1

Background

In this chapter we begin by reviewing the main definitions and theorems from the basic theory of functional analysis, linear operators and geometry of Banach spaces. It is not our intention to summarize the whole of analysis within a few pages, but we do supply the necessary background to the results used later in the book. This material is very standard and likely to be met in any basic course on functional analysis, and so we give just the essentials of the subject, without proofs.

In the last sections of this chapter, we also recall some basic facts of function theory. In particular we discuss the fundamental properties of Hardy spaces, which are Banach spaces of holomorphic functions defined in the unit disc and extended to the unit circle $\mathbb{T}$. We also briefly review the definitions of the disc algebra, functions of bounded mean oscillation, and the Hilbert transform of real functions defined on the unit circle.

1.1 Functional analysis

1.1.1 Weak topology

The term *weak topology* is most commonly used for the topology of a normed vector space or topological vector space induced by its (continuous) dual.

One may call subsets of a topological vector space *weakly closed* (respectively, *compact* etc.) if they are closed (respectively, compact etc.) in the weak topology. Likewise, functions are sometimes called *weakly continuous* (respectively, *differentiable, analytic* etc.) if they are continuous (respectively, differentiable, analytic etc.) in the weak topology.

The strong and weak topologies

Let $\mathcal{X}$ be a topological vector space; then in particular $\mathcal{X}$ is a topological space carrying a topology as part of its definition. (For example, a normed vector space $\mathcal{X}$ is, by using the norm to measure distances, a metric space, and hence also a topological vector space.) This topology is also called the *strong topology* on $\mathcal{X}$.

The weak topology on $\mathcal{X}$ is defined using the continuous dual space $\mathcal{X}^*$. This dual space consists of all linear functions from $\mathcal{X}$ into the base field $\mathbb{R}$ or $\mathbb{C}$ which are continuous with respect to the strong topology. The *weak topology* on $\mathcal{X}$ is the weakest topology (the topology with the fewest open sets) such that all elements of $\mathcal{X}^*$ remain continuous. Explicitly, a sub-base for the weak topology is the collection of sets of the form $\phi^{-1}(\mathcal{U})$ where $\phi \in \mathcal{X}^*$ and $\mathcal{U}$ is an open subset of the base field $\mathbb{R}$ or $\mathbb{C}$. In other words, a subset of $\mathcal{X}$ is open in the weak topology if and only if it can be written as a union of (possibly infinitely many) sets, each of which is an intersection of finitely many sets of the form $\phi^{-1}(\mathcal{U})$.

The weak topology is characterized by the following condition: a net $(x_\lambda)_\lambda$ in $\mathcal{X}$ converges in the weak topology to the element x of $\mathcal{X}$ if and only if $(\phi(x_\lambda))_\lambda$ converges to $\phi(x)$ in $\mathbb{R}$ or $\mathbb{C}$ for all ϕ in $\mathcal{X}^*$.

In particular, if $(x_n)_n$ is a sequence in $\mathcal{X}$, then $(x_n)_n$ converges weakly to x if

$$\phi(x_n) \to \phi(x) \qquad \text{as} \quad n \to \infty,$$

for all ϕ in $\mathcal{X}^*$. In this case, it is customary to write

$$x_n \xrightarrow{w} x$$

or, sometimes,

$$x_n \rightharpoonup x.$$

If $\mathcal{X}$ is equipped with the weak topology, then addition and scalar multiplication remain continuous operations, and $\mathcal{X}$ is a locally convex topological vector space.

The weak* topology

A normed space $\mathcal{X}$ can be embedded into $\mathcal{X}^{**}$ by

$$x \mapsto Tx,$$

where

$$Tx(\phi) = \phi(x).$$

In fact, $T : \mathcal{X} \to \mathcal{X}^{**}$ is an injective linear mapping. In the particular case where T is surjective, one says that $\mathcal{X}$ is *reflexive*.

The *weak* topology* on $\mathcal{X}^*$ is the weak topology induced by the image of $T : \mathcal{X} \to \mathcal{X}^{**}$. A net $(\phi_\lambda)_\lambda$ in $\mathcal{X}^*$ is convergent to ϕ in the weak* topology if it converges pointwise:

$$\phi_\lambda(x) \to \phi(x),$$

for all x in $\mathcal{X}$. In particular, a sequence $(\phi_n)_n \in \mathcal{X}^*$ converges to ϕ provided that

$$\phi_n(x) \to \phi(x),$$

for all x in X. In this case, one writes

$$\phi_n \stackrel{w^*}{\to} \phi$$

as $n \to \infty$.

This weak* topology is sometimes called the topology of *simple convergence* or the topology of *pointwise convergence*. Indeed, it coincides with the topology of pointwise convergence of linear functionals.

By definition, the weak* topology is weaker than the weak topology on $\mathcal{X}^*$. An important fact about the weak* topology is the *Banach–Alaoglu theorem*: if $\mathcal{X}$ is normed, then the unit ball in $\mathcal{X}^*$ is weak*-compact. Moreover, the unit ball in a normed space $\mathcal{X}$ is compact in the weak topology if and only if $\mathcal{X}$ is reflexive.

If a normed space $\mathcal{X}$ is separable, then the weak* topology is metrizable on (norm-)bounded subsets of $\mathcal{X}^*$.

1.1.2 Hahn–Banach theorem

Suppose that $\mathcal{X}$ is a normed space (real or complex). Then the *norm* of a continuous linear map $f : \mathcal{X} \to \mathbb{C}$ (a continuous linear *functional*) is given by

$$\|f\| = \sup\{|f(x)| : \|x\| \le 1\}.$$

Sometimes, for convenience, we shall use the alternative notation $\langle x, f\rangle$ instead of $f(x)$.

The *dual space*, $\mathcal{X}^*$, is the space of linear functionals equipped with the above norm.

Suppose now that $\mathcal{X}$ and $\mathcal{Y}$ are two normed spaces with $\mathcal{X} \subset \mathcal{Y}$. Then an element $g \in \mathcal{Y}^*$ clearly determines a unique element $g_{|\mathcal{X}}$ of $\mathcal{X}^*$ by restricting its action to $\mathcal{X}$. Moreover, $\|g_{|\mathcal{X}}\| \le \|g\|$. The Hahn–Banach theorem is concerned with the converse situation: the extension of a linear functional to a larger normed space. In its most common form it is stated as follows.

Theorem 1.1.1 (Hahn–Banach) *If $\mathcal{X}$ and $\mathcal{Y}$ are normed spaces with $\mathcal{X} \subset \mathcal{Y}$ and $f \in \mathcal{X}^*$, then there exists a functional $\tilde{f} \in \mathcal{Y}^*$ such that $\tilde{f}(x) = f(x)$ for all $x \in \mathcal{X}$, and such that $\|\tilde{f}\|_{\mathcal{Y}^*} = \|f\|_{\mathcal{X}^*}$.*

Another form of the Hahn–Banach theorem is more geometrical.

Definition 1.1.2 A set $\mathcal{S}$ in a normed space is *convex* if, for all $s, t \in \mathcal{S}$, the line segment joining s and t is contained in $\mathcal{S}$, i.e.,

$$\lambda s + (1 - \lambda)t \in \mathcal{S} \quad \text{for all} \quad 0 \le \lambda \le 1.$$

A set $\mathcal{S}$ in a normed space is *absolutely convex* if it is convex and if, in addition,

$$\lambda s \in \mathcal{S} \quad \text{for all} \quad s \in \mathcal{S}, \quad |\lambda| \le 1,$$

λ being real or complex, as appropriate.

The theorem of the separating hyperplane can now be stated.

Theorem 1.1.3 (separating hyperplane theorem) *Let $\mathcal{X}$ be a normed space, $\mathcal{S}$ a closed absolutely convex subset of $\mathcal{X}$, and x a point of $\mathcal{X}$ which is not in $\mathcal{S}$. Then there exists a functional $f \in \mathcal{X}^*$ such that $|f(s)| \le 1$ for all $s \in \mathcal{S}$, and $|f(x)| > 1$.*

A further application of the Hahn–Banach theorem is the following (see Rudin [180, Theorem 3.4]).

Theorem 1.1.4 *Let A and B be disjoint non-empty convex open subsets of a normed space $\mathcal{X}$. Then there are a functional $\Lambda \in \mathcal{X}^*$ and a real number γ such that*

$$\operatorname{Re} \Lambda(x) < \gamma < \operatorname{Re} \Lambda(y)$$

for all $x \in A$ and $y \in B$.

1.1.3 Stone–Weierstrass theorem

Let K be a compact metric space, for example $[0, 1]$ or the unit circle $\mathbb{T}$. We write $C(K, \mathbb{R})$ for the space of continuous real-valued functions on K, and $C(K)$ for the space of continuous complex-valued functions. Each is a normed space over the appropriate field, with the (supremum) norm

$$\|f\| = \sup\{|f(x)| : x \in K\}.$$

The classical *Weierstrass approximation theorem* states that the polynomials are dense in $C([0,1],\mathbb{R})$, that is, that a real continuous function can be uniformly approximated by polynomials on the interval $[0,1]$.

The Stone–Weierstrass theorem is a generalization of this, and requires us to consider an *algebra* of functions, that is, a set of functions that forms a vector space and is also closed under multiplication. So, for example, the polynomials form an algebra, as do the trigonometric polynomials (polynomials in e^{it} and e^{-it}).

An algebra $\mathcal{A}$ of functions is said to *separate points* if, given any two distinct points $x, y \in K$, there is a function $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. Over the reals we then have the simplest form of the Stone–Weierstrass theorem, as follows.

Theorem 1.1.5 (Stone–Weierstrass over $\mathbb{R}$) *If $\mathcal{A}$ is a real algebra of continuous functions on a compact metric space K, which separates points and contains the constant functions, then $\mathcal{A}$ is dense in $C(K,\mathbb{R})$. In other words, every function in $C(K,\mathbb{R})$ can be approximated arbitrarily closely (in the uniform norm) by functions in $\mathcal{A}$.*

Over the complex numbers the above form of the Stone–Weierstrass theorem does not hold, since, for example, the function $f(z) = \overline{z}$ cannot be approximated arbitrarily closely on $\mathbb{T}$ by polynomials in z, since it is not analytic. However, taking into account this special case, we obtain a complex form of the theorem: it can be deduced from the real form by taking real and imaginary parts.

Theorem 1.1.6 *If $\mathcal{A}$ is a complex algebra of continuous functions on a compact metric space K, which separates points, contains the constant functions, and is closed under complex conjugation, then $\mathcal{A}$ is dense in $C(K)$.*

1.1.4 Banach–Steinhaus theorem

Given two normed spaces $\mathcal{X}$ and $\mathcal{Y}$, we write $\mathcal{L}(\mathcal{X},\mathcal{Y})$ for the space of all bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$. Moreover, we write $\mathcal{L}(\mathcal{X})$ for $\mathcal{L}(\mathcal{X},\mathcal{X})$.

The Banach–Steinhaus theorem (or *uniform boundedness theorem*) may be stated as follows.

Theorem 1.1.7 *Suppose that $\mathcal{X}$ is a Banach space and $\mathcal{Y}$ a normed space. Then a collection of operators $\mathcal{S} \subset \mathcal{L}(\mathcal{X},\mathcal{Y})$ is* uniformly bounded in norm, *i.e.,*

$$\sup\{\|T\| : T \in \mathcal{S}\} < \infty,$$

if and only if it is pointwise bounded*, that is,*

$$\sup\{\|Tx\| : T \in \mathcal{S}\} < \infty \qquad \textit{for each} \quad x \in \mathcal{X}.$$

We may deduce from Theorem 1.1.7 that, when $\mathcal{X}$ is a Banach space, a sufficient condition for a family $\mathcal{S} \subset \mathcal{L}(\mathcal{X})$ to be uniformly bounded in norm is that

$$\sup\{|\langle Tx, x^*\rangle| : T \in \mathcal{S}\} < \infty, \qquad \text{for each} \quad x \in \mathcal{X} \quad \text{and} \quad x^* \in \mathcal{X}^*.$$

1.1.5 Complex measures

Definition

Definition 1.1.8 We define a *σ-ring* of subsets of a set X to be a non-empty collection Σ of sets closed under taking countable unions $\bigcup_{n=1}^{\infty} A_n$ and complements $A \setminus B$.

A *real positive measure* μ on a *measure space* (X, Σ) is a function

$$\mu : \Sigma \to [0, \infty],$$

defined on a σ-ring Σ, which is *σ-additive*; that is, for any sequence $(A_n)_n$ of disjoint sets in Σ one has

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

where the sum on the right may be finite or may diverge to ∞.

We shall also require more general (complex) measures, but here we shall always suppose that they are finite.

Definition 1.1.9 A *finite complex measure* μ on a *measure space* (X, Σ) is a function

$$\mu : \Sigma \to \mathbb{C},$$

defined on a σ-ring Σ, which is σ-additive in the sense that for any sequence $(A_n)_n$ of disjoint sets in Σ one has

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

where the sum on the right converges absolutely.

The set of complex measures defined on the unit circle $\mathbb{T}$ will be denoted by $\mathcal{M}(\mathbb{T})$.

Integration with respect to a complex measure

One can define the integral of a complex-valued measurable function with respect to a complex measure in the same way as the Lebesgue integral of a real-valued measurable function with respect to a non-negative measure, by approximating a measurable function with simple functions, beginning with the formula

$$\int_X \left(\sum_{j=1}^N a_j \chi_{A_j} \right) d\mu = \sum_{j=1}^N a_j \mu(A_j) \quad (a_1, \dots, a_N \in \mathbb{C}, \quad A_1, \dots, A_N \in \Sigma),$$

where χ_A is the *characteristic function* or *indicator function* of A, defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in X \setminus A. \end{cases}$$

Just as in the case of ordinary integration, this more general complex-valued integral may fail to exist, or its value may be infinite.

Another approach is to start with the concept of the integral of a real-valued function with respect to a non-negative measure. To that end, it is easy to verify that the real and imaginary parts μ_1 and μ_2 of a complex measure μ are finite-valued signed measures. One can apply the *Hahn–Jordan decomposition* to these measures to split them as

$$\mu_1 = \mu_1^+ - \mu_1^-$$

and

$$\mu_2 = \mu_2^+ - \mu_2^-,$$

where μ_1^+, μ_1^-, μ_2^+, μ_2^- are finite-valued non-negative measures (unique in some sense). Then, for a real-valued measurable function f, one can define

$$\int_X f\, d\mu = \left(\int_X f\, d\mu_1^+ - \int_X f\, d\mu_1^- \right) + i \left(\int_X f\, d\mu_2^+ - \int_X f\, d\mu_2^- \right)$$

as long as the expression on the right-hand side is defined, that is, all four integrals exist and when adding them up one does not encounter the indeterminate $\infty - \infty$.

Now, given a complex-valued measurable function, one can integrate its real and imaginary components separately as above and define, as expected,

$$\int_X f\, d\mu = \int_X \mathrm{Re}(f)\, d\mu + i \int_X \mathrm{Im}(f)\, d\mu.$$

We shall have recourse several times to *Fubini's theorem*, which deals with integration of a function $f : X \times Y \to \mathbb{C}$, supposed measurable with respect to a product measure $\mu \times \nu$. It is necessary to restrict ourselves to the case when μ and ν are *σ-finite* measures: this means that X is the countable union of a sequence of sets with finite μ-measure, and similarly for Y. Now, provided that we have the absolute convergence condition

$$\int_X \left(\int_Y |f(x, y)|\, d\nu(y) \right) d\mu(x) < \infty$$

(or, indeed, the same condition with the roles of X and Y interchanged), we may conclude that

$$\int_X \left(\int_Y f(x, y)\, d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y)\, d\mu(x) \right) d\nu(y). \quad (1.1)$$

We recall that for a measure space (X, μ), where μ is a positive measure, the spaces $L^p(X, \mu)$, for p satisfying $1 \le p < \infty$, consist of the measurable functions f such that

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{1/p} < \infty,$$

two functions being identified if they are equal *almost everywhere* (*a.e.*) (i.e., except on a set of measure 0). The space $L^\infty(X, \mu)$ consists of all functions such that

$$\|f\|_\infty := \inf\{K \ge 0 : \ |f| \le K \quad \text{a.e.}\} < \infty,$$

with the same identification of functions equal a.e. There is a duality formula, namely $(L^p)^* = L^q$, where $1 \le p < \infty$, $1 < q \le \infty$ and $1/p + 1/q = 1$, which arises from *Hölder's inequality*,

$$\|fg\|_1 \le \|f\|_p \|g\|_q. \quad (1.2)$$

If neither f nor g is identically zero, then the equality in (1.2) occurs if and only if there exists $c > 0$ such that

$$|f|^p = c|g|^q \quad \text{a.e.} \quad (1.3)$$

Let μ be a positive measure. We recall that a sequence (f_n) of measurable functions *converges in measure* to a function f if

$$\mu\{x \in X : |f_n(x) - f(x)| > \varepsilon\} \to 0$$

for all $\varepsilon > 0$. This condition is implied by convergence in the L^p norm.

We shall also require *Lusin's theorem*, in the following form.

Theorem 1.1.10 *Let f be a measurable function on a closed bounded interval $[a, b]$. Then, for every $\varepsilon > 0$ there is a closed subset $K_\varepsilon \subset [a, b]$ with $\mu([a, b] \setminus K_\varepsilon) < \varepsilon$, such that the restriction of f to K_ε is continuous.*

The idea underlying the proof of this result is that every bounded measurable function is the pointwise limit of a sequence of continuous functions, and indeed in this case the sequence converges uniformly on the complement of some set of measure less than ε.

Variation of a complex measure and polar decomposition

For a complex measure μ, one defines its *variation* or *absolute value* $|\mu|$ by the formula

$$|\mu|(A) = \sup \sum_{n=1}^{\infty} |\mu(A_n)|,$$

where A is in Σ and the supremum runs over all sequences of disjoint sets $(A_n)_n$ whose union is A. Taking only finite partitions of the set A into measurable subsets, one obtains an equivalent definition.

It turns out that $|\mu|$ is a non-negative finite measure. In the same way that a complex number can be represented in a polar form, one has a *polar decomposition* for a complex measure: there exists a measurable function θ with real values such that

$$d\mu = e^{i\theta} d|\mu|,$$

meaning that

$$\int_X f \, d\mu = \int_X f e^{i\theta} \, d|\mu|$$

for any absolutely integrable measurable function f, i.e., f satisfying

$$\int_X |f| \, d|\mu| < \infty.$$

One can use the Radon–Nikodým theorem (see below) to prove that the variation is a measure and the existence of the polar decomposition.

The space of complex measures

The sum of two complex measures is a complex measure, as is the product of a complex measure by a complex number. That is to say, the set of all complex measures on a measure space (X, Σ) forms a vector space. Moreover, the *total variation* $\|\cdot\|$, defined by

$$\|\mu\| = |\mu|(X),$$

is a norm, with respect to which the space of complex measures of finite variation is a Banach space.

Absolute continuity

Definition 1.1.11 Let ν and μ be positive, σ-finite measures on (X, Σ).

- The measure μ is *absolutely continuous* with respect to ν if, for all $A \in \Sigma$ such that $\nu(A) = 0$, we have also $\mu(A) = 0$. We denote this by
$$\mu \ll \nu.$$
- The measure μ is *supported* by $E \in \Sigma$ if, for all $A \in \Sigma$, we have
$$\mu(A) = \mu(A \cap E) \quad \text{or, equivalently,} \quad \mu(A \setminus E) = 0.$$
- The measures μ and ν are *mutually singular* if there exists $E \in \Sigma$ such that μ is supported by E and ν is supported by its complement E^c. We denote this by
$$\mu \perp \nu.$$

The Radon–Nikodým theorem and its consequences

Theorem 1.1.12 (Radon–Nikodým theorem) *Let ν and μ be positive σ-finite measures on (X, Σ). Then we have*

1. *There exists a unique pair of measures μ_1 and μ_2 such that*
   - $\mu = \mu_1 + \mu_2$*;*
   - $\mu_1 \ll \nu$*;*
   - $\mu_2 \perp \nu$*.*

   The measures μ_1 and μ_2 are positive and σ-finite.
2. *There exists a unique (ν almost everywhere) positive ν-integrable function h, such that, for all $A \in \Sigma$, we have*
$$\mu_1(A) = \int_A h \, d\nu = \int_X \chi_A h \, d\nu.$$

We now mention various consequences of the Radon–Nikodým theorem.

Definition 1.1.13 Let ν be a positive σ-finite measure on (X, Σ) and let μ be a positive σ-finite measure on (X, Σ). One says that μ has a *density* h with respect to ν if h is a positive ν-integrable function, such that for all $A \in \Sigma$ we have
$$\mu(A) = \int_A h \, d\nu = \int_X \chi_A h \, d\nu.$$