Chapter 1

Preliminary Background

This chapter is composed of two parts: Basic Analysis and Intermediate Analysis.

The first part is a review of some of the basic background that is required from the first 2 years of a standard program in mathematics. There are program differences so that some areas may receive more coverage while others receive less. Our purpose is to provide a reference point for the reader and establish notation. In a few important cases, we have provided proofs of major theorems. In other cases, we delayed a proof when a more general result is proven in a later chapter.

In the second part of this chapter, we include some intermediate to advanced material that is required later. In most cases, motivation is given along with additional proof detail and specific references.

Part I: Basic Analysis

The first part of this chapter is devoted to a brief discussion of the circle of ideas required for advanced parts of analysis and the basics of operator theory. Those with a strong background in theoretical chemistry or physics but little or no formal training in analysis will find Reed and Simon (vol.1) to be an excellent copilot (see below).

T.L. Gill, W.W. Zachary, *Functional Analysis and the Feynman Operator Calculus*, DOI 10.1007/978-3-319-27595-6_1

General references for this section are Dunford and Schwartz [DS], Jones [J], Reed and Simon [RS], Royden [RO], and Rudin [RU].

1.1. Analysis

1.1.1. Sets. Let X be a nonempty set, let $\emptyset$ be the emptyset, and let $\mathcal{P}(X)$ be the power set of X (i.e., the set of all subsets of X).

Definition 1.1. Let $A, B, A_n \in \mathcal{P}(X), n \in \mathbb{N}$, then

(1) $A^c = \{a \in X : a \notin A\}$, the compliment of A.

(2) $A \backslash B = A \cap B^c$.

(3) (De Morgan's Laws)

$$\left[\bigcup_{k=1}^{\infty} A_k\right]^c = \bigcap_{k=1}^{\infty} A_k^c, \qquad \left[\bigcap_{k=1}^{\infty} A_k\right]^c = \bigcup_{k=1}^{\infty} A_k^c.$$

We define the lim inf and lim sup for sets by:

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k \qquad \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k.$$

Theorem 1.2. *Let $\{A_n\} \subset \mathcal{P}(X), n \in \mathbb{N}$, then the* lim inf and lim sup *satisfy:*

(1)

$$\liminf_{n\to\infty}\ A_n \subset \limsup_{n\to\infty}\ A_n.$$

(2)

$$\limsup_{n\to\infty}\ A_n = \{a : a \in A_k \text{ for infinitely many } k\}.$$

(3)

$$\liminf_{n\to\infty}\ A_n = \{a : a \in A_k \text{ for all but finitely many } k\}.$$

(4)

$$(\limsup_{n\to\infty}\ A_n)^c = \liminf_{n\to\infty}\ A_n^c.$$

(5) *If $A_n \supset A_{n+1}$, then*

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = \lim_{n\to\infty}\ A_n = \bigcap_{k=1}^{\infty} A_k.$$

(6) *If* $A_n \subset A_{n+1}$, *then*

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = \lim_{n\to\infty} A_n = \bigcup_{k=1}^{\infty} A_k.$$

Definition 1.3. Let $A, B \subset X$. (We assume they are nonempty.)

(1) The cartesian product, denoted $A \times B$, is defined by

$$A \times B = \{(a, b) : a \in A,\ b \in B\}.$$

In general, $A \times B \neq B \times A$, so that the order matters. If $\{A_k\}$ is a countable collection of subsets of X, we define the cartesian product by:

$$\prod_{k=1}^{\infty} A_k = \{(a_1, a_2, \ldots) : a_k \in A_k\}.$$

Definition 1.4. A map $f : A \to B$ (a function, or a transformation), with domain $D(f) \subset A$ and range $R(f) \subset B$ is a subset $f \subset A \times B$ such that, for each $x \in A$, there is one and only one $y \in B$, with $(x, y) \in f$. We write $y = f(x)$ and call $f(A) = \{f(x) : x \in A\} \subset B$, the image of f and, call $f^{-1}(B) = \{x : f(x) \in B\} \subset A$, the inverse image of B. We say that f is one to one or injective, if for all $x_1 \neq x_2 \in A$, we have that $y_1 = f(x_1) \neq y_2 = f(x_2) \in B$. We say that f is onto or surjective if, for each $y \in B$, there is a $x \in A$, with $y = f(x)$.

1.1.2. Topology. We only consider Hausdorff spaces or spaces with the Hausdorff topology (see below). For an elementary introduction to topology, we recommend Mendelson [ME]. Dugundji [DU] is more advanced, but is also worth consulting.

Definition 1.5. Let X be a nonempty set and let τ be a set of subsets of X. We say that τ defines a Hausdorff topology on X, or that X is Hausdorff, if

(1) X and $\emptyset \in \tau$.

(2) If $O_1, \ldots, O_n$ is a finite collection of sets in τ, then $\bigcap_{i=1}^{n} O_i \in \tau$.

(3) If Γ is a index set and, for each $\gamma \in \Gamma$, there is a set $O_\gamma \in \tau$, then $\bigcup_{\gamma\in\Gamma} O_\gamma \in \tau$.

(4) If $x, y \in X$ are any two distinct points, there are two disjoint sets $O_1, O_2 \in \tau$ (i.e., $O_1 \cap O_2 = \emptyset$), such that $x \in O_1$ and $y \in O_2$.

We call the collection τ the open sets of the topology for X. A set $N \in \tau$ is called a neighborhood for each point $x \in N$, and the set $\tau_x \subset \tau$ of all neighborhoods for x is called a complete neighborhood basis for x. Thus, any set O, containing x, also contains some neighborhood basis set $N(x) \in \tau_x$.

A set P is said to be closed if P^c is open. It follows that, if Γ is any index set and, for each $\gamma \in \Gamma$, there is a closed set $P_\gamma \in \tau$, then by De Morgan's Law, $\bigcap_{\gamma\in\Gamma} P_\gamma$ is also closed. Thus, we can also define the same topology τ, using closed sets.

Let $M \neq \emptyset$, be a subset of X.

(1) The interior of M, denoted $\mathrm{int}(M)$, is the union of all $O \in \tau$ such that $O \subset M$. If $x \in \mathrm{int}(M)$, we say that x is an interior point of M.

(2) The closure of M, which we denote by $\overline{M}$, is the set of all $x \in X$ such that, for all $N(x) \in \tau_x$, $N(x) \cap M \neq \emptyset$.

(3) We say that M is dense in X if $\overline{M} = X$. If M is also countable, we say that X is separable.

If M and N are any two subsets of X, then $\overline{M \cup N} = \overline{M} \cup \overline{N}$ and, $\overline{M} = M$ if and only if M is closed.

We say that $x_0 \in X$ is a limit point of $M \subset X$, if $x_0 \in \overline{M\backslash\{x_0\}}$ or equivalently, for every $N(x_0) \in \tau_{x_0}$, there is a $y \in N(x_0)$ and $y \notin M$.

Definition 1.6. Let (X_1, τ_1) and (X_2, τ_2) be two Hausdorff spaces. A function f, with $D(f) = X_1$ and $R(f) \subset X_2$, is said to be continuous at a point $x \in X_1$ if, for each neighborhood basis set $N[f(x)] \in \tau_{2,x}$, there is a neighborhood basis set $N(x) \in \tau_{1,x}$ such that $f[N(x)] \subset N[f(x)]$. In terms of inverse images, this says that $f^{-1}\{N[f(x)]\}$ is open in X_1 for each $N[f(x)]$ in X_2. (A little reflection shows that the above definition may be translated to the one we learned in elementary calculus, using ε's and δ's, when $X_1 = X_2 = \mathbb{R}$.) We say that f is continuous on X_1 if it is continuous at each point of X_1.

The topological space (X, τ) is said to be connected if it is not the disjoint union of two open sets. In a connected space X and $\emptyset$ are the only two sets that are both open and closed.

If Γ is a index set, $\{A_\gamma : \gamma \in \Gamma\} \subset X$ is called a cover of $M \subset X$, if $M \subset \bigcup_{\gamma\in\Gamma} A_\gamma$. If each $A_\gamma \in \tau$, we call $\{A_\gamma : \gamma \in \Gamma\}$ an open cover of M. If in addition Γ is finite, we call it a finite open cover of M.

We say that M is compact if, for every open cover $\{A_\gamma : \gamma \in \Gamma\}$, there always exists a finite subset of Γ, $\gamma_1, \ldots, \gamma_n$ such that $M \subset \bigcup_{k=1}^n A_{\gamma_k}$.

Definition 1.7. Let (X_1, τ_1) and (X_2, τ_2) be two topological spaces, with $X_1 \cap X_2 = \emptyset$. The coproduct space $(X, \tau) = (X_1, \tau_1) \oplus (X_2, \tau_2)$ is the unique topological space, with the property that each open set $O \subset X$ is of the form $O = O_1 \cup O_2$, where $O_1 \in \tau_1$ and $O_2 \in \tau_2$.

(X, τ) is also known as the disjoint union space or direct sum space. (If (X_1, τ_1) and (X_2, τ_2) are Hausdorff, then it is easy to see that (X, τ) is Hausdorff.)

1.1.3. σ-Algebras.

Definition 1.8. Let $\mathcal{A} \subset \mathcal{P}(X)$ be a collection of subsets of $X \neq \emptyset$. We say that $\mathcal{A}$ is an algebra if the following holds:

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

(2) If $A, B \in \mathcal{A}$ then $A^c, B^c \in \mathcal{A}$ and $A \cup B \in \mathcal{A}$.
It is easy to verify that:

(3) $A \cap B \in \mathcal{A}$ and $A \setminus B \in \mathcal{A}$.

(4) If n is finite and $\{A_k\} \subset \mathcal{A}$, $1 \leq k \leq n$, then

$$\bigcup_{k=1}^n A_k \in \mathcal{A}, \qquad \bigcap_{k=1}^n A_k \in \mathcal{A}.$$

Definition 1.9. Let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra. We say that $\mathcal{A}$ is a σ-algebra if

$$\bigcup_{k=1}^\infty A_k \in \mathcal{A},$$

for any countable family of sets $\{A_k\} \in \mathcal{A}$. It is also easy to see that

$$\bigcap_{k=1}^\infty A_k \in \mathcal{A},$$

along with

$$\liminf_{n \to \infty} A_n \in \mathcal{A}$$

and

$$\limsup_{n \to \infty} A_n \in \mathcal{A}.$$

Definition 1.10. If Σ is a nonempty class of subsets of X, the smallest σ-algebra $\mathcal{A}$, with $\Sigma \subset \mathcal{A}$ is called the σ-algebra generated by Σ and is written $\mathcal{A}(\Sigma)$.

Remark 1.11. Since $\Sigma \subset \mathcal{P}(X)$, there is at least one σ-algebra containing Σ.

Lemma 1.12. *If J is an index set and for each $\alpha \in J$, $\mathcal{A}_\alpha$ is σ-algebra, then $\mathcal{A} = \bigcap_{\alpha \in J} \mathcal{A}_\alpha$ is a σ-algebra.*

Definition 1.13. If $\mathcal{A}$ is a σ-algebra of subsets of a nonempty set X, we call the couple $(X, \mathcal{A})$ a measurable space.

Definition 1.14. If $\mathcal{A}$ is a σ-algebra of subsets of a nonempty set X, we call a sequence $\{A_k\} \subset \mathcal{A}$ a partition of X if the sequence is disjoint and $\bigcup_{k=1}^{\infty} A_k = X$.

Definition 1.15. If X is a topological space and Σ is the class of open sets of X, then $\mathcal{A}(\Sigma) = \mathfrak{B}(X)$ is called the Borel σ-algebra of X.

1.1.4. Measure Spaces.

Definition 1.16. Let X be a nonempty set. An outer measure ν^* is a function on $\mathcal{P}(X) \to [0, \infty]$, such that

(1) $\nu^*(\emptyset) = 0$.
(2) If $B \subset A$, then $\nu^*(B) \le \nu^*(A)$.
(3) If $A \subset \bigcup_{k=1}^{\infty} A_k$, then

$$\nu^*(A) \le \sum_{k=1}^{\infty} \nu^*(A_k).$$

If for each sequence of disjoint sets $\{A_k\} \subset \mathcal{A}$,

$$\nu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \nu(A_k),$$

we say that ν is a measure. We also say that ν is σ-additive and call the triple $(X, \mathcal{A}, \nu)$ a measure space.

Definition 1.17. Let $(X, \mathcal{A})$ be a measurable space and let $\nu(A) \in \mathbb{C}$, the complex numbers, for each $A \in \mathcal{A}$. We say that ν is a complex measure if $\nu(\emptyset) = 0$ and for each disjoint countable union $\bigcup_{k=1}^{\infty} A_k$ of sets in $\mathcal{A}$, we have

$$\nu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \nu(A_k),$$

where the convergence on the right is absolute.

Definition 1.18. Let $(X, \mathcal{A}, \nu)$ a measure space.

(1) We say that ν is a finite measure if $\nu(X) < \infty$.

(2) We say that ν is concentrated on a set $A \in \mathcal{A}$, if $A = U^c$ and U is the largest open set with the property that $\nu(U) = 0$. We also call A the support of ν.

(3) We say that ν is a regular measure if given $A \in \mathcal{A}$, for each $\varepsilon > 0$, there is a open set O and a closed set K such that: $K \subset A \subset O$ and $\nu(O \setminus K) < \varepsilon$.

(4) We say that ν is a σ-finite measure if there is a sequence $\{A_k\} \subset \mathcal{A}$, with

$$X = \bigcup_{k=1}^{\infty} A_k, \text{ and } \nu(A_k) < \infty.$$

(5) We say that ν is a Radon measure, if the set K in (3) can be chosen as compact or the sequence $\{A_k\} \subset \mathcal{A}$ in (4) can be chosen with each A_k is compact.

(6) We say that ν is a complete measure if $A \in \mathcal{A}$, with $B \subset A$ and $\nu(A) = 0$ then $B \in \mathcal{A}$ and $\nu(B) = 0$.

(7) We say that ν is a probability measure if $\nu(X) = 1$.

(8) We say that a complex measure ν is of bounded variation if

$$|\nu|(X) = \sup \sum_{k=1}^{\infty} |\nu(A_k)| < \infty,$$

where the supremum is taken over all partitions of X. We call $|\nu|(X)$ the total variation of ν.

(9) We say that the complex measure ν is a signed measure if both $|\nu| + \nu$ and $|\nu| - \nu$ are real valued. In this case, we define the positive part and the negative part by: $\nu^+ = \frac{1}{2}(|\nu| + \nu)$ and $\nu^- = \frac{1}{2}(|\nu| - \nu)$. We call this the Jordan Decomposition.

Theorem 1.19 (The Hahn Decomposition Theorem)**.** *Let ν be a signed measure on $(X, \mathcal{A})$. Then there exists a partition X_1, X_2 of X such that, for every $A \in \mathcal{A}$:*

$$\nu^+(A) = \nu(A \cap X_1) \text{ and } \nu^-(A) = -\nu(A \cap X_2).$$

Theorem 1.20 (The Jordan Decomposition Theorem)**.** *Let ν be a signed measure on $(X, \mathcal{A})$. If μ_1 and μ_2 are positive measures and $\nu = \mu_1 - \mu_2$, then $\nu^+ \leq \mu_1$ and $\nu^- \leq \mu_2$.*

Thus, the Jordan decomposition $\nu = \nu^+ - \nu^-$, has the above minimal property. If ν is complex, this decomposition becomes $\nu = \nu_1^+ - \nu_1^- + i(\nu_2^+ - \nu_2^-)$, for two positive measures, ν_1 and ν_2.

Definition 1.21. We say that X is an Abelian group if for each pair $x, y \in X$, $x \oplus y \in X$ and

(1) $x \oplus y = y \oplus x$. (The Abelian property.)
(2) For all $x, y, z \in X$ $(x \oplus y) \oplus z = x \oplus (y \oplus z)$.
(3) There is an element $0 \in X$ called the identity and $x \oplus 0 = 0 \oplus x = 0$, for all $x \in X$.
(4) For each $y \in X$, there is a unique element $y^- \in X$, such that $y \oplus y^- = y^- \oplus y = 0$.
(5) We say that Y is a subgroup of X if $Y \subset X$ and for all $y_1, y_2 \in Y$, $y_1 \oplus y_2 \in Y$, satisfying conditions (1)–(4) above.

The real or complex numbers form an Abelian group with addition (or multiplication if we exclude zero). The rational numbers (real or complex) form a subgroup, with the same exception for multiplication.

When X is an Abelian group (with $\oplus = +$) and $(X, \mathcal{A}, \nu)$ is a measure space, we say that $\mathfrak{T}$ is an admissible translation invariance group for $(X, \mathcal{A}, \nu)$ if $\mathfrak{T}$ is a subgroup of X and $\nu(A - t) = \nu(A)$, for all $t \in \mathfrak{T}$. If $\mathfrak{T} = X$, we say that ν is translation invariant on X.

1.1.5. Integral.

Let $(X, \mathcal{A}, \nu)$ a measure space.

Definition 1.22. Let f be a function on X, $f : X \to K$, where $K = \mathbb{R}$ or $\mathbb{C}$.

(1) We say that f is measurable if $f^{-1}(B) \in \mathcal{A}$, for every set $B \in \mathfrak{B}[K]$, the Borel algebra on K. In this case, we say that $f \in \mathcal{M}[X]$ or $\mathcal{M}$, when X is understood.
(2) We say that two functions f and g are equal almost everywhere and write $f(x) = g(x)$, ν-(a.e.), if they have the same domain and $\nu\{x : f(x) \neq g(x)\} = 0$. In general, a property is said to hold ν-(a.e.) on X if the set of points where this property fails has ν-measure zero.

Definition 1.23. A (nonnegative) simple function s is defined on X by

$$s(x) = \sum_{k=1}^{n} a_k \chi_{A_k}(x),$$

where the $a_k \in [0, \infty)$ and the family of measurable sets $\{A_k\}$ form a (finite) partition of X (i.e., $\nu(A_i \cap A_j) = 0$, $i \neq j$ and $\bigcup_{k=1}^n A_k = X$).

(By convention, if need be, we can always add a set A_{n+1} to the collection and define $a_{n+1} = 0$ so that the union is always X.)

Lemma 1.24. *If $0 \leq f \in \mathcal{M}$, then there is a sequence of simple functions $\{s_n\}$, with $s_n \leq s_{n+1}$ and $s_n \to f$ (a.e.) at each point of X, as $n \to \infty$.*

Definition 1.25. If $f : X \to [0, \infty]$ is a measurable function and $A \in \mathfrak{B}(X)$, we define the integral of f over A by:

$$\int_A f(x) d\nu = \lim_{n\to\infty} \int_A s_n(x)\, d\nu,$$

where $\{s_n\}$ is any increasing family of simple functions converging to $f(x)$.

Theorem 1.26. *If f, g are nonnegative measurable functions and $0 \leq c < \infty$, we have:*

(1) *$\int_X f(x) d\nu(x)$ is independent of the family of simple functions used;*

(2) *$0 \leq \int_X f(x) d\nu(x) \leq \infty$;*

(3) *$\int_X cf(x) d\nu(x) = c \int_X f(x) d\nu(x)$;*

(4)
$$\int_X [f(x) + g(x)] d\nu(x) = \int_X f(x) d\nu(x) + \int_X g(x) d\nu(x).$$

(5) *If $f \leq g$, then $\int_X f(x) d\nu(x) \leq \int_X g(x) d\nu(x)$.*

Theorem 1.27 (Fatou's Lemma)**.** *Let $\{f_n\} \subset \mathcal{M}$ be a nonnegative family of functions, then:*

$$\int_X \left(\liminf_{n\to\infty} f_n(x) \right) d\nu(x) \leq \liminf_{n\to\infty} \int_X f_n(x) d\nu(x).$$

Theorem 1.28 (Monotone Convergence Theorem)**.** *Let $\{f_n\} \subset \mathcal{M}$ be a nonnegative family of functions, with $f_n \leq f_{n+1}$. Then:*

$$\lim_{n\to\infty} \int_X f_n(x) d\nu(x) = \int_{\mathcal{B}} \left(\lim_{n\to\infty} f_n(x) \right) d\nu(x).$$

Definition 1.29. If $f \in \mathcal{M}$, we define

$$\int_X f(x) d\nu(x) = \int_X f_+(x) d\nu(x) - \int_X f_-(x) d\nu(x),$$

where $f_+(x) = \frac{1}{2}(|f(x)| + f(x))$ and $f_-(x) = \frac{1}{2}(|f(x)| - f(x))$. We say that f is integrable whenever both integrals on the right are finite. The set of all integrable functions is denoted by $\mathcal{L}^1[X, \mathfrak{B}(X), \nu] = \mathcal{L}^1[X]$.

Remark 1.30. As is carefully discussed in elementary analysis, the functions in $\mathcal{L}^1[X]$ are not uniquely defined. Following tradition, we let $L^1[X]$ denote the set of equivalence classes of functions in $\mathcal{L}^1[X]$ that differ by a set of ν-measure zero. By a slight abuse, we will identify an integrable function f as measurable (in $\mathcal{L}^1[X]$) and its equivalence class in $L^1[X]$. The same convention also applies to functions in $L^p[X]$ and will be used later without further comment.

Theorem 1.31 (Dominated Convergence Theorem)**.** *Let* $f_n \in \mathcal{M}[X, \nu]$*,* $n \in \mathbb{N}$*,* $g \in L^1(X)$*, with* $g \geq 0$ *and* $|f_n(x)| \leq g(x)$*,* ν-(a.e.). *If* $\lim_{n\to\infty} f_n(x)$ *exists* ν-(a.e.), *then* $\lim_{n\to\infty} f_n \in L^1[X]$ *and*

$$\lim_{n\to\infty} \int_X f_n(x) d\nu(x) = \int_X \left(\lim_{n\to\infty} f_n(x) \right) d\nu(x).$$

1.2. Functional Analysis

In this section, we include a few basic background results from functional analysis and Banach space theory. Detailed discussions can be found in Dunford and Schwartz [DS], Hille and Phillips [HP], Lax [L1], Reed and Simon [RS], Rudin [RU], or Yosida [YS].

1.2.1. Topological Vector Spaces.

Definition 1.32. A vector space $\mathfrak{X}$ over $\mathbb{C}$ is an Abelian group under addition that is closed under multiplication by elements of $\mathbb{C}$. That is:

(1) For each $x, y \in \mathfrak{X}$, $x + y \in \mathfrak{X}$.

(2) For all $x, y, z \in \mathfrak{X}$, $x + y = y + x$ and $(x + y) + z = x + (y + z)$.

(3) There is a unique element $0 \in \mathfrak{X}$ called zero and $x + 0 = 0 + x = x$ for all $x \in \mathfrak{X}$.

(4) For all $x \in \mathfrak{X}$, there is a unique element $-x \in \mathfrak{X}$ and $x + (-x) = (-x) + x = 0$.

(5) For all $x, y \in \mathfrak{X}$ and $a, b \in \mathbb{C}$, $ax \in \mathfrak{X}$, $1x = x$, $(ab)x = a(bx)$ and $a(x + y) = ax + ay$. We call $b \in \mathbb{C}$ a scalar.

If $\mathfrak{X}$ is a vector space over $\mathbb{C}$, a mapping $\rho(\cdot) : \mathfrak{X} \to [0, \infty)$ is a seminorm on $\mathfrak{X}$ if:

(1) For each $x,\ y \in \mathfrak{X},\ \rho(x) \geq 0$ and $\rho(x+y) \leq \rho(x) + \rho(y)$.

(2) For each $\lambda \in \mathbb{C}$ and each $x \in \mathfrak{X},\ \rho(\lambda x) = |\lambda|\, \rho(x)$.

Definition 1.33. Let V be a subset of $\mathfrak{X}$.

(1) We say that V is a convex subset of $\mathfrak{X}$ if for each $x, y \in V,\ \alpha x + (1-\alpha)y \in V$, for all $\alpha \in [0,1]$.

(2) We say that V is an balanced subset of $\mathfrak{X}$ if for each $x \in V$ and $\alpha \in \mathbb{C}$, with $|\alpha| \leq 1,\ \alpha x \in V$.

(3) We say that V is an absolutely convex subset of $\mathfrak{X}$ if it is both convex and balanced.

(4) We say that V is a absorbent subset of $\mathfrak{X}$ if for each $x \in \mathfrak{X},\ \alpha x \in V$, for some $\alpha > 0$. Thus, every point in $x \in \mathfrak{X}$ is in αV for some positive α.

Definition 1.34. A locally convex topological vector space is a vector space with its topology defined by a family of semi-norms $\{\rho_\gamma\}$, where γ is in some index set Γ. Given any $x \in \mathfrak{X}$, a base of ε-neighborhoods about x is a set of the form $V_{\Gamma_0,\varepsilon}(x)$, where Γ_0 is a finite subset of Γ and

$$V_{\Gamma,\varepsilon}(x) = \{\, y \in \mathfrak{X} : \rho_\gamma(x-y) < \varepsilon,\ \gamma \in \Gamma \,\}.$$

Definition 1.35. A locally convex topological vector space $\mathfrak{X}$ is a Fréchet space if it satisfies the following:

(1) $\mathfrak{X}$ is a Hausdorff space.

(2) The neighborhood base about each $x \in \mathfrak{X}$ is induced by a countable number of seminorms (i.e., Γ is a countable set).

(3) $\mathfrak{X}$ is a complete relative to the family of seminorms.

Theorem 1.36. *The vector space $\mathfrak{X}$ is a Fréchet space if and only if:*

(1) *$\mathfrak{X}$ is a locally convex.*

(2) *There is a metric $d : \mathfrak{X} \times \mathfrak{X} \to [0,\infty)$ such that, for all $x, y, z \in \mathfrak{X},\ d(x+z, y+z) = d(x,y)$.*

(3) *$\mathfrak{X}$ is a complete relative to the metric $d(\cdot,\cdot)$.*

Remark 1.37. If the index Γ for the family of semi-norms is countable, then we can define a metric $d(x,y)$ by:

$$d(x,y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\rho_n(x-y)}{1+\rho_n(x-y)}.$$

A sequence $\{x_n\}$ in a metric space $\mathfrak{X}$ converges to a limit $x \in \mathfrak{X}$ if and only if $\lim_{n\to\infty} d(x_n, x) = 0$. In this case, by the triangle inequality

$$d(x_n, x_m) \leq d(x_n, x) + d(x_m, x).$$

We say that a sequence satisfies the Cauchy convergence condition, or is a Cauchy sequence if

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

A metric space is said to be complete if every Cauchy sequence converges to a point in the space.

1.2.2. Separable Banach Spaces. Hilbert and Banach spaces are discussed further in Chaps. 4 and 5. Let $\mathcal{B}$ be a vector space over $\mathbb{R}$ or $\mathbb{C}$. We say that $\mathcal{B}$ is separable if it contains a countable dense subset.

Definition 1.38. A norm on a vector space $\mathcal{B}$ is a mapping $\|\cdot\|_{\mathcal{B}}$: $\mathcal{B} \to [0, \infty]$, such that

(1) $\|x\|_{\mathcal{B}} = 0$ if and only if $x = 0$.
(2) $\|ax\|_{\mathcal{B}} = |a| \, \|x\|_{\mathcal{B}}$ for all $x \in \mathcal{B}$ and $a \in \mathbb{C}$.
(3) $\|x + y\|_{\mathcal{B}} \leq \|x\|_{\mathcal{B}} + \|y\|_{\mathcal{B}}$, for all $x, y \in \mathcal{B}$.
(4) We say that $\mathcal{B}$ is uniformly convex if, for each $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) > 0$ such that, for all $x, y \in \mathcal{B}$ with

$$\max(\|x\|, \|y\|) \leqslant 1, \ \ \|x - y\| \geqslant \varepsilon \ \Rightarrow \tfrac{1}{2} \ \|x + y\| \leqslant 1 - \delta.$$

The topology on $\mathcal{B}$ is generated by the metric defined by:

$$d(x, y) = \|x - y\|_{\mathcal{B}},$$

so that $\{x : \|x - y\|_{\mathcal{B}} < r\}$ is an open ball about y of radius r.

The space $\mathcal{B}$ is complete if every Cauchy sequence in the above norm converges to an element in $\mathcal{B}$. A complete normed space is called a Banach space.

Definition 1.39. Let $\mathcal{B}$ be a Banach space and let A be a transformation on $\mathcal{B}$, with domain $D(A)$ (i.e., $A : D(A) \subset \mathcal{B} \to \mathcal{B}$).

(1) We say that A is a linear operator on $\mathcal{B}$, if $A(ax + by) = aAx + bAy$, for all $a, b \in \mathbb{C}$ and all $x, y \in D(A)$.
(2) We say that A is densely defined if $D(A)$ is dense in $\mathcal{B}$.

(3) We say that A is a closed linear operator if and only if the following condition is satisfied: $\{x_n\} \subset D(A), x_n \to x$ and $Ax_n \to z$ always implies that $x \in D(A)$ and $z = Ax$.

(4) We say that A is a bounded linear operator if and only if $D(A) = \mathcal{B}$ and

$$\sup_{\|x\|_{\mathcal{B}} \leqslant 1} \|Ax\|_{\mathcal{B}} < \infty.$$

In this case we define the norm of A, $\|A\|_{\mathcal{B}}$, by the above supremum.

1.2.2.1. *Dual Spaces.*

Definition 1.40. Let $\mathcal{B}$ be a Banach space.

(1) The dual space $\mathcal{B}'$ is the set of all bounded linear operators $x^* : \mathcal{B} \to \mathbb{C}$ (called bounded linear functionals on $\mathcal{B}$). The norm of x^* is defined by:

$$\|x^*\|_{\mathcal{B}'} = \sup_{\|x\|_{\mathcal{B}} \leq 1} |x^*(x)| = \sup_{\|x\|_{\mathcal{B}} \leq 1} |\langle x, x^* \rangle| \,.$$

With this norm $\mathcal{B}'$ is a Banach space. We write $\mathcal{B}'$ as $\mathcal{B}'_s$ and call it the strong dual. The topology is known as the strong topology.

(2) The weak and weak* topology are defined on $\mathcal{B}$ and $\mathcal{B}'$ respectively in the following manner:
- A sequence $\{x_n\} \subset \mathcal{B}$ is said to converge in the weak topology to $x \in \mathcal{B}$ if and only if, for each bounded linear functional $y^* \in \mathcal{B}'$,

$$\lim_{n\to\infty} y^*(x_n) = y^*(x).$$

We also write $w - \lim_{n\to\infty} x_n = x$.
- A sequence $\{x_n^*\} \subset \mathcal{B}'$ is said to converge in the weak* topology to $x^* \in \mathcal{B}'$ if and only if, for each $y \in \mathcal{B}$,

$$\lim_{n\to\infty} x_n^*(y) = x^*(y).$$

We also write $w^* - \lim_{n\to\infty} x_n^* = x^*$.

(3) If $\mathcal{B} = \mathcal{B}''$, we say that $\mathcal{B}$ is reflexive.

(4) A duality map $\mathcal{J} : \mathcal{B} \mapsto \mathcal{B}'$ is a set

$$\mathcal{J}(u) = \left\{ u^* \in \mathcal{B}' \,\middle|\, u^*(u) = \langle u, u^* \rangle = \|u\|^2 = \|u^*\|^2 \right\}, \text{ for all } u \in \mathcal{B}.$$

Remark 1.41. The following remarks are important.

(1) In the definition, we used x^* to represent an element in $\mathcal{B}'$. The notation used varies with the tradition of the particular topical area. To the extent possible, we will try to be consistent within topics studied and the tradition of the field so that the reader will see some correspondence when consulting references for different topics.

(2) It is easy to see that

$$|y^*(x_n) - y^*(x)| \leq \|x_n - x\|_{\mathcal{B}} \|y^*\|_{\mathcal{B}'}$$

for all $y^* \in \mathcal{B}$, so that norm convergence in $\mathcal{B}$ always implies weak convergence. It is also easy to see that

$$|x_n^*(y) - x^*(y)| \leq \|x_n^* - x^*\|_{\mathcal{B}'} \|y\|_{\mathcal{B}},$$

for all $y \in \mathcal{B}$, so that norm convergence in $\mathcal{B}'$ always implies weak* convergence. However (in both cases), the reverse is not true (see Lax [L1, p. 106]).

(3) It is known that every uniformly convex Banach space is reflexive. Furthermore, when $\mathcal{B}$ is uniformly convex, the duality set $\mathcal{J}(u)$, is single valued and uniquely defined by u. However, if $\mathcal{B}$ is not uniformly convex, the duality set $\mathcal{J}(u)$ can have the power of the continuum.

The following examples will help one see what is possible in concrete cases.

(1) If λ_n is Lebesgue measure on $\mathbb{R}^n$, $u \in L^p[\mathbb{R}^n]$, $1 < p < \infty$ and q is such that $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\mathcal{J}(u)(x) = \|u\|_p^{2-p} |u(x)|^{p-2} u(x) = u^* \in L^q\mathbb{R}^n],$$

and

$$\langle u, u^* \rangle = \|u\|_p^{2-p} \int_{\mathbb{R}^n} |u(x)|^p \, d\lambda_n(x) = \|u\|_p^2 = \|u^*\|_q^2.$$

Thus, it is easy to see that $(L^p[\mathbb{R}^n])'' = L^p[\mathbb{R}^n]$, so that $L^p[\mathbb{R}^n]$ is reflexive for $1 < p < \infty$.

(2) The space $L^1[\mathbb{R}^n]$ is not reflexive, for if $u \in L^1[\mathbb{R}^n]$, then

$$\mathcal{J}(u)(x) = \{ v \in L^\infty[\mathbb{R}^n]| \ : v(x) \in \{\|u\|_1 \operatorname{sign}[u(x)]\}\},$$

where

$$\operatorname{sign}[u(x)] = \begin{cases} 1, & u(x) > 0, \\ -1, & u(x) < 0, \\ [-1,1], & u(x) = 0. \end{cases}$$

It follows that $\mathcal{J}(u)(x)$ is uncountable for each $u \in L^1[\mathbb{R}^n]$.

The transpose matrix on $\mathbb{R}^n$ or the transpose conjugate matrix on $\mathbb{C}^n$ has its parallel for Banach spaces. In this case, they are known as dual operators. They are also known as adjoint operators, but we will reserve this term for a special class of operators on Banach spaces, discussed in Chap. 5. We will also use adjoint for the same class defined on Hilbert spaces in the next section and explain the distinction.

Definition 1.42. Let $A : D(A) \to \mathcal{B}$ be a closed linear operator on $\mathcal{B}$ with a dense domain $D(A)$. The dual of A, A' is defined on $\mathcal{B}'$ as follows. Its domain $D(A')$ is the set of all $y^* \in \mathcal{B}'$ for which there exists an $x^* \in \mathcal{B}'$ such that

$$\langle Ax, y^* \rangle = \langle x, x^* \rangle ,$$

for all $x \in D(A)$; in this case we define $A'y^* = x^*$.

A proof of the following theorem can be found in [HP] or [YS].

Theorem 1.43. *Let $A : D(A) \to \mathcal{B}$ be a closed linear operator on $\mathcal{B}$ with a dense domain $D(A)$.*

(1) *Then $A' : D(A') \to \mathcal{B}'$ is a closed linear operator on $\mathcal{B}'$ and its domain $D(A')$ is dense in $\mathcal{B}'$.*
(2) *If, in addition, $\|A\|_{\mathcal{B}} < \infty$, then $D(A') = \mathcal{B}'$ and $\|A'\|_{\mathcal{B}'} = \|A\|_{\mathcal{B}}$.*

1.2.2.2. *Hilbert Space.*

Definition 1.44. An inner product on $\mathcal{B} = \mathcal{H}$ is a bilinear mapping $(\cdot,\cdot)_{\mathcal{H}} : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$, such that

(1) $(x,x)_{\mathcal{H}} \geq 0$ and $(x,x)_{\mathcal{H}} = 0$ if and only if $x = 0$.
(2) $(ax + by, z)_{\mathcal{H}} = a(x,z)_{\mathcal{H}} + b(y,z)_{\mathcal{H}}$ and $(w, ax + by)_{\mathcal{H}} = a^c(w,x)_{\mathcal{H}} + b^c(w,y)_{\mathcal{H}}$.

If $(\cdot,\cdot)_{\mathcal{H}}$ is a inner product, it induces a norm on $\mathcal{H}$ by

$$\|x - y\|_{\mathcal{H}} = \sqrt{(x-y, x-y)_{\mathcal{H}}}.$$

If $\mathcal{H}$ is complete with this norm, we call it a Hilbert space.

If $(\cdot,\cdot)_{\mathcal{H}}$ is the inner product on the Hilbert space $\mathcal{H}$, then the same Cauchy–Schwarz inequality from $\mathbb{R}^n$ still holds, $|(x,y)_{\mathcal{H}}| \le \|x\|_{\mathcal{H}}\,\|y\|_{\mathcal{H}}$.

The following polarization identity also holds for a general Hilbert space:

$$(x,y)_{\mathcal{H}} = \frac{1}{4}\left(\|x+y\|_{\mathcal{H}}^2 - \|x-y\|_{\mathcal{H}}^2\right),$$

if the field of $\mathcal{H}$ is $\mathbb{R}$ and

$$(x,y)_{\mathcal{H}} = \frac{1}{4}\left\{\left(\|x+y\|_{\mathcal{H}}^2 - \|x-y\|_{\mathcal{H}}^2\right) + i\left(\|x+iy\|_{\mathcal{H}}^2 - \|x-iy\|_{\mathcal{H}}^2\right)\right\},$$

if the field of $\mathcal{H}$ is $\mathbb{C}$.

Definition 1.45. Let $A : D(A) \to \mathcal{H}$ be a closed linear operator on $\mathcal{H}$ with a dense domain $D(A)$. The adjoint of A, A^* is defined on $\mathcal{H}$ as follows. Its domain $D(A^*)$ is the set of all $y \in \mathcal{H}$ for which there exists an $x \in \mathcal{H}$ such that

$$(Ax, y)_{\mathcal{H}} = (x, A^*y)_{\mathcal{H}}.$$

We will always call A^* the adjoint of A when it is defined on the same space and A', the dual of A when it is defined on the dual space. In Chap. 5, we will see that the adjoint is also possible for a certain class of Banach spaces, which include the uniformly convex ones.

Theorem 1.43 can be slightly modified to show that $D(A^*)$ is dense in $\mathcal{H}$ and, if $\|A\|_{\mathcal{H}} < \infty$, then $D(A^*) = \mathcal{H}$ and $\|A^*\|_{\mathcal{H}} = \|A\|_{\mathcal{H}}$.

Recall that, two functions $f, g \in \mathcal{H}$ are orthogonal, if $(f,g)_{\mathcal{H}} = 0$ and they are orthonormal if in addition, $\|f\|_{\mathcal{H}} = \|g\|_{\mathcal{H}} = 1$. A set $\{\phi_n\} \subset \mathcal{H}$ is an orthonormal basis for $\mathcal{H}$ if they are orthonormal and each $x \in \mathcal{H}$ can be written as $x = \sum_{k=1}^{\infty} a_k\phi_k$, for a unique family of scalars $\{a_n\} \subset \mathbb{C}$.

Definition 1.46. Let A be a linear operator defined on $\mathcal{H}$.

(1) We say that A is a projection operator if $A^2x = Ax$ for all $x \in \mathcal{H}$.

(2) We say that A is the self-adjoint if $D(A) = D(A^*)$ and $Ax = A^*x$, for all $x \in D(A)$.

(3) We say that a bounded linear operator A is the compact, if for every bounded sequence $\{x_n\} \subset \mathcal{H}$, the sequence $\{Ax_n\}$ has a convergent subsequence.

(4) We say that a compact operator A is trace class if, for some orthonormal basis $\{\phi_n\}$ of $\mathcal{H}$, the trace of A, $tr[A]$ is finite,

where

$$tr[A] = \sum_{n=1}^{\infty} (A\phi_n, \phi_n).$$

It is easy to check that the trace (if it exists) is independent of the basis used.

1.2.3. The Hahn–Banach Theorem.

Theorem 1.47. *Let $\mathcal{B}$ be a Banach space over $\mathbb{C}$ and let $p : \mathcal{B} \to \mathbb{R}$ be such that, for all $x, y \in \mathcal{B}$*

$$p(ax + by) \leq |a|\, p(x) + |a|\, p(y), \text{ whenever } |a| + |b| = 1. \tag{1.1}$$

If $\bar{L}$ is a linear functional defined on a subspace $\mathcal{D} \subset \mathcal{B}$, with $\left|\bar{L}(x)\right| \leq p(x)$, for all $x \in \mathcal{D}$, then $\bar{L}$ can be extended to a linear functional L on $\mathcal{B}$ such that $|L(x)| \leq p(x),\ x \in \mathcal{B}$ and $L(x) = \bar{L}(x)$ on $\mathcal{D}$.

Proof. We first assume that the field is $\mathbb{R}$. Suppose that $x \in \mathcal{B}$ but $x \notin \mathcal{D}$. Let $\mathcal{E} = (x, \mathcal{D})$ be the vector space spanned by x and $\mathcal{D}$. If we have an extension L of $\bar{L}$ from $\mathcal{D}$ to $\mathcal{E}$, it must satisfy

$$L\,(ax + by) = \lambda L(x) + \bar{L}(y),\ y \in \mathcal{D}.$$

and from (1.1), $|a| + |b| = 1$ implies that

$$p(ax + by) \leq |a|\, p(x) + |b|\, p(y).$$

Suppose that $y_1, y_2 \in \mathcal{D},\ a, b > 0,\ a + b = 1$. Then

$$\begin{aligned} &a\bar{L}(y_1) + b\bar{L}(y_2) = \bar{L}(ay_1 + by_2) \leqslant p[a(y_1 - \tfrac{1}{a}x) + b(y_2 + \tfrac{1}{b}x)] \\ &\leqslant ap(y_1 - \tfrac{1}{a}x) + bp(y_2 + \tfrac{1}{b}x). \end{aligned}$$

We see that for all $y_1, y_2 \in \mathcal{D}$ and all $a, b > 0,\ a + b = 1$, we have

$$\tfrac{1}{a}\left[-p[y_1 - ax) + \bar{L}(y_1)\right] \leqslant \tfrac{1}{b}\left[p(y_2 + bx) - \bar{L}(y_2)\right].$$

It now follows that we must be able to find a number c such that for all $a > 0$,

$$\sup_{y \in D} \tfrac{1}{a}\left[-p[y - ax) + \bar{L}(y)\right] \leqslant c \leqslant \inf_{y \in D} \tfrac{1}{a}\left[p(y + ax) - \bar{L}(y)\right].$$

We can define $L(x) = c$. It is easy to check that $L(x) \leq p(x)$, for all $x \in \mathcal{E}$. We now appeal to Zorn's Lemma (see Yosida [YS, p. 2]), to show that $\bar{L}$ can be extended to all of $\mathcal{B}$, when the field is $\mathbb{R}$.

To extend our result to complex linear functionals, let $\bar{L}$ be given on $\mathcal{D}$ and define $L'(x) = \mathrm{Re}\{\bar{L}(x)\}$, so that it is a real linear functional on $\mathcal{D}$. Since

$$L'(ix) = \mathrm{Re}\{i\bar{L}(x)\} = -\mathrm{Im}\{\bar{L}(x)\},$$

we see that $\bar{L}(x) = L'(x) - iL'(ix)$. Furthermore, since L' is real, it has an extension L, to all of $\mathcal{B}$ such that $L(x) \leq p(x)$, for all $x \in \mathcal{B}$. We can now define $F(x) = L(x) - iL(ix)$. It is easy to check that F is a complex linear functional.

Since $|a| = 1$ implies that $p(ax) = p(x)$, we can set $\theta = \mathrm{Arg}\{F(x)\}$. If we now use the fact that $\mathrm{Re}\{F\} = L$, we have

$$|F(x)| = e^{-i\theta}F(x) = F(e^{-i\theta}x) = L(e^{-i\theta}x) \leqslant p(e^{-i\theta}x) = p(x),$$

we are done. □

Theorem 1.48. *Let M be a linear subspace of $\mathcal{B}$. If $x_0 \in \mathcal{B}$, with $0 < c = d(M, x_0)$, then there exists a bounded linear functional $L(\cdot)$ defined on $\mathcal{B}$ such that*

$$L(x_0) = 1, \ \|L\|_{\mathcal{B}} = \frac{1}{c}, \ L(x) = 0, \text{ for all } x \in M.$$

Proof. Let $M_1 = (M, x_0)$ be the subspace spanned by M and x_0. Thus, each point $z \in M_1$ is of the form $z = y + \lambda x_0$, where $y \in M$ and $\lambda \in \mathbb{C}$ are uniquely determined by z. Define $F(\cdot)$ on $\mathcal{B}$ by $F(y + \lambda x_0)$ $= \lambda$. Clearly F is a bounded linear functional, and if $\lambda \neq 0$ then

$$\|y + \lambda x_0\|_{\mathcal{B}} = \left\|\frac{y}{\lambda} + x_0\right\|_{\mathcal{B}} |\lambda| \geqslant c\,|\lambda|\,.$$

It follows that $|F(z)| \leqslant \frac{1}{c}\|z\|_{\mathcal{B}}$, so that $\|F\|_{\mathcal{B}'} \leqslant \frac{1}{c}$. If $\{z_n\} \subset M$, $\|x_0 - z_n\| \to c$, then

$$1 = F(x_0 - z_n) \leqslant \|q\|_{\mathcal{B}} \|x_0 - z_n|\_\mathcal{B} \to c\,\|F\|_{\mathcal{B}}\,.$$

Thus, $\|F\|_{\mathcal{B}} = \frac{1}{c}$. Thus, by Theorem 1.47 with L replacing F finishes the proof. □

The following is a consequence of the last two results.

Theorem 1.49. *If $\mathcal{B}$ is a Banach space, we have*

(1) *For any $x \in \mathcal{B}$, $x \neq 0$, there exists a linear functional $L \in \mathcal{B}'$ such that $\|L\|_{\mathcal{B}'} = 1$ and $L(x) = \|x\|_{\mathcal{B}}$.*

(2) *If $x \neq y$, there exists a linear functional $L \in \mathcal{B}'$ with $L(x) \neq L(y)$.*

(3) *For* $x \in \mathcal{B}$,

$$\|x\|_{\mathcal{B}} = \sup_{L \neq 0} \frac{|L(x)|}{\|L\|_{\mathcal{B}'}} = \sup_{\|L\|=1} |L(x)| \,.$$

(4) *If* M *is a subspace of* $\mathcal{B}$ *and* $x_0 \in \mathcal{B}$, $x_0 \notin \overline{M}$, *then there exists a linear functional* $L \in \mathcal{B}'$ *such that* $L(x_0) = 1$ *and* $L(x) = 0$, *for all* $x \in \overline{M}$.

1.2.4. The Baire Category Theorem. In this section we introduce Baire's Theorem and some of its consequences. First we need a definition.

Definition 1.50. Let $\mathcal{B}$ be a Banach space. A subset $E \subset \mathcal{B}$ is said to be nowhere dense if its closure has empty interior. A set is said to be meager (or of the first category) in $\mathcal{B}$ if it is a countable union of nowhere dense sets. A set in $\mathcal{B}$ that is not meager (not of the first category) in $\mathcal{B}$ is said to be nonmeager (of the second category) in $\mathcal{B}$.

Theorem 1.51. (Baire's Theorem) *If* $\mathcal{B}$ *is a Banach space, then the intersection of every countable collection of dense open subsets of* $\mathcal{B}$ *is a dense set in* $\mathcal{B}$.

Proof. Let $\{U_1, U_2, U_3, \ldots\}$ be any countable collection of dense open subsets of $\mathcal{B}$. If T_0 is any ball in $\mathcal{B}$ of radius 1, choose a ball T_1 of radius $\frac{1}{2}$ such that the closure of T_1, $\overline{T}_1 \subset U_1 \cap T_0$. (Check that this is possible.) Continue this process, so that at step n, we choose a ball T_n of radius $\frac{1}{n}$ such that $\overline{T}_n \subset U_n \cap T_{n-1}$ and define K by:

$$K = \bigcap_{n=1}^{\infty} \overline{T}_n.$$

It is easy to see that the centers of our nested balls form a Cauchy sequence that converges to a point in K, so that K is nonempty. Since $K \subset T_0$ and $K \subset T_n$ for each n, we see that the intersection of T_0 with $\bigcap_{n=1}^{\infty} U_n$ is nonempty. □

The following two lemmas are required for our proof of the Banach–Steinhaus Theorem in the next section. The second lemma is true for an arbitrary index set, but for our use the restriction of the index set to $\mathbb{R}^+$ is sufficient.

Lemma 1.52. *Let $\mathcal{B}$ be a Banach space. Suppose that $\{V_1, V_2, V_3, \ldots\}$ is a countable collection of closed subsets of $\mathcal{B}$ with* $\operatorname{int}(V_n) = \emptyset$. *Then* $V = \bigcap_{n=1}^{\infty} V_n = \emptyset$.

Proof. Since V is meager and $\operatorname{int}(V) \subset V$, it follows that $\operatorname{int}(V)$ is meager. By Baire's Theorem, we see that $\operatorname{int}(V) = \emptyset$. □

Lemma 1.53. *Let $\mathcal{B}$ be a Banach space. Suppose that $\{f_t\}$, $t \in \mathbb{R}^+$ is a pointwise bounded family of continuous real-valued functions on $\mathcal{B}$. Then the family is uniformly bounded on some nonempty open subset of $\mathcal{B}$.*

Proof. Suppose that $|f_t(\varphi)| \leq c_\varphi$ for all $t \in \mathbb{R}^+$ and define

$$V_n^t = \{\varphi \in B \mid |f_t(\varphi)| \leq n\}.$$

It is clear that V_n^t is closed in $\mathcal{B}$, since f_t is continuous. Therefore the set:

$$V_n = \bigcap_{n=1}^{\infty} \{\varphi \in B \mid |f_t(\varphi)| \leq n\},$$

defined for each n, is also closed in $\mathcal{B}$. Since f_t is pointwise bounded, we have that $\mathcal{B} = \bigcup_{n=1}^{\infty} V_n$. If $\operatorname{int}(V_m) = \emptyset$ for all m, then from Lemma 1.52, $\bigcup_{n=1}^{\infty} V_n$ is meager. Since $\mathcal{B}$ is of the second category, this is a contradiction. Therefore, $\operatorname{int}(V_m) \neq \emptyset$ for some m. If we set $M = m$ and $U = \operatorname{int}(V_m)$, it follows that $\{f(t)\}$, $t \in \mathbb{R}^+$ is uniformly bounded on U. □

1.2.5. The Banach–Steinhaus Theorem. The next important result, known in the early literature as the Banach–Steinhaus Theorem, is much better known now as the principle of uniform boundedness.

Theorem 1.54 (Uniform Boundedness Theorem)**.** *Let $\{T(t)\}$ be a family of continuous mappings on the Banach space $\mathcal{B}$ for $t \in \mathbb{R}^+$. If for each $\varphi \in \mathcal{B}$, the family $\{\|T(t)\varphi\|_{\mathcal{B}}\}$ is bounded for all $t \in \mathbb{R}^+$, then the $\{\|T(t)\|_{\mathcal{B}}\}$ is a bounded family.*

Proof. For each $t \in \mathbb{R}^+$ define $f_t : \mathcal{B} \to \mathbb{R}^+$ by $f_t(\varphi) = \|T(t)\varphi\|_{\mathcal{B}}$. Since the norm is continuous, we see that f_t is also continuous. From Lemma 1.53, there is a nonempty open set $V_{n_0} \subset \mathcal{B}$ and $f_t(\varphi) \leq n_0$ for all $t \in \mathbb{R}^+$ and all $\varphi \in V_{n_0}$. Without loss of generality, we can assume that $U = \{\varphi | \; \|\varphi\|_{\mathcal{B}} < r\} \subset V_{n_0}$ for some $r > 0$. It follows that, for

$\varphi_0 \in U$, $\|f_t(\varphi + r\phi)\|_{\mathcal{B}} \leqslant n_0$ for all $t \in \mathbb{R}^+$ and all ϕ with $\|\phi\|_{\mathcal{B}} < 1$. This implies that

$$r\|T\|_{\mathcal{B}} = r \sup_{\|\phi\|_{\mathcal{B}} \leqslant 1} \|T(t)\phi\|_{\mathcal{B}} = \sup_{\|\phi\|_{\mathcal{B}} \leqslant 1} \|T(t)\left[\varphi_0 + r\phi\right] - T(t)\varphi_0\|_{\mathcal{B}}$$
$$\leqslant n_0 + \|T(t)\varphi_0\|_{\mathcal{B}} < \infty.$$

□

The next result (the open mapping theorem) is one of the important theorems in functional analysis. It is used to prove the two theorems that follow. The first of the two will be used in the next section, while the second is fundamental for Chaps. 4 and 5.

Theorem 1.55 (Open Mapping Theorem)**.** *Let $\mathcal{B}_1, \mathcal{B}_2$ be two Banach spaces and let A be a continuous linear surjective mapping of $\mathcal{B}_1 \to \mathcal{B}_2$. Then, whenever U is an open set in $\mathcal{B}_1$, $A[U]$ is an open set in $\mathcal{B}_2$.*

Proof. It suffices to show that, for every open ball U about zero in $\mathcal{B}_1$, $A[U]$ contains an open ball about zero in $\mathcal{B}_2$. Hence, fix U and let $\{U_0, U_1, U_2, \ldots\}$ be a sequence of open balls of radius $r/2^n$, $(n = 0,\ 1,\ 2, \ldots)$, where r is chosen so that $U_0 \subset U$. We are done if we can prove that there is an open set W such that:

$$W \subset \overline{A(U_1)} \subset A(\mathbf{U}),$$

where $\overline{A(U_1)}$ is the closure of $A(U_1)$. Since $U_2 - U_2 \subset U_1$, we first need to prove that $W \subset \overline{A(U_1)}$. To do this, note that:

$$\overline{A(U_1)} \supset \overline{A(U_2) - A(U_2)} \supset \overline{A(U_2)} - \overline{A(U_2)}.$$

We will be done with this part of the proof if we show that the interior of $\overline{A(U_2)}$ is nonempty. But

$$A(\mathcal{B}_1) = \bigcup_{m=1}^{\infty} mA(U_2),$$

since U_2 is a ball centered at zero and A is a surjection. Therefore, at least one of the $mA(U_2)$ is of the second category in $\mathcal{B}_2$. But, as the mapping $\varphi \to m\varphi$ is a homeomorphism of $\mathcal{B}_2$ onto $\mathcal{B}_2$, $A(U_2)$ is nonmeager in $\mathcal{B}_2$. Therefore, there exists an open set $W \subset \overline{A(U_2)}$.

To prove that $\overline{A(U_1)} \subset A(U)$, let $\varphi_1 \in \overline{A(U_1)}$ be fixed, and observe by the first part that

$$\left(\varphi_1 - \overline{A(U_2)}\right) \bigcap A(U) \neq \emptyset.$$

Thus, there is a $y_1 \in U_1$ with $A(y_1) \in \varphi_1 - \overline{A(U_2)}$. Now, for any $n \geq 1$, $\overline{A(U_n)}$ contains an open neighborhood of zero. Hence, assume that $\varphi_n \in \overline{A(U_n)}$ has been chosen with

$$\left(\varphi_n - \overline{A(U_{n+1})}\right) \bigcap A(U_n) \neq \emptyset.$$

This means there is a $y_n \in U_n$ such that $A(y_n) \in \varphi_n - \overline{A(U_{n+1})}$. Set $y_{n+1} = y_n - A(\varphi_n)$. Then $y_{n+1} \in \overline{A(U_{n+1})}$ and we continue the construction. It is easy to see that the sums $\varphi_1 + \varphi_2 + \varphi_3 + \cdots + \varphi_n$ form a convergent Cauchy sequence which converges to some $\varphi \in \mathcal{B}$, and $\|\varphi\| < r$. It follows that $\varphi \in U$ and, as

$$\sum_{n=1}^{m} A(\varphi_n) = \sum_{n=1}^{m} (y_n - y_{n+1}) = y_1 - y_{m+1},$$

we see that $y_{m+1} \to 0$ since A is continuous. Thus, $y_1 = A(\varphi_1) \in A(U)$. Since φ_1 was arbitrary, we see that $\overline{A(U_1)} \subset A(U)$. □

Theorem 1.56 (Inverse Mapping Theorem)**.** *Let $\mathcal{B}_1, \mathcal{B}_2$ be two Banach spaces and let A be a continuous bijective linear mapping of $\mathcal{B}_1 \to \mathcal{B}_2$. Then $A^{-1} : \mathcal{B}_2 \to \mathcal{B}_1$ is continuous.*

Proof. Since A is continuous, injective, and surjective, it is an open mapping. As A^{-1} exists and, since $A^{-1}\{A[\mathcal{O}]\} = \mathcal{O}$ for all open sets, A^{-1} is continuous. □

Theorem 1.57 (Closed Graph Theorem)**.** *Let $\mathcal{B}$ be a Banach space and let A a closed linear operator on $\mathcal{B}$. If $D(A) = \mathcal{B}$, then $A \in L[\mathcal{B}]$.*

Proof. By definition, $G(A)$ is closed and is a Banach space in the norm $\|(\varphi, A\varphi)\| = \|\varphi\|_{\mathcal{B}} + \|A\varphi\|_{\mathcal{B}}$. Consider the two continuous mappings: $\pi_1 : (\varphi, A\varphi) \to \varphi$, $\pi_2 : (\varphi, A\varphi) \to A\varphi$. Since π_1 is a bijection, π_1^{-1} is continuous so $A = \pi_2 \circ \pi_1^{-1}$ is also continuous. □

Part II: Intermediate Analysis

In this second part of this chapter, we introduce a number of topics that are rarely covered in the first 2 years of the standard graduate programs. These topics will be used at a number of points in the book and are collected here for reference as needed.

S-Basis. In this section, we review a few results that belong to Banach space theory proper. We provide a few proofs, but all the results can be found in Carothers [CA]. Let $\mathcal{B}$ be a separable Banach space.

Definition 1.58. A sequence $(x_n) \in \mathcal{B}$ is called a Schauder basis (S-basis) for $\mathcal{B}$ if $\|x_n\|_{\mathcal{B}} = 1$ for all n and, for each $f \in \mathcal{B}$, there is a unique sequence (a_n) of scalars such that

$$x = \lim_{k\to\infty} \sum_{n=1}^{k} a_n x_n = \sum_{n=1}^{\infty} a_n x_n.$$

All spaces of interest in this book have an S-basis. However, it is known that there are separable Banach spaces without an S-basis (see Carothers [CA] or Diestel [DI]).

Example 1.59. *Let $\mathcal{B} = \ell_p,\ 1 < p < \infty$, where*

$$\ell_p = \left\{ x = (x_1, \cdots) : \sum_{k=1}^{\infty} |x_k|^p < \infty \right\}.$$

The set of vectors $\{e_k\}$, where

$$e_k = \left(0, 0, \ldots, \overset{k}{1}, 0, \ldots \right),$$

form a norm-one S-basis for this space (see [CA]).

If $\Omega = [0,1]$ and $\mathcal{B} = L^p[\Omega],\ 1 < p < \infty$, the family of vectors

$$\{1, \cos(2\pi t), \sin(2\pi t) \cos(4\pi t), \sin(4\pi t), \ldots\}$$

is a norm-one S-basis for $\mathcal{B}$ (see [CA]).

It is easy to see that every Banach space with an S-basis is separable. Let $\mathcal{P}_n x = \sum_{k=1}^{n} a_k x_k$ and define a new norm on $\mathcal{B}$ by

$$|||x|||_{\mathcal{B}} = \sup_n \|\mathcal{P}_n x\|_{\mathcal{B}} = \sup_n \left\| \sum_{k=1}^{n} a_k x_k \right\|_{\mathcal{B}}.$$

Example 1.60. *Let $\Omega = [0,1]$ and $\mathcal{B} = L^p[\Omega]$ over the complex numbers. If $x(t) \in \mathcal{B}$, define*

$$c_k = \int_0^1 e^{-2\pi i k t} x(t) dt, \quad k = 0, \pm 1, \pm 2, \ldots$$

It is easy to see that,

$$|||x|||_{\mathcal{B}} = \sup_n \int_0^1 \left| \sum_{k=-n}^{n} c_k e^{2\pi i t} \right|^2 dt$$

defines a norm on $\mathcal{B}$.

Theorem 1.61. *The norm $|||\cdot|||_{\mathcal{B}}$, is an equivalent norm on $\mathcal{B}$ and*

$$\|x\|_{\mathcal{B}} \leq |||x|||_{\mathcal{B}} = \sup_n \|\mathcal{P}_n x\|_{\mathcal{B}}.$$

Proof. Since $\lim_n \mathcal{P}_n x \to x$, it is clear that $|||x|||_{\mathcal{B}}$ is finite for all $x \in \mathcal{B}$. Since the identity map of $(\mathcal{B}, |||\cdot|||_{\mathcal{B}}) \to (\mathcal{B}, ||\| \cdot \||_{\mathcal{B}})$ is continuous, by the Inverse Mapping Theorem (Theorem 1.56), we are done if we can show that this map has a continuous inverse. It suffices to show that $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$ is complete (i.e., a Banach space).

For this, suppose that let (z_k) be a Cauchy sequence in $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$. Then $(\mathcal{P}_n z_k)$ is a Cauchy sequence in $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$, since $\|\mathcal{P}_n z_i - \mathcal{P}_n z_j\|_{\mathcal{B}} \leqslant |||z_i - z_j|||_{\mathcal{B}}$, for all n (uniformly Cauchy). Thus, if $y_n = \lim_{k\to\infty} \mathcal{P}_n z_k$ then $\lim_{k\to\infty} \|\mathcal{P}_n z_k - y_n\|_{\mathcal{B}} = 0$, uniformly in n.

It now follows using the standard $\frac{\varepsilon}{3}$ argument that (y_n) is a Cauchy sequence in $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$. If $y = \lim_{n\to\infty} y_n$ in $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$, we are done if we show that $y = \lim_{k\to\infty} z_k$ in $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$.

Since there is a unique sequence of scalars (a_i) such that $y = \sum_{i=1}^{\infty} a_i x_i$, we see that $y_n = \sum_{i=1}^{n} a_i x_i$, so that $\mathcal{P}_n y = y_n$ and

$$|||z_k - y|||_{\mathcal{B}} = \sup_n \|\mathcal{P}_n z_k - y_n\|_{\mathcal{B}} \to 0 \text{ as } k \to \infty.$$

□

Since, $x \in \mathcal{B}$ implies that $\|\mathcal{P}_n x\|_{\mathcal{B}} < \infty$ for all n, By the Uniform Boundedness Theorem 1.54, we see that $\sup_n \|\mathcal{P}_n\|_{\mathcal{B}} < \infty$.

Definition 1.62. The set $\{\mathcal{P}_n\}$ is called the natural family of projections associated with the S-basis $\{x_n\}$ and $\sup_n \|\mathcal{P}_n\|_{\mathcal{B}} = K$ is called the basis constant of $\{x_n\}$. In terms of the equivalent norm of the last theorem, $K = 1$.

Definition 1.63. Let $\mathcal{B}$ be a Banach space with an S-basis and let x_n^* be the linear functional on $\mathcal{B}$ defined by $x_n^*(x) = a_n$, where $x = \sum_{k=1}^{\infty} a_k x_k$. Since $x_n^*(x_m) = \delta_{mn}$, we say that the sequence of pairs $\{x_n^*, x_n\}$ are biorthogonal. We call the family $\{x_n^*\}$ the coordinate functionals.

Definition 1.64. We define the span of a set of vectors $\{x_n\}$, in a vector space $\mathcal{B}$, written span($\{x_n\}$), to be the set of all finite linear combinations of subsets of $\{x_n\}$. When $\mathcal{B}$ is a Banach space, we let $[x_n]$ represent the closed subspace of $\mathcal{B}$ generated by span($\{x_n\}$).

A proof of the next result can be found in Carothers (see [CA, pp. 67–71]).

Theorem 1.65. *If $\mathcal{B}$ is a reflexive Banach space and $\{x_n\}$ is an* S-*basis for $\mathcal{B}$, then $\{x_n^*\}$ is an S-basis for $\mathcal{B}'$. Furthermore, the natural embedding $j : \mathcal{B} \to \mathcal{B}''$ defined by $x^{**}(y^*) = y^*(x)$ for all $y^* \in \mathcal{B}'$, is an isometric isomorphism.*

1.3. Distributions and Sobolev Spaces

References for this section are Strichartz [SZ], Yosida [YS], Leoni [GL], Reed and Simon [RS], Rudin [RU1], and Evans [EV]. The purpose of this section is to establish the basic ideas for use in Chaps. 2 and 3. However, neither this section nor the material in Chaps. 2 and 3 is a substitute for a complete introduction to the subject. Those with no background should at least consult Strichartz [SZ].

1.3.1. The Test Functions and Distributions.

Definition 1.66. Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ be a multi-index of non-negative integers, with $|\alpha| = \sum_{k=1}^n \alpha_k$. We define the operators D_n^α and $D_{\alpha,n}$ by

$$D_n^\alpha = \prod_{k=1}^n \frac{\partial^{\alpha_k}}{\partial x^{\alpha_k}} \quad D_{\alpha,n} = \prod_{k=1}^n \left(\frac{1}{2\pi i}\frac{\partial}{\partial x_k}\right)^{\alpha_k}$$

Let $\mathbb{C}_c(\mathbb{R}^n)$ be the class of infinitely differentiable functions on $\mathbb{R}^n$ with compact support and impose the natural locally convex topology τ on $\mathbb{C}_c(\mathbb{R}^n)$ to obtain $\mathfrak{D}(\mathbb{R}^n)$. A definition in terms of neighborhoods can be found in Leoni [GL] (see also Yosida [YS] and Reed and Simon [RS]).

Definition 1.67. A sequence $\{f_m\}$ converges to $f \in \mathfrak{D}(\mathbb{R}^n)$ with respect to the compact sequential limit topology if and only if there exists a compact set $K \subset \mathbb{R}^n$, which contains the support of $f_m - f$ for each m and $D_n^\alpha f_m \to D_n^\alpha f$ uniformly on K, for every multi-index $\alpha \in \mathbb{N}^n$.

Let $u \in \mathbb{C}^1(\mathbb{R}^n)$ and suppose that $\phi \in \mathbb{C}_c^\infty(\mathbb{R}^n)$ has its support in a ball B_r, of radius $r > 0$. Integration by parts gives:

$$\int_{\mathbb{R}^n} (\phi u_{y_i}) d\lambda_n = \int_{\partial B_r} (u\phi)\nu_i d\mathbf{S} - \int_{\mathbb{R}^n} (u\phi_{y_i}) d\lambda_n,\ 1 \le i \le n,$$

where $\boldsymbol{\nu}$ is the unit outward normal to B_r. Since ϕ vanishes on the ∂B_r, the above reduces to:

$$\int_{\mathbb{R}^n} (\phi u_{y_i})\ d\lambda_n = -\int_{\mathbb{R}^n} (u\phi_{y_i}) d\lambda_n,\ 1 \le i \le n.$$

In the general case, for any $u \in \mathbb{C}^m[\mathbb{R}^n]$ and any multi-index $\alpha = (\alpha_1, \ldots, \alpha_n),\ |\alpha| = \sum_{\alpha=1}^n \alpha_i = m$, we have

$$\int_{\mathbb{R}^n} \phi(D^\alpha u) d\lambda_n = (-1)^m \int_{\mathbb{R}^n} u(D^\alpha \phi) d\lambda_n. \tag{1.2}$$

We now observe that the right-hand side of Eq. (1.2) makes sense, even if $D^\alpha u$ does not exist according to our normal definition. This is the basic idea behind the notion of a distributional derivative. Before giving the formal definition, recall that a function $u \in L^1_{\text{loc}}[\mathbb{R}^n]$ if it is Lebesgue integrable on every compact subset of $\mathbb{R}^n$.

Definition 1.68. If α is a multi-index and $u,\ v \in L^1_{\text{loc}}[\mathbb{R}^n]$, we say that v is the α^{th}-weak (or distributional) partial derivative of u and write $D^\alpha u = v$ provided that

$$\int_{\mathbb{R}^n} u(D^\alpha \phi) d\lambda_n = (-1)^{|\alpha|} \int_{\mathbb{R}^n} \phi v\ d\lambda_n$$

for all functions $\phi \in \mathbb{C}_c^\infty[\mathbb{R}^n]$. Thus, v is in the dual space $\mathcal{D}'[\mathbb{R}^n]$ of $\mathcal{D}[\mathbb{R}^n]$.

The next result is easy.

Lemma 1.69. *If a weak αth-partial derivative exists for u, then it is unique λ_n-(a.e.).*

Definition 1.70. If $m \geq 0$ is fixed and $1 \leq p \leq \infty$, we define the Sobolev space $W^{m,p}[\mathbb{R}^n]$ to be the set of all locally integrable functions $u:\ \mathbb{R}^n \to \mathbb{R}$ such that, for each multi-index α with $|\alpha| \leqslant m,\ D^\alpha u$ exists in the weak sense and belongs to $L^p[\mathbb{R}^n]$.

Extensions and Decompositions. We need an extension theorem for functions defined on a domain of $\mathbb{R}^n$ and a result which shows that a domain in $\mathbb{R}^n$ can be written as a union of nonoverlapping closed cubes. (Proofs of these results can be found in Evans [EV] and Stein [STE], respectively.)

Let $\mathbb{D}$ be a bounded open connected set of $\mathbb{R}^n$ (a domain) with boundary $\partial\mathbb{D}$ and closure $\overline{\mathbb{D}}$.

Definition 1.71. Let k be a positive integer. We say that $\partial\mathbb{D}$ is of class $\mathbf{C}^k$ if, for every point $\mathbf{x} \in \partial\mathbb{D}$, there is a homeomorphism ϕ of a neighborhood U of $\mathbf{x}$ into $\mathbb{R}^n$ such that both ϕ and ϕ^{-1} have k continuous derivatives with

$$\varphi(\mathbb{D} \cap U) \subset \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_n > 0\}$$

and

$$\varphi(\partial\mathbb{D} \cap U) \subset \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_n = 0\}.$$

Theorem 1.72. *Let $\mathbb{D}$ be a domain in $\mathbb{R}^n$ with $\partial\mathbb{D}$ of class $\mathbb{C}^1$. Let $\mathbb{U}$ be any bounded open set such that $\overline{\mathbb{D}}$, the closure of $\mathbb{D} \subset\subset \mathbb{U}$ (i.e., the closure of $\mathbb{D}$ is a compact subset of $\mathbb{U}$). Then there is a linear operator $\mathfrak{E}$ mapping functions on $\mathbb{D}$ to functions on $\mathbb{R}^n$ such that:*

(1) *The operator $\mathfrak{E}$ maps $W^{1,p}[\mathbb{D}]$ continuously into $W^{1,p}[\mathbb{R}^n]$ for all $1 \le p \le \infty$.*

(2) $\mathfrak{E}(f)\,|_{\mathbb{D}} = f$ *(e.g., $\mathfrak{E}(\cdot)$ is an extension operator).*

(3) $\mathfrak{E}(f)(x) = 0$ *for $x \in \mathbb{U}^c$ (e.g., $\mathfrak{E}(f)$ has support inside $\mathbb{U}$).*

Theorem 1.73. *Let $\mathbb{D}$ be a domain in $\mathbb{R}^n$. Then $\mathbb{D}$ is the union of a sequence of closed cubes $\{\mathbb{D}_k\}$ whose sides are parallel to the coordinate axes and whose interiors are mutually disjoint.*

Thus, if a function f is defined on a domain in $\mathbb{R}^n$, by Theorem 1.72 it can be extended to the whole space. On the other hand, without loss of generality, by Theorem 1.73, we can assume that the domain is a cube with sides parallel to the coordinate axes.

Definition 1.74. If $\mathbb{D}$ is a domain in $\mathbb{R}^n$, we define $W_0^{m,p}[\mathbb{D}]$ to be the closure of $\mathbb{C}_c^\infty(\mathbb{D})$ in $W^{m,p}[\mathbb{D}]$.

Remark 1.75. Thus, $W_0^{m,p}[\mathbb{D}]$ contains those functions $u \in W^{m,p}[\mathbb{D}]$ such that, for all $|\alpha| \le m-1$, $D^\alpha u = 0$ on the boundary of $\mathbb{D}$, $\partial\mathbb{D}$.

We also note that, when $p = 2$ it is standard to use $H^m(\mathbb{D}) = W^{m,2}(\mathbb{D})$ and $H_0^m(\mathbb{D}) = W_0^{m,2}(\mathbb{D})$.

1.4. Tensor Products

Tensor products of Banach spaces are not a part of the normal graduate program. This section is an introduction to the finite theory that is background for the infinite tensor product theory in Chap. 6. At this point, it is assumed that the reader has at least studied Chap. 4 or is already familiar with the material from some other source.

Since tensor products of Banach spaces have a bad reputation, we should at least comment on this "public relations problem." This reputation is due to questions and studies unrelated to partial differential equations, path integrals, stochastic processes, analysis (proper), and the many possible applications in science and engineering. We approach the subject from a more natural point of view, so that its usefulness for these important and equally interesting areas will be transparent.

1.4.1. Elementary Background. For those with no background in tensor products, we begin with $\mathbb{R}^3$, a space which is well known from calculus (any finite dimension will do). There are a number of ways to patch together two copies of $\mathbb{R}^3$ to obtain a new space. The first is called the direct sum:

$$\mathbb{R}^3 \oplus \mathbb{R}^3 = \{(a_1, a_2, a_3, b_1, b_2, b_3) : (a_1, a_2, a_3), (b_1, b_2, b_3) \in \mathbb{R}^3\}.$$

It is clear that $\mathbb{R}^3 \oplus \mathbb{R}^3$ is isomorphic to $\mathbb{R}^6$. There are also two ways we can define a product on $\mathbb{R}^3$; the first is the dot product

$$\begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \sum_{i=1}^{3} b_i a_i,$$

which takes two vectors and produces a scalar. The other is the tensor product

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{bmatrix},$$

which takes two vectors and produces a 3×3 matrix. It is easy to see that we can write the resulting matrix as a vector in $\mathbb{R}^9 = \mathbb{R}^{3^2}$. Thus, the tensor product of $\mathbb{R}^3$ with itself, written as $\mathbb{R}^3 \otimes \mathbb{R}^3$, is isomorphic to $\mathbb{R}^9$.

Implicit in our use of the dot product is the assumption that the norm induced is the natural one generated by the dot product on $\mathbb{R}^3$:

$$\|\mathbf{a}-\mathbf{b}\|_{\mathbb{R}^3} = \sqrt{(\mathbf{a}-\mathbf{b})\cdot(\mathbf{a}-\mathbf{b})} = \sqrt{\sum_{i=1}^{3}(a_i-b_i)^2}.$$

However, there are a number of other norms possible on $\mathbb{R}^3$ which are not induced by a dot product. For example:

$$\begin{aligned} \|\mathbf{a}-\mathbf{b}\|_p &= \left[\sum_{i=1}^{3}|a_i-b_i|^p\right]^{1/p}, \quad 1\leqslant p<\infty,\ p\neq 2, \\ \|\mathbf{a}-\mathbf{b}\|_\infty &= \max_{1\leqslant i\leqslant 3}|a_i-b_i|. \end{aligned} \tag{1.3}$$

We will discuss this later. However, the case $p=2$ is the standard one because it is the only (unique) one generated by a dot product (even in infinite dimensions). Let $\ell_p\left(\mathbb{R}^3\right)$ represent $\mathbb{R}^3$ with the norm $\|\cdot\|_p$. It is easy to check that $\ell_p\left(\mathbb{R}^3\right)$ is a Banach space for $p\neq 2$ and that $\ell_2\left(\mathbb{R}^3\right)$ is a Hilbert space.

We know that $\mathbb{R}^3\otimes\mathbb{R}^3=\mathbb{R}^9$. The basic question is, how do we define the norm, so that $\ell_p(\mathbb{R}^3)\otimes\ell_p(\mathbb{R}^3)=\ell_p(\mathbb{R}^9)$. It is known that, on $\mathbb{R}^n$, all norms are equivalent. That is, for any pair p,q, there exists constants $c_{p,q},C_{p,q}$, such that, for any vector $\mathbf{a}$,

$$c_{p,q}\|\mathbf{a}\|_q\leqslant\|\mathbf{a}\|_p\leqslant C_{p,q}\|\mathbf{a}\|_q.$$

Thus, we can define the dot product on $\mathbb{R}^9$ and use the norm equivalence to obtain all the others. However, in the infinite-dimensional case $(\ell_p(\mathbb{R}^\infty))$, this is no longer true and each of the norms in Eq. (1.3) generates distinct Banach spaces.

Example 1.76. *Let us see what the norm looks like for $\mathbb{R}^2\otimes\mathbb{R}^2$. A direct computation shows that*

$$\begin{aligned} \mathbf{a}\otimes\mathbf{b} &= \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}\begin{bmatrix} b_1 & b_2 \end{bmatrix} \\ &= \begin{bmatrix} a_1b_1 & a_1b_2 \\ a_2b_1 & a_2b_1 \end{bmatrix} = \left[a_1b_1, a_1b_2, a_2b_1, a_2b_1\right]. \end{aligned}$$

and

$$
\begin{aligned}
&\|\mathbf{a}\|_2\|\mathbf{b}\|_2 \\
&= \left[\sum_{i=1}^{2} a_i^2\right]^{1/2}\left[\sum_{k=1}^{2} b_k^2\right]^{1/2} = \left[\left(\sum_{i=1}^{2} a_i^2\right)\left(\sum_{k=1}^{2} b_k^2\right)\right]^{1/2} \qquad (1.4) \\
&= \left[a_1^2b_1^2 + a_1^2b_2^2 + a_2^2b_1^2 + a_2^2b_2^2\right]^{1/2} = \|\mathbf{a}\otimes\mathbf{b}\|_4.
\end{aligned}
$$

This is a special property called the crossnorm (i.e., $\|\mathbf{a}\otimes\mathbf{b}\|_4 = \|\mathbf{a}\|_2\|\mathbf{b}\|_2$*).*

1.4.2. General Background. We begin with a few concrete examples and ideas that reveal the landscape. Let $f(x) \in \mathbb{C}(\Omega_1)$, $g(y) \in \mathbb{C}(\Omega_2)$, where Ω_1 and Ω_2 are compact sets in $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively. If we let $F(x,y) = f(x)g(y)$, then $F(x,y) \in \mathbb{C}(\Omega_1 \times \Omega_2)$. It is clear that:

(1)

$$\frac{\partial^2}{\partial x \partial y}F(x,y) = \left[\frac{d}{dx}f(x)\right]\left[\frac{d}{dy}g(y)\right]$$

and

(2)

$$\int_{\Omega_1\times\Omega_2} F(x,y)dxdy = \int_{\Omega_1} f(x)dx \int_{\Omega_2} g(y)dy.$$

If we let S be the set of all finite sums, $S = \{F_m(x,y)\},\ m \in \mathbb{N}$, where

$$F_m(x,y) = \sum\nolimits_{i=1}^{m} f_i(x)g_i(y), \quad m \in \mathbb{N},$$

it is easy to see that S is dense in $\mathbb{C}(\Omega_1 \times \Omega_2)$. We can now ask the natural question; What norm should we use on S so that the completion of S, $\bar{S} = \mathbb{C}(\Omega_1 \times \Omega_2)$? It is not hard to show that the appropriate norm is

$$\|F_m\|_\infty = \sup_{x\in\Omega_1}\sup_{y\in\Omega_2}\left|\sum_{i=1}^{m} f_i(x)g_i(y)\right|. \qquad (1.5)$$

On the other hand, we could replace Eq. (1.5) with

$$\|F_m\|_p = \left[\int_{\Omega_2}\int_{\Omega_1}\left|\sum_{i=1}^{m} f_i(x)g_i(y)\right|^p dxdy\right]^{1/p},$$

where $1 \le p \le \infty$ and ask the same question. If we use this norm on S, we clearly do not expect to get $\mathbb{C}(\Omega_1 \times \Omega_2)$. The theory of tensor products of Banach spaces is designed to make the above precise and reveal the nature of the resulting space.

1.4.3. Tensor Products of Hilbert Spaces. Tensor products of Hilbert spaces are the easiest Banach spaces to study, because as noted earlier, there exists only one norm that will make the result another Hilbert space.

Let $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ be three Hilbert spaces over $\mathbb{C}$. A mapping $T : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H}_3$ is said to be bilinear if for all $a, b \in \mathbb{C}$ and all $x \in \mathcal{H}_1,\ y \in \mathcal{H}_2$,

(1)
$$T(ax_1 + bx_2, y) = aT(x_1, y) + bT(x_2, y)$$
and

(2)
$$T(x, ay_1 + by_2) = a^c T(x, y_1) + b^c T(x, y_2).$$

Definition 1.77. The complete tensor product of the Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ is a Hilbert space $\mathcal{H}_3$ and a bilinear mapping $T : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H}_3$, such that

(1) The closed linear span of all the vectors $T(x, y),\ x \in \mathcal{H}_1,\ y \in \mathcal{H}_2$ is equal to $\mathcal{H}_3$.

(2) The inner product for $\mathcal{H}_3$ satisfies:
$$(T(x_1, y_1), T(x_2, y_2))_3 = (x_1, x_2)_1 (y_1, y_2)_2, \tag{1.6}$$

for all pairs $x_1, x_2 \in \mathcal{H}_1$ and $y_1, y_2 \in \mathcal{H}_2$. We call the pair $(\mathcal{H}_3, T)$ the (complete) tensor product of $\mathcal{H}_1$ and $\mathcal{H}_2$. We denote the linear span of the Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ by $\mathcal{H}_1 \otimes \mathcal{H}_2$, replace $\mathcal{H}_3$ with $\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2$ and $T(x, y)$ by $x \otimes y$, which is the standard representation. If we let $x_1 = x_2,\ y_1 = y_2$ in Eq. (1.6), it now reads $\|x \otimes y\|_3^2 = \|x\|_1^2 \|y\|_2^2$ or $\|x \otimes y\|_3 = \|x\|_1 \|y\|_2$. This is the crossnorm relationship we saw in Eq. (1.4).

The tensor product $x \otimes y$ is a bilinear mapping of $\mathcal{H}_1 \times \mathcal{H}_2$ to $\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2$, we can also view it as a functional in the space $B(\mathcal{H}_1, \mathcal{H}_2, \mathbb{C}) = B(\mathcal{H}_1, \mathcal{H}_2)$, of bilinear mappings on $\mathcal{H}_1 \times \mathcal{H}_2$ to $\mathbb{C}$. That is, from Eq. (1.6),
$$(x_1 \otimes y_1, x_2 \otimes y_2)_3 = (x_1, x_2)_1 (y_1, y_2)_2.$$

We will use this interpretation later in the section to define the tensor product of two Banach spaces.

Theorem 1.78. *If the family $\{\phi_n\}$ is a orthonormal basis for $\mathcal{H}_1$ and the family $\{\psi_m\}$ is a orthonormal basis for $\mathcal{H}_2$, the family $\{\phi_n \otimes \psi_m\}$ is a orthonormal basis for $\mathcal{H}_3$.*

Proof. Since $\|\phi_n \otimes \psi_m\|_3 = \|\phi_n\|_1 \|\psi_m\|_2 = 1$, they are normal. Furthermore,

$$(\phi_n \otimes \psi_m, \phi_i \otimes \psi_j)_3 = (\phi_n, \phi_i)_1 (\psi_m, \psi_j)_2 = \delta_{n,i}\delta_{m,j},$$

so they are also orthogonal.

Thus, we are done if we can show that the closure of the linear span of $\{\phi_n \otimes \psi_m\}$ is all of $\mathcal{H}_3$. Let $f \otimes g \in \mathcal{H}_3$, so that $f \in \mathcal{H}_1$ and $g \in \mathcal{H}_2$. Then there exist unique families of constants $\{a_i\}, \{b_j\}$ such that: $f = \sum_{i=1}^{\infty} a_i \phi_i$ and $g = \sum_{j=1}^{\infty} b_j \psi_j$. It is now easy to show that $f \otimes g = \sum_{i,j}^{\infty} a_i b_j \phi_i \otimes \psi_j$. A little reflection will convince the reader that every vector $u \in \mathcal{H}_3$ can be written as $u = \sum_{i,j}^{\infty} a_{ij} \phi_i \otimes \psi_j$, for some (unique) constants $\{a_{ij}\}$, so that the family $\{\phi_i \otimes \psi_j\}$ is a basis for $\mathcal{H}_3$. $\square$

1.5. Tensor Products of Banach Spaces

If $\mathcal{B}_1$ and $\mathcal{B}_2$ are two Banach spaces, the algebraic tensor product of $\mathcal{B}_1$ and $\mathcal{B}_2$ is denoted by $\mathcal{B}_1 \otimes \mathcal{B}_2$, and every element $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$ may be written as $\phi = \sum_{i=1}^{n} \phi_1^i \otimes \phi_2^i$, where $\{\phi_1^i\} \in \mathcal{B}_1, \{\phi_2^i\} \in \mathcal{B}_2$ and n is some nonnegative integer. We denote by $B(\mathcal{B}_1, \mathcal{B}_2) = B(\mathcal{B}_1, \mathcal{B}_2, \mathbb{C})$, the space of all continuous bilinear functionals on $\mathcal{B}_1 \times \mathcal{B}_2$. If l is a bilinear form on $\mathcal{B}_1 \times \mathcal{B}_2$, it generates a natural linear functional $\hat{l}$ on $\mathcal{B}_1 \otimes \mathcal{B}_2$ defined by evaluation:

$$\left\langle \varphi \otimes \psi, \hat{l} \right\rangle = l(\varphi, \psi),\ (\varphi, \psi) \in \mathcal{B}_1 \times \mathcal{B}_2,\ l \in B\left(\mathcal{B}_1, \mathcal{B}_2\right).$$

Also, $\langle \mathcal{B}_1 \otimes \mathcal{B}_2, \mathcal{B}'_1 \otimes \mathcal{B}'_2 \rangle$ defines a (strong) dual system by:

$$\langle \varphi \otimes \psi, \varphi^* \otimes \psi^* \rangle = \langle \varphi, \varphi^* \rangle \langle \psi, \psi^* \rangle,\ (\varphi, \psi) \in \mathcal{B}_1 \times \mathcal{B}_2,\ (\varphi^*, \psi^*) \in \mathcal{B}'_1 \times \mathcal{B}'_2.$$

It follows that we can consider $\mathcal{B}_1 \otimes \mathcal{B}_2$ as the space of continuous bilinear functionals on $\mathcal{B}'_1 \times \mathcal{B}'_2$ ($\mathcal{B}_1 \otimes \mathcal{B}_2 \subset B\left(\mathcal{B}'_1, \mathcal{B}'_2\right)$), and $\mathcal{B}'_1 \otimes \mathcal{B}'_2$ as the space of continuous bilinear functionals on $\mathcal{B}_1 \times \mathcal{B}_2$, so that $\mathcal{B}'_1 \otimes \mathcal{B}'_2 \subset B\left(\mathcal{B}_1, \mathcal{B}_2\right)$.

For notation consistent with the field, when studying one of the L^p-type spaces (with $1 \leq p \leq \infty$), we will use $\Delta_p(\cdot)$ in place of $\|\cdot\|_p$. Although there are many norms that may be defined on $\mathcal{B}_1 \otimes \mathcal{B}_2$ such that the completion is a Banach space, we will always use the one that is natural for the spaces under consideration. This means that we will restrict our attention to spaces of direct interest for analysis, applied mathematics, mathematical physics, and probability theory. (Those with interest in the general theory and other approaches should consult the nice books by Defant and Floret [DOF] and Ryan [RA], along with the references therein.)

Let $A_i,\ i = 1, 2$ be closed linear operators with domains $D_i \subset \mathcal{B}_i$, $A_i : D_i \subset \mathcal{B}_i \to \mathcal{B}_i$, $i = 1, 2$. The mapping $(\phi_1, \phi_2) \to A_1\phi_1 \otimes A_2\phi_2$ is bilinear from $D(A_1) \times D(A_2) \to \mathcal{B}_1 \otimes \mathcal{B}_2$. The corresponding linear mapping of $D(A_1) \otimes D(A_2)$ into $\mathcal{B}_1 \otimes \mathcal{B}_2$ is denoted by $A_1 \otimes A_2$, and is called the tensor product of the operators A_1 and A_2.

Definition 1.79. Let α be a norm (written $\|\cdot\|_\alpha$) on $\mathcal{B}_1 \otimes \mathcal{B}_2$.

(1) We say that α is a crossnorm if for $\phi_1 \in \mathcal{B}_1,\ \phi_2 \in \mathcal{B}_2$, we have that:

$$\alpha(\phi_1 \otimes \phi_2) = \|\phi_1 \otimes \phi_2\|_\alpha = \|\phi_1\|_{\mathcal{B}_1} \|\phi_2\|_{\mathcal{B}_2}. \tag{1.7}$$

(2) The greatest crossnorm γ on $\mathcal{B}_1 \otimes \mathcal{B}_2$ can be defined on the unit ball in $B(\mathcal{B}_1, \mathcal{B}_2)$. For $\phi = \sum_{i=1}^n \phi_1^i \otimes \phi_2^i \in \mathcal{B}_1 \otimes \mathcal{B}_2$,

$$\|\phi\|_\gamma = \sup_{l \in B(\mathcal{B}_1, \mathcal{B}_2)} \left\{ |\langle \phi, l\rangle| : \sum_{i=1}^n \phi_1^i \otimes \phi_2^i = \sum_{k=1}^m \psi_1^i \otimes \psi_2^i \right\}.$$

This norm is equivalent to:

$$\|\phi\|_\gamma = \inf \left\{ \sum_{k=1}^m \|\psi_1^i\|_{\mathcal{B}_1} \|\psi_2^i\|_{\mathcal{B}_2} : \sum_{i=1}^n \phi_1^i \otimes \phi_2^i = \sum_{k=1}^m \psi_1^i \otimes \psi_2^i \right\}.$$

(3) The least crossnorm λ is the norm induced on $\mathcal{B}_1 \otimes \mathcal{B}_2$ by the topology of bi-equicontinuous convergence in $B(\mathcal{B}'_1, \mathcal{B}'_2)$. That is, for $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$ and $(F_1, F_2) \in \mathcal{B}'_1 \times \mathcal{B}'_2$,

$$\|\phi\|_\lambda = \sup \left\{ |\langle \phi, F_1 \otimes F_2\rangle| : \|F_1\|_{\mathcal{B}'_1} \leqslant 1,\ \|F_2\|_{\mathcal{B}'_2} \leqslant 1 \right\}. \tag{1.8}$$

Remark 1.80. For the spaces we are interested in, the least crossnorm $\lambda = \Delta_\infty$, while the greatest crossnorm $\gamma = \Delta_1$.

Definition 1.81. Let α be a given crossnorm on $\mathcal{B}_1 \otimes \mathcal{B}_2$. We say that:

(1) The crossnorm α is a reasonable crossnorm if the dual norm α' induced by the dual of $\mathcal{B}_1 \otimes^\alpha \mathcal{B}_2$ is a crossnorm on $\mathcal{B}'_1 \otimes \mathcal{B}'_2$.

(2) The crossnorm α is uniform relative to $\mathcal{B}_1$ and $\mathcal{B}_2$ if, for $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$ and $A_1,\ A_2 \in L[\mathcal{B}_1], L[\mathcal{B}_2]$:

$$\sup_{\|\phi\|_\alpha \leqslant 1} \|(A_1 \otimes A_2)\phi\|_\alpha \leqslant \|A_1\|_{\mathcal{B}_1} \|A_2\|_{\mathcal{B}_2}. \tag{1.9}$$

(3) If conditions (1) and (2) are satisfied, we say that the crossnorm α is a relative tensor norm.

If α is reasonable, then the norm α' on $\mathcal{B}'_1 \otimes \mathcal{B}'_2$ induced by $(\mathcal{B}_1 \otimes^\alpha \mathcal{B}_2)'$ is also reasonable. We denote by $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$ the completion of $\mathcal{B}_1 \otimes \mathcal{B}_2$ with respect to α, and by $\mathcal{B}'_1 \hat{\otimes}^{\alpha'} \mathcal{B}'_2$ the completion of $\mathcal{B}'_1 \otimes \mathcal{B}'_2$ with respect to α'. In general, $\mathcal{B}'_1 \hat{\otimes}^{\alpha'} \mathcal{B}'_2$ can be identified with a closed subspace of $(\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2)'$ (cf. Schatten [S], in Chap. 6).

Remark 1.82. Our definition of a relative uniform norm depends on the spaces under consideration. This is a restriction of the conventional definition, which is independent of the spaces, and is called a uniform norm. We refer to Defant and Floret [DOF] for a complete discussion of the standard case. They follow Grothendieck and replace the notion of a uniform norm with the condition that α has the metric approximation property. This, coupled with the first condition, leads to the definition of a tensor norm.

We now give some examples of the norms and standard spaces of interest. Let Ω be a compact domain in $\mathbb{R}^n$ and let $\mathbb{C}[\Omega]$ be the set of bounded continuous functions on Ω, let $L^p[\Omega, \mathfrak{B}(\Omega), m] = L^p[\Omega]$ be the space of Lebesgue integrable functions on Ω that have finite L^p norm, where m is a measure on Ω, and let $\mathfrak{B}(\Omega)$ be the Borel σ-algebra generated by the open sets of Ω. For any Banach space $\mathcal{B}$, it is easy to see that $\mathbb{C}[\Omega, \mathcal{B}] = \mathbb{C}[\Omega] \hat{\otimes}^\lambda \mathcal{B}$.

Example 1.83. *The following are elementary and most will be proved later. We present them because they are what one would naturally expect:*

(1) *Let* $\mathcal{B} = \mathbb{C}[\Omega]$ *as above, so that* $\mathbb{C}[\Omega] \hat{\otimes}^\lambda \mathbb{C}[\Omega] = \mathbb{C}[\Omega \times \Omega]$. *If* $A_1 = d/dx$, $A_2 = d/dy$, *then*

$$A_1 \hat{\otimes}^\lambda A_2 = \partial^2/\partial x \partial y = d/dx \hat{\otimes}^\lambda d/dy = (d/dx \hat{\otimes}^\lambda I)(I \hat{\otimes}^\lambda d/dy)$$

(see Ichinose [IC70], in Chap. 6).

(2) *Let* $\mathcal{B}_3 = \mathcal{B}_4 = L^1[\Omega]$, *then* $L^1[\Omega] \hat{\otimes}^\gamma L^1[\Omega] = L^1[\Omega \times \Omega]$ *(see Dunford and Schatten [DSH], in Chap. 6).*

(3) *Let* $\mathcal{B}_5 = \mathcal{B}_6 = L^p[\Omega]$, *then* $L^p[\Omega]\hat{\otimes}^{\Delta_p}L^p[\Omega] = L^p[\Omega \times \Omega]$ *for* $1 \leqslant p < \infty$ *(see Schatten [S], in Chap. 6), where*

$$\Delta_p\left(\sum\nolimits_{i=1}^{n} \phi_i \otimes \varphi_i\right) \equiv^{def} \left\{\iint\limits_{\Omega\times\Omega} \left|\sum\nolimits_{i=1}^{n} \phi_i(x) \otimes \varphi_i(y)\right|^p dxdy\right\}^{1/p}. \quad (1.10)$$

Remark 1.84. Note that $\Delta_1 = \gamma$, $\Delta_\infty = \lambda$, so that Δ_p is always a tensor norm relative to $L^p[\Omega]$ $(1 \le p \le \infty)$. It is easy to show that $L^\infty[\Omega]\,\hat{\otimes}^{\lambda}L^\infty[\Omega] \subset L^\infty[\Omega \times \Omega]$ with the inclusion proper (see Dunford and Schatten [DSH], in Chap. 6). [Similar results show that Δ_p is also a tensor norm relative to the various Sobolev spaces $W^{m,p}[\Omega]$ (see Adams [A], in Chap. 6).] Finally, we can allow that Ω be a locally compact group or complete separable metric space with minor adjustments.

1.5.1. Basic Results. In this section, we prove a number of basic results (including some of those mentioned above) about the tensor product of spaces in the Δ_p norm, $1 \le p \le \infty$.

Theorem 1.85. *If* $\mathcal{B}_1$ *and* $\mathcal{B}_2$ *are Banach spaces, then both* $\gamma = \Delta_1$ *and* $\lambda = \Delta_\infty$ *provide norms on* $\mathcal{B}_1 \otimes \mathcal{B}_2$, *with* $\Delta_\infty(\phi) \le \Delta_1(\phi)$ *for all* $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$.

Proof. We prove that Δ_1 is a norm and $\Delta_\infty(\phi) \le \Delta_1(\phi)$ for all $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$. The proof that Δ_∞ is a norm is left as an exercise.

It is clear that $\Delta_1(a\phi) = |a|\,\Delta_1(\phi)$. To prove the triangle inequality, let $\varepsilon > 0$ be given and choose $\phi = \sum_{i=1}^{n} \phi_1^i \otimes \phi_2^i$, $\psi = \sum_{i=1}^{m} \psi_1^i \otimes \psi_2^i$, so that

$$\sum_{i=1}^{n} \left\|\phi_1^i\right\|_{\mathcal{B}_1} \left\|\phi_2^i\right\|_{\mathcal{B}_2} \leqslant \Delta_1(\phi) + \frac{\varepsilon}{2}$$

and

$$\sum_{i=1}^{m} \left\|\psi_1^i\right\|_{\mathcal{B}_1} \left\|\psi_2^i\right\|_{\mathcal{B}_2} \leqslant \Delta_1(\psi) + \frac{\varepsilon}{2}.$$

By definition, this implies that

$$\begin{aligned}\Delta_1(\phi + \psi) &\leqslant \sum_{i=1}^{n} \left\|\phi_1^i\right\|_{\mathcal{B}_1} \left\|\phi_2^i\right\|_{\mathcal{B}_2} + \sum_{i=1}^{m} \left\|\psi_1^i\right\|_{\mathcal{B}_1} \left\|\psi_2^i\right\|_{\mathcal{B}_2} \\ &\leqslant \Delta_1(\phi) + \Delta_1(\psi) + \varepsilon.\end{aligned}$$

Since this is true for all $\varepsilon > 0$, $\Delta_1(\phi + \psi) \le \Delta_1(\phi) + \Delta_1(\psi)$.

Suppose that $\Delta_1(\phi) = 0$. Then, for each $\varepsilon > 0$, there exists a representation $\phi = \sum_{i=1}^n \phi_1^i \otimes \phi_2^i$ such that $\sum_{i=1}^n \|\phi_1^i\|_{\mathcal{B}_1} \|\phi_2^i\|_{\mathcal{B}_2} \leq \varepsilon$. It follows that, for all $(F_1 \otimes F_2) \in \mathcal{B}'_1 \otimes \mathcal{B}'_2$,

$$\left|(F_1 \otimes F_2)\left(\sum_{i=1}^n \phi_1^i \otimes \phi_2^i\right)\right| \leq \left|\sum_{i=1}^n F_1(\phi_1^i)F_2(\phi_2^i)\right| \leqslant \varepsilon \|F_1\|_{\mathcal{B}'_1} \|F_2\|_{\mathcal{B}'_2}.$$

Since $\mathcal{B}'_1 \otimes \mathcal{B}'_2$ is fundamental for $\mathcal{B}_1 \otimes \mathcal{B}_2$, we must have $\phi = 0$.

In order to show that $\Delta_1(\phi_1 \otimes \phi_2) = \|\phi_1\|_{\mathcal{B}_1} \|\phi_2\|_{\mathcal{B}_2}$, first note that $\Delta_1(\phi_1 \otimes \phi_2) \leq \|\phi_1\|_{\mathcal{B}_1} \|\phi_2\|_{\mathcal{B}_2}$, so we need to only prove the opposite relation. Let $F_1 \otimes F_2 \in \mathcal{B}'_1 \otimes \mathcal{B}'_2$ satisfy $F_1(\phi_1) = \|\phi_1\|_{\mathcal{B}_1}$ and $F_2(\phi_2) = \|\phi_2\|_{\mathcal{B}_1}$ (duality maps). Since

$$\begin{aligned} &\left|(F_1 \otimes F_2)\left(\sum_{i=1}^n \phi_1^i \otimes \phi_2^i\right)\right| \\ &\leqslant \sum_{i=1}^n \left|(F_1 \otimes F_2)\left(\phi_1^i \otimes \phi_2^i\right)\right| \\ &= \sum_{i=1}^n \left|F_1(\phi_1^i)F_2(\phi_2^i)\right| \leqslant \sum_{i=1}^n \|\phi_1^i\|_{\mathcal{B}_1} \|\phi_2^i\|_{\mathcal{B}_2}, \end{aligned} \tag{1.11}$$

we see that $|(F_1 \otimes F_2)\phi| \leqslant \Delta_1(\phi)$, where $\phi = \sum_{i=1}^n \phi_1^i \otimes \phi_2^i$. Thus, $(F_1 \otimes F_2)(\phi_1 \otimes \phi_2) = \|\phi_1\|_{\mathcal{B}_1} \|\phi_2\|_{\mathcal{B}_2} \leq \Delta_1(\phi_1 \otimes \phi_2)$.

From Eq. (1.11) we see that $|(F_1 \otimes F_2)\phi| \leqslant \Delta_1(\phi)$ for all $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$. It follows from the definition of Δ_∞ that $\Delta_\infty(\phi) \leq \Delta_1(\phi)$. □

Let Ω be a domain in $\mathbb{R}^n$, let μ be a measure on Ω, and let $\mathcal{B}$ be a separable Banach space with a Schauder basis.

Theorem 1.86. *The completion of $L^1[\Omega, \mu] \otimes \mathcal{B}$ with the Δ_1 norm, $L^1[\Omega, \mu]\hat{\otimes}^{\Delta_1} B$, is isometrically isomorphic to $L^1[\Omega, \mu; B]$, the space of Bochner integrable functions on Ω with values in $\mathcal{B}$.*

Proof. Let $J : L^1[\Omega, \mu] \times \mathcal{B} \to L^1[\Omega, \mu; \mathcal{B}]$, via $(\phi_1, \phi_b) \to \phi = \phi_1 \otimes \phi_b$. By linearization, this induces a norm one mapping

$$J : L^1[\Omega, \mu]\hat{\otimes}^{\Delta_1} \mathcal{B} \to L^1[\Omega, \mu; \mathcal{B}].$$

It follows that $\|J\phi\|_1 \leq \Delta_1(\phi)$ for all $\phi \in L^1[\Omega,\mu]\hat{\otimes}^{\Delta_1}\mathcal{B}$. However, if $\phi \in L^1[\Omega,\mu;\mathcal{B}]$ is a simple function, then $\phi = \sum_{k=1}^n \chi_{A_k}\phi_b^k$ and $J\phi = \sum_{k=1}^n \chi_{A_k} \otimes \phi_b^k$. Thus,

$$\Delta_1(\phi) \leq \sum_{k=1}^n \mu(A_k)\left\|\phi_b^k\right\| = \|J\phi\|_1 .$$

It follows that $\Delta_1(\phi) = \|J\phi\|_1$ for all simple functions. Since the class S of simple functions is dense in $L^1[\Omega,\mu]$, we see that $S \otimes \mathcal{B}$ is dense in $L^1[\Omega,\mu]\hat{\otimes}^{\Delta_1}\mathcal{B}$. Since the norm closure of the class of Bochner integrable functions is all of $L^1[\Omega,\mu;\mathcal{B}]$, we see that J is surjective. Furthermore, since $\Delta_1(\phi) = \|J\phi\|_1$ on $S \otimes \mathcal{B}$, the extension to $L^1[\Omega,\mu]\hat{\otimes}^{\Delta_1}B$ is both injective and isometric. Thus J is an isometry. □

Corollary 1.87. *If* Ω_1, Ω_2 *are domains in* $\mathbb{R}^n$ *with measures* μ_1, μ_2, *then*

$$L^1[\Omega_1,\mu_1]\hat{\otimes}^{\Delta_1}L^1[\Omega_2,\mu_2] = L^1[\Omega_1 \times \Omega_2, \mu_1 \otimes \mu_2].$$

Definition 1.88. Let Ω_1, Ω_2 be compact domains in $\mathbb{R}^n$, then

$$\mathbb{C}(\Omega_1)\otimes\mathbb{C}(\Omega_2) =: \{\phi(x,y) \in \mathbb{C}(\Omega_1\times\Omega_2) \mid \exists n \in \mathbb{N},$$
$$\left\{\phi_1^k(x)\right\}_{k=1}^n \subset \mathbb{C}(\Omega_1),\ \left\{\phi_2^k(y)\right\}_{k=1}^n \subset \mathbb{C}(\Omega_2) \text{ and } \phi(x,y) = \sum\nolimits_{k=1}^n \phi_1^k(x)\phi_2^k(y)\right\}.$$

(If $\Omega_1 = \mathbb{R}^n$, $\Omega_2 = \mathbb{R}^m$ for some n, m, use the one point compactification and the result still applies.)

This is why the notation $(\phi_1 \otimes \phi_2)(x,y) = \phi_1(x)\phi_2(y)$ is used to denote products of functions of two variables (in this case). By the Weierstrass Approximation Theorem, we see that $\mathbb{C}(\Omega_1)\otimes\mathbb{C}(\Omega_2)$ is dense in $\mathbb{C}(\Omega_1\times\Omega_2)$.

Theorem 1.89. $\mathbb{C}(\Omega_1)\hat{\otimes}^{\Delta_\infty}\mathbb{C}(\Omega_2) = \mathbb{C}(\Omega_1\times\Omega_2)$.

Theorem 1.90. *Let* $\mathcal{B}_1$, $\mathcal{B}_2$ *be separable Banach spaces with a Schauder basis.*

(1) *The norm* α *is a reasonable crossnorm on* $\mathcal{B}_1\otimes\mathcal{B}_2$ *if and only if*

$$\Delta_\infty(\phi) \leq \alpha(\phi) \leq \Delta_1(\phi) \text{ for all } \phi \in \mathcal{B}_1\otimes\mathcal{B}_2.$$

(2) *If* α *is a reasonable crossnorm, then the norm* α' *on* $\mathcal{B}'_1\otimes\mathcal{B}'_2$ *induced by* $(\mathcal{B}_1\otimes^\alpha\mathcal{B}_2)'$ *is also a reasonable crossnorm.*

Proof. To begin, we let $\mathbb{B}_{\mathcal{B}'_i}$ denote the unit ball in $\mathcal{B}'_i$, $i = 1, 2$.

If α is a reasonable crossnorm on $\mathcal{B}_1 \otimes \mathcal{B}_2$, then for any representation of $\phi = \sum_{i=1}^{n} \phi_1^i \otimes \phi_2^i \in \mathcal{B}_1 \otimes \mathcal{B}_2$ we have

$$\alpha(\phi) \leqslant \sum_{i=1}^{n} \alpha(\phi_1^i \otimes \phi_2^i) = \sum_{i=1}^{n} \left\|\phi_1^i\right\|_{\mathcal{B}_1} \left\|\phi_2^i\right\|_{\mathcal{B}_2},$$

so that $\alpha(\phi) \le \Delta_1(\phi)$. To see that $\Delta_\infty(\phi) \le \alpha(\phi)$, let $F_1 \otimes F_2$ and $F = \sum_{i=1}^{n} F_1^i \otimes F_2^i$ be in $\mathcal{B}'_1 \otimes \mathcal{B}'_2$. Then

$$\begin{aligned} &\Delta_\infty(\phi) \\ &= \sup\left\{|(F_1 \otimes F_2)\phi| : F_1 \otimes F_2 \in \mathbb{B}_{\mathcal{B}'_1} \otimes \mathbb{B}_{\mathcal{B}'_2}\right\} \\ &\leqslant \sup\left\{|(F)\phi| : F = \sum_{i=1}^{n} F_1^i \otimes F_2^i \in \mathbb{B}_{\mathcal{B}'_1} \otimes \mathbb{B}_{\mathcal{B}'_2}\right\} = \alpha(\phi). \end{aligned}$$

On the other hand, if α is a norm on $\mathcal{B}_1 \otimes \mathcal{B}_2$ with

$$\Delta_\infty(\phi) \le \alpha(\phi) \le \Delta_1(\phi), \text{ for all } \phi \in \mathcal{B}_1 \otimes \mathcal{B}_2,$$

then $\Delta_\infty(\phi \otimes \psi) = \alpha(\phi \otimes \psi) = \Delta_1(\phi \otimes \psi)$, so that α is a crossnorm. To see that α' is a crossnorm on $\mathcal{B}'_1 \otimes \mathcal{B}'_2$, use $\Delta_\infty \le \alpha \le \Delta_1$ to get that, for $\phi = \sum_{i=1}^{n} \phi_1^i \otimes \phi_2^i \in \mathcal{B}_1 \otimes \mathcal{B}_2$,

$$\begin{aligned} &\|F_1\|_{\mathcal{B}'_1} \|F_2\|_{\mathcal{B}'_2} \\ &= \sup\left\{|(F_1 \otimes F_2)\phi| : \phi \in \mathcal{B}_1 \otimes \mathcal{B}_2,\ \Delta_1(\phi) \leqslant 1\right\} \\ &\leqslant \alpha'(F_1 \otimes F_2) = \sup\left\{|(F_1 \otimes F_2)\phi| : \phi \in \mathcal{B}_1 \otimes \mathcal{B}_2,\ \alpha(\phi) \leqslant 1\right\} \\ &\leqslant \sup\left\{|(F_1 \otimes F_2)\phi||\ \phi \in \mathcal{B}_1 \otimes \mathcal{B}_2,\ \Delta_\infty(\phi) \leqslant 1\right\} = \|F_1\|_{\mathcal{B}'_1} \|F_2\|_{\mathcal{B}'_2}. \end{aligned}$$

It follows that α' is a reasonable crossnorm. □

We denote by $\mathcal{B}_1 \hat{\otimes}^{\alpha} \mathcal{B}_2$ the completion of $\mathcal{B}_1 \otimes \mathcal{B}_2$ with respect to α, and by $\mathcal{B}'_1 \hat{\otimes}^{\alpha'} \mathcal{B}'_2$ the completion of $\mathcal{B}'_1 \otimes \mathcal{B}'_2$ with respect to α'. In general, $\mathcal{B}'_1 \hat{\otimes}^{\alpha'} \mathcal{B}'_2$ can be identified with a closed subspace of $(\mathcal{B}_1 \hat{\otimes}^{\alpha} \mathcal{B}_2)'$ (cf. Schatten [S], in Chap. 6).

Theorem 1.91. *If Ω is a domain in $\mathbb{R}^n$ and μ is a measure on $\mathfrak{B}$, then Δ_p is a reasonable crossnorm on $L^p[\Omega, \mathfrak{B}(\Omega), \mu] \otimes \mathcal{B} =: L^p[\Omega] \otimes \mathcal{B}$, $1 \le p \le \infty$, for any separable Banach space $\mathcal{B}$, and $L^p[\Omega] \hat{\otimes}^{\Delta_p} \mathcal{B} = L^p[\Omega; \mathcal{B}]$ for $1 \le p < \infty$.*

Proof. The proof for $p = 1$ was given in Theorem 6.8, so we need to only consider the case for $1 < p \le \infty$. Let $J : L^p[\Omega] \otimes \mathcal{B} \to L^p[\Omega; \mathcal{B}]$

be defined by $J[f \otimes \phi] = f(\cdot)\phi$. This is clearly an injective mapping. Let $g = \sum_{k=1}^{n} f_k \otimes \phi_i$ and define

$$\Delta_p[g] = \left[\int_\Omega \|g(\omega)\|_{\mathcal{B}}^p \, d\mu(\omega) \right]^{1/p}.$$

It is clear that

$$\Delta_p[f \otimes \phi] = \left[\int_\Omega \|f(\omega)\phi\|_{\mathcal{B}}^p \, d\mu(\omega) \right]^{1/p} = \|\phi\|_{\mathcal{B}} \|f\|_p ,$$

so that $\Delta_p[\cdot]$ is a crossnorm. To see that $\Delta_p[\cdot] \le \Delta_1[\cdot]$, note that

$$\begin{aligned} &\Delta_p[g] \\ &= \left[\int_\Omega \left\| \sum_{k=1}^{n} f_k(\omega)\phi_k \right\|_{\mathcal{B}}^p d\mu(\omega) \right]^{1/p} \\ &\leqslant \left[\sum_{k=1}^{n} \|\phi_k\|_{\mathcal{B}}^p \int_\Omega |f_k(\omega)|^p \, d\mu(\omega) \right]^{1/p} \\ &\leqslant \sum_{k=1}^{n} \|\phi_k\|_{\mathcal{B}} \|f_k\|_p, \end{aligned}$$

so that $\Delta_p[g] \leqslant \Delta_1[g]$. To see that $\Delta_\infty[\cdot] \le \Delta_p[\cdot]$, let $F \otimes \Phi \in [L^p]' \otimes \mathcal{B}'$ be in the respective unit balls (i.e., $\|F\|_{p'} \leqslant 1,\ \|\Phi\|_{B'} \leqslant 1,\ \frac{1}{p'} = 1 - \frac{1}{p}$). Then

$$\begin{aligned} &|\langle F \otimes \Phi, g \rangle| \\ &= \left| \int_\Omega F(\omega) \langle \Phi, g(\omega) \rangle \, d\mu(\omega) \right| \\ &\le \sup_{\|\Phi\| \leqslant 1,\ \|F\|_{p'} \leqslant 1} |\langle F \otimes \Phi, g \rangle| = \Delta_\infty[g] \\ &\leqslant \|F\|_{p'} \left[\int_\Omega |\langle \Phi, g(\omega) \rangle|^p \, d\mu(\omega) \right]^{1/p} \\ &\leqslant \sup_{\|\Phi\| \leqslant 1} \left[\int_\Omega |\langle \Phi, g(\omega) \rangle|^p \, d\mu(\omega) \right]^{1/p} = \Delta_p[g]. \end{aligned}$$

Thus, $\Delta_p[\cdot]$ is a reasonable crossnorm on $L^p[\Omega] \otimes \mathcal{B}$ for any p, $1 \le p \le \infty$. If $p < \infty$, then the (equivalence class of) step functions

$$S(\mu) \otimes B = \left\{ \sum_{k=1}^{n} \chi_{A_k} \otimes \phi_k : \ n \in \mathbb{N},\ \mu(A_k) < \infty, \phi_k \in B \right\}$$

is dense in $L^p[\Omega;\mathcal{B}]$. This implies that $L^p[\Omega]\otimes\mathcal{B}$ is dense in $L^p[\Omega;\mathcal{B}]$. It follows that

$$L^p[\Omega]\hat{\otimes}^{\Delta_p} B = L^p[\Omega; B].$$

□

Corollary 1.92. *Let $(\Omega_1, \mathfrak{B}_1, \mu_1)$ and $(\Omega_2, \mathfrak{B}_2, \mu_2)$ be σ-finite measure spaces. Then $L^p[\Omega_1 \times \Omega_2,\ \mathfrak{B}_1 \times \mathfrak{B}_2,\ \mu_1 \times \mu_2] = L^p[\Omega_1]\hat{\otimes}^{\Delta_p} L^p[\Omega_2]$ for $1 \le p < \infty$, where $\mathfrak{B}_1 \times \mathfrak{B}_2$ is the σ-algebra generated by $\mathfrak{B}_1$ and $\mathfrak{B}_2$.*

Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be separable Banach spaces, and let $A_1 \in L[\mathcal{B}_1]$, $A_2 \in L[\mathcal{B}_2]$.

Theorem 1.93. *The operator $A_1 \otimes A_2 : \mathcal{B}_1 \otimes \mathcal{B}_2 \to \mathcal{B}_1 \otimes \mathcal{B}_2$ has an unique extension to both $\mathcal{B}_1\hat{\otimes}^{\Delta_1}\mathcal{B}_2$ and $\mathcal{B}_1\hat{\otimes}^{\Delta_\infty}\mathcal{B}_2$ as a bounded linear operator and*

$$\left\|A_1\hat{\otimes}A_2\right\| = \|A_1\|\,\|A_2\|\,.$$

Proof. We first prove it for Δ_1. Let $\sum_{i=1}^n \phi_1^i \otimes \phi_2^i$ be a representation for $\phi \in \mathcal{B}_1\hat{\otimes}^{\Delta_1}\mathcal{B}_2$. Then

$$\Delta_1[(A_1 \otimes A_2)\,\phi] = \Delta_1\left[\sum_{i=1}^n A_1\phi_1^i \otimes A_2\phi_2^i\right] \leqslant \|A_1\|\,\|A_2\| \sum_{i=1}^n \left\|\phi_1^i\right\|_{\mathcal{B}_1} \left\|\phi_2^i\right\|_{\mathcal{B}_2},$$

so that $\Delta_1[(A_1 \otimes A_2)\,\phi] \le \|A_1\|\,\|A_2\|\,\Delta_1[\phi]$. It follows that $\|A_1 \otimes A_2\| \le \|A_1\|\,\|A_2\|$. However, from $(A_1 \otimes A_2)(\phi_1 \otimes \phi_2) = (A_1\phi_1) \otimes (A_2\phi_2)$, we see that (using the crossnorm property of Δ_1)

$$\begin{aligned}
&\|A_1\|\,\|A_2\| \\
&= \sup \frac{\|(A_1\phi_1)\|_{\mathcal{B}_1}\,\|(A_2\phi_2)\|_{\mathcal{B}_2}}{\|\phi_1\|_{\mathcal{B}_1}\,\|\phi_2\|_{\mathcal{B}_2}} \\
&= \sup \frac{\|(A_1\phi_1 \otimes A_2\phi_2)\|_{\Delta_1}}{\|\phi_1\|_{\mathcal{B}_1}\,\|\phi_2\|_{\mathcal{B}_2}} \\
&= \sup \frac{\|(A_1 \otimes A_2)\,(\phi_1 \otimes \phi_2)\|_{\Delta_1}}{\|(\phi_1 \otimes \phi_2)\|_{\Delta_1}} \leqslant \|A_1 \otimes A_2\|\,.
\end{aligned}$$

It follows that $\|A_1 \otimes A_2\| = \|A_1\|\,\|A_2\|$. It is clear that this equality holds for the unique extension $A_1\hat{\otimes}A_2$ of $A_1 \otimes A_2$ to all of $\mathcal{B}_1\hat{\otimes}^{\Delta_1}\mathcal{B}_2$.

To prove the result for Δ_∞, let $\phi = \sum_{i=1}^{n} \phi_1^i \otimes \phi_2^i$. Then

$$\begin{aligned}
&\Delta_\infty[(A_1 \otimes A_2)\,\phi] \\
&= \sup\left\{ \sum_{i=1}^{n} F_1(A_1\phi_1^i)F_2(A_2\phi_2^i) :\ F_1 \in \mathcal{B}'_1,\ F_2 \in \mathcal{B}'_2;\ \|F_1\| \leqslant 1,\ \|F_2\| \leqslant 1 \right\} \\
&= \sup\left\{ \sum_{i=1}^{n} (A_1'F_1)(\phi_1^i)(A_2'F_2)(\phi_2^i) :\ F_1 \in \mathcal{B}'_1,\ F_2 \in \mathcal{B}'_2;\ \|F_1\| \leqslant 1,\ \|F_2\| \leqslant 1 \right\} \\
&\leqslant \left\|A_1'\right\| \left\|A_2'\right\| \Delta_\infty[\phi] = \|A_1\|\,\|A_2\|\,\Delta_\infty[\phi].
\end{aligned}$$

Thus, $A_1 \otimes A_2$ has a bounded extension to $\mathcal{B}_1 \hat{\otimes}^{\Delta_\infty} \mathcal{B}_2$. Now let $\varepsilon > 0$ and choose $\phi_1 \in \mathcal{B}_1,\ \phi_2 \in \mathcal{B}_2$ with $\|\phi_1\|_{\mathcal{B}_1} \leq 1,\ \|\phi_2\|_{\mathcal{B}_2} \leq 1$ and, such that

$$\|A_1\phi_1\|_{\mathcal{B}_1} \geqslant (1-\varepsilon)\,\|A_1\|_{\mathcal{B}_1}\,; \quad \|A_2\phi_2\|_{\mathcal{B}_2} \geqslant (1-\varepsilon)\,\|A_2\|_{\mathcal{B}_2}\,.$$

Thus, $\Delta_\infty(\phi_1 \otimes \phi_2) \leq 1$ and

$$\|A_1\phi_1\|_{\mathcal{B}_1}\,\|A_2\phi_2\|_{\mathcal{B}_2} = \Delta_\infty[(A_1 \otimes A_2)\,(\phi_1 \otimes \phi_2)] \geqslant (1-\varepsilon)^2\,\|A_1\|\,\|A_2\|\,.$$

Since ε is arbitrary, $\|A_1 \otimes A_2\| = \|A_1\|\,\|A_2\|$. It follows that the same is true for the unique extension $A_1 \hat{\otimes} A_2$ of $A_1 \otimes A_2$ to all of $\mathcal{B}_1 \hat{\otimes}^{\Delta_\infty} \mathcal{B}_2$. □

From Theorem 1.93, we see that Δ_1 and Δ_∞ are uniform for all Banach space couples (tensor norms). The following example shows that, for $1 < p < \infty$, we cannot expect Δ_p to be uniform for all Banach space couples.

Let $L^2[\mathbb{R}]$ and $\ell_1(\mathbb{R})$ have the standard definitions, and let $\mathfrak{F}$ be the Fourier transform on $L^2[\mathbb{R}]$, which is an isometry, and let $\mathbf{I}_1$ be the identity on $\ell_1(\mathbb{R})$. If $\mathcal{B}_1 = L^2(\mathbb{R})$, $\mathcal{B}_2 = \ell_1(\mathbb{R})$ and $\alpha = \Delta_2$, we have

$$\Delta_2(\sum_{m=1}^{n} \varphi_m \otimes \psi_m) \equiv^{def} \left\{ \int_{-\infty}^{\infty} \left\| \sum_{m=1}^{n} \varphi_m(x)\psi_m(y) \right\|_{\mathcal{B}_2}^2 dx \right\}^{1/2}.$$

Example 1.94. *Set $f_n = \sum_{m=1}^{n} \chi_{[m,m+1)} \otimes e_m$, where $\chi_{[m,m+1)}(x)$ is the characteristic function of the interval $[m, m+1)$, and e_m is the mth unit basis vector of $\ell_1(\mathbb{R})$. Then*

$$\Delta_2(f_n) = \left\{ \int_{-\infty}^{\infty} \left\| \sum_{m=1}^{n} \chi_{[m,m+1)}(x) e_m \right\|_{l_1}^2 dx \right\}^{1/2} = \sqrt{n}.$$

However, if we look at the norm of $(\mathfrak{F}\otimes \mathbf{I}_1)f_n$*, we get:*

$$\begin{aligned}
&\|(\mathfrak{F}\otimes \mathbf{I}_1)f_n\|_{\Delta_2} \\
&= \left\{\int_{-\infty}^{\infty}\left\|\tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-ixy}\left[\sum_{m=1}^{n}\chi_{[m,m+1]}(y)e_m\right]dy\right\|_{l_1}^{2}dx\right\}^{1/2} \\
&\quad = \tfrac{1}{\sqrt{2\pi}}\left\{\int_{-\infty}^{\infty}\left\|\tfrac{1}{x}\left[\sum_{m=1}^{n}\{e^{[-i(m+1)x]}-e^{[-imx]}\}e_m\right]\right\|_{l_1}^{2}dx\right\}^{1/2} \\
&\quad = \tfrac{1}{\sqrt{2\pi}}\left\{\int_{-\infty}^{\infty}\tfrac{1}{x^2}\left[\sum_{m=1}^{n}\left|e^{[-i(m+1)x]}-e^{[-imx]}\right|\right]^{2}dx\right\}^{1/2} \\
&\quad = \frac{n}{\sqrt{2\pi}}\left\{\int_{-\infty}^{\infty}\left|\frac{\exp(-ix)-1}{x}\right|^{2}dx\right\}^{1/2} = n.
\end{aligned}$$

It follows that $\mathfrak{F}\otimes \mathbf{I}_1$ *cannot extend to a bounded operator on* $L^2[\mathbb{R}]\hat{\otimes}^{\Delta_2}\ell_1(\mathbb{R})$*. Thus,* Δ_2 *is not uniform with respect to* $L^2[\mathbb{R}]$ *and* $\ell_1(\mathbb{R})$*, so that* Δ_2 *is not a tensor norm. However, it is a relative tensor norm for the right space. To see this in the above case, replace* $\ell_1(\mathbb{R})$ *by* $\ell_2(\mathbb{R})$ *and note that, if* e_m *is the* m*th unit basis vector of* $\ell_2(\mathbb{R})$*, then* $f_n\in L^2[\mathbb{R}]\hat{\otimes}^{\Delta_2}\ell_2(\mathbb{R})$ *and we have:*

$$\Delta_2(f_n)=\left\{\int_{-\infty}^{\infty}\left\|\sum_{m=1}^{n}\chi_{[m,m+1)}(x)e_m\right\|_{l_2}^{2}dx\right\}^{1/2}=\sqrt{n}.$$

If we now look at the norm of $(\mathfrak{F}\otimes \mathbf{I}_2)f_n$*, we get (the expected result):*

$$\begin{aligned}
&\|(\mathcal{F}\otimes \mathbf{I}_2)f_n\|_{\Delta_2} \\
&= \left\{\int_{-\infty}^{\infty}\left\|\tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-ixy}\left[\sum_{m=1}^{n}\chi_{[m,m+1]}(y)e_m\right]dy\right\|_{l_2}^{2}dx\right\}^{1/2} \\
&\qquad = \tfrac{1}{\sqrt{2\pi}}\left\{\int_{-\infty}^{\infty}\left\|\tfrac{1}{x}\left[\sum_{m=1}^{n}\{e^{[-i(m+1)x]}-e^{[-imx]}\}e_m\right]\right\|_{l_2}^{2}dx\right\}^{1/2}
\end{aligned}$$

$$= \tfrac{1}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} \tfrac{1}{x^2} \sum_{m=1}^{n} \left| e^{[-i(m+1)x]} - e^{[-imx]} \right|^2 dx \right\}^{1/2}$$

$$= \frac{\sqrt{n}}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} \left| \frac{\exp(-ix) - 1}{x} \right|^2 dx \right\}^{1/2} = \sqrt{n}.$$

It follows that $(\mathfrak{F} \otimes \mathbf{I}_2)$ *can be extended to a bounded operator on* $L^2[\mathbb{R}]\hat{\otimes}^{\Delta_2} \ell_2(\mathbb{R})$.

By a theorem of Kwapień [KP], in Chap. 6, $\mathcal{B}$ is isomorphic to a Hilbert space if and only if $(\mathfrak{F} \otimes \mathbf{I}_{\mathcal{B}})$ is continuous on $\mathbf{L}^2[\mathbb{R}]\hat{\otimes}^{\Delta_2}\mathcal{B}$. The point is that Δ_2 is a relative tensor norm which is not a tensor norm. On the other hand, if α is any tensor norm, $\mathfrak{F} \otimes \mathbf{I}_1$ has an extension to a bounded linear operator on $\mathbf{L}^2[\mathbb{R}]\hat{\otimes}^{\alpha} l_1(\mathbb{Z})$ (see [DOF, p. 147]).

We now show that Δ_p is uniform relative to the tensor product of L^p spaces.

Theorem 1.95. *Let* $(\Omega_1, \mathfrak{B}_1, \mu_1)$ *and* $(\Omega_2, \mathfrak{B}_2, \mu_2)$ *be* σ*-finite measure spaces. Let* $A_1 : L^p[\Omega_1] \to L^p[\Omega_1]$ *and* $A_2 : L^p[\Omega_2] \to L^p[\Omega_2]$*. Then, for* $1 < p < \infty$*, the operator*

$$A_1 \otimes A_2 : L^p[\Omega_1] \otimes L^p[\Omega_2] \to L^p[\Omega_1] \otimes L^p[\Omega_2]$$

has an unique extension to a bounded linear operator

$$A_1 \hat{\otimes} A_2 : L^p[\Omega_1]\hat{\otimes}^{\Delta_p} L^p[\Omega_2] \to L^p[\Omega_1 \times \Omega_2],$$

and $\left\| A_1 \hat{\otimes} A_2 \right\| = \|A_1\| \, \|A_2\|$.

Proof. We first show that $I_1 \otimes A_2$ is bounded as an operator mapping $L^p[\Omega_1] \otimes L^p[\Omega_2] \to L^p[\Omega_1 \times \Omega_2]$.

Let $\{\phi_2^i\}$ be a Schauder basis for $L^p[\Omega_2]$ and, for $1 \le i \le n$, $n \in \mathbb{N}$, let $\psi_2^i = A_2 \phi_2^i$. Then, for all scalars $a_1, \dots, a_n$, we have

$$\left\| \sum_{i=1}^{n} a_i \psi_2^i \right\|_p \leqslant \|A_2\| \left\| \sum_{i=1}^{n} a_i \phi_2^i \right\|_p .$$

It follows that, for arbitrary functions $a_1(\cdot), \dots, a_n(\cdot) \in L^p[\Omega_1]$,

$$\int_{\Omega_2} \left| \sum_{i=1}^{n} a_i(x) \psi_2^i(y) \right|^p \mu_2(dy) \leqslant \|A_2\|^p \int_{\Omega_2} \left| \sum_{i=1}^{n} a_i(x) \phi_2^i(y) \right|^p \mu_2(dy).$$

Integrating both sides with respect to μ_1, we see that

$$\begin{aligned}&\int_{\Omega_1}\int_{\Omega_2}\left|\sum_{i=1}^{n} a_i(x)\psi_2^i(y)\right|^p \mu_2(dy)\mu_1(dx)\\&\leqslant \|A_2\|^p \int_{\Omega_1}\int_{\Omega_2}\left|\sum_{i=1}^{n} a_i(x)\phi_2^i(y)\right|^p \mu_2(dy)\mu_1(dx).\end{aligned}\tag{1.12}$$

Since μ_1 and μ_2 are σ-finite, we can use Fubini's Theorem to get

$$\begin{aligned}&\int_{\Omega_1}\int_{\Omega_2}\left|\sum_{i=1}^{n} a_i(x)\psi_2^i(y)\right|^p \mu_2(dy)\mu_1(dx)\\&= \int_{\Omega_2}\left\|\sum_{i=1}^{n} a_i(\cdot)\psi_2^i(y)\right\|_p^p \mu_2(dy).\end{aligned}\tag{1.13}$$

If we set $\Phi(x) = \left|\sum_{i=1}^{n} a_i(x)\psi_2^i(\cdot)\right|^p, \ \ x \in \Omega_1$, then

$$\begin{aligned}&\int_{\Omega_2}\left\|\sum_{i=1}^{n} a_i(\cdot)\psi_2^i(y)\right\|_p^p \mu_2(dy)\\&= \left\|\int_{\Omega_1}\Phi(x)\mu_1(dx)\right\|_p^p \leqslant \int_{\Omega_1}\|\Phi(x)\|_p^p\, \mu_1(dx).\end{aligned}\tag{1.14}$$

If we combine Eqs. (6.8)–(6.10), we get

$$\int_{\Omega_2}\left\|\sum_{i=1}^{n} a_i(\cdot)\psi_2^i(y)\right\|_p^p \mu_2(dy) \leqslant \|A_2\|^p \int_{\Omega_2}\left\|\sum_{i=1}^{n} a_i(\cdot)\phi_2^i(y)\right\|_p^p \mu_2(dy).$$

It follows that $\|I_1 \otimes A_2\| \leqslant \|A_2\|^p$.

Since

$$\begin{aligned}\left|\sum_{i=1}^{n} a_i(x)\psi_2^i(y)\right| &= \left|A_2\left[\sum_{i=1}^{n} a_i(x)\phi_2^i(y)\right]\right|\\&= \left|\left((I_1 \otimes A_2)\left[\sum_{i=1}^{n} a_i \otimes \phi_2^i\right]\right)(x,y)\right|,\end{aligned}$$

we see that

$$\left\| \left[A_2 \left(\sum_{i=1}^{n} a_i(\cdot)\psi_2^i(\cdot) \right) \right] \right\|_p = \left\| \sum_{i=1}^{n} a_i \otimes A_2\phi_2^i \right\|_{\Delta_p}$$

$$= \left\| \left((I_1 \otimes A_2) \left[\sum_{i=1}^{n} a_i \otimes \phi_2^i \right] \right) \right\|_{\Delta_p} \leqslant \| I_1 \otimes A_2 \| \left\| \sum_{i=1}^{n} a_i \otimes \phi_2^i \right\|_{\Delta_p}.$$

Thus we see that $\|I_1 \otimes A_2\| = \|A_2\|$.

The same proof (with minor adjustments) shows that $(A_1 \otimes I_2)$ is also bounded as an operator mapping $L^p[\Omega_1]\otimes L^p[\Omega_2] \to L^p[\Omega_1 \times \Omega_2]$. Since $A_1 \otimes A_2 = (I_1 \otimes A_2)(A_1 \otimes I_2)$, we see that $\|A_1 \otimes A_2\| = \|A_1\| \, \|A_2\|$. □

Thus, we see that Δ_p is always a tensor norm relative to $\mathbf{L}^p[\Omega]$ $(1 \leqslant p \leq \infty)$.

Theorem 1.96. *[Schatten* [S]*, in Chap. 6] The norms λ, γ are tensor norms on $\mathcal{B}_1 \otimes \mathcal{B}_2$ and $\lambda \leqslant \gamma$. Furthermore, if α is any norm with $\lambda \leqslant \alpha \leqslant \gamma$, then α is a reasonable crossnorm which is a relative tensor norm that need not be a tensor norm, and $\gamma' \leqslant \alpha' \leqslant \lambda'$ (i.e., α' is a crossnorm on $\mathcal{B}'_1 \otimes \mathcal{B}'_2$, and $\gamma' = \lambda$, $\lambda' = \gamma$).*

Definition 1.97. A relative tensor norm α is said to be faithful if the natural linear mapping of $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$ into $\mathbb{L}_s(\mathcal{B}'_1, \mathcal{B}'_2)$, obtained by extending the identity $\mathbf{I}_1 \otimes \mathbf{I}_2$ on $\mathcal{B}_1 \otimes \mathcal{B}_2 \subset \mathcal{B}_1 \hat{\otimes}^\lambda \mathcal{B}_2$ by continuity to the entire space $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$, is one-to-one.

To say that α is faithful means that, if an element of $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$ vanishes on $\mathcal{B}'_1 \otimes \mathcal{B}'_2$, it is the zero function. For all of the above spaces, the relative tensor norm is faithful. Indeed, it has been shown by Gelbaum and Gil de Lamadrid [GG], in Chap. 6, that, if both $\mathcal{B}_1$ and $\mathcal{B}_2$ have Schauder bases and α is a relative tensor norm, then $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$ has a Schauder basis so that α is faithful. The following result is due to Ichinose [IC70], in Chap. 6.

Theorem 1.98. *Let A_1 and A_2 be closed densely defined linear operators on $\mathcal{B}_1$ and $\mathcal{B}_2$ respectively, and let α be a faithful relative tensor norm. Unless one of the extended spectra $\sigma_e(A_1)$ and $\sigma_e(A_2)$ contains* 0 *while the other contains ∞,*

$$(A_1 \hat{\otimes}^\alpha I_2)(I_1 \hat{\otimes}^\alpha A_2) = (I_1 \hat{\otimes}^\alpha A_2)(A_1 \hat{\otimes}^\alpha I_2) = A_1 \hat{\otimes}^\alpha A_2. \tag{1.15}$$

References

[CA] N.L. Carothers, *A Short Course on Banach Space Theory*. London Mathematical Society Student Texts, vol. 64 (Cambridge University Press, Cambridge, 2005)

[DOF] A. Defant, K. Floret, *Tensor Norms and Operator Ideals*. Mathematics Studies, vol. 176 (North-Holland, New York, 1993)

[DI] J. Diestel, *Sequences and Series in Banach Spaces*. Graduate Texts in Mathematics (Springer, New York, 1984)

[DU] J. Dugundji, *Linear Operators Part I: General Theory*. Wiley Classics Edition (Wiley, New York, 1988)

[DS] N. Dunford, J.T. Schwartz, *Linear Operators Part I: General Theory*. Wiley Classics Edition (Wiley, New York, 1988)

[EV] L.C. Evans, *Partial Differential Equations*. Graduate Studies in Mathematics, vol. 18 (American Mathematical Society, Providence, 1998)

[HP] E. Hille, R.S. Phillips, *Functional Analysis and Semigroups*. Colloquium Publications, vol. 31 (American Mathematical Society, Providence, 1957)

[J] F. Jones, *Lebesgue Integration on Euclidean Space*, revised edn. (Jones and Bartlett, Boston, 2001)

[L1] P.D. Lax, *Functional Analysis*, 2nd edn. (Wiley, New York, 2002)

[GL] G. Leoni, *A First Course in Sobolev Spaces*. Graduate Studies in Mathematics, vol. 105 (American Mathematical Society, Providence, 2009)

[ME] B. Mendelson, *Introduction to Topology* (Dover, New York, 1980)

[RS] M. Reed, B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic, New York, 1972)

[RO] H.L. Royden, *Real Analysis*, 2nd edn. (Macmillan, New York, 1968)

[RU1] W. Rudin, *Functional Analysis* (McGraw-Hill, New York, 1973)

[RU] W. Rudin, *Real and Complex Analysis*, 3rd edn. (McGraw-Hill, New York, 1987)

[RA] R. Ryan, *Introduction to Tensor Products of Banach Spaces* (Springer, New York, 2002)

[STE] E. Stein, *Singular Integrals and Differentiability Properties of Functions*. Princeton Mathematical Series, vol. 30 (Princeton University Press, Princeton, 1970)

[SZ] R.S. Strichartz, *A Guide to Distribution Theory and Fourier Transforms* (World Scientific, River Edge, 1994)

[YS] K. Yosida, *Functional Analysis*, 2nd edn. (Springer, New York, 1968)