of  $\mathcal{B}$ .

We shall *assume* that all of our topological spaces  $X$  are *Hausdorff*, i.e., that they satisfy the following property:

$$\forall x, y \in X, x \neq y, \exists U_x, U_y \in \mathcal{T} \text{ such that } x \in U_x, y \in U_y, \text{ and } U_x \cap U_y = \emptyset.$$

Let  $(X, \mathcal{T})$  be a topological space and let  $Y \subseteq X$ . Define  $\mathcal{T}_Y = \{U \cap Y : U \in \mathcal{T}\}$ . As such,  $(Y, \mathcal{T}_Y)$  is a topological space, and  $\mathcal{T}_Y$  is the *induced topology*

on  $Y$  from  $(X, \mathcal{T})$ . Among other natural situations, the concept of induced topology allows us to discuss the Borel algebra  $\mathcal{B}(Y)$  in terms of  $\mathcal{B}(X)$ , as well as Borel measurable functions on  $Y$  when we are given  $\mathcal{B}(X)$ , e.g., see Sections 2.4 and 8.7.

A set  $Y \subseteq X$  is *dense* in  $X$  if for each  $x \in X$  and each open set  $U$  containing  $x$  there is a point  $y \in Y \cap U$ . According to Definition 1.2.11a,  $K \subseteq X$  is *compact* if every covering of  $K$  by open sets contains a finite subcovering; and  $K \subseteq X$  is *relatively compact* if its *closure* (the smallest closed set containing it) is compact. A topological space  $X$  is *locally compact* if every point has at least one compact neighborhood, i.e., if

$$\forall x \in X, \exists K \subseteq X, \text{ compact, and } \exists V \in \mathcal{T} \text{ such that } x \in V \subseteq K.$$

Recall that a function  $f : X_1 \rightarrow X_2$  is a *bijection* if  $f$  is one-to-one (injective) and onto (surjective). Two topological spaces  $(X_i, \mathcal{T}_i)$ ,  $i = 1, 2$ , are *homeomorphic* if there is a bijection  $f : X_1 \rightarrow X_2$  such that

$$\forall U \in \mathcal{T}_1, f(U) \in \mathcal{T}_2 \quad \text{and} \quad \forall V \in \mathcal{T}_2, f^{-1}(V) \in \mathcal{T}_1.$$

In this case,  $f$  is an *homeomorphism*. These two conditions define the *continuity* of  $f^{-1}$  and  $f$  on  $X_2$  and  $X_1$ , respectively, cf., the equivalent definition of continuity for metric spaces  $X$  and  $Y$  in Definition A.4.2. This latter definition emphasizes the local nature of continuity by defining continuity at a point.

### Theorem A.1.3. Urysohn lemma

*Let  $X$  be a locally compact Hausdorff space. If  $K \subseteq X$  is compact and  $U \subseteq X$  is an open set containing  $K$ , then there is a continuous function  $f : X \rightarrow [0, 1]$  such that  $f = 1$  on  $K$  and  $f = 0$  on  $U^c$ .*

A topological space  $X$  is *connected* if it cannot be represented as a disjoint union of two non-empty closed sets.  $X$  is *locally connected* if it has the following property at each  $x \in X$ : Every neighborhood of  $x$  contains a connected neighborhood of  $x$ . If  $X$  is a locally compact Hausdorff space and if for every two points  $x, y \in X$  there exists a continuous function  $p : [0, 1] \rightarrow X$  such that  $p(0) = x$  and  $p(1) = y$ , we say that  $X$  is *pathwise connected*.

Standard references for topological spaces include [271], [298].

### Definition A.1.4. Metric space

**a.** A *metric space*  $X$  is a pair  $(X, \rho)$ , where  $X$  is a non-empty set and  $\rho : X \times X \rightarrow \mathbb{R}^+$  satisfies the conditions:

- i.  $\forall x, y \in X, \rho(x, y) \geq 0$ ,
- ii.  $\forall x, y \in X, \rho(x, y) = \rho(y, x)$ ,
- iii.  $\forall x, y, z \in X, \rho(x, z) \leq \rho(x, y) + \rho(y, z)$  (triangle inequality),
- iv.  $\forall x, y \in X, \rho(x, y) = 0 \iff x = y$ .

$\rho$  is a *metric*. If only the first three conditions are satisfied we say that  $\rho$  is a *pseudometric*.

b. The *open ball*  $B(x, r)$ , with center  $x$  and radius  $r$ , in a metric space  $X$  is

$$B(x, r) = \{y \in X : \rho(x, y) < r\}.$$

A metric space is a topological space and  $U$  is defined to be *open* if

$$\forall x \in U, \exists B(x, r) \subseteq U.$$

Equivalently, we can define a basis  $\mathcal{B}$  for the topology in a metric space  $X$  to be  $\{B(x, r) : x \in X, r > 0\}$ .

In particular, metric spaces are Hausdorff.

c. A sequence  $\{x_n : n = 1, \dots\} \subseteq X$ , where  $X$  is a metric space, is *Cauchy* if

$$\forall \varepsilon > 0, \exists N \text{ such that } \forall m, n > N, \rho(x_m, x_n) < \varepsilon.$$

If  $X$  is a metric space in which every Cauchy sequence  $\{x_n : n = 1, \dots\}$  converges to some element  $x$ , i.e.,  $\rho(x_n, x) \rightarrow 0$ , then  $X$  is *complete*.

d. Two metric spaces  $(X_i, \rho_i)$ ,  $i = 1, 2$ , are *isometric* if there is a bijection  $f : X_1 \rightarrow X_2$  such that

$$\forall x, y \in X_1, \rho_1(x, y) = \rho_2(f(x), f(y)).$$

In this case,  $f$  is an *isometry*.

e. Let  $(X, \rho)$  be a metric space. A subset  $V \subseteq X$  is *closed* if, whenever  $\{x_n : n = 1, \dots\} \subseteq V$  and  $\rho(x_n, x) \rightarrow 0$  for some  $x \in X$ , we can conclude that  $x \in V$ . The *closure*  $\bar{Y}$  of a subset  $Y \subseteq X$  is the set of all elements  $x \in X$  for which there is a sequence  $\{x_n : n = 1, \dots\} \subseteq Y$  such that  $\rho(x_n, x) \rightarrow 0$ . The complement of a closed set  $V$  is open and vice-versa.

f. The *diameter* of a subset  $Y$  of a metric space  $(X, \rho)$ , denoted by  $\text{diam}(Y)$ , is

$$\text{diam}(Y) = \sup\{\rho(x, y) : x, y \in Y\}.$$

### Theorem A.1.5. Compact metric spaces

Let  $(X, \rho)$  be a metric space. A subset  $K \subseteq X$  is *compact* if and only if every sequence has a convergent subsequence

**Remark.** If a topological space  $X$  has a countable basis then  $X$  is said to satisfy the *second axiom of countability*. If  $X$  is a locally compact Hausdorff space, then the second axiom of countability is equivalent to the existence of a metric  $\rho$  on  $X$  and a sequence of compact sets  $F_n$  such that  $X = \bigcup_{n=1}^{\infty} F_n$ ,  $F_n \subseteq \text{int } F_{n+1}$ , see [270], Theorem I.5.3.

### Example A.1.6. Hilbert cube

An important example of a metric space is the *Hilbert cube*  $[0, 1]^{\aleph_0}$ . It is defined to be the Cartesian product of countably infinitely many copies of  $[0, 1]$  equipped with the metric

$$\rho(x, y) = \left( \sum_{i=1}^{\infty} (\min(1/i, |x_i - y_i|))^2 \right)^{1/2},$$

for  $x = \{x_i : i = 1, \dots\}$  and  $y = \{y_i : i = 1, \dots\}$ .

The following result should be compared with our construction of the *Cantor function* in Example 1.2.7d.

**Proposition A.1.7.** *There exists a continuous function  $f : C \rightarrow [0, 1]^{\aleph_0}$  that maps the Cantor set  $C$  onto the Hilbert cube  $[0, 1]^{\aleph_0}$ .*

An excellent reference for metric spaces is [192].

It is sometimes necessary to consider topological vector spaces where the topology cannot be described by a metric, e.g., in the theory of distributions, see Chapter 7, Example A.6.5, [235], [415], or [39], Chapter 2. In such cases we would still like to have a notion of completeness and this is accomplished through the theory of uniform spaces, e.g., [271].

**Definition A.1.8. Uniform space**

A *uniform structure* on a set  $X$  is a family  $\mathcal{X}$  of subsets of  $X \times X$  which satisfies the conditions:

- i.  $\forall V \in \mathcal{X}, \{(x, x) : x \in X\} \subseteq V$ ,
- ii.  $\forall V \in \mathcal{X}, \{(y, x) : (x, y) \in V\} \in \mathcal{X}$ ,
- iii.  $\forall V \in \mathcal{X}, \exists V' \in \mathcal{X}$  such that

$$\{(x, y) : \exists z \in X \text{ such that } (x, z), (z, y) \in V'\} \in \mathcal{X},$$

- iv.  $\forall V, V' \in \mathcal{X}, V \cap V' \in \mathcal{X}$ ,
- v.  $\forall V \subseteq X \times X$ , for which  $\exists V' \in \mathcal{X}$  and  $V' \subseteq V$ , we have  $V \in \mathcal{X}$ .

A *uniform space*  $(X, \mathcal{X})$  is a topological space  $(X, \mathcal{T})$  with the topology  $\mathcal{T}$  defined by sets of the form

$$U = \{x : x \in X \text{ and } \exists y \in A \subseteq X \text{ such that } (y, x) \in V \in \mathcal{X}\},$$

for all subsets  $A \subseteq X$  and for all sets  $V \in \mathcal{X}$ .

A uniform structure  $\mathcal{X}$  is *pseudometrizable* if its corresponding topology  $\mathcal{T}$  has a countable basis.

If  $\{\rho_i, i = 1, \dots\}$  is a family of pseudometrics, respectively, metrics, on a non-empty set  $X$ , consider the uniform structure  $\mathcal{X}$  on  $X$  defined by the collection of sets

$$\{(x, y) \in X \times X : \rho_i(x, y) < \varepsilon\}, \quad \varepsilon > 0, i = 1, \dots$$

If the topology  $\mathcal{T}$  corresponding to this uniform structure  $\mathcal{X}$  has a countable basis, then there exists a pseudometric, respectively, metric,  $\rho$  which induces the same topology as  $(X, \mathcal{X})$ .

Standard references for uniform spaces are [271], [71].

**Definition A.1.9. Normed vector space and Banach space**

Let  $X$  be a vector space over  $\mathbb{F}$ ,  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{F} = \mathbb{C}$ .  $X$  is a *normed vector space* if there is a function  $\|\dots\| : X \rightarrow \mathbb{R}^+$  such that

- i.  $\forall x \in X, \|x\| = 0 \iff x = 0$ ,
- ii.  $\forall x, y \in X, \|x + y\| \leq \|x\| + \|y\|$  (triangle inequality),
- iii.  $\forall a \in \mathbb{F}, \forall x \in X, \|ax\| = |a|\|x\|$ .

$\|\dots\|$  is a *norm*. A normed vector space is a metric space with metric  $\rho(x, y) = \|x - y\|$ .

A complete normed vector space is a *Banach space*.

Let  $X$  be a normed vector space.  $\sum x_n$  converges to  $x \in X$ , for  $x_n \in X$ ,  $n = 1, \dots$ , if

$$\lim_{N \rightarrow \infty} \left\| x - \sum_{n=1}^N x_n \right\| = 0.$$

$\sum x_n$  is *absolutely convergent* if

$$\sum_{n=1}^{\infty} \|x_n\| < \infty.$$

We have the following characterization of Banach spaces.

**Proposition A.1.10.** *A normed vector space  $X$  is a Banach space if and only if every absolutely convergent series is convergent.*

*Proof.* ( $\implies$ ) Take  $\{x_n : n = 1, \dots\} \subseteq X$  for which  $\sum \|x_n\| < \infty$ , and choose  $\varepsilon > 0$ . If  $\sum_{n=N}^{\infty} \|x_n\| < \varepsilon/2$ , then, for each  $n > m \geq N$ ,

$$\left\| \sum_{j=m}^n x_j \right\| < \varepsilon.$$

Thus,  $\sum x_n$  converges to some  $x \in X$  since  $X$  is complete.

( $\impliedby$ ) Let  $\{x_n : n = 1, \dots\} \subseteq X$  be a Cauchy sequence in  $X$ . Hence, for each  $k$  there is  $n_k \in \mathbb{N}$  such that

$$\forall m, n \geq n_k, \|x_m - x_n\| < \frac{1}{2^k};$$

we can also choose  $n_{k+1} > n_k$ . Set  $y_k = x_{n_k} - x_{n_{k-1}}$  for  $k = 1, \dots$ , where  $x_{n_0} = 0$ . Therefore,  $\sum y_k$  is absolutely convergent, so that by hypothesis and the fact that

$$\sum_{k=1}^m y_k = x_{n_m},$$

$\{x_{n_m} : m = 1, \dots\}$  converges to some  $x \in X$ . It is easy to check that

$$\lim_{n \rightarrow \infty} \|x - x_n\| = 0.$$

□

Let  $X$  be a Banach space over  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{F} = \mathbb{C}$ . A subset  $V \subseteq X$  is a *linear subspace* of  $X$  if  $V$  is a vector space over  $\mathbb{F}$ . If  $V \subseteq X$  is a linear subspace then its closure  $\overline{V}$  in  $X$  is also a linear subspace.

The *span* of a subset  $Z \subseteq X$ , designated  $\text{span } Z$ , is the set of all finite linear combinations  $x = \sum c_n x_n$ , where  $c_n \in \mathbb{F}$  and  $x_n \in Z$ . (The notion of  $\text{span } Z$  can be defined in any vector space.) Clearly,  $\text{span } Z$  is a linear subspace of  $X$  and its closure is designated by  $\overline{\text{span } Z}$ .

**Definition A.1.11. Hilbert space**

Let  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{F} = \mathbb{C}$ . A *Hilbert space*  $H$  is a Banach space with a function  $\langle \dots, \dots \rangle : H \times H \rightarrow \mathbb{F}$  which satisfies the conditions:

- i.  $\forall x, y \in H, \quad \overline{\langle x, y \rangle} = \langle y, x \rangle,$
  - ii.  $\forall x, y, z \in H, \quad \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle,$
  - iii.  $\forall a \in \mathbb{F} \text{ and } \forall x, y \in H, \quad \langle ax, y \rangle = a \langle x, y \rangle,$
  - iv.  $\forall x \in H, \quad \|x\| = \sqrt{\langle x, x \rangle}.$
- $\langle \dots, \dots \rangle$  is an *inner product*.

The following result is straightforward to verify, and it does not require completeness.

**Proposition A.1.12.** *Let  $H$  be a Hilbert space. Then,*

$$\forall x, y \in H, \quad |\langle x, y \rangle| \leq \|x\| \|y\| \quad (\text{A.1})$$

and

$$\forall x, y \in H, \quad \|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2). \quad (\text{A.2})$$

**Remark.** (A.1) is the *Schwarz inequality*, which in the case of  $H = L^2_\mu(X)$  is the Hölder inequality, see Theorem 5.5.2b and Example A.2.3. Of course, this does not mean there is a simple proof of the Hölder inequality by means of the elementary inequality (A.1). In fact, the Schwarz inequality *assumes* the existence of an inner product; and the Hölder inequality *shows* the existence of an inner product for  $H$ .

(A.2) is the *parallelogram law*, see Example A.2.5.

Excellent references for Banach and Hilbert spaces are [19], [194], [379], [393], [451], [502].

## A.2 Examples

1. **a.** Let  $X$  be a topological space and let  $C(X)$  be the vector space of continuous functions  $f : X \rightarrow \mathbb{C}$ .  $C_b(X)$  denotes the vector space of functions  $f \in C(X)$  such that

$$\|f\|_\infty = \sup_{x \in X} |f(x)| < \infty, \quad (\text{A.3})$$

i.e.,  $f$  is bounded on  $X$ .

**b.** Now let  $X$  be a locally compact Hausdorff space and let  $C_0(X)$  be the vector space of functions  $f \in C_b(X)$  such that

$$\forall \varepsilon > 0, \exists K_f \subseteq X, \text{ compact, for which } \forall x \notin K_f, |f(x)| < \varepsilon. \quad (\text{A.4})$$

Intuitively,  $f$  “vanishes at infinity”.  $C_c(X)$  is the vector space of functions  $f \in C_0(X)$  such that  $f$  vanishes outside of some compact set  $K_f \subseteq X$ . (A.3) defines a norm on  $C_b(X)$ , and, with this norm,  $C_b(X)$  and  $C_0(X)$  are Banach spaces.  $C_0(X)$  is a closed subspace of  $C_b(X)$ . With this norm on  $C_0(X)$ , the Urysohn lemma gives

$$\overline{C_c(X)} = C_0(X).$$

In fact, for  $f \in C_0(X)$ ,  $\varepsilon > 0$ , and  $K_f$  as in (A.4), choose  $g \in C_c(X)$  with  $0 \leq g \leq 1$  and  $g = 1$  on  $K_f$  by Theorem A.1.3, set  $h = fg \in C_c(X)$ , and obtain  $\|f - h\|_\infty < \varepsilon$ .

If  $X$  is compact we write  $C(X) = C_b(X)$ .

**c.** Let  $(X, \mathcal{A}, \mu)$  be a measure space which is also a topological space. Since the uniform limit of continuous functions is continuous,  $C_b(X)$  can be regarded as a closed subspace of  $L_\mu^\infty(X)$ , defined in Definition 2.5.9. It is for this reason we use the notation  $\|\dots\|_\infty$  from Definition 2.5.9 in (A.3).

2.  $L_\mu^p(X)$ ,  $1 \leq p < \infty$ , with  $L^p$ -norm  $\|\dots\|_p$  defined in Definition 5.5.1, is a Banach space (Theorem 5.5.2). Further, the set of simple functions  $\sum_{j=1}^n a_j \mathbb{1}_{A_j}$ ,  $\mu(A_j) < \infty$ , is dense in  $L_\mu^p(X)$  (Theorem 5.5.3). In the case that  $X$  is a locally compact Hausdorff space and  $\mu$  is a regular Borel measure, we noted that

$$\overline{C_c(X)} = L_\mu^p(X)$$

(Theorem 7.2.6).

3. For any measure space  $(X, \mathcal{A}, \mu)$ ,  $L_\mu^2(X)$  is a Hilbert space with inner product

$$\langle f, g \rangle = \int_X f(x) \overline{g(x)} d\mu(x).$$

The fact that the integral is defined follows from the Hölder inequality (Theorem 5.5.2b). The structurally important converse is: *Let  $H$  be a non-zero Hilbert space; then there is a set  $X$  and a linear bijection,*

$$L : H \rightarrow \ell^2(X),$$

such that

$$\langle x, y \rangle = \sum_{t \in X} (L(x))(t) \overline{(L(y))(t)}.$$

The fundamental elementary results of Hilbert space theory are used to prove this fact and to determine  $\text{card } X$  uniquely in terms of the cardinality of *orthonormal sets*, see Definition A.12.1.

4. If the measure space  $(X, \mathcal{A}, \mu)$  is also a compact Hausdorff space and if  $\mathcal{A}$  contains the Borel algebra, then

$$C(X) \subseteq L_\mu^\infty(X) \subseteq \dots \subseteq L_\mu^p(X) \subseteq L_\mu^r(X) \subseteq \dots \subseteq L_\mu^1(X), \quad 1 \leq r \leq p.$$

In  $(X, \mathcal{P}(X), c)$ , where  $X$  is topologized with the metric  $\rho(x, y) = 0$  if  $x = y$  and  $\rho(x, y) = 1$  if  $x \neq y$ , we have

$$\ell^1(X) \subseteq \dots \subseteq \ell^p(X) \subseteq \ell^r(X) \subseteq \dots \subseteq \ell^\infty(X) = C_b(X), \quad 1 \leq p \leq r.$$

In both cases we have the inequality  $\| \cdot \|_p \geq \| \cdot \|_r$  so that the corresponding injection is continuous (continuous functions are defined in Definition A.4.2).

5. **a.** A Banach space is a Hilbert space if and only if the *parallelogram law*,  $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$ , is valid, see Proposition A.1.12. Using this fact we see that there are Banach spaces which are not Hilbert spaces.

**b.** Next, we give a standard example of a non-trivial complete metric vector space which is not a Banach space. Let  $X$  be the space of  $C^\infty$ -functions on  $[0, 1]$ . Define the metric  $\rho$  by

$$\rho(f, g) = \sum_{k=0}^{\infty} \frac{\|f - g\|_{(k)}}{2^k(1 + \|f - g\|_{(k)})},$$

where

$$\|f\|_{(k)} = \sup_{0 \leq j \leq k} \|f^{(j)}\|_\infty.$$

As such  $X$  is complete. If the complete metric space  $X$  is a normed vector space with norm  $\| \cdot \|$ , it is possible to show that

$$\forall n = 1, \dots, \exists C_n \text{ such that } \forall f \in X, \text{ for which } \|f\| \leq 1, \\ \|f^{(n)}\|_\infty \leq C_n.$$

It is then not difficult to find  $f \in X$  such that

$$\forall n = 1, \dots, \|f^{(n)}\|_\infty > nC_n,$$

from which we obtain the desired contradiction to the hypothesis that  $X$  is normed.

6. If  $(X, \rho)$  is a metric space there is a complete metric space  $(\tilde{X}, \tilde{\rho})$  such that  $X \subseteq \tilde{X}$ ,  $\tilde{\rho} = \rho$  on  $X \times X$ , and  $X$  is dense in  $\tilde{X}$ .  $\tilde{X}$  is the set of equivalence classes of Cauchy sequences from  $X$ , where  $\{x_n : n = 1, \dots\}$  is said to be *equivalent* to  $\{y_n : n = 1, \dots\}$  if  $\rho(x_n, y_n) \rightarrow 0$ . Let  $\{x_n\} \subseteq X$  be a Cauchy sequence, and let  $\{\{x_n\}\}$  be the equivalence class of all Cauchy sequences  $\{z_n\} \subseteq X$  equivalent to  $\{x_n\}$ , i.e.,  $\lim_{n \rightarrow \infty} \rho(x_n, z_n) = 0$ . For two equivalence classes,  $\{\{x_n\}\}$  and  $\{\{y_n\}\}$ ,  $\tilde{\rho}$  is defined by



$$\tilde{\rho}(\{\{x_n\}\}, \{\{y_n\}\}) = \lim_{n \rightarrow \infty} \rho(x'_n, y'_n),$$

where  $\{x'_n : n = 1, \dots\}$  and  $\{y'_n : n = 1, \dots\}$  are any representatives of the equivalence classes  $\{\{x_n\}\}$  and  $\{\{y_n\}\}$ , respectively.  $(\tilde{X}, \tilde{\rho})$  is the completion of  $(X, \rho)$ . A relevant theorem using this concept is the following. *Define*

$$\forall f, g \in C([a, b]), \quad \rho(f, g) = R \int_a^b |f - g|;$$

then  $\tilde{C}([a, b]) = L^1_m([a, b])$  and  $\tilde{\rho}(f, g) = \int_a^b |f - g|$ , cf., Theorem 7.1.1 and the Remark at the end of Section 7.3.

An even more basic example of the completion of a metric space is the construction of real numbers from rational numbers mentioned in Chapter 1.

For an alternative way to describe the completion of a metric space  $(X, \rho)$ , let  $B(X)$  be the Banach space of bounded real functions on  $X$  with metric  $\sigma(f, g) = \sup\{|f(x) - g(x)| : x \in X\}$ . Fix  $x_0 \in X$  and define the function  $F : X \rightarrow B(X)$ ,  $x \mapsto f_x$ , where

$$f_x(y) = \rho(x, y) - \rho(x_0, y).$$

Then,  $F$  is an isometry  $X \rightarrow F(X) \subseteq B(X)$  and  $\tilde{X} = \overline{F(X)}$ .

7. Let  $p \geq 2$  be a prime number. Any  $x \in \mathbb{Q} \setminus \{0\}$  has the unique factorization  $x = p^r q$ , where  $r \in \mathbb{Z}$  and where the numerator and denominator of  $q \in \mathbb{Q}$  are both relatively prime to  $p$ . The  $p$ -adic norm  $\|x\|$  of  $x$  is  $\|x\| = p^{-r}$  and we define  $\|0\| = 0$ . It is elementary to check that

$$\forall x, y \in \mathbb{Q}, \quad \|x + y\| \leq \max(\|x\|, \|y\|) \quad \text{and} \quad \|xy\| = \|x\| \|y\|. \quad (\text{A.5})$$

The function  $\rho_p(x, y) = \|x - y\|$  defines the  $p$ -adic metric on  $\mathbb{Q}$ , and the completion of  $\mathbb{Q}$  with respect to  $\rho_p$  is the field  $\mathbb{Q}_p$  of  $p$ -adic numbers. The completion  $\mathbb{Z}_p$  of  $\mathbb{Z}$  with respect to  $\rho_p$  is the ring of  $p$ -adic integers. Note the analogy with the construction of  $\mathbb{R}$  from  $\mathbb{Q}$ , as the completion of  $\mathbb{Q}$  with respect to the usual absolute value norm. For one entry into  $p$ -adic analysis, see [385], [370].

We point out that  $\mathbb{Q}_p$  consists of all formal Laurent series in  $p$  with coefficients  $0, 1, \dots, p-1$ , with addition and multiplication as usual for Laurent series, except with carrying of digits. For example, in  $\mathbb{Q}_5$ , we have

$$(3 + 2 \cdot 5) + (4 + 3 \cdot 5) = 2 + 1 \cdot 5 + 1 \cdot 5^2.$$

$\mathbb{Q}_p$  is a locally compact Abelian group under addition, with topology induced by the  $p$ -adic norm, see Appendix B.9 for a definition of a locally compact group. However, an important distinction between  $\mathbb{R}$  and  $\mathbb{Q}_p$ , driven by (A.5), is the fact that  $\mathbb{Z}_p$  is a compact open subgroup of  $\mathbb{Q}_p$ . This property leads to fascinating analysis with far-reaching applications in subjects as diverse as number theory, quantum field theory, and wavelet theory. In this last area, see [40].

### A.3 Separability

A topological space is *separable* if it contains a countable dense subset. It is not difficult to prove the following theorem.

**Theorem A.3.1. Separability of some  $L^p$ -spaces,  $p \in [1, \infty)$**

Let  $(X, \mathcal{M}(\mathbb{R}^d), m^d)$  be a measure space, where  $X \subseteq \mathbb{R}^d$  and where  $m^d$  is Lebesgue measure on  $X$ . If  $p \in [1, \infty)$ , then  $L_{m^d}^p(X)$  is separable.

**Example A.3.2.  $L_m^\infty([0, 1])$  is not separable**

We shall prove that  $L_m^\infty([0, 1])$  is not separable. Let  $\{f_n : n = 1, \dots\}$  be an arbitrary sequence in  $L_m^\infty([0, 1])$  and write

$$(0, 1] = \bigcup_{n=1}^{\infty} \left( \frac{1}{2^n}, \frac{1}{2^{n-1}} \right] = \bigcup_{n=1}^{\infty} E_n.$$

If

$$\operatorname{ess\,sup}_{x \in E_n} |f_n(x)| \leq \frac{1}{2},$$

define  $g = 1$  on  $E_n$ . ("ess sup" was defined in Definition 2.5.9.) Otherwise set  $g = 0$  on  $E_n$ . Consequently,  $g \in L_m^\infty([0, 1])$  and

$$\forall n = 1, \dots, \quad \|f_n - g\|_\infty \geq \frac{1}{2}.$$

Thus,  $\{f_n : n = 1, \dots\}$  is not dense in  $L_m^\infty([0, 1])$ . Since  $\{f_n : n = 1, \dots\}$  is arbitrary,  $L_m^\infty([0, 1])$  cannot be separable.

**Example A.3.3. Non-separability of some  $L^p$ -spaces,  $p \in [1, \infty)$**

The fact that a given space  $X$  is separable has no bearing on the separability of  $L_\mu^p(X)$ ,  $1 \leq p < \infty$ . Take  $([0, 1], \mathcal{P}([0, 1]), c)$ , where  $c$  is counting measure. If  $f \in L_c^1([0, 1])$ , then  $f = 0$  outside of a countable set. Thus, if  $\{f_n : n = 1, \dots\} \subseteq L_c^1([0, 1])$ , then there is  $y \in [0, 1]$  such that  $f_n(y) = 0$  for each  $n$ . Define  $g = \mathbf{1}_{\{y\}} \in L_c^1([0, 1])$  so that

$$\forall n = 1, \dots, \quad \|f_n - g\|_1 \geq 1.$$

**Theorem A.3.4. Sequential pointwise convergence of simple functions**

Let  $(X, \mathcal{A}, \mu)$  be a measure space, let  $Y$  be a separable complete metric space, and let  $(Y, \mathcal{C}, \nu)$  be a measure space, where  $\mathcal{B}(Y) \subseteq \mathcal{C}$ . If  $f : X \rightarrow Y$  is measurable, then there is a sequence  $\{g_k : k = 1, \dots\}$  of simple functions  $X \rightarrow Y$  such that  $\{g_k : k = 1, \dots\}$  converges pointwise to  $f$ .

Historically, a separable complete metric space is *Polish*.

The next theorem states that the Hilbert cube is "universal" for separable metric spaces.

**Theorem A.3.5. Urysohn theorem**

Every separable metric space  $X$  is homeomorphic with a subset of the Hilbert cube  $[0, 1]^{\aleph_0}$ .

**Corollary A.3.6.** For every separable metric space  $X$  there exists a subset  $A$  of the Cantor set  $C$  and a continuous surjective function  $f : A \rightarrow X$ . If  $X$  is compact, then  $A$  can be chosen to be closed.

Corollary A.3.6 is an extension of Proposition A.1.7 and Example 1.2.7d.

**A.4 Moore–Smith and Arzelà–Ascoli theorems**

Let  $(X, \rho)$  be a metric space. We say that  $\{x_{m,n}\} \rightarrow x$ , i.e.,  $\lim_{m,n \rightarrow \infty} x_{m,n} = x$ , if

$$\forall \varepsilon > 0, \exists N \text{ such that } \forall m, n > N, \quad \rho(x_{m,n}, x) < \varepsilon.$$

The following result can be generalized to uniform spaces with essentially the same proof.

**Theorem A.4.1. Moore–Smith theorem**

Let  $\{x_{m,n} : m, n = 1, \dots\}$  be a sequence in a complete metric space  $(X, \rho)$ . Assume

- i.  $\exists \lim_{n \rightarrow \infty} x_{m,n} = y_m$  uniformly in  $m$ ,
- ii.  $\forall n = 1, \dots, \exists \lim_{m \rightarrow \infty} x_{m,n} = z_n$ .

Then,  $\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} x_{m,n}$ ,  $\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} x_{m,n}$ , and  $\lim_{m,n \rightarrow \infty} x_{m,n}$  all exist and are equal.

*Proof.* Assumption i means that

$$\forall \varepsilon > 0, \exists K > 0 \text{ such that } \forall n > K \text{ and } \forall m, \quad \rho(y_m, x_{m,n}) < \varepsilon.$$

Using i and ii we show that  $\{y_m : m = 1, \dots\}$  is Cauchy by computing

$$\rho(y_m, x_{m,n}) < \frac{\varepsilon}{4} \quad \text{and} \quad \rho(z_k, x_{p,k}) < \frac{\varepsilon}{4}.$$

Since  $X$  is complete,  $y_m \rightarrow w \in X$ ; and it is easy to check that

$$\lim_{m,n \rightarrow \infty} x_{m,n} = w. \tag{A.6}$$

Thus,

$$\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} x_{m,n} = \lim_{m,n \rightarrow \infty} x_{m,n} = w.$$

Finally, in order to prove that  $\lim_n z_n = w$ , take  $\varepsilon > 0$  and write

$$\rho(z_n, w) \leq \rho(z_n, x_{m,n}) + \rho(x_{m,n}, w).$$

Since (A.6) holds,

$$\exists N \in \mathbb{N} \text{ such that } \forall m, n > N, \quad \rho(x_{m,n}, w) < \varepsilon;$$

and so

$$\forall n > N, \quad \rho(z_n, w) \leq \overline{\lim}_{m \rightarrow \infty} \rho(z_n, x_{m,n}) + \varepsilon = \varepsilon.$$

□

**Definition A.4.2. Continuity and equicontinuity**

Let  $(X, \rho)$  and  $(Y, \theta)$  be metric spaces. A function  $F : X \rightarrow Y$  is *continuous at*  $x \in X$  if

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ such that } \rho(z, x) < \delta \implies \theta(f(z), f(x)) < \varepsilon;$$

and  $f$  is *continuous on*  $X$  if it is continuous at each  $x \in X$ .

A sequence  $\{f_n : n = 1, \dots\}$  of continuous functions is *equicontinuous at*  $x \in X$  if

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ such that } \forall n, \quad \rho(z, x) < \delta \implies \theta(f_n(z), f_n(x)) < \varepsilon.$$

$\{f_n : n = 1, \dots\}$  is *equicontinuous on*  $X$  if it is equicontinuous at each  $x \in X$ .

The notion of equicontinuity was introduced by ASCOLI in 1883 [13] and the following theorem was proved by ARZELÀ in 1895 and 1899 [11], [12], cf., [230], [231]. Clearly, the theorem generalizes the Bolzano–Weierstrass property of  $\mathbb{R}$ .

**Theorem A.4.3. Arzelà–Ascoli theorem**

Let  $X$  be a separable metric space, let  $Y$  be a compact metric space, and let  $\{f_n : n = 1, \dots\}$  be an equicontinuous sequence of functions  $X \rightarrow Y$ . Then, there is a subsequence of  $\{f_n : n = 1, \dots\}$  which converges pointwise to a continuous function.

*Proof.* Let  $\{x_n : n = 1, \dots\} \subseteq X$  be dense. Since  $Y$  is compact,

$$\exists J_1 \subseteq \mathbb{N} \text{ such that } \{f_n(x_1) : n \in J_1\} \text{ is convergent.}$$

Pick  $J_2 \subseteq J_1$  such that  $\{f_n(x_2) : n \in J_2\}$  is convergent, and continue in this way. Consequently,

$$\forall j = 1, \dots, \exists \lim_{k \rightarrow \infty} f_{n_k}(x_j) = g(x_j),$$

where  $n_k \in J_k$  and  $\lim_{k \rightarrow \infty} n_k = \infty$ . Let  $z \in X \setminus \{x_n : n = 1, \dots\}$  with  $x_{q_p} \rightarrow z$  as  $p \rightarrow \infty$ . Then,

$$\lim_{p \rightarrow \infty} f_{n_k}(x_{q_p}) = f_{n_k}(z), \quad \text{uniformly in } k = 1, \dots,$$

by the equicontinuity hypothesis. Also,

$$\forall p = 1, \dots, \quad \lim_{k \rightarrow \infty} f_{n_k}(x_{q_p}) = g(x_{q_p}).$$

Consequently, by the Moore–Smith theorem,

$$\exists \lim_{k \rightarrow \infty} f_{n_k}(z) = g(z).$$

The continuity of  $g$  is straightforward to check. □

Obviously the result is still true if, instead of assuming that  $Y$  is compact, we assume that the range of each  $f_n$  is compact in  $Y$ . It is also easy to prove that the convergence of  $\{f_{n_k} : k = 1, \dots\}$  is uniform on compact subsets of  $X$ .

**Remark.** The notion of equicontinuity of a sequence can be generalized to an equicontinuous set by replacing “ $\forall n$ ” in Definition A.4.2 with “for all elements of the set”. Specifically, if  $X$  is a compact set and  $S \subseteq C(X)$ , then Theorem A.4.3 can be formulated as follows: *If  $S$  is pointwise bounded and equicontinuous, then  $S$  is relatively compact in the sup norm topology on  $C(X)$ , and every sequence in  $S$  has a uniformly convergent subsequence.*

## A.5 Uniformly continuous functions

### Definition A.5.1. Uniform continuity

Let  $(X, \rho)$  and  $(Y, \theta)$  be metric spaces.  $f : X \rightarrow Y$  is *uniformly continuous* if

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ such that } \rho(x, y) < \delta \implies \theta(f(x), f(y)) < \varepsilon.$$

**Remark.** If  $X$  is a compact metric space and  $f : X \rightarrow \mathbb{R}$  is continuous then  $f$  is uniformly continuous.  $f(x) = \sin(1/x)$  is a bounded continuous function  $(0, 1] \rightarrow [-1, 1]$  which is not uniformly continuous. Observe that

$$\begin{aligned} f : [0, 1) &\rightarrow [0, \infty) \\ x &\mapsto \frac{x}{1-x} \end{aligned}$$

is bijective and bicontinuous, i.e., a homeomorphism, whereas the Cauchy sequence  $\{1 - (1/n) : n = 1, \dots\}$  in  $[0, 1)$  is transformed into the sequence  $\{n - 1 : n = 1, \dots\}$ , which is not Cauchy. In this case the range space is complete and  $[0, 1)$  is not complete. Such a phenomenon leads us to distinguish between topological properties, dealing with homeomorphisms, and uniform properties, dealing with Cauchyness, uniform continuity, and completeness.

Generally there are no relations between these two categories except the following: *Let  $X$  be a metric space;  $X$  is compact if and only if it is complete and totally bounded ( $(X, \rho)$  is totally bounded if*

$$\forall \varepsilon > 0, \exists x_1, \dots, x_n \in X \text{ such that } X \subseteq \bigcup_{j=1}^n B(x_j, \varepsilon).$$

The proof of this theorem can be obtained by means of a circular chain of implications in which Theorem A.1.5 is also proved, e.g., [192], pages 267–268. In any case, for perspective, recall from Definition 1.2.11 that the compact subsets of  $\mathbb{R}^d$  are precisely the closed and bounded subsets of  $\mathbb{R}^d$ .

**Theorem A.5.2. Unique uniformly continuous extensions**

Let  $X$  be a metric space and let  $Y$  be a complete metric space. Assume that  $Z \subseteq X$  and that  $f : Z \rightarrow Y$  is a uniformly continuous function. Then,  $f$  has a unique uniformly continuous extension to  $\bar{Z}$ .

**Definition A.5.3. Absolute continuity**

Let  $(X, \rho)$  and  $(Y, \theta)$  be metric spaces and let  $f : X \rightarrow Y$  be a continuous function.  $f$  is *absolutely continuous* if

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ such that } \forall \{x_1, \dots, x_n\} \subseteq X, \\ \sum_{j=1}^{n-1} \rho(x_j, x_{j+1}) < \delta \implies \sum_{j=1}^{n-1} \theta(f(x_j), f(x_{j+1})) < \varepsilon.$$

**Remark.** Let  $(X, \rho)$  and  $(Y, \theta)$  be metric spaces and let  $f : X \rightarrow Y$  be absolutely continuous. If  $\sigma : X \times X \rightarrow \mathbb{R}$  is defined by  $\sigma(x, y) = \rho(x, y) + \theta(f(x), f(y))$ , then  $(X, \rho)$  and  $(X, \sigma)$  have the same topologies and

$$f : (X, \sigma) \rightarrow (Y, \theta)$$

is absolutely continuous. For  $X = Y = \mathbb{R}$ , taken with the absolute value metric, this definition of absolute continuity characterizes the class of Lipschitz functions, which, in turn, is properly contained in the class of absolutely continuous functions on  $\mathbb{R}$  as defined in Chapter 4.

**Example A.5.4. Comparison between absolute and uniform continuity**

Let  $(X, \rho)$  and  $(Y, \theta)$  be metric spaces and let  $f : X \rightarrow Y$  be a continuous function. We shall show that it is not generally possible to find metrics  $\sigma$  and  $\tau$  on  $X$  and  $Y$ , respectively, so that  $f : (X, \sigma) \rightarrow (Y, \tau)$  is absolutely and uniformly continuous. Take  $f : (0, 1] \rightarrow [1, \infty)$ ,  $f(x) = 1/x$ , with the usual metrics. Assume we can find  $\sigma, \tau$  which yield both absolute and uniform continuity. Then, from Theorem A.5.2,  $f$  has a unique uniformly continuous extension  $[0, 1] \rightarrow [1, \infty)$ , and this is obviously false.

## A.6 Baire category theorem

An excellent reference for the Baire category theorem is [352]. A metric space is *Baire* if every countable intersection of open dense sets is dense. Since  $\mathbb{R}$  is

Insert the following Remark immediately before  
Example A.10.4 and just after the ~~short~~ <sup>6/29/08</sup>  
outline of proof of Theorem A.10.3

IMA

Remark Let  $L \in \mathcal{L}(X, Y)$  be injective. If  
 $L$  is not surjective, ~~then we would have~~  
~~that  $L(X) = Y$~~  then  $L(X)$  is of first  
category in  $Y$ . To verify this fact, we  
assume  $L(X)$  is of second category in  $Y$ . Then,  
it is not difficult to verify that

(\*)  $\forall r > 0, \exists R > 0$  such that  $\overline{B(0, R)} \subseteq L(\overline{B(0, r)})$ .

no P Take ~~any~~ any  $y \in Y$ . For any  $r > 0$ , choose  $R$  as  
in (\*). There is  $N_R$  such that

$$\forall n > N_R, \frac{1}{n}y \in \overline{B(0, R)} \subseteq L(\overline{B(0, r)}).$$

Then, there is  $x \in \overline{B(0, r)}$  ~~such that~~

(for any such  $n$ ),  
for which  $L(nx) = \frac{1}{n}y$ . We conclude  
that  $L(X) = Y$ .

a complete metric space, Theorem A.6.1 and Theorem A.6.2b yield the fact that  $\mathbb{R}$  is not a set of first category, see (1.12).

**Theorem A.6.1. Baire category theorem I**

*Every complete metric space  $X$  is Baire.*

*Proof. i.* We give CANTOR's necessary conditions for the completeness of a metric space, as promised in Section 2.1. (The converse is true and easy.) Take  $\{A_n : n = 1, \dots\} \subseteq X$  where each  $A_n$  is closed, non-empty, and  $A_1 \supseteq A_2 \supseteq \dots$ . Assuming that  $\lim_{n \rightarrow \infty} \sup\{\rho(y, z) : y, z \in A_n\} = 0$  we verify that  $\bigcap A_n = \{x\} \subseteq X$  for some  $x \in X$ .

For all  $n$ , let  $x_n \in A_n$ . The sequence  $\{x_n : n = 1, \dots\}$  is Cauchy, for if  $m \geq n$ , then

$$\rho(x_m, x_n) \leq \sup\{\rho(y, z) : y, z \in A_n\} = \text{diam } A_n \rightarrow 0, \quad n \rightarrow \infty.$$

Here, for  $A \subseteq \mathbb{R}^d$ , we write

$$\text{diam } (A) = \sup \{|x - y| : x, y \in A\}.$$

By the completeness of  $X$  there is a point  $x \in X$  such that  $\rho(x_n, x) \rightarrow 0$ . Now, for each  $n$ ,  $x_m \in A_n$  when  $m$  is sufficiently large. Consequently,  $x \in \bigcap A_n$  since  $A_n$  is closed. If  $y \in \bigcap A_n$ , then  $\rho(x, y) \leq \text{diam } A_n$  for each  $n$  so that, by hypothesis,  $\rho(x, y) = 0$ . Thus,  $x = y$ .

*ii.* Let  $U_n$  be an open and dense subset of  $X$ . Thus,  $A_n = U_n^c$  is *nowhere dense*, i.e.,  $\text{int } A_n = \emptyset$ .  $(\bigcap U_n)^c = \bigcup A_n = A$  where each  $A_n$  is closed. We prove that if  $V$  is open, then

$$V \cap \left( \bigcap_{n=1}^{\infty} U_n \right) \neq \emptyset.$$

Choose an open set  $V_1$  such that  $\bar{V}_1 \subseteq V$  and  $\text{diam } \bar{V}_1 < 1$ . Since  $V_1$  is not a subset of  $A_1$ ,

$$V_1 \cap U_1 \neq \emptyset,$$

and  $V_1 \cap U_1$  is open.

Choose an open set  $V_2$  such that  $\bar{V}_2 \subseteq V_1 \cap U_1$  and  $\text{diam } \bar{V}_2 < 1/2$ . Generally, then, we choose open sets  $V_n$  with  $\bar{V}_n \subseteq V_{n-1} \cap U_{n-1}$  and  $\text{diam } \bar{V}_n < 1/n$ . The hypotheses of part *i* are satisfied for  $\bar{V}_n$ , and hence  $\bigcap \bar{V}_n = \{x\}$ . Therefore,

$$x \in \bigcap_{n=1}^{\infty} (V_n \cap U_n) \subseteq V_1 \cap \left( \bigcap_{n=1}^{\infty} U_n \right).$$

□



Let  $X$  be a metric space.  $A \subseteq X$  is a set of *first category* if it is the countable union of nowhere dense sets, i.e., sets having empty interior. Any other subset of  $X$  is a set of *second category*. RENÉ BAIRE introduced these notions in 1899. Among other results, he proved that the countable intersection of open dense sets (in  $\mathbb{R}$ ) is dense, and this is our definition of a Baire metric space. The following is straightforward to prove.

**Theorem A.6.2. Baire category theorem II**

*The following are equivalent for a metric space  $X$ .*

- a.  $X$  is Baire.
- b. Every countable union of closed nowhere dense sets has empty interior.
- c. Every non-empty open set is of second category.
- d. If  $\bigcup A_n$ ,  $A_n$  closed, contains an open set, then some  $A_j$  contains an open set.
- e. The complement of every set of first category is dense in  $X$ .

**Example A.6.3. Sets of first and second category**

a. First category sets are not necessarily nowhere dense. In fact, take  $\mathbb{Q} \subseteq \mathbb{R}$  noting that  $\mathbb{Q}$  is of first category and  $\overline{\mathbb{Q}} = \mathbb{R}$ .

b. Let  $S \subseteq \mathbb{R}$ . If  $\{x - y : x, y \in S\}$  is a set of first category, then  $S$  is a set of first category; and so, if  $S$  is of second category, then  $\{x - y : x, y \in S\}$  is of second category.

c. It is easy to construct a first category set of Lebesgue measure 1 in  $[0, 1]$ . Let  $E_n$  be a perfect symmetric set with  $m(E_n) \geq 1 - (1/n)$ . Then,  $E = \bigcup E_n$  does the trick, cf., Problem 2.9 and Problem 2.10.

d. Clearly,  $[0, 1]$  does not contain a countable dense  $\mathcal{G}_\delta$ ,  $D = \bigcap U_j$ . In fact, if  $D = \{d_j : j = 1, \dots\}$  were such a set, then  $V_j = U_j \setminus (\bigcup_{n=1}^j d_n)$  is open and dense, and  $\bigcap V_j = \emptyset$ . This contradicts Baire category theorem I.

**Example A.6.4. Open coverings of accessible points**

Let  $E \subseteq [0, 1]$  be any perfect symmetric set. As such, it is associated with a countable set  $A$  of accessible points. ( $a \in A \subseteq \mathbb{R}$  is *accessible* if it is the endpoint of a contiguous open interval.) Note that if  $\{U_\alpha\}$  is an open covering of  $A$ , it does not necessarily follow that  $E \subseteq \bigcup U_\alpha$ . For example, if  $x \in E \setminus A$  consider  $[0, x) \cup (x, 1]$ . For each  $a_n \in A$  let  $\{I_{m,n} : m = 1, \dots\}$  be a sequence of open intervals about  $a$  whose lengths tend to 0. Then,  $\{a_n : n = 1, \dots\} = \bigcap_{m=1}^{\infty} I_{m,n}$ . Now let  $V_m = \bigcup_{n=1}^{\infty} I_{m,n}$  so that  $E \cap V_m$  is open and dense in  $E$ . Observe that  $A \subseteq \bigcap_{m=1}^{\infty} (E \cap V_m)$ , properly. To prove this note that  $U_m = E \cap V_m \setminus \{a_1, \dots, a_m\}$  is open and dense in  $E$  so that  $\bigcap U_m$  is dense. On the other hand,  $A \cap (\bigcap U_m) = \emptyset$  and  $\bigcap (E \cap V_m) = A \cup (\bigcap U_m)$ .

**Example A.6.5. A complete non-metric space**

Let  $C_c(\mathbb{R})$  be the vector space of continuous functions  $f : \mathbb{R} \rightarrow \mathbb{C}$  which vanish outside of some compact set, depending on  $f$ . We define sequential convergence in  $C_c(\mathbb{R})$  as follows:

$$f_n \rightarrow f \text{ in } C_c(\mathbb{R}), f_n, f \in C_c(\mathbb{R}), \text{ if } \|f_n - f\|_\infty \rightarrow 0 \text{ and} \\ \exists r > 0 \text{ such that } \forall n, f_n = 0 \text{ on } [-r, r]^\sim.$$

We shall prove that, with this convergence,  $C_c(\mathbb{R})$  cannot be a complete metric space  $(C_c(\mathbb{R}), \rho)$ . If such a metric  $\rho$  exists, then  $C_c(\mathbb{R})$  is a Baire space. We shall show that  $(C_c(\mathbb{R}), \rho)$  is of first category to obtain the contradiction.

First, note that

$$C_c(\mathbb{R}) = \bigcup_{n=1}^{\infty} C_{c,n}, \quad C_{c,n} = \{f \in C_c(\mathbb{R}) : f = 0 \text{ on } [-n, n]^\sim\}.$$

Clearly,  $\overline{C_{c,n}} = C_{c,n}$ , and it is sufficient to check that  $\text{int } C_{c,n} = \emptyset$ . Assume not, and let  $V \subseteq C_{c,n}$  be an open neighborhood of 0 in  $C_c(\mathbb{R})$ . Choose  $f_k \in C_{c,n+1} \setminus C_{c,n}$  such that  $\rho(f_k, 0) \rightarrow 0$ . Consequently,  $f_k \in V \subseteq C_{c,n}$ , and this contradicts the definition of  $f_k$ . There is, in fact, a (completely regular) topology on  $C_c(\mathbb{R})$  whose uniform structure renders  $C_c(\mathbb{R})$  complete and whose sequential convergence is that given above, cf., Section 7.3.

**Example A.6.6. Everywhere continuous nowhere differentiable functions**

In Chapter 1 we discussed everywhere continuous nowhere differentiable functions. The soft analysis proof of their existence uses Baire category theorem I. Take  $C([0, 1])$  with the  $\|\dots\|_\infty$  norm so that  $C([0, 1])$  is complete with the metric  $\rho(f, g) = \|f - g\|_\infty$ . Define

$$F_n = \left\{ f \in C([0, 1]) : \exists x \in [0, 1] \text{ such that } \forall h > 0, \left| \frac{f(x+h) - f(x)}{h} \right| < n \right\}.$$

Each  $F_n$  is closed and nowhere dense, and so  $C([0, 1]) \neq \bigcup F_n$ . Consequently, the set of continuous nowhere differentiable functions is dense in  $C([0, 1])$ .

**Example A.6.7. Sets  $A$  such that  $0 < \mu(A \cap I) < \mu(I)$  for all  $I$**

The proof of Problem 2.45b is elementary. First, let  $A_1 \subseteq [0, 1]$  be a perfect symmetric set of measure  $1/4$ . Then, let  $A_2 = \bigcup_{j=1}^{\infty} A_{j,2}$ , where the measure of  $A_2$  is  $1/8$  and where each  $A_{j,2}$  is a perfect symmetric set of positive measure in the  $j$ th contiguous interval of  $A_1$ . Define all of the  $A_j$  in this way, and set  $A = \bigcup_{j=1}^{\infty} A_j$ . A generalization of this result is due to R. B. KIRK [277]: *Let the measure space  $(X, \mathcal{A}, \mu)$  be a separable metric space with metric  $\rho$ , assume  $\mu$  is continuous, and suppose  $\mathcal{B}(X) \subseteq \mathcal{A}$ ; then there is  $A \in \mathcal{B}(X)$  such that for each open set  $I$  of positive measure we have*

$$0 < \mu(A \cap I) < \mu(I).$$

**A.7 Uniform Boundedness Principle and Schur lemma**

The Uniform Boundedness Principle, also known as the Banach–Steinhaus theorem, is one of fundamental results in functional analysis. The first result of this type was proved by BANACH and STEINHAUS in 1927 [24].

**Theorem A.7.1. Banach–Steinhaus Uniform Boundedness Principle**

Let  $(X, \rho)$  be a complete metric space and let  $\mathcal{F}$  be a set of continuous functions  $X \rightarrow \mathbb{C}$ . Assume

$$\forall x \in X, \exists M_x > 0 \text{ such that } \forall f \in \mathcal{F}, |f(x)| \leq M_x.$$

Then, there is a non-empty open set  $U \subseteq X$  and a constant  $M$  such that

$$\forall x \in U \text{ and } \forall f \in \mathcal{F}, |f(x)| \leq M.$$

*Proof.* For each  $f \in \mathcal{F}$  and  $m \in \mathbb{N}$ , define

$$A_{m,f} = \{x : |f(x)| \leq m\} \text{ and } A_m = \bigcap_{f \in \mathcal{F}} A_{m,f}.$$

Since  $f$  is continuous,  $A_m$  is closed. We show that  $X = \bigcup A_m$ . In fact, if  $x \in X$  choose  $m = M_x$ , so that  $x \in A_m$ .

Consequently, from Baire category theorems I and II,  $U = \text{int } A_n \neq \emptyset$  for some  $n$  and we take  $M = n$ .

□

See Theorem A.8.6 for a statement of the Uniform Boundedness Principle in terms of Banach spaces.

As noted after Definition 6.3.1, sequential weak convergence in  $L^1_\mu(X)$  is actually sequential convergence for a certain topology (called the *weak topology*) on  $L^1_\mu(X)$ . We shall discuss the weak topology generally in Appendix A.9 but for now consider a special result for the case of  $\ell^1(\mathbb{N})$ . This result, Theorem A.7.3 (the Schur lemma), is studied in greater detail in Chapter 6. The proof we outline uses the Baire category theorems. To formulate the Schur lemma we need the following definition.

**Definition A.7.2. The weak topology  $\sigma(\ell^1(\mathbb{N}), \ell^\infty(\mathbb{N}))$** 

Let  $F \subseteq \ell^\infty(\mathbb{N})$  be a finite set and let  $x \in \ell^1(\mathbb{N})$ . Define

$$U(F, x, \varepsilon) = \{y \in \ell^1(\mathbb{N}) : \forall x' \in F, |x'(y - x)| < \varepsilon\},$$

where, if  $x' = \{x'_j\} \subseteq \mathbb{C}$ ,  $y = \{y_j\} \subseteq \mathbb{C}$ , and  $x = \{x_j\} \subseteq \mathbb{C}$ , then the operation of  $x'$  on  $(y - x)$  is  $x'(y - x) = \sum_{j=1}^\infty x'_j(y_j - x_j)$ . The family  $\{U(F, x, \varepsilon) : F \subseteq \ell^\infty(\mathbb{N}), x \in \ell^1(\mathbb{N}), \text{ and } \varepsilon > 0\}$  is a basis for a topology  $\sigma(\ell^1(\mathbb{N}), \ell^\infty(\mathbb{N}))$  on  $\ell^1(\mathbb{N})$ . This is the *weak topology* for  $\ell^1(\mathbb{N})$ .

If  $U$  is open for the  $\|\dots\|_1$  topology on  $\ell^1(\mathbb{N})$ , then  $U \in \sigma(\ell^1(\mathbb{N}), \ell^\infty(\mathbb{N}))$ . It is not difficult to verify that

$$S = \left\{ x \in \ell^1(\mathbb{N}) : \sum_{j=1}^\infty |x_j| = 1 \right\} \text{ is } \|\dots\|_1 \text{ closed}$$

and

the weak closure of  $S$  is the  $\|\dots\|_1$  closure of  $B(0, 1)$ ,

where the ball  $B(0, 1)$  is defined in terms of  $\|\dots\|_1$ . Consequently, the topology  $\sigma(\ell^1(\mathbb{N}), \ell^\infty(\mathbb{N}))$  is *strictly* weaker than the norm topology on  $\ell^1(\mathbb{N})$ . Note that sequential weak convergence for  $\ell^1(\mathbb{N})$  is precisely the analogue for  $\mathbb{N}$  of that defined for  $[0, 1]$  immediately after Theorem A.7.1.

The Schur lemma [411] tells us that  $\sigma(\ell^1(\mathbb{N}), \ell^\infty(\mathbb{N}))$  and  $\|\dots\|_1$  yield the same convergent sequences in  $\ell^1(\mathbb{N})$ , see [19], pages 137–139, cf., [210], Section 3.2, and [451], pages 327–329, for selected results from [411].

**Theorem A.7.3. Schur lemma**

If  $\{x^{(n)} : n = 1, \dots\} \subseteq \ell^1(\mathbb{N})$  converges to 0 in  $\sigma(\ell^1(\mathbb{N}), \ell^\infty(\mathbb{N}))$ , then  $\|x^{(n)}\|_1 \rightarrow 0$ .

*Proof.* Let  $Y = \{x' \in \ell^\infty(\mathbb{N}) : \sup |x'_j| \leq 1\}$  and define

$$\forall x', y' \in Y, \quad \rho(x', y') = \sum_{j=1}^{\infty} \frac{|x'_j - y'_j|}{2^j}.$$

$(Y, \rho)$  is complete, and sets of the form

$$\forall x' \in Y, \quad S_{J, \delta} = \{y' : |x'_j - y'_j| < \delta, |j| \leq J\}$$

are a basis at  $x'$  for the topology of  $(Y, \rho)$ . Next we define

$$\forall \varepsilon > 0 \text{ and } \forall m, \quad A_m = \left\{ y' \in Y : \forall n \geq m, \left| \sum_{j=1}^{\infty} y'_j x_j^{(n)} \right| \leq \varepsilon, \right\}.$$

It can be shown that  $A_m$  is closed in  $(Y, \rho)$  and  $Y = \bigcup A_m$ , so that by Baire category theorems I and II there is  $m$  such that  $\text{int } A_m \neq \emptyset$ . From this point it is straightforward to prove that  $\|x^{(n)}\|_1 \rightarrow 0$ . □

## A.8 Hahn–Banach theorem

Our presentation of the Hahn–Banach theorem (Theorem A.8.3) is standard. There are basically three distinct parts to the proof. The first and crucial step is Lemma A.8.2 which allows us to extend continuous linear functionals from a closed subspace  $Y$  to the closed subspace generated by  $Y$  and an element  $x$  (the setting here is necessarily with real vector spaces). Second, an axiom of choice argument is used to expand this finite procedure to extend maps in the infinite dimensional case. Finally, an ingenious trick due to HENRI F. BOHNENBLUST and SOBECZYK yields the result for the complex case.

**Remark.** Let  $X$  and  $Y$  be vector spaces over  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{F} = \mathbb{C}$ , let  $V \subseteq X$  be a linear subspace, and let  $L : V \rightarrow Y$  be a linear function. Then, there exists  $K : X \rightarrow Y$  such that  $K$  is linear on  $X$  and  $K = L$  on  $V$ . In infinite dimensions the proof requires the axiom of choice, usually in the form of the Zorn lemma, see Section 2.6.1. The general problem of extending continuous linear functions  $L : Z \rightarrow Y$ ,  $Z \subseteq X$ , is usually intractable.

Our setting for Appendix A.8 – Appendix A.11 will be non-zero normed vector spaces, although many of the results are true for Hausdorff locally convex topological vector spaces. There are just a few cases where we need this generality, e.g., in Chapter 7, so we have chosen to be efficient space-wise (sic), at least in this case, and not write-out all the details. See BOURBAKI's *Espaces Vectorielles Topologiques* or [235] for classic and classical presentations.

**Proposition A.8.1.** *Let  $X$  and  $Y$  be normed vector spaces, and let  $L : X \rightarrow Y$  be a linear function.*

- a. *Either  $L$  is continuous at every  $x \in X$  or at no  $x \in X$ .*
- b.  *$L$  is continuous on  $X$  if and only if*

$$\exists C > 0 \text{ such that } \forall x \in X, \quad \|L(x)\| \leq C\|x\|.$$

*Proof.* Part a and the sufficient condition for continuity in part b are straightforward.

For the necessary condition in part b, assume  $L$  is continuous at 0. Thus,

$$\exists \delta > 0 \text{ such that } \|x\| < \delta \implies \|L(x)\| < 1.$$

If  $x \neq 0$ , let  $x_\delta = (\delta x)/(2\|x\|)$ , and so  $\|L(x_\delta)\| < 1$ . Hence,

$$\|L(x)\| = \frac{2}{\delta} \|x\| \|L(x_\delta)\| < \frac{2}{\delta} \|x\|,$$

and so we set  $C = 2/\delta$ . □

If  $X$  and  $Y$  are normed vector spaces,  $Z \subseteq X$  is a linear subspace, and  $L : Z \rightarrow Y$  is linear, we define

$$\|L\| = \sup\{\|L(x)\| : \|x\| \leq 1, x \in Z\}. \quad (\text{A.7})$$

Thus,  $\|L\|$  is the smallest constant  $C$  such that  $\|L(x)\| \leq C\|x\|$  for all  $x \in Z$ . If  $\|L\| < \infty$  then, because of Proposition A.8.1, we say that  $L$  is a *continuous* or *bounded linear function*  $Z \rightarrow Y$ . Clearly,  $\|L\|$  depends on the subspace  $Z$ . This is important in what follows.

The space of continuous linear functions  $X \rightarrow \mathbb{C}$  is denoted by  $X'$ .  $X'$  is the *dual space* of  $X$ , and its elements are usually called *continuous* or *bounded linear functionals*.

Hilbert spaces  $H$  have the property that

$$H' = H. \quad (\text{A.8})$$

(A.8) is the Riesz representation theorem for the case of Hilbert spaces, see the last comment in Section 7.1.

**Lemma A.8.2.** *Let  $X$  be a real normed vector space,  $Y \subseteq X$  a closed linear subspace, and  $Z$  the closed linear subspace of  $X$  generated by  $Y$  and some  $x \in X \setminus Y$ . If  $L : Y \rightarrow \mathbb{R}$  is linear and continuous, then there is a continuous linear functional  $K : Z \rightarrow \mathbb{R}$  such that  $K = L$  on  $Y$  and  $\|K\| = \|L\|$ .*

*Proof.* If  $x, y \in Y$ , then

$$L(x) - L(y) \leq \|L\|\|x + z\| + \|L\|\|y + z\|;$$

and so

$$\sup_{u \in Y} (-\|L\|\|u + z\| - L(u)) = a \leq b = \inf_{u \in Y} (\|L\|\|u + z\| - L(u)).$$

For fixed  $c \in [a, b]$  we define  $K(y + rz) = L(y) + rc$ , where  $r \in \mathbb{R}$ ,  $y \in Y$ , and  $\{y + rz : r \in \mathbb{R}, y \in Y\} = Z$ . □

### Theorem A.8.3. Hahn–Banach theorem

**a.** *Let  $Y \subseteq X$  be a linear subspace of the normed vector space  $X$ , and assume  $L : Y \rightarrow \mathbb{C}$  is linear and continuous. Then, there is  $K \in X'$  such that  $K = L$  on  $Y$  and  $\|K\| = \|L\|$ .*

**b.** *If  $Y \subseteq X$  is a closed linear subspace of the normed vector space  $X$  and  $z \notin Y$ , then there is  $L \in X'$  such that  $L(z) \neq 0$  and  $L = 0$  on  $Y$ .*

*Proof.* **a.i.** We choose  $Y$  to be closed without any loss of generality. In fact, it is easy to extend  $L$  to  $\bar{Y}$  by Theorem A.5.2.

**ii.** We now prove part **a** for the real case, assuming that  $Y$  is closed and that  $Y \subseteq X$  properly.

Let  $\mathcal{L}$  be the family of all continuous linear functions  $K : Z \rightarrow \mathbb{R}$  such that  $Y \subseteq Z$ ,  $K = L$  on  $Y$ , and  $\|K\| = \|L\|$ . From Lemma A.8.2,  $\mathcal{L}$  is non-trivial. We order  $\mathcal{L}$  by setting  $K \leq K_1$  if  $Z \subseteq Z_1$  and  $K_1 = K$  on  $Z$ . From the Zorn lemma (Section 2.6.1), i.e., the axiom of choice, there is a maximal element  $K : Z \rightarrow \mathbb{R}$  and we easily check that  $Z = X$ .

**iii.** Let  $W$  be a complex vector space. If  $K : W \rightarrow \mathbb{C}$  is real linear then  $K$  is complex linear if and only if  $K(ix) = iK(x)$ . Let  $L : Y \rightarrow \mathbb{C}$  be complex linear, as in part **a**. Set  $L_1 = \operatorname{Re} L$ ,  $L_2 = \operatorname{Im} L$ , and note that  $L$  is real linear. Thus,  $L(iy) = iL(y)$  on  $Y$ , and, using this fact, we compute that

$$\forall y \in Y, \quad L_2(y) = -L_1(iy).$$

Because of part **ii** we can extend  $L_1$  to  $K_1$  on  $X$ , considered as a real vector space, such that  $\|L_1\| = \|K_1\|$ . Set  $K(x) = K_1(x) - iK_1(ix)$  on  $X$ .

Similar computations show that  $K$  has the desired properties.

**b.** Part *b* is a consequence of part *a*. Indeed, define  $L_z(y + rz) = r$ , where  $y + rz$ , for  $y \in Y$  and  $r \in \mathbb{R}$ , is a typical element of the closed linear subspace generated by  $Y$  and  $z$ . Note that

$$a = \inf_{y \in Y} \|z + y\| > 0$$

and

$$|L_z(y + rz)| \leq \frac{1}{a} \|y + rz\|.$$

Thus, we apply part *a* directly. □

**Remark.** The Hahn–Banach theorem allows us to assert that if  $\{x_n : n = 1, \dots\} \subseteq X$ , a Banach space, then  $\overline{\text{span}} \{x_n\} = X$  if and only if, whenever  $L(x_n) = 0$  for all  $n$  for any given  $L \in X'$ , we can conclude that  $L = 0$ .

By (A.8), the equivalent assertion for a Hilbert space  $H$  is that  $\overline{\text{span}} \{x_n\} = H$  if and only if, whenever  $\langle y, x_n \rangle = 0$  for all  $n$  and any given  $y \in H$ , we can conclude that  $y = 0$ .

**Example A.8.4.  $L^p$ -duality**

**a.** Let  $1 \leq p < \infty$ , let  $1/p + 1/q = 1$ , and let  $\mu$  be a  $\sigma$ -finite measure on  $\mathbb{R}$ . Then,  $(L_\mu^p(\mathbb{R}))' = L_\mu^q(\mathbb{R})$ , where  $g : L_\mu^p \rightarrow \mathbb{C}$  is well-defined by  $g(f) = \int_{\mathbb{R}} f(t)\overline{g(t)} d\mu(t)$  for all  $f \in L_\mu^p(\mathbb{R})$ , see Theorem 5.5.5. In particular, the Hilbert space  $H = L_\mu^2(\mathbb{R})$  has the property that  $(L_\mu^2(\mathbb{R}))' = L_\mu^2(\mathbb{R})$ .

**b.** Let  $(\mathbb{R}, \mathcal{A}, \mu)$  be a measure space. Then, according to Theorem 5.5.7,  $(L_\mu^\infty(\mathbb{R}))'$  is the space of complex valued finitely additive bounded set functions on  $\mathcal{A}$ , see also Example A.11.3 as well as [146], part I, Chapter IV, Section 8.

**Example A.8.5. Sufficiently many elements in  $X'$**

Let  $X$  be a normed vector space, and choose  $x, y \in X$ ,  $x \neq y$ . By the Hahn–Banach theorem we see that there is  $L \in X'$  such that  $L(x) \neq L(y)$ . In fact, let  $Y$  be the linear subspace generated by  $x - y$ , define  $K(r(x - y)) = r\|x - y\|$ , observe that  $\|K\| = 1$ , and use Theorem A.8.3.

The contrapositive equivalent assertion for  $x \in X$  is the following:

$$\forall L \in X', \quad L(x) = 0 \implies x = 0.$$

The following restatement of Theorem A.7.1 does not require the Hahn–Banach theorem, but it does use the terminology defined in this section.

**Theorem A.8.6. Uniform Boundedness Principle for Banach spaces**

Let  $X$  be a Banach space, let  $Y$  be a normed vector space, and let  $\mathcal{L}$  be a set of continuous linear functions  $X \rightarrow Y$ . Assume that

$$\forall x \in X, \exists C_x > 0 \text{ such that } \forall L \in \mathcal{L}, \quad \|L(x)\| \leq C_x.$$

Then,

$$\exists C > 0 \text{ such that } \forall L \in \mathcal{L}, \quad \|L\| \leq C. \quad (\text{A.9})$$

**Corollary A.8.7.** *Let  $X$  be a Banach space, let  $Y$  be a normed vector space, and let  $\{L_n : n = 1, \dots\}$  be a sequence of continuous linear functions  $X \rightarrow Y$ . Assume*

$$\forall x \in X, \exists L(x) \in Y \text{ such that } \lim_{n \rightarrow \infty} \|L_n(x) - L(x)\| = 0.$$

Then,  $L : X \rightarrow Y$  is a continuous linear function.

**Corollary A.8.8.** *Let  $X$  be a Banach space and let  $\{y_k : k = 1, \dots\} \subseteq X'$ ,  $y \in X'$ . The following are equivalent.*

- a.  $\sup_{k \geq 1} \|y_k\| < \infty$  and  $y_k \rightarrow y$  on a dense subset of  $X$ .
- b.  $y_k \rightarrow y$  uniformly on each compact subset of  $X$ .
- c.  $\forall x \in X, y_k(x) \rightarrow y(x)$ .

A typical application of Theorem A.8.6 is for the case  $Y = \mathbb{C}$  and  $\mathcal{L} \subseteq X'$ .

Theorems A.7.1 and A.8.6, as well as Corollaries A.8.7 and A.8.8, can be formulated in somewhat different settings, e.g., [134], page 83, [393], pages 43–46. Corollary A.8.7 is a useful form of the Banach–Steinhaus theorem. BANACH and STEINHAUS' original assertion in 1927 is more general, see [19], pages 79–80.

**Remark.** Given the setting but not the assumption of Theorem A.8.6. Then, the Uniform Boundedness Principle is the dichotomous assertion: Either (A.9) holds or there is a non-empty set  $Z \subseteq X$  for which  $\bar{Z} = X$  and

$$\forall x \in Z, \sup_{L \in \mathcal{L}} \|L(x)\| = \infty.$$

$Z$  is also the intersection of a countable family of open sets.

**Example A.8.9. Computation of  $\|L\|$**

**a.i.** Let  $X$  and  $Y$  be normed vector spaces, let  $Z \subseteq X$  be a linear subspace, and let  $L : Z \rightarrow Y$  be a non-zero linear function. The quantity  $\|L\|$  defined by (A.7) can also be written as

$$\|L\| = \sup\{\|L(x)\| : \|x\| = 1, x \in Z\} \quad (\text{A.10})$$

and

$$\|L\| = \sup\left\{\frac{\|L(x)\|}{\|x\|} : x \in Z \setminus \{0\}\right\}. \quad (\text{A.11})$$

**a.ii.** We shall verify the assertions of part **a.i.** If  $x \in Z \setminus \{0\}$ , then

$$\frac{\|L(x)\|}{\|x\|} = \left\|L\left(\frac{x}{\|x\|}\right)\right\| \leq \sup\{\|L(z)\| : \|z\| = 1, z \in Z\};$$



and, if  $x \in Z$ ,  $\|x\| = 1$ , then

$$\|L(x)\| = \frac{\|L(x)\|}{\|x\|} \leq \sup \left\{ \frac{\|L(y)\|}{\|y\|} : y \in Z \setminus \{0\} \right\}.$$

Thus, the right sides of (A.10) and (A.11) are equal.

Next, label either of these suprema as  $r > 0$ , and let  $\|L\|$  be defined by (A.7). For any  $r > \varepsilon > 0$ , we can choose  $y \in Z \setminus \{0\}$  such that  $(r - \varepsilon)\|y\| < \|L(y)\|$  by (A.11). Thus,  $(r - \varepsilon) \leq \|L\|$  and so  $r \leq \|L\|$ .

We shall assume  $r < \|L\|$  and obtain a contradiction. Since  $\|L\| - r = p > 0$  we have  $r < \|L\| - p/2$  so that

$$\forall x \in Z \setminus \{0\}, \quad \frac{\|L(x)\|}{\|x\|} \leq r < \|L\| - \frac{p}{2},$$

i.e.,  $\|L(x)\| < (\|L\| - p/2)\|x\|$ . This contradicts the definition of  $\|L\|$  as the smallest constant  $C$  for which (A.7) holds.

**b.** The following situation frequently arises and the result is useful. Let  $X$  be a Banach space over  $\mathbb{C}$ , let  $Z \subseteq X$  be a dense linear subspace, and let  $L : Z \rightarrow \mathbb{C}$  be a linear function for which

$$r = \sup \left\{ \frac{|L(x)|}{\|x\|} : x \in Z \setminus \{0\} \right\} < \infty.$$

Then,  $L \in X'$  and  $\|L\| = r$ . The proof is not difficult and first requires proving that  $L$  is a well-defined linear function  $X \rightarrow \mathbb{C}$ . The hypotheses to the claim can also be weakened.

**c.** Let  $X = \mathbb{C}$  and let  $Z = \mathbb{C} \setminus \{z \in \mathbb{C} : |z| = 1\}$ . Let  $L \in X' \setminus \{0\}$ . Note that with  $\|L\|$  defined by (A.7), we have

$$\|L\| = \sup_{y \in Z \setminus \{0\}} \frac{|L(y)|}{\|y\|} \neq \sup_{y \in Z, \|y\|=1} |L(y)|$$

since  $\{y \in Z : \|y\| = 1\} = \emptyset$ .

**Example A.8.10. Hilbert–Schmidt operators and Schur lemma**

**a.i.** Let  $(X, \mathcal{A}, \mu)$  be a  $\sigma$ -finite measure space, and let  $\mu \times \mu$  be the corresponding product measure on  $X \times X$ . If  $K \in L^2_{\mu \times \mu}(X \times X)$ , then we define the operator  $L$  as

$$(L(f))(x) = \int_X K(x, y)f(y) d\mu(y).$$

It is not difficult to prove that  $L \in \mathcal{L}(L^2_\mu(X))$ , the space of continuous linear functions  $L^2_\mu(X) \rightarrow L^2_\mu(X)$ , where  $L^2_\mu(X)$  is given the  $L^2$ -norm, see Appendix A.10. In fact, one makes the estimate,

$$\|L\|^2 \leq \int_X \int_X |K(x, y)|^2 d\mu(x)d\mu(y),$$

using the definition of  $\|L\|$ .  $L$  is a *Hilbert–Schmidt integral operator*.

**a.ii.** A natural generalization of the notion of Hilbert–Schmidt integral operators are *Hilbert–Schmidt operators* acting on a separable Hilbert space  $H$ . We say that  $A \in \mathcal{L}(H)$ , the space of continuous linear functions on  $H$ , is a *Hilbert–Schmidt operator* if there exists an ONB  $\{e_n : n = 1, \dots\}$  for  $H$  such that

$$\sum_{n=1}^{\infty} \|A(e_n)\|^2 < \infty.$$

We define the *Hilbert–Schmidt norm* of  $A$  to be

$$\|A\|_{HS} = \left( \sum_{n=1}^{\infty} \|A(e_n)\|^2 \right)^{1/2}.$$

**b.i.** We have seen versions of the Schur lemma in Theorem 6.2.1 and Theorem A.7.3. In this realm of ideas, SCHUR proved the following result. If  $\{c_{m,n} : m, n \in \mathbb{Z}\}$  is a double sequence of complex numbers with the properties,

$$\sum_{n \in \mathbb{Z}} |c_{m,n}| \leq C_1, \quad \text{independent of } m$$

and

$$\sum_{m \in \mathbb{Z}} |c_{m,n}| \leq C_2, \quad \text{independent of } n,$$

then the linear operator  $L$  defined by the matrix,  $(c_{m,n})_{m,n \in \mathbb{Z}}$ , is an element of  $\mathcal{L}(\ell^2(\mathbb{Z}))$ . In fact,

$$\|L\|^2 \leq C_1 C_2.$$

**b.ii.** Using the Schur lemma from part *b.i* and the Fubini theorem, we shall prove that  $L \in \mathcal{L}(L^2_\mu(X))$ , where the hypothesis on  $K$  from part *a* is replaced by the conditions,

$$\exists C_1 > 0 \text{ such that } \forall x \in X, \quad \int_X |K(x, y)| \, d\mu(y) \leq C_1$$

and

$$\exists C_2 > 0 \text{ such that } \forall y \in X, \quad \int_X |K(x, y)| \, d\mu(x) \leq C_2.$$

The proof is based on the following calculation for  $f, g \in L^2_\mu(X)$ .

$$\begin{aligned}
& \left( \int_X \int_X |K(x, y)| |f(x)| |g(y)| d\mu(x) d\mu(y) \right)^2 \\
& \leq \left( \int_X \int_X |K(x, y)| |f(x)|^2 d\mu(x) d\mu(y) \right) \left( \int_X \int_X |K(x, y)| |g(y)|^2 d\mu(y) d\mu(x) \right) \\
& = \left( \int_X \left[ \int_X |K(x, y)| d\mu(y) \right] |f(x)|^2 d\mu(x) \right) \\
& \times \left( \int_X \left[ \int_X |K(x, y)| d\mu(x) \right] |g(y)|^2 d\mu(y) \right) \leq C_1 C_2 \|f\|_2^2 \|g\|_2^2.
\end{aligned}$$

The following result depends on the compactness criteria in terms of sets being totally bounded (Appendix A.5) as well as the Uniform Boundedness Principle.

**Theorem A.8.11. Compact subsets of a Banach space**

Let  $X$  be a Banach space. A set  $Y \subseteq X$  is compact in  $X$  if and only if for every sequence of linear functionals  $L_n : X \rightarrow \mathbb{C}$ , for which

$$\forall x \in X, \quad L_n(x) \rightarrow 0, \quad (\text{A.12})$$

we have

$$L_n \rightarrow 0 \text{ uniformly on } Y. \quad (\text{A.13})$$

*Proof.* We prove the necessary conditions for compactness, which only require  $X$  to be a normed vector space. By the Uniform Boundedness Principle, (A.12) implies that

$$\exists M \text{ such that } \forall n \in \mathbb{N}, \quad \|L_n\| \leq M,$$

and, in particular, each  $L_n \in X'$ .

Since  $Y$  is compact and, hence, totally bounded, we have

$$\forall \varepsilon > 0, \exists y_1, \dots, y_m \in Y \text{ such that}$$

$$\forall y \in Y, \exists y_{j(y)} \in \{y_1, \dots, y_m\} \text{ for which } \|y - y_{j(y)}\| < \frac{\varepsilon}{2M}.$$

By hypothesis,

$$\exists N = N(\varepsilon) \text{ such that } \forall n \geq N \text{ and } \forall k = 1, \dots, m, \quad \|L_n(y_k)\| < \frac{\varepsilon}{2}.$$

Thus,

$$\begin{aligned}
\forall n \geq N \text{ and } \forall y \in Y, \quad \|L_n(y)\| & \leq \|L_n(y_{j(y)})\| + \|L_n(y - y_{j(y)})\| \\
& < \frac{\varepsilon}{2} + M \frac{\varepsilon}{2M}.
\end{aligned}$$

This is the desired uniform convergence in  $Y$ . □

See [267], pages 300-301, for a proof of the sufficient conditions, cf., the Arzelà-Ascoli (Theorem A.4.3) and Kolmogorov compactness (Theorem 6.6.1) theorems.

### A.9 The weak and weak \* topologies

Let  $X$  be a normed vector space.  $X'$  is a Banach space normed by

$$\forall x' \in X', \quad \|x'\| = \sup\{|x'(x)| : \|x\| \leq 1\}. \quad (\text{A.14})$$

As such,  $X'$  is the *dual* of  $X$ . We then consider  $(X')' = X''$ , normed analogously, noting that  $X''$  is a Banach space and that  $X$  can be embedded isometrically and algebraically isomorphically onto a linear subspace of  $X''$ . The mapping defining this isomorphism is given by

$$\forall x \in X, \quad x(x') = x'(x).$$

It should be pointed out that the proof that the natural mapping  $X \rightarrow (X')^*$ ,  $x \mapsto L_x$ , defined by  $L_x(x') = x'(x)$ , is injective requires the Hahn–Banach theorem in the form of Example A.8.5. ( $(X')^*$  is the space of linear functions (functionals)  $X' \rightarrow \mathbb{C}$ .)

$X$  is *reflexive* if  $X = X''$  under this canonical mapping.

#### Theorem A.9.1. The norm in terms of the dual space

Let  $X$  be a normed vector space, and let

$$B' = \{x' \in X' : \|x'\| \leq 1\}.$$

Then,

$$\forall x \in X, \quad \|x\| = \sup\{|x'(x)| : x' \in B'\}.$$

In particular, for each fixed  $x \in X$ , the linear functional  $L_x : X' \rightarrow \mathbb{C}$ ,  $x' \mapsto x'(x)$ , is continuous so that  $L_x \in X''$  and  $\|L_x\| = \|x\|$ .

*Proof.* Let  $x \in X$ . It is a consequence of the Hahn–Banach theorem (Theorem A.8.3) that

$$\exists y' \in B' \text{ such that } y'(x) = \|x\|.$$

Also,

$$\forall x' \in B', \quad |x'(x)| \leq \|x\| \|x'\| \leq \|x\|.$$

Thus,

$$\|x\| = y'(x) = |y'(x)| \leq \sup\{|x'(x)| : x' \in B'\} \leq \|x\|,$$

and we have the result. □

#### Definition A.9.2. The weak and weak \* topologies

Let  $X$  be a normed vector space.

The *weak topology* on  $X$ , denoted by  $\sigma(X, X')$ , has a basis at  $0 \in X$  given by sets of the form

$$\{x \in X : |x'_j(x)| < \varepsilon, j = 1, \dots, n\},$$

where  $\varepsilon > 0$  and  $\{x'_1, \dots, x'_n\}$  is an arbitrary finite subset of  $X'$ . Similarly, we define  $\sigma(X', X'')$ . See [451], pages 149, 151–154, 227–231, for a clear rationale and exposition of the weak topology.

The *weak \* topology* on  $X'$  denoted by  $\sigma(X', X)$ , is defined analogously with corresponding sets

$$\{x' \in X' : |x'(x_j)| < \varepsilon, j = 1, \dots, n\},$$

where  $\varepsilon > 0$  and  $x_j \in X, j = 1, \dots, n$ . Clearly,  $\sigma(X', X)$  is generally weaker than  $\sigma(X', X'')$ , i.e.,  $\sigma(X', X) \subseteq \sigma(X', X'')$ .

The following theorem is a consequence of the Hahn–Banach theorem, a finite dimensional algebraic result, and the definitions of the weak and weak \* topologies, see [386], pages 31–33, for a most efficient proof. It is also true, with analogous proof, for Hausdorff locally convex topological vector spaces (LCTVSSs).

**Theorem A.9.3. Weak and weak \* dual spaces**

Let  $X$  be a normed vector space with dual space  $X'$ .

- a. The dual space of  $X$  taken with the weak topology  $\sigma(X, X')$  is  $X'$ .
- b. The dual space of  $X'$  taken with the weak \* topology  $\sigma(X', X)$  is  $X$ .

$K \subseteq X$ , a vector space, is *convex* if, for each  $x, y \in K$  and  $0 \leq r \leq 1$ ,

$$rx + (1 - r)y \in K.$$

An important application of Theorem A.8.3 is the following fact.

**Theorem A.9.4. Equivalent norm and weak closures**

Let  $X$  be a normed vector space and let  $K \subseteq X$  be convex. Then,  $K$  has the same norm and  $\sigma(X, X')$  closure.

BANACH proved the following result for the case of separable spaces in 1932 [19], Chapter VIII, Theorem 3. The general version was obtained by LEONIDAS ALAOGU in 1940 [3].

**Theorem A.9.5. Banach–Alaoglu theorem**

Let  $X$  be a normed vector space. Then,  $B'$  is weak \* compact.

*Proof.* For each  $x \in X$  define

$$D_x = \{z \in \mathbb{C} : |z| \leq \|x\|\}.$$

Clearly,  $B' \subseteq D = \prod_{x \in X} D_x$ . Since the product of compact spaces is compact (this statement is equivalent to the axiom of choice and it is called the *Tychonov theorem*) and since it is easy to check that  $B'$  is closed in  $D$ ,  $B'$  is a compact subset of  $D$ .

It is immediate from definition that the induced product topology on  $B'$  is its weak \* topology. (For the definition of the product topology see, e.g., [271].)

□

In this regard we note the following fact.

**Theorem A.9.6. Characterization of weak \* compactness**

Let  $X$  be a Banach space.  $Y \subseteq X'$  is weak \* compact if and only if  $Y$  is weak \* closed and norm bounded.

*Proof.* The sufficient condition for weak \* compactness follows from Theorem A.9.5.

For the necessary condition we must verify that weak \* boundedness implies norm boundedness (since weak \* compactness yields weak \* boundedness). This follows from Theorem A.7.1 or Theorem A.8.6, noting that  $X$  is complete. □

Since weak \* boundedness implies norm boundedness in a Banach space, we see that every weak \* convergent sequence is norm bounded.

**Remark.** We require  $X$  to be complete in Theorem A.9.6. For a counterexample let  $X$  be the vector space of all finite sequences of complex numbers normed by  $\|x\| = \sup\{|x_n|\}$ ,  $x = \{x_n : n = 1, \dots\}$ . Set  $x'_n(x) = n|x_n|$  and  $Y = \{0\} \cup \{x'_n : n = 1, \dots\} \subseteq X'$ .  $x'_n \rightarrow 0$  in  $\sigma(X', X)$ , whereas  $\|x'_n\| = n$ . The situation is corrected by the following result: *Let  $X$  be a normed vector space and let  $Y \subseteq X'$  be weak \* compact;  $Y$  is norm bounded if and only if the weak \* closure of the smallest convex set containing  $Y$  is weak \* compact.*

A useful result, e.g., [32], page 141, concerning weak \* closures is the *Krein-Smulian theorem*: *Let  $X$  be a Banach space and let  $K \in X'$  be convex; by definition, a net  $\{x'_\alpha\} \subseteq X'$  converges to 0 in the Krein-Smulian topology if  $x'_\alpha \rightarrow 0$  uniformly on compact sets of  $X$  ([271], page 65); then the Krein-Smulian and weak \* closures of  $K$  are identical.*

Note, of course, that the finite subsets of  $X$  are compact.

**Example A.9.7. A weak \* closure of characteristic functions**

Define  $Y$  to be the space of functions  $f : [0, 1] \rightarrow [0, 1]$  having the form  $f = \mathbb{1}_A$ , where the subset  $A \subseteq [0, 1]$  is a finite disjoint union of intervals. The weak \* closure of  $Y$ , as a subset of  $L^\infty_m([0, 1])$ , is

$$\{f \in L^\infty_m([0, 1]) : 0 \leq f \leq 1\}.$$

**Theorem A.9.8. Sequential weak \* compactness**

Let  $X$  be a separable normed vector space. Then,  $B'$  is sequentially compact in the  $\sigma(X', X)$  topology, cf., Theorem A.9.5.

*Proof.* Let  $\{x'_k : k = 1, \dots\} \subseteq B'$  and let  $\{x_n : n = 1, \dots\}$  be a countable dense subset of  $X$ . By the expected diagonal argument there is a subsequence  $\{x'_{k_j} : j = 1, \dots\} \subseteq \{x'_k : k = 1, \dots\}$  such that

$$\forall n \in \mathbb{N}, \quad \lim_{j \rightarrow \infty} x'_{k_j}(x_n) = x'(x_n).$$

For  $x \in X$ , let  $x_{n_p} \rightarrow x$ , so that

$$\lim_{p \rightarrow \infty} x'_{k_j}(x_p)$$

exists uniformly in  $j$ . We complete the proof by the Moore–Smith theorem.  $\square$

Note that a compact topological space is metrizable if and only if it has a countable basis. Thus, *if a normed vector space  $X$  is separable, then  $B'$  with the weak \* topology is metrizable* (by definition of the weak \* topology). However,  $X'$  is never metrizable in its weak \* topology if  $X$  is infinite dimensional.

If  $X$  is a normed vector space, then, as noted at the beginning of this section,  $X'$  is a Banach space with the norm defined by (A.14).

Convergence criteria, compatible with the weak topology, require nets. However, by definition, a sequence  $\{x_n : n = 1, \dots\} \subseteq X$  converges to  $0 \in X$  if for every weak neighborhood  $U$  of  $0$ , there is  $N_U$  such that  $x_n \in U$  for all  $n \geq N_U$ . Thus, by the definition of  $U$ , we not only have

$$\|x_n\| \rightarrow 0 \implies x_n \rightarrow 0 \text{ in } \sigma(X, X'),$$

but we also have the following satisfying result, cf., Section 6.3.

**Theorem A.9.9. Sequential weak convergence**

*Let  $X$  be a normed vector space, and given a sequence  $\{x_n : n = 1, \dots\} \subseteq X$ . Then,  $x_n \rightarrow 0$  in the weak topology  $\sigma(X, X')$  if and only if*

$$\forall x' \in X', \quad \lim_{n \rightarrow \infty} x'(x_n) = 0.$$

An immediate corollary of the Uniform Boundedness Principle (Theorem A.8.6) for the Banach space  $X'$  is the following result.

**Theorem A.9.10. Boundedness of weakly convergent sequences**

*Let  $X$  be a normed vector space and assume  $x_n \rightarrow x$  in  $\sigma(X, X')$ . Then,  $\{\|x_n\| : n = 1, \dots\}$  is bounded.*

Using Theorem A.9.8 we can prove the following “converse” to Theorem A.9.10.

**Theorem A.9.11. Weak convergence of norm bounded sequences**

*Let  $X$  be a reflexive Banach space and let  $\{x_n : n = 1, \dots\}$  be a norm bounded sequence in  $X$ . Then, there is a subsequence which converges to some  $x \in X$  in the  $\sigma(X, X')$  topology, cf., Theorem 6.5.5.*

If “subsequence” is replaced by “subnet” in Theorem A.9.11 the result is immediate from the Banach–Alaoglu theorem. It is interesting to compare

this result with Theorem 6.3.2 noting that  $L_m^1([0, 1])$  is not reflexive. Because of Theorem A.9.11 it is easy to check that reflexive Banach spaces are sequentially weakly complete.

By Theorem A.9.3 and since Theorem A.8.3*b* is valid for Hausdorff locally convex topological vector spaces, we obtain the following result.

**Theorem A.9.12. Hahn–Banach in the weak \* setting**

*Let  $X$  be a normed vector space and let  $Y \subseteq X'$  be a  $\sigma(X', X)$  closed linear subspace. If  $y' \in X' \setminus Y$ , then there is  $x \in X$  such that  $y'(x) \neq 0$  and*

$$\forall x' \in Y, \quad x'(x) = 0.$$

## A.10 Linear maps

If  $X$  and  $Y$  are Banach spaces,  $\mathcal{L}(X, Y)$  denotes the space of continuous linear functions  $X \rightarrow Y$ . If  $X = Y$  we write  $\mathcal{L}(X)$ .

**Proposition A.10.1.** *If  $X$  and  $Y$  are Banach spaces over  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{F} = \mathbb{C}$  ( $X$  a normed vector space is sufficient), then  $\mathcal{L}(X, Y)$  is a Banach space over  $\mathbb{F}$  where  $L = c_1L_1 + c_2L_2$  is defined by  $L(x) = c_1L_1(x) + c_2L_2(x)$ ,  $x \in X$  and  $c_1, c_2 \in \mathbb{F}$ , and where  $\|L\|$  is defined by (A.7).*

By Theorem A.5.2 we have the following result.

**Proposition A.10.2. a.** *Let  $X$  and  $Y$  be Banach spaces and let  $L \in \mathcal{L}(X, Y)$ . Then,  $L$  is uniformly continuous.*

**b.** *Let  $X$  and  $Y$  be Banach spaces and let  $Z \subseteq X$  be a linear subspace of  $X$ . If  $L \in \mathcal{L}(Z, Y)$ , then  $L$  has a unique continuous linear extension to  $\overline{Z}$ .*

Parts *a* and *b* of the following result are the *Banach open mapping theorem* and *Banach closed graph theorem*, respectively.

**Theorem A.10.3. Banach open mapping and closed graph theorems**

**a.** *Let  $L \in \mathcal{L}(X, Y)$  be bijective. Then,  $L^{-1} \in \mathcal{L}(X, Y)$ .*

**b.** *Let  $X$  and  $Y$  be Banach spaces and let  $L : X \rightarrow Y$  be linear. Assume that*

$$\|x_n - x\| \rightarrow 0 \quad \text{and} \quad \|L(x_n) - y\| \rightarrow 0 \quad (\text{A.15})$$

*imply  $y = L(x)$ . Then,  $L \in \mathcal{L}(X, Y)$ .*

The proof of part *a* depends on the Baire category theorem. Part *b* is clear from part *a* by applying part *a* to the setting

$$\begin{aligned} X \times L(X) &\rightarrow X \\ (x, L(x)) &\mapsto x, \end{aligned}$$

where the norm on  $X \times L(X)$  is given by  $\|(x, L(x))\| = \|x\| + \|L(x)\|$ . Criterion (A.15) is used to check that  $X \times L(X)$  is complete.



**Example A.10.4. What the Banach closed graph theorem asserts**

The Banach closed graph theorem does not say that if  $X \times L(X)$  is closed in  $X \times Y$  then  $L$  is continuous. It asserts the continuity of  $L$  if each  $(x, y) \in \overline{X \times L(X)} \subseteq X \times Y$  can be approximated by  $\{(x_n, L(x_n)) : n = 1, \dots\}$ , for some sequence  $\{x_n\} \subseteq X$ .

Assume  $\{x_\alpha\} \subseteq X$  and  $\{y_\alpha\} \subseteq Y$  are Hamel bases with  $\|x_\alpha\| \leq 1$ ,  $\sup \|y_\alpha\| = \infty$ . Taking  $\text{card } X = \text{card } Y$  we define  $L(x_\alpha) = y_\alpha$ , and extend  $L$  linearly to all of  $X$ . Then,  $L$  is a linear surjection and  $X \times L(X) = X \times Y$ , but  $X \times L(X)$  does not satisfy (A.15). Clearly,  $L$  is not continuous.

**Example A.10.5. Discontinuous identity mappings on  $\ell^\infty(\mathbb{N})$**

We shall put two norms on  $\ell^\infty(\mathbb{N})$  so that  $\ell^\infty(\mathbb{N})$  is a Banach space for each norm but such that neither identity mapping  $\ell^\infty(\mathbb{N}) \rightarrow \ell^\infty(\mathbb{N})$  is continuous. Choose  $\|\dots\|_\infty$  for the first norm. To define the second norm first observe that

$$\text{card } \ell^1(\mathbb{N}) = \text{card } \ell^\infty(\mathbb{N}). \tag{A.16}$$

To prove (A.16) consider the injection

$$\begin{aligned} \ell^\infty(\mathbb{N}) &\rightarrow \ell^1(\mathbb{N}) \\ \{x_n : n = 1, \dots\} &\mapsto \{x_n/2^n : n = 1, \dots\}. \end{aligned}$$

Thus,  $\text{card } \ell^\infty(\mathbb{N}) \leq \text{card } \ell^1(\mathbb{N})$ . On the other hand,  $\text{card } \ell^1(\mathbb{N}) \leq \text{card } \ell^\infty(\mathbb{N})$  since  $\ell^1(\mathbb{N}) \subseteq \ell^\infty(\mathbb{N})$ . (A.16) follows from the Schröder–Bernstein theorem, e.g., Problem 1.6. Consequently, if  $H_p$  is a Hamel basis for  $\ell^p(\mathbb{N})$ , then  $\text{card } H_\infty = \text{card } H_1$ , and so we choose any bijection  $b : H_\infty \rightarrow H_1$ . We extend  $b$  by linearity to a bijection  $L : \ell^\infty(\mathbb{N}) \rightarrow \ell^1(\mathbb{N})$ . By Theorem A.10.3a, the non-separability of  $\ell^\infty(\mathbb{N})$ , and the separability of  $\ell^1(\mathbb{N})$ , we see that  $L \notin \mathcal{L}(\ell^\infty(\mathbb{N}), \ell^1(\mathbb{N}))$ .

The second norm on  $\ell^\infty(\mathbb{N})$  is then defined by

$$\|x\| = \|L(x)\|_1.$$

It is easy to check that  $\ell^\infty(\mathbb{N})$  with this norm is complete.

The following was given by LENNART CARLESON with regard to an interpolation problem [87].

**Theorem A.10.6. Carleson open mapping theorem**

Let  $X$  and  $Y$  be Banach spaces with norms  $\|\dots\|_X$  and  $\|\dots\|_Y$ , respectively. Assume  $Y \subseteq X$  and  $\|\dots\|_Y \geq \|\dots\|_X$  on  $Y$ . If

$$\begin{aligned} \exists M > 0 \text{ and } \exists \{x_n : n = 1, \dots\} \subseteq Y \text{ such that,} \\ \forall x' \in X', \quad \|x'\| \leq M \sup_{n \in \mathbb{N}} |x'(x_n)|, \end{aligned}$$

then  $X = Y$  and  $\|\dots\|_Y \leq M \|\dots\|_X$ .

Let  $X$  and  $Y$  be Banach spaces. If  $L \in \mathcal{L}(X, Y)$ , then the *adjoint*,  $L'$ , of  $L$  is the element of  $\mathcal{L}(Y', X')$  defined by

$$\forall x \in X \text{ and } \forall y' \in Y', \quad (L'(y'))(x) = y'(L(x)).$$

$L'$  is an *open mapping* if  $L'(U) \subseteq X'$  is open for every open set  $U \subseteq Y'$ , i.e., if

$$\exists C > 0 \text{ such that } \forall y' \in Y', \quad \|y'\|_{Y'} \leq C \|L'(y')\|_{X'}.$$

**Theorem A.10.7. Surjectivity consequences of the Banach open mapping theorem**

Let  $X$  and  $Y$  be Banach spaces and assume  $L \in \mathcal{L}(X, Y)$  is injective and  $\overline{L(X)} = Y$ . The following are equivalent:

- a.  $L(X) = Y$ ,
- b.  $L'$  is an open mapping,
- c.  $L'(Y') = X'$ .

Part a of the following result is true when  $X$  and  $Y$  are normed vector spaces. As in Theorem A.10.7, it depends on the Banach open mapping theorem.

**Theorem A.10.8. Injectivity and surjectivity duality**

Let  $X$  and  $Y$  be Banach spaces and let  $L \in \mathcal{L}(X, Y)$ .

- a.  $L'(Y') = X' \iff L^{-1}$  exists and  $L^{-1} \in \mathcal{L}(L(X), X)$ .
- b.  $L(X) = Y \iff (L')^{-1}$  exists and  $(L')^{-1} \in \mathcal{L}(L'(Y'), Y')$ . Further, if  $L^{-1}$  exists then  $L^{-1} \in \mathcal{L}(L(X), X)$ .

## A.11 Embeddings of dual spaces

Let  $B_1 \subseteq B_2$ , where  $B_1$  and  $B_2$  are normed vector spaces, and let  $Id : B_1 \rightarrow B_2$  be the identity mapping with adjoint  $Id' : B_2' \rightarrow B_1'$  acting between the dual Banach spaces. By definition,

$$\forall x \in B_1 \text{ and } \forall y \in B_2', \quad (Id'(y))(x) = y(x),$$

i.e.,  $Id'(y) = y$  on  $B_1 \subseteq B_2$ .

Assume  $Id$ , and hence  $Id'$ , are continuous. Note that if  $y \in B_2'$ , then  $y|_{B_1}$ , the restriction of  $y$  to  $B_1$ , is an element of  $B_1'$ . To see this first note that since  $B_1 \subseteq B_2$  and  $y \in B_2'$ , then  $y|_{B_1}$  is linear on  $B_1$ .  $y|_{B_1}$  is also continuous on  $B_1$  because of the continuity of  $Id$ . In fact, since  $y$  is continuous on  $B_1$  with the induced topology from  $B_2$ , then it is continuous on  $B_1$  with its given norm convergence because this latter topology is stronger (finer) than the  $B_2$  criterion. (*Continuity of a function for a given topology on its domain implies continuity for any stronger topology on that domain.*)

**Definition A.11.1. Embedding of dual spaces**

$B'_2$  is embedded in  $B'_1$ , in which case we write  $B'_2 \subseteq B'_1$ , if  $Id'$  is a continuous injection. This means that whenever  $Id'(y) = 0 \in B'_1$ , then  $y = 0$ , i.e.,  $y(x) = 0$  for all  $x \in B_2$ .

With the above assumptions, we further assume that  $\overline{B_1} = B_2$ . Let  $y \in B'_2$  have the property that  $Id'(y) = 0 \in B'_1$ . Suppose  $x \in B_2$  and  $\lim_{n \rightarrow \infty} \|x_n - x\|_{B_2} = 0$ , where  $\{x_n\} \subseteq B_1$ . Then,  $\lim_{n \rightarrow \infty} y(x_n) = y(x)$ , and  $y(x_n) = (Id'(y))(x_n) = 0$ . Thus,  $y(x) = 0$ , and so  $y \in B'_2$  is the 0-element. Hence,  $Id'$  is a continuous injection.  $Id'$  is also the identity function, i.e., for all  $y \in B'_2$ ,  $Id'(y) = y$  on a dense linear subspace of  $B_2$ .

We can summarize what has been said by the following embedding theorem.

**Theorem A.11.2. Embedding theorem**

Let  $B_1$  and  $B_2$  be normed vector spaces. If  $B_1 \subseteq B_2$  in the sense that the identity mapping  $Id : B_1 \rightarrow B_2$  is continuous, and if  $\overline{B_1} = B_2$ , then  $B'_2 \subseteq B'_1$ .

**Example A.11.3.  $C_b(\mathbb{R})$ ,  $M_b(\mathbb{R})$ , and duality**

**a.** Let  $C_b(\mathbb{R})$  be the Banach space of continuous bounded functions on  $\mathbb{R}$  taken with the  $L^\infty$ -norm  $\|\dots\|_\infty$ ; and let  $C_0(\mathbb{R})$  be the closed linear subspace of  $C_b(\mathbb{R})$  whose elements  $f$  satisfy the condition that  $\lim_{|t| \rightarrow \infty} f(t) = 0$ . Recall from Theorem 7.2.7 (RRT) that  $(C_0(\mathbb{R}))' = M_b(\mathbb{R})$ .

$C_b(\mathbb{R})$  is a closed linear subspace of  $L^\infty(\mathbb{R})$ , and so it is natural to describe the relation between  $(C_b(\mathbb{R}))'$  and the dual space of  $L^\infty(\mathbb{R})$ , see Theorem 5.5.7. To this end, let  $\mathcal{A}$  be the algebra generated by the closed subsets of  $\mathbb{R}$ . Then, let  $FR(X) \subseteq F(X)$ , defined before Theorem 5.5.7, be the set of those elements  $\nu \in F(X)$  for which  $|\nu|$  is regular in the sense of Definition 2.5.12. It is not difficult to prove that  $(C_b(\mathbb{R}))' = FR(X)$ , see [146], Part I, Chapter IV, Section 6.

**b.** Since the continuous identity mapping  $Id : C_0(\mathbb{R}) \rightarrow C_b(\mathbb{R})$  is not dense, we can *not* conclude that  $(C_b(\mathbb{R}))' \subseteq M_b(\mathbb{R})$ , as is apparent from the characterizations of  $(C_0(\mathbb{R}))'$  and  $(C_b(\mathbb{R}))'$ , cf., Theorem A.11.2.

**c.** The characterizations of  $(C_0(\mathbb{R}))'$  and  $(C_b(\mathbb{R}))'$  do imply, however, that

$$M_b(\mathbb{R}) \subseteq (C_b(\mathbb{R}))'. \quad (\text{A.17})$$

In this regard, if  $\mu \in (C_0(\mathbb{R}))'$ , then  $\mu$  extends to an element  $\mu_e \in (C_b(\mathbb{R}))'$  by the Hahn-Banach theorem. Of course, there is no à priori guarantee of a unique extension.

On the other hand, and without invoking the characterization of  $(C_b(\mathbb{R}))'$ , we can see the validity of (A.17) in the following way.

Let  $\mu \in M_b(\mathbb{R})$ . If  $f \in C_b(\mathbb{R})$  we can choose  $\{f_n\} \subseteq C_0(\mathbb{R})$  for which  $\lim_{n \rightarrow \infty} f_n = f$  pointwise on  $\mathbb{R}$  and  $\sup_n \|f_n\|_\infty = \|f\|_\infty < \infty$ . Then, we apply LDC for  $L^1_{|\mu|}(\mathbb{R})$ , which allows us to assert that  $f \in L^1_{|\mu|}(\mathbb{R})$  and  $\lim_{n \rightarrow \infty} \|f_n - f\|_1 = 0$ . The integral  $\mu(f)$  is well defined, i.e., it is independent of the sequence  $\{f_n\} \subseteq C_0(\mathbb{R})$ . Further,  $\mu : C_b(\mathbb{R}) \rightarrow \mathbb{C}$  is linear. To

prove the continuity of  $\mu$  on  $C_b(\mathbb{R})$ , let  $f \in C_b(\mathbb{R})$ , let  $\varepsilon > 0$ , and choose  $\{f_n\}$  as above. Then,

$$\exists N > 0 \text{ such that } \forall n \geq N, \quad |\mu(f - f_n)| < \varepsilon;$$

and so, for such  $n$ ,

$$|\mu(f)| \leq \varepsilon + |\mu(f_n)| \leq \varepsilon + \|\mu\|_1 \|f\|_\infty.$$

This is true for all  $\varepsilon > 0$ , and so  $\mu \in (C_b(\mathbb{R}))'$ . We designate  $\mu$  so defined on  $C_b(\mathbb{R})$  by  $\mu^*$ .

The inclusion (A.17) is accomplished by the mapping  $\mu \mapsto \mu^*$ . The fact that many extensions  $\mu_e$  of  $\mu$  exist does not contradict (A.17). In fact,  $\nu_e = \mu_e - \mu^* \in (C_b(\mathbb{R}))'$  vanishes on  $C_0(\mathbb{R})$ ; and if  $\nu_e$  is not identically 0 on  $C_b(\mathbb{R})$ , then  $\mu_e$  is not countably additive on  $\mathcal{B}(\mathbb{R})$  and so it does not correspond to an element of  $M_b(\mathbb{R})$ .

## A.12 Hilbert spaces

### Definition A.12.1. Orthonormal set and orthonormal basis (ONB)

**a.** Let  $H$  be a Hilbert space. Elements  $x, y \in H$  are *orthogonal* if  $\langle x, y \rangle = 0$ ; and this property is denoted by  $x \perp y$ . An element  $x \in H$  is *orthogonal to the set*  $S \subseteq H$ , denoted by  $x \perp S$ , if  $\langle x, y \rangle = 0$  for all  $y \in S$ . A set  $S \subseteq H$  is an *orthogonal set* if  $x \perp y$  for all  $x, y \in S$  for which  $x \neq y$ . A set  $S \subseteq H$  is an *orthonormal set* if it is orthogonal and if  $\|x\| = 1$  for each  $x \in S$ .

**b.** A countable orthonormal set  $S = \{x_n : n = 1, \dots\}$  is an *orthonormal basis* (ONB) for  $H$  if

$$\forall x \in H, \exists \{c_n : n = 1, \dots\} \subseteq \mathbb{C} \text{ such that } x = \sum_{n=1}^{\infty} c_n x_n \text{ in } H.$$

**Proposition A.12.2.** Let  $S = \{x_\alpha\}$  be an orthonormal set in a separable Hilbert space  $H$ . Then,  $S$  is a countable set.

*Proof.* By separability, let  $D = \{y_n : n = 1, \dots\}$  be a countable dense subset of  $H$ . Since  $S$  is orthonormal, we can assert that

$$\forall \alpha, \beta \text{ for which } \alpha \neq \beta, \quad \|x_\alpha - x_\beta\| = \sqrt{2}. \quad (\text{A.18})$$

Using the density, we have

$$\forall \alpha, \exists n = n(\alpha) \in \mathbb{N}, \text{ such that } \|x_\alpha - y_{n(\alpha)}\| < \frac{\sqrt{2}}{2}. \quad (\text{A.19})$$

We are forced into choosing a different  $n(\alpha)$  for each  $\alpha$ , for, otherwise, if  $y_{n(\alpha)}$  corresponds to both  $x_\alpha$  and  $x_\beta$  in the sense of (A.19), then

$$\|x_\alpha - x_\beta\| \leq \|x_\alpha - y_{n(\alpha)}\| + \|y_{n(\alpha)} - x_\beta\| < \sqrt{2},$$

and this contradicts (A.18). Thus, (A.19) gives rise to an injective mapping  $S \rightarrow D$ , and, hence,  $S$  is countable. □

**Example A.12.3. Hilbert spaces and ONBs**

**a.**  $H = L^2(\mathbb{T}_{2\Omega})$  is a Hilbert space with inner product defined by  $\langle F, G \rangle = \int_{\mathbb{T}_{2\Omega}} F(x)\overline{G(x)} dx$ , where  $\mathbb{T}_{2\Omega} = \mathbb{R}/(2\Omega\mathbb{Z})$ ,  $F$  and  $G$  are  $2\Omega$ -periodic on  $\mathbb{R}$ , and  $\int_{\mathbb{T}_{2\Omega}} F(x) dx$  is defined as the Lebesgue integral  $\frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(x) dx$ . The sequence  $\{e^{-\pi i n x / \Omega} : n \in \mathbb{Z}\}$  is an ONB for  $L^2(\mathbb{T}_{2\Omega})$ , e.g., Proposition B.8.1.

**b.**  $H = \ell^2(\mathbb{Z}^d)$  is defined to be the vector space of all sequences  $f : \mathbb{Z}^d \rightarrow \mathbb{C}$  with the property that

$$\|f\|_{\ell^2(\mathbb{Z}^d)} = \left( \sum_{n \in \mathbb{Z}^d} |f[n]|^2 \right)^{1/2} < \infty.$$

With this norm,  $\ell^2(\mathbb{Z}^d)$  is a Hilbert space, and its inner product is given by

$$\forall f, g \in \ell^2(\mathbb{Z}^d), \quad \langle f, g \rangle = \sum_{n \in \mathbb{Z}^d} f[n]\overline{g[n]}.$$

Let  $u_n \in \ell^2(\mathbb{Z}^d)$  be defined by  $u_n[m] = \delta(m, n)$ , for  $m, n \in \mathbb{Z}^d$ , where

$$\delta(m, n) = \begin{cases} 1, & \text{if } m = n, \\ 0, & \text{if } m \neq n. \end{cases}$$

It is easy to check that the sequence  $\{u_n\}$  is an ONB for  $\ell^2(\mathbb{Z}^d)$ .

**c.**  $H = L^2(\mathbb{R})$  is a Hilbert space with inner product defined by  $\langle f, g \rangle = \int_{\mathbb{R}} f(t)\overline{g(t)} dt$ . The *Hermite functions*,  $h_n(x) = e^{-\pi x^2} H_n(2\sqrt{\pi}x)$ ,  $n = 0, \dots$ , where

$$\forall n = 0, \dots, \quad H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2},$$

are an ONB for  $L^2(\mathbb{R})$ , see, e.g., [490], [35], and Remark 2.4.11 in [39]. The concept of a multiresolution analysis in wavelet theory leads to the construction of many other ONBs for  $L^2(\mathbb{R})$ , e.g., [114], [338], cf., [461].

**d.**  $H = PW_\Omega = \{f \in L^2(\mathbb{R}) : \text{supp } \hat{f} \subseteq [-\Omega, \Omega]\}$  is a closed linear subspace of  $L^2(\mathbb{R})$ , and is the so-called *Paley-Wiener space* of  $\Omega$ -bandlimited functions.  $\hat{f}$  designates the Fourier transform of  $f$ , see Appendix B.  $PW_\Omega$  is a Hilbert space with inner product induced from  $L^2(\mathbb{R})$ . The sequence  $\{(\frac{1}{\sqrt{2\Omega}})\tau_{n/(2\Omega)}(d_{2\pi\Omega}) : n \in \mathbb{Z}\}$  is an ONB for  $PW_\Omega$ , where  $\tau_x(f)(y) = f(y-x)$  and where

$$d_{2\pi\Omega}(t) = \frac{\sin(2\pi\Omega t)}{\pi t}.$$

In light of our mention of multiresolution in part *c*, we note that  $PW_\Omega$  can be considered as part of a multiresolution analysis of  $L^2(\mathbb{R})$  for the so-called Shannon wavelet system, e.g., [114], [338], [108], [109], cf. [461]

The following is an immediate, useful consequence of the Schwarz inequality.

**Proposition A.12.4. Continuity of the inner product**

Let  $H$  be a Hilbert space. The inner product is continuous on  $H \times H$ , i.e., if  $\{x_n : n = 1, \dots\} \subseteq H$  converges to  $x \in H$  and  $\{y_n : n = 1, \dots\} \subseteq H$  converges to  $y \in H$ , then

$$\lim_{n \rightarrow \infty} \langle x_n, y_n \rangle = \langle x, y \rangle.$$

**Theorem A.12.5. Consequences of orthonormality**

Let  $H$  be a Hilbert space and let  $\{x_n : n = 1, \dots\}$  be an orthonormal sequence.

*a. Bessel inequality. The mapping*

$$\begin{aligned} L : H &\longrightarrow \ell^2(\mathbb{N}) \\ y &\longmapsto \{\langle y, x_n \rangle\} \end{aligned} \quad (\text{A.20})$$

is well-defined, linear, and continuous; in fact

$$\forall y \in H, \quad \sum_{n=1}^{\infty} |\langle y, x_n \rangle|^2 \leq \|y\|^2.$$

*b. For each  $y \in H$ ,  $\sum \langle y, x_n \rangle x_n$  converges in  $H$ .*

*c.  $\sum c_n x_n$  converges in  $H$  if and only if  $c = \{c_n : n = 1, \dots\} \in \ell^2(\mathbb{N})$ .*

*d. If  $y = \sum c_n x_n$  converges in  $H$ , then each  $c_n = \langle y, x_n \rangle$ .*

*Proof.* *i.* Let  $y \in H$ , let  $F \subseteq \mathbb{N}$  be finite, and suppose  $\{c_n : n \in F\} \subseteq \mathbb{C}$ . Using orthonormality, two direct calculations yield

$$\left\| \sum_{n \in F} c_n x_n \right\|^2 = \sum_{n \in F} |c_n|^2 \quad (\text{A.21})$$

and

$$0 \leq \left\| y - \sum_{n \in F} \langle y, x_n \rangle x_n \right\|^2 = \|y\|^2 - \sum_{n \in F} |\langle y, x_n \rangle|^2. \quad (\text{A.22})$$

*ii.* The Bessel inequality (part *a*) is immediate from (A.22). In particular,  $\{\langle y, x_n \rangle\} \in \ell^2(\mathbb{N})$ . Since  $H$  is complete, to prove part *b* we need only show that

$$\{s_N\} = \left\{ \sum_{n=1}^N \langle y, x_n \rangle x_n \right\}, \quad N > 0,$$

is a Cauchy sequence in  $H$ . This is a consequence of (A.21) and the fact that  $\{\langle y, x_n \rangle\} \in \ell^2(\mathbb{N})$ , which, in turn, was a consequence of part *a*. Part *c* also follows from (A.21).

*iii.* To prove part *d*, we use the orthonormality and the continuity of inner products (Proposition A.12.4) to compute

$$\langle y, x_n \rangle = \lim_{N \rightarrow \infty} \left\langle \sum_{m=1}^N c_m x_m, x_n \right\rangle = c_n. \quad \square$$

□

The following result is also elementary to verify. One efficient route is to prove the implications: *a* implies *b* implies *c* implies *d* implies *e* implies *a*.

**Theorem A.12.6. Parseval formula and ONB**

Let  $H$  be a Hilbert space and let  $\{x_n : n = 1, \dots\}$  be an orthonormal sequence. The following are equivalent.

- a.*  $\{x_n\}$  is an ONB for  $H$ .
- b.* Parseval formula.

$$\forall x, y \in H, \quad \langle x, y \rangle = \sum_{n=1}^{\infty} \langle x, x_n \rangle \overline{\langle y, x_n \rangle}.$$

*c.* The mapping  $L$  of (A.20) (in Theorem A.12.5a) is a linear surjective isometry, and, in fact,

$$\forall y \in H, \quad \|y\| = \left( \sum_{n=1}^{\infty} |\langle y, x_n \rangle|^2 \right)^{1/2}.$$

- d.*  $\overline{\text{span}} \{x_n\} = H$ .
- e.* If  $\langle y, x_n \rangle = 0$  for each  $n \in \mathbb{N}$ , then  $y = 0$ .

Because of Example A.12.3a, the coefficients  $\langle y, x_n \rangle$  for an ONB  $\{x_n\} \subseteq H$  are called the *Fourier coefficients* of  $y \in H$ , cf., Definition B.5.1.

**Theorem A.12.7. Hilbert space Fourier series**

Let  $H$  be a Hilbert space and let  $\{x_n : n = 1, \dots\}$  be an ONB for  $H$ . Then,

$$\forall y \in H, \quad y = \sum_{n=1}^{\infty} \langle y, x_n \rangle x_n \quad \text{in } H.$$

*Proof.* If  $y \in H$ , then  $\sum \langle y, x_n \rangle x_n = x$  in  $H$  for some  $x \in H$  by Theorem A.12.5b. Hence,

$$\forall n \in \mathbb{Z}^d, \quad \langle x, x_n \rangle = \langle y, x_n \rangle$$

by Theorem A.12.5d. The result follows by Theorem A.12.6, using either the equivalence of parts *a* and *c* or of parts *a* and *e*.

□

**Remark. a.** By the definition of an ONB, if  $H$  contains an ONB, then  $H$  is separable. The converse is also true: If  $H$  is a separable Hilbert space, then  $H$  contains an ONB. The proof of the converse has four elementary steps. First, if  $S = \{x_n : n = 1, \dots\}$  is a countable dense subset of  $H$ , then  $\overline{\text{span}} \{x_n\} = H$ . Next, we choose a linearly independent subset  $\{y_n\}$  of  $\{x_n\}$ , which also has the property that  $\overline{\text{span}} \{y_n\} = H$ . This can be accomplished both constructively and iteratively by throwing-out those  $x_n$  which are linear combinations of finite sets  $\{x_j : j \in F \text{ and } j \neq n\}$ . Third, the *Gram-Schmidt orthogonalization* procedure, e.g., [194], pages 21–22, constructs  $\{u_n\}$  in terms of  $\{y_n\}$  with the properties that  $\{u_n\}$  is orthonormal and  $\overline{\text{span}} \{u_n\} = H$ . Finally, we invoke Theorem A.12.6 to complete the proof.

**b.** Let  $X$  be a Banach space over  $\mathbb{C}$ . A sequence  $\{x_n : n = 1, \dots\} \subseteq X$  is a *Schauder basis* for  $X$  if each  $x \in X$  has a unique representation  $x = \sum_{n=1}^{\infty} c_n(x)x_n$ , where each  $c_n(x) \in \mathbb{C}$  and where the series converges in  $X$  in the ordinary sense that

$$\lim_{N \rightarrow \infty} \sum_{n=1}^N c_n(x)x_n = x.$$

If  $\{x_n\}$  is a Schauder basis for  $X$  and if we consider  $\{c_n(x)\}$  as a sequence of mappings  $c_n : X \rightarrow \mathbb{C}$ ,  $x \mapsto c_n(x)$ , then each  $c_n \in X'$ , e.g., [430], page 20. The situation in part *a* leads to the question (the *basis problem*), posed by BANACH in 1932 [19], of whether or not every separable Banach space contains a Schauder basis. Using Walsh functions and lacunary Fourier series, PER ENFLO proved in 1973 that there are separable Banach spaces having no Schauder basis [154], see [406], especially Sections 1.1 and 5.6.

#### Definition A.12.8. Direct sum and orthogonal complement

**a.** Let  $B$  be a Banach space, and let  $X, Y \subseteq B$  be linear subspaces of  $B$  for which  $X \cap Y = \{0\}$  and  $X + Y = \{x + y : x \in X, y \in Y\} = B$ . We denote this situation by

$$B = X \oplus Y,$$

and  $B$  is the *direct sum* of  $X$  and  $Y$ .

**b.** Suppose  $B = X \oplus Y$ . Let  $z \in B$  and assume  $z = x_1 + y_1 = x_2 + y_2$ , where  $x_j \in X$  and  $y_j \in Y$ . Then,  $x_1 - x_2 = y_2 - y_1$ . Thus,  $x_1 - x_2, y_2 - y_1 \in X \cap Y$  and so  $x_1 = x_2$  and  $y_1 = y_2$ . Therefore, if  $B = X \oplus Y$ , then each  $z \in B$  has a unique representation  $z = x + y$  for some  $x \in X$  and  $y \in Y$ .

**c.** Let  $H$  be a Hilbert space, and let  $X \subseteq H$  be a subset of  $H$ . The *orthogonal complement*  $X^\perp$  of  $X$  is the set  $\{y \in H : \forall x \in X, x \perp y\}$ .

**d.** Let  $X$  be a closed linear subspace of  $H$ . It is not difficult to prove that for each  $z \in H$  there are unique elements  $x \in X$  and  $y \in X^\perp$  such that  $z = x + y$ . The proof requires the following two results:

*i.* For each  $z \in H$  there is a unique element  $x \in X$  such that

$$\|z - x\| = \inf\{\|z - w\| : w \in X\};$$



ii. If  $z \in H$  and  $x \in X$ , then  $\langle z - y, w \rangle = 0$  for all  $w \in X$  if and only if

$$\|z - x\| = \inf\{\|z - u\| : u \in X\}.$$

e. From part d, we see that if  $X \neq \{0\}$  is a closed linear subspace of  $H$ , then  $X^\perp$  is a closed linear subspace of  $H$ ,

$$H = X \oplus X^\perp,$$

and  $(X^\perp)^\perp = X$ . We refer to  $X \oplus X^\perp$  as an *orthogonal complement direct sum*.

### A.13 Operators on Hilbert spaces

In the case of Hilbert spaces  $H_1$  and  $H_2$  over  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{F} = \mathbb{C}$ , we write

$$\forall x \in H_1 \text{ and } \forall y \in H_2, \quad \langle L(x), y \rangle_{H_2} = \langle x, L'(y) \rangle_{H_1} \quad (\text{A.23})$$

to define the *adjoint*  $L'$  of  $L \in \mathcal{L}(H_1, H_2)$ . The adjoint was defined for Banach spaces after Theorem A.10.6. In (A.23) we have used the fact that Hilbert spaces  $H$  have the property that  $H' = H$  of (A.8), cf., Example A.8.4b.

We shall now make use of the orthogonal complement (Definition A.12.8) and of the range and kernel of an operator  $L \in \mathcal{L}(H_1, H_2)$ . The *range* of  $L$ , also called the *image* of  $L$ , is defined as  $\mathcal{R}(L) = \{Lx : x \in H_1\} \subseteq H_2$ ; and the *kernel* of  $L$ , also called the *null space* of  $L$ , is defined as the closed linear subspace  $\ker L = \{x \in H_1 : Lx = 0 \in H_2\}$ .

#### Theorem A.13.1. Kernel and range properties for Hilbert space operators

Let  $H_1$  and  $H_2$  be Hilbert spaces over  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{F} = \mathbb{C}$ , and let  $L \in \mathcal{L}(H_1, H_2)$ .

a.  $L'(H_2) = H_1$  if and only if  $L^{-1}$  exists and  $L^{-1} \in \mathcal{L}(L(H_1), H_1)$ .

b.  $L(H_1) = H_2$  if and only if  $(L')^{-1}$  exists and  $(L')^{-1} \in \mathcal{L}(L'(H_2), H_2)$ .

Further, if  $L^{-1}$  exists, then it is in  $\mathcal{L}(L(H_1), H_1)$ .

c. If  $L^{-1}$  exists and  $L^{-1} \in \mathcal{L}(H_2, H_1)$ , then  $(L')^{-1} \in \mathcal{L}(H_1, H_2)$  and

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

d.  $\ker L = (\mathcal{R}(L'))^\perp$ ,  $\ker L' = (\mathcal{R}(L))^\perp$ ,  $\overline{\mathcal{R}(L)} = (\ker L')^\perp$ , and  $\overline{\mathcal{R}(L')} = (\ker L)^\perp$ .

Let  $H$  be a Hilbert space over  $\mathbb{C}$  and let  $X \neq \{0\}$  be a closed linear subspace of  $H$ .  $P : H \rightarrow H$  is the *orthogonal projection onto*  $X$  if

$$\forall x \in X \text{ and } \forall y \in X^\perp, \quad P(x + y) = x.$$

The *orthogonal projection onto*  $X$  is an element of  $\mathcal{L}(H)$  and, in fact,  $\|P\| = 1$ .  $L \in \mathcal{L}(H)$  is *self-adjoint* or *Hermitian* if  $L' = L$ . The orthogonal projections

of  $H$  are the building blocks for the theory of self-adjoint operators in the sense that every  $L \in \mathcal{L}(H)$  is the limit in norm of a sequence of linear combinations of orthogonal projections, see [194].

An elementary calculation yields the following result.

**Proposition A.13.2.** *Let  $H$  be a Hilbert space over  $\mathbb{C}$ .  $P \in \mathcal{L}(H)$  is an orthogonal projection onto some closed linear subspace of  $H$  if and only if  $P$  is self-adjoint and  $P^2 = P$ .*

Let  $X \subseteq H \setminus \{0\}$  be a closed linear subspace of the Hilbert space  $H$  over  $\mathbb{C}$ , and let  $L \in \mathcal{L}(H)$ .  $X$  is an  $L$ -invariant subspace if  $L(X) \subseteq X$ . There is the following relationship between invariant subspaces and orthogonal projections.

**Theorem A.13.3. Invariant subspaces and orthogonal projections**  
*Let  $X \subseteq H \setminus \{0\}$  be a closed linear subspace of the Hilbert space  $H$  over  $\mathbb{C}$ , and let  $L \in \mathcal{L}(H)$ .  $X$  is  $L$ -invariant if and only if  $LP = PLP$ , where  $P$  is the orthogonal projection onto  $X$ .*

**Remark.** The *invariant subspace problem* is to determine whether or not, for any given Hilbert space  $H$  over  $\mathbb{C}$ , every  $L \in \mathcal{L}(H)$  has a non-trivial  $L$ -invariant subspace. There is a spectacular positive solution due to VICTOR LOMONOSOV, which is even valid for Banach spaces for the case of compact operators. (Compact operators  $L \in \mathcal{L}(H)$  are those for which any sequence  $\{x_n : n = 1, \dots\} \subseteq H$  of unit norm elements has the property that  $\{L(x_n) : n = 1, \dots\}$  has a convergent subsequence in  $H$ .) There has been progress since LOMONOSOV but the general problem is open, see [359].

**Remark.** It is elementary to check that  $L \in \mathcal{L}(H)$  is self-adjoint if and only if  $\langle L(x), x \rangle \in \mathbb{R}$  for all  $x \in H$ . One direction is immediate:  $L' = L$  implies  $\langle L(x), x \rangle = \langle x, L(x) \rangle = \overline{\langle L(x), x \rangle}$ . Conversely,  $\langle L(z), z \rangle \in \mathbb{R}$  implies  $\langle L(x + cy), x + cy \rangle = \langle x + cy, L(x + cy) \rangle$  for all  $x, y \in H$  and  $c \in \mathbb{C}$ ; and using the hypothesis again on this equality we can calculate that  $\text{Im} \langle cL(y), x \rangle = \text{Im} \langle cy, L(x) \rangle$ , which in turn gives  $L' = L$  by considering  $c = 1$  and  $c = i$ .

We shall say that  $L \in \mathcal{L}(H)$  is *positive* if  $\langle L(x), x \rangle \geq 0$  for all  $x \in H$ . By the above observation we see that if  $H$  is a complex Hilbert space and  $L \in \mathcal{L}(H)$  is positive, then  $L$  self-adjoint.

Let  $H_1, H_2$  be Hilbert spaces over  $\mathbb{C}$ .  $U \in \mathcal{L}(H_1, H_2)$  is an *isometry* if  $\|U(x)\|_{H_2} = \|x\|_{H_1}$  for all  $x \in H_1$ . If  $H = H_1 = H_2$  and  $U \in \mathcal{L}(H)$  is a surjective isometry, then  $U$  is a *unitary operator*.

**Proposition A.13.4.** *Let  $H_1$  and  $H_2$  be Hilbert spaces over  $\mathbb{C}$ , and let  $U \in \mathcal{L}(H_1, H_2)$ .*

- a. *The following are equivalent:*
  - i.  *$U$  is an isometry;*
  - ii.  *$U'U$  is the identity mapping  $Id$  on  $H_1$ ;*

$$\text{iii. } \forall x, y \in H_1, \langle U(x), U(y) \rangle_{H_2} = \langle x, y \rangle_{H_1}.$$

Further, if  $U$  is a surjective isometry then  $U'U$  is the identity mapping  $Id$  on  $H_2$ .

b. Let  $H = H_1 = H_2$ .  $U \in \mathcal{L}(H)$  is unitary if and only if  $U^{-1}$  exists on  $H$  and  $U^{-1} = U'$ . Thus, unitary operators  $U$  are characterized by the property that

$$UU' = Id = U'U.$$

**Example A.13.5. Unitary operators**

The Fourier transform mapping  $\mathcal{F}$  on  $L^2_{m^d}(\mathbb{R}^d)$ , the DFT mapping  $\mathcal{F}_N$  on  $L^2_c(\mathbb{Z}_N)$ , and the Hilbert transform mapping  $\mathcal{H}$  on  $L^2(\mathbb{R})$  are all unitary operators, see Appendix B.

**A.14 Potpourri and tittillation**

1. At the beginning of the Preface we referred to this book as a paean to 20th century real analysis. This development of real variables, measure theory, and integration theory was one of several interleaving intellectual threads through the century. One such journey is the theory of frames.

With Theorem A.12.6 as a backdrop we make the following definition, which, at first blush, may seem to be an effete generalization of an ONB. Let  $H$  be a separable Hilbert space. A sequence  $\{x_n : n = 1, \dots\} \subseteq H$  is a *frame* for  $H$  if there are  $A, B > 0$  such that

$$\forall x \in H, \quad A\|x\|^2 \leq \sum_{n=1}^{\infty} |\langle x, x_n \rangle|^2 \leq B\|x\|^2.$$

The constants  $A$  and  $B$  are *frame bounds*, and a frame is *tight* if  $A = B$ . A frame is an *exact frame* if it is no longer a frame whenever any of its elements is removed. The following is the basic decomposition theorem for frames.

**Theorem A.14.1. Frame decomposition**

Let  $\{x_n : n = 1, \dots\} \subseteq H$  be a frame for  $H$ , and define the mapping

$$\begin{aligned} S : H &\rightarrow H \\ x &\mapsto \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n. \end{aligned}$$

Then,  $S$  is a continuous bijection onto  $H$ , and

$$\forall x \in H, \quad x = \sum_{n=1}^{\infty} \langle x, S^{-1}(x_n) \rangle x_n = \sum_{n=1}^{\infty} \langle x, x_n \rangle S^{-1}(x_n). \quad (\text{A.24})$$

In Theorem A.14.1, the assertion that  $S$  is a continuous bijection onto  $H$  implies  $S^{-1} : H \rightarrow H$  is continuous by Theorem A.10.3. Generally, we refer to any continuous bijection  $L : H \rightarrow H$  as a *topological isomorphism*, see [45], Chapters 3 and 7.

Expositions of the theory of frames are found in [503], [114], [45], and [102]. The theory was explicitly formulated by RICHARD J. DUFFIN (1909-1996) and ALBERT CHARLES SCHAEFFER (1952) [143]. What is truly remarkable is the genuine applicability of the theory of frames in addressing sampling problems, erasure problems associated with the internet, quantization problems arising in audio, image processing problems, and a host of other problems, e.g., see [98], [99], [311], [452]. A reason for this applicability is the effectiveness of frames in providing numerically stable, robust, and generally “inexpensive” decompositions; and this reason is due to the fact that frames are generally not ONBs or even Schauder bases, even though there are representations such as (A.24).

In order to describe some of the early developments of frames, we first expand on the definition of a Schauder basis in the setting of a separable Hilbert space  $H$ . A Schauder basis  $\{x_n : n = 1, \dots\}$  for  $H$  is an *unconditional basis* for  $H$  if

$$\begin{aligned} &\exists C > 0 \text{ such that } \forall F \subseteq \mathbb{N}, \text{ where } \text{card } F < \infty, \\ &\text{and } \forall b_n, c_n \in \mathbb{C}, \text{ where } n \in F \text{ and } |b_n| < |c_n|, \\ &\left\| \sum_{n \in F} b_n x_n \right\| \leq C \left\| \sum_{n \in F} c_n x_n \right\|. \end{aligned}$$

An unconditional basis is a *bounded unconditional basis* for  $H$  if

$$\exists A, B > 0 \text{ such that } \forall n \in \mathbb{N}, \quad A \leq \|x_n\| \leq B.$$

Finally, a Schauder basis  $\{x_n : n = 1, \dots\}$  for  $H$  is a *Riesz basis* for  $H$  if there is a topological isomorphism on  $H$  mapping  $\{x_n : n = 1, \dots\}$  onto an ONB for  $H$ .

In 1936 KÖTHER [292] proved that bounded unconditional bases are exact frames, and the converse is straightforward. Also, the category of Riesz bases is precisely that of exact frames. Thus, the following three notions are equivalent: Riesz bases, exact frames, and bounded unconditional bases. Besides the article by DUFFIN and SCHAEFFER, BARI’s characterization of Riesz bases (1951) [26] is fundamental in this realm of ideas. From our point of view, her work has all the more impact because it was motivated in part by her early research, with others in the Russian school, in analyzing RIEMANN’s sets of uniqueness for trigonometric series, see Section 3.8.4.

From a functional analytic point of view, in 1921 VITALI [473] proved that if  $\{x_n : n = 1, \dots\}$  is a tight frame with  $A = B = 1$  and with  $\|x_n\| = 1$  for all  $n$ , then  $\{x_n : n = 1, \dots\}$  is an ONB. Actually, VITALI’s result is stronger for the setting  $H = L^2([a, b])$  in which he dealt.

Frames have also been studied in terms of the celebrated Naimark dilation theorem (1943), a special case of which asserts that any frame can be obtained by “compression” from a basis. The rank 1 case of NAIMARK’s theorem is the previous assertion for tight frames. The finite decomposition rank 1 case of NAIMARK’s theorem antedates NAIMARK’s paper, and it is due to HADWIGER (1940) and GASTON JULIA (1942). This is particularly interesting in light of modern applications of finite unit norm tight frames in communications theory. In this context, we mention CHANDLER DAVIS’ use of Walsh functions to give explicit constructions of dilations [121]. DAVIS [122] also provides an in-depth perspective of the results referred to in this paragraph.

Other applications of NAIMARK’s theorem in the context of frames include feasibility issues for von Neumann measurements in quantum signal processing.

The general theory of frames was inspired by the study of non-harmonic Fourier series and Fourier frames. Just as we define Fourier series in Example 3.3.4 and Appendix B, we define *non-harmonic Fourier series* to be of the form  $\sum_{\lambda \in A} c_\lambda e_\lambda$ , where  $A \subseteq \mathbb{R}$  is countable and  $e_\lambda = e^{-2\pi i x \lambda}$ . Typically, we investigate the elements of  $L^2([-R, R])$  which can be represented in  $L^2$ -norm by such series in a manner analogous to Theorem A.14.1. As such, Fourier frames can be thought of as going back to DINI (1880) and his book on Fourier series [136], pages 190 ff. There he gives Fourier expansions in terms of the set  $\{e_\lambda : \lambda \in A\}$  of harmonics, where each  $\lambda$  is a solution of the equation

$$x \cos(\pi x) + a \sin(\pi x) = 0. \quad (\text{A.25})$$

Equation (A.25) was chosen because of a problem in mathematical physics from RIEMANN’s and later RIEMANN–WEBER’s classical treatise [374], pages 158–167. DINI returned to this topic in 1917, just before his death, with a significant generalization including Fourier frames that are not ONBs [137].

The inequalities defining a Fourier frame  $\{e_\lambda : \lambda \in A\}$  for  $L^2([-R, R])$  (of which our definition of a frame is a natural generalization) were explicitly written by PALEY and WIENER [353], page 115, inequalities (30.56). The book by PALEY and WIENER (1934), and to a lesser extent a stability theorem by G. D. BIRKHOFF (1917), had tremendous influence on mid-20th century harmonic analysis. Although non-harmonic Fourier series expansions were developed, the major effort in the study of Fourier systems emanating from [353] addressed completeness problems of sequences  $\{e_\lambda : \lambda \in A\} \subseteq L^2([-R, R])$ , i.e., on determining when the closed linear span of  $\{e_\lambda : \lambda \in A\}$  is all of  $L^2([-R, R])$ . This culminated in the profound work of BEURLING and MALLIAVIN in 1962 and 1966 [59], [60], [286], see [43], Chapter 1, for a technical overview.

A landmark on the road to the results of BEURLING and MALLIAVIN is the article by DUFFIN and SHAEFFER. In retrospect, their paper was underappreciated when it appeared in 1952. The authors defined Fourier frames as well as the general notion of a frame for a Hilbert space  $H$ . They emphasized that

frames  $\{x_n : n = 1, \dots\} \subseteq H$  provide discrete representations  $x = \sum_{n=1}^{\infty} c_n x_n$  in norm, as opposed to the previous emphasis on completeness. These discrete representations for Fourier frames provide a natural setting for non-uniform sampling, e.g., [43], Chapter 1, [223], [500], [165], [216]. DUFFIN and SHAEFFER understood that the Paley-Wiener theory for Fourier systems is equivalent to the theory of exact Fourier frames. (We noted above that PALEY and WIENER used precisely the inequalities defining Fourier frames.) DUFFIN and SHAEFFER also knew that generally they were dealing with overcomplete systems, a useful feature in noise reduction problems and the other applications we have mentioned.

The next step on this path created by DUFFIN and SHAEFFER is the article by DAUBECHIES, GROSSMANN, and MEYER [115]. From the point of view of the affine and Heisenberg groups, and inspired by DUFFIN and SHAEFFER, this article establishes the basic theory of wavelet and Gabor frames. About 1990, DUFFIN expressed satisfactory surprise to one of the authors that the theory of frames had risen like a phoenix almost 40 years after its creation.

