Part I

Introduction

This introductory part contains background material. It is not intended to be a course in any subject. It is simply a collection of definitions and facts given without proof (with one exception). We provide references to works which contain detailed exposition, full proofs, examples and motivation. In a sense this whole part is a quick reference guide to results which will be used in the later parts. The introduction is divided into two chapters, Chapter I.A describing what we will need from general functional analysis and Chapter I.B which contains results about concrete spaces and operators. Since I.A is really a review of a standard course in functional analysis, references are given only at the end of the chapter. In I.B we give references after each paragraph.

The references given in this part are usually to standard textbooks and monographs, not to original works. If a particular result or subject cannot be easily located using the table of contents or index we try to provide more detailed information (sections or pages). Sometimes we formulate a result in a form which is more convenient to us but different from the one given in the reference. Usually in such a case it is easy to derive our formulation from the one given in the reference.

Cambridge University Press
978-0-521-56675-9 - Banach Spaces for Analysts
P. Wojtaszczyk
Excerpt
More information

I.A. Functional analysis

1. A *linear topological space* X is a linear space over the real or complex numbers endowed with a topology τ such that the map $(x,y) \mapsto x+y$ is continuous from $(X,\tau)\times(X,\tau)$ into (X,τ) and the map $(t,x)\mapsto tx$ is continuous from $\mathbb{R}\times X$ (or $\mathbb{C}\times X$) into X. Such a topology is fully described by a basis of neighbourhoods of 0. A subset $V\subset X$ is called *convex* if whenever $x_1, x_2 \in V$ then the whole interval $\alpha x_1 + (1-\alpha)x_2$ for $0\le\alpha\le 1$ is in V. A linear topological space is called *locally convex* if it has a basis of convex neighbourhoods of 0. A *functional* on X is a continuous linear map from X into scalars. The set of all functionals on X will be denoted X^*, and called the *dual space*. A *linear operator* (or just *operator*) $T\colon X\to Y$ (where X and Y are linear topological spaces) is a continuous linear map. A *subspace* of X will always (unless explicitly stated otherwise) denote a closed linear subspace. Given a set $V\subset X$ by $\operatorname{span} V$ we denote the closure of the set of all linear combinations of elements from V (i.e. the subspace of X spanned by V).

2. A linear topological space X is called an *F-space* if its topology is given by a metric ρ such that $\rho(x,y)=\rho(x-y,0)$ and X is complete with respect to this metric. A *quasi-norm* on a linear space X is a function q from X into the nonnegative reals satisfying

(a) $q(x)=0$ if and only if $x=0$,

(b) $q(\lambda x)=|\lambda|q(x)$ for all scalars λ and all $x\in X$,

(c) there exists a constant C_X such that $q(x+y)\le C_X(q(x)+q(y))$ for all $x,y\in X$.

One easily checks that each quasi-norm defines a linear topology on X. The basis of neighbourhoods of the point x consists of 'balls' around x, i.e. sets $\{y\in X\colon q(x-y)<\varepsilon\}, \varepsilon>0$. A linear space X with a quasi-norm will be called a *quasi-normed space*. A very important special case of quasi-norm is a *norm*. This is a quasi-norm on X for which the constant C_X in (c) above equals 1 (one gets from (b) that $C_X\ge 1$). The usual notation for the norm of x is $\|x\|$. Every norm $\|\cdot\|$ on X defines a metric $\rho(x,y)=\|x-y\|$. A *Banach space* X is a linear space equipped with a norm $\|\cdot\|$ and such that X is complete with respect to the metric ρ. Clearly every Banach space is an F-space. The symbol B_X will denote the closed unit ball of X, i.e. $B_X = \{x \in X: \|x\| \le 1\}$. Its interior is an open convex set, so every Banach space is locally convex. This need not be true for general quasi-norms. Note also that a subspace of a Banach space (resp. quasi-normed space) is a Banach space (resp. quasi-normed space). We simply have to restrict the norm (resp. quasi-norm).

3. Let X and Y be two Banach spaces and let $T: X \to Y$ be a linear map. Then T is continuous if and only if $\|T\| = \sup\{\|Tx\|: \|x\| \le 1, x \in X\} < \infty$. The quantity $\|T\|$ is a norm on the linear space $L(X, Y)$ of all operators from X into Y. The space $L(X, Y)$ with the above defined norm is a Banach space. Unless otherwise indicated the convergence of operators will be understood in this norm.

4. In particular, for a Banach space X the space X^* is also a Banach space with the norm of a functional $x^* \in X^*$ defined as $\|x^*\| = \sup\{|x^*(x)|: x \in X, \|x\| \le 1\}$.

5. Open Mapping Theorem. Let X and Y be Banach spaces and let $T: X \to Y$ be a linear operator such that $\overline{T(B_X)}$ contains some open ball in Y. Then $T(X) = Y$ and there exists a positive number r such that $T(B_X) \supset r \cdot B_Y = \{y \in Y: \|y\| < r\}$.

6. Closed Graph Theorem. Suppose that $T: X \to Y$ is a linear map (not assumed to be continuous, but defined everywhere on X) from an F-space X into an F-space Y. Assume that $\{(x, Tx): x \in X\} \subset X \times Y$ is closed in the product topology. Then T is continuous.

7. Banach-Steinhaus Theorem. Suppose $(T_\gamma)_{\gamma \in \Gamma}$ is a family of linear operators from a Banach space X into a Banach space Y. Assume that for every $x \in X$ we have $\sup\{\|T_\gamma x\|: \gamma \in \Gamma\} < \infty$. Then $\sup\{\|T_\gamma\|: \gamma \in \Gamma\} < \infty$.

In particular we get that the pointwise limit of a sequence of linear operators (if it exists everywhere) is a linear operator.

8. Hahn-Banach Theorem. Let X be a linear space over the real numbers (without any topology) and let $Y \subset X$ be a linear subspace. Assume also that we have a function $p: X \to \mathbb{R}$ such that $p(x + y) \le p(x) + p(y)$ for all $x, y \in X$ and $p(tx) = tp(x)$ for all $x \in X$ and $t \in \mathbb{R}, t \ge 0$. Assume moreover that we have a linear map $f: Y \to \mathbb{R}$ such that $f(y) \le p(y)$ for all $y \in Y$. Then there exists a linear map $F: X \to \mathbb{R}$ such that $F \mid Y = f$ and $-p(-x) \le F(x) \le p(x)$ for all $x \in X$.

This general algebraic theorem has many very important special cases. Some of them are stated below in **9-11**. Note that despite the fact that **8** is true only for real spaces, the consequences listed below are true also for spaces over the complex scalars.

9. Taking in **8** $p(y) = \|y\|$ we obtain:

If X is a Banach space and $Y \subset X$ is a subspace and if $y^* \in Y^*$, then there exists $x^* \in X^*$ such that $\|x^*\| = \|y^*\|$ and $x^* \mid Y = y^*$. In particular we get

$$\|x\| = \sup\{|x^*(x)| : x^* \in X^*, \|x^*\| \leq 1\}.$$

10. A judicious choice of p yields also the following:

If X is a locally convex space and $A, B \subset X$ are disjoint closed convex sets with A being compact, then there exists a continuous linear functional f on X and a real number α such that $Re\ f(A) < \alpha$ and $Re\ f(B) > \alpha$.

In particular we see that X^* separates the points of X.

11. We also have the following version of **10**.

If X is a locally convex space and $A, B \subset X$ are disjoint convex sets with A open, then there exists a continuous linear functional f on X and a real number α such that $Re\ f(A) < \alpha$ and $Re\ f(B) \geq \alpha$.

12. If $T: X \to Y$ is a linear operator then it induces an operator $T^*: Y^* \to X^*$, called the *adjoint* (or *dual*) operator and defined as $T^*(y^*)(x) = y^*(Tx)$. One easily checks that $\|T\| = \|T^*\|$.

13. If $T: X \to Y$ is onto then T^* is 1-1 and there exists a constant $c > 0$ such that $\|T^* y^*\| \geq c\|y^*\|$ for all $y^* \in Y^*$.

14. If $T: X \to Y$ is such that there exists $c > 0$ such that $\|Tx\| \geq c\|x\|$ for all $x \in X$ then T^* maps Y^* onto X^*.

15. A linear map $T: X \to Y$ (X, Y Banach spaces) is *compact* if the set $\overline{T(B_X)}$ is compact in the norm topology of Y. One easily checks that each compact map is automatically continuous. The set of all compact operators from X into Y, denoted $K(X, Y)$, is a subspace (i.e. closed and linear) of $L(X, Y)$. Also an operator $T: X \to Y$ is compact if and only if $T^*: Y^* \to X^*$ is compact. It is easy to see that if $T \in K(X, Y)$ and $S_1 \in L(Z_1, X)$ and $S_2 \in L(Y, Z_2)$ then $S_2 T S_1 \in K(Z_1, Z_2)$.

16. An operator $T: X \to X$ is called *power-compact* if T^n is compact for some $n \in \mathbb{N}$.

17. An operator $T: X \to Y$ is called *invertible* if there exists an operator $S: Y \to X$ (usually we write T^{-1} instead of S) such that $ST = id_X$ and $TS = id_Y$. By id_X (or id_Y) we mean the identity operator on X (or on Y), i.e. $id_X(x) = x$ for all $x \in X$. It is important that we need both conditions $ST = id_X$ and $TS = id_Y$; one of them is not enough.

18. If X is a complex Banach space (i.e. it is a linear space over the complex numbers) and $T: X \to X$ is a linear operator, then the *spectrum* of T, denoted $\sigma(T)$, is the set of all $\lambda \in \mathbb{C}$ such that $(\lambda id_X - T)$ is not an invertible operator. The set $\sigma(T)$ is a non-empty, compact subset of $\mathbb{C}$. A number $\lambda \in \mathbb{C}$ is called an *eigenvalue* of T if there exists a vector $x \in X, x \neq 0$, called an *eigenvector* associated with the eigenvalue λ, such that $Tx = \lambda x$. Clearly each eigenvalue of T belongs to $\sigma(T)$. With each eigenvalue λ we associate its *spectral manifold* $E_\lambda = E_\lambda(T) = \bigcup_{n \geq 1} \ker(\lambda id_X - T)^n$.

Clearly E_λ is an increasing union of subspaces of X. The number $\dim E_\lambda$ (possibly ∞) is called the *multiplicity* of the eigenvalue λ.

19. If an operator $T: X \to Y$ is power-compact then $\sigma(T)$ is finite or consists of a sequence of points tending to zero together with zero itself. Every point $\lambda \in \sigma(T)$, except possibly zero, is an eigenvalue of T of finite multiplicity.

20. An operator $P: X \to X$ is called a *projection* if $P^2 = P$. Then $P(X)$ is a closed subspace of X and $Px = x$ for $x \in P(X)$. A subspace $Y \subset X$ which equals $P(X)$ for some projection $P: X \to X$ is called *complemented.*

21. Let V be a convex set in a locally convex topological vector space X. A point $v \in V$ is called an *extreme point* if it is not in the interior of any closed interval contained in V, i.e. if $v_1, v_2 \in V$ and $v = \alpha v_1 + (1-\alpha)v_2$ with $0 < \alpha < 1$ then $v = v_1 = v_2$. If A is any subset of X then the convex hull of A, denoted conv A, equals

$$\left\{ x \in X: x = \sum_{j=1}^{n} \alpha_j a_j \text{ with } \sum_{j=1}^{n} \alpha_j = \sum_{j=1}^{n} |\alpha_j| = 1 \right.$$
$$\left. \text{and } a_j \in A \text{ for } j = 1, 2, \ldots, n,\ n \in \mathbb{N} \right\}.$$

22. Krein-Milman Theorem. If V is a compact, convex subset of a locally convex, topological vector space, then V equals the closure of the convex hull of its extreme points.

In particular this theorem implies that each convex, compact subset of a locally convex, topological vector space has an extreme point.

References. Everything said above can be found in most textbooks on functional analysis. In particular everything can be found in Dunford-Schwartz [1958] or Edwards [1965] and everything except power-compact operators can be found in Rudin [1973].

I.B. Examples of spaces and operators

1. Whenever we consider a measure space (Ω, Σ, μ) we assume that the measure μ is complete and that there are no atoms of infinite measure. We say that the measure space (Ω, Σ, μ) is *separable* if there exists a countable family of sets $(A_j)_{j=1}^{\infty} \subset \Sigma$ such that the smallest complete σ-field containing $(A_j)_{j=1}^{\infty}$ is Σ. The following characterizes separable measure spaces.

Suppose that (Ω, Σ, μ) is a separable non-atomic measure space, with μ a positive measure and $\mu(\Omega) = 1$ (i.e. μ is a probability measure). Then (Ω, Σ, μ) is isomorphic to the unit interval [0,1] with the Lebesgue measure.

This can be found in Halmos [1950] §41. A general classification theorem for arbitrary measure spaces is due to Maharam [1942] and can also be found in Lacey [1974] §14.

2. For any measure space (Ω, Σ, μ) with μ positive we define $L_p(\Omega, \Sigma, \mu)$, $0 < p \leq \infty$, to be the space of Σ-measurable functions (more precisely of classes of functions where we identify functions which are equal μ-a.e.) such that $\int_\Omega |f(\omega)|^p d\mu(\omega) < \infty$. For $p = \infty$ we mean supess $|f(\omega)| < \infty$. Usually the σ-field Σ is clear so we will use the notation $L_p(\Omega, \mu)$ and when either the measure or the set are clear from the context we will suppress them also. We will use the notation $\|f\|_p = (\int_\Omega |f(\omega)|^p d\mu(\omega))^{\frac{1}{p}}$ for $0 < p < \infty$ and $\|f\|_\infty = \text{supess}|f(\omega)|$. If $1 \leq p \leq \infty$ then $\|\cdot\|_p$ is a norm and for $0 < p < 1$ it is a quasi-norm and it satisfies the inequality

$$\|f + g\|_p \leq (\|f\|_p^p + \|g\|_p^p)^{\frac{1}{p}}.$$

Thus $L_p(\Omega, \mu)$, $0 < p \leq \infty$, are linear spaces. For $1 \leq p \leq \infty$ they are Banach spaces. For $0 < p < 1$ we can introduce the natural metric on $L_p(\Omega, \mu)$ by $\rho(f, g) = \|f - g\|_p^p$. $L_p(\Omega, \mu)$ is complete with respect to this metric and the topology induced by this metric makes $L_p(\Omega, \mu)$ an F-space. It is easy to see that $L_p(\Omega, \mu)$ is not locally convex if $0 < p < 1$. By $L_0(\Omega, \mu)$ we mean the space of all (classes of) measurable, almost everywhere finite functions on Ω with the topology of convergence in measure. If $\mu(\Omega)$ is finite, this topology can be given by the metric $\rho(f, g) = \int_\Omega (\frac{|f-g|}{(1+|f-g|)}) d\mu$. When (Ω, μ) is σ-finite we write $\Omega = \bigcup_{j=1}^{\infty} E_j$ with the E_j's disjoint and $\mu(E_j) < \infty$ and define $\rho(f,g) = \sum_{j=1}^{\infty} 2^{-j}\mu(E_j)^{-1}\int_{E_j}(\frac{|f-g|}{(1+|f-g|)})d\mu$. Clearly $L_0(\Omega,\mu)$ (at least for a σ-finite measure space) is an F-space.

All this can be found in Dunford-Schwartz [1958], Edwards [1965], Banach [1932] and in most textbooks on functional analysis.

3. For a number p, $1 \le p \le \infty$ we will denote by p' the *conjugate exponent*, i.e. p' satisfies $\frac{1}{p} + \frac{1}{p'} = 1$ (so clearly $p' = \frac{p}{(p-1)}$) with the convention $1' = \infty$ and $\infty' = 1$. The following inequality, called **Hölder's inequality**, is of fundamental importance:

For any p, $1 \le p \le \infty$ and any two functions f, g we have

$$\int_\Omega |f(\omega)g(\omega)|d\mu(\omega) \le \left(\int_\Omega |f(\omega)|^p d\mu(\omega)\right)^{\frac{1}{p}} \left(\int_\Omega |g(\omega)|^{p'} d\mu(\omega)\right)^{\frac{1}{p'}}$$

or in short $\|f \cdot g\|_1 \le \|f\|_p \|g\|_{p'}$.

If this inequality is actually an equality then $g(\omega) = c \cdot \overline{\operatorname{sgn}(f(\omega))}$ $|f(\omega)|^{p-1}$ μ-a.e. where c is a constant and $\operatorname{sgn}(\alpha)$ is the complex signum of α, i.e. $\operatorname{sgn}\alpha = \alpha/|\alpha|$.

The proof can be found in almost every textbook of functional analysis or measure theory, e.g. Dunford-Schwartz [1958], Edwards [1965].

4. The duals of $L_p(\Omega,\mu)$ spaces for $1 \le p < \infty$ are described as follows (this is really a restatement of Hölder's inequality): $L_p(\Omega,\mu)^* = L_{p'}(\Omega,\mu)$ and the duality is given as $g(f) = \int_\Omega g(\omega)f(\omega)d\mu(\omega)$ for $g \in L_{p'}(\Omega,\mu)$ and $f \in L_p(\Omega,\mu)$. The space $L_\infty(\Omega,\mu)^*$ is (unless (Ω,μ) consists of a finite number of atoms) much bigger than $L_1(\Omega,\mu)$ and is impossible to describe explicitly. For $p < 1$ the dual depends on the measure space. If (Ω,μ) has no atoms then 0 is the only continuous linear functional on $L_p(\Omega,\mu)$. If (Ω,μ) is purely atomic and each atom has measure 1 then $L_p(\Omega,\mu)^* = L_\infty(\Omega,\mu)$ for $0 < p < 1$.

All this can be found in Dunford-Schwartz [1958] or Edwards [1965] or in most textbooks on functional analysis.

5. The following notational conventions will be used throughout this book. When the measure on the set Ω and the σ-field are clear from the context, we will use the notation $L_p(\Omega)$, in particular $L_p(\mathbb{R})$ and $L_p[0,1]$ will mean that we consider the usual Lebesgue measure and $L_p(\mathbb{T})$ will be with respect to Lebesgue measure on $\mathbb{T}$, but normalized to make the measure of the whole circle 1. The Lebesgue measure of a set A is denoted by $|A|$. If the measure space consists of a certain number of atoms each having measure 1, then $L_p(\Omega)$ will be denoted $\ell_p(\Omega)$. If those atoms are naturally identified with the integers $\mathbb{Z}$ or the natural numbers $\mathbb{N}$ we will use the notation ℓ_p. If the measure space Ω consists of n atoms each of measure 1, then $\ell_p(\Omega)$ will be denoted by ℓ_p^n. All this applies to $0 < p \leq \infty$.

6. The spaces $L_p(\Omega, \mu)$ for various values of p are clearly related. One of the expressions of this relationship are interpolation theorems. In this book we will use only the simplest cases of the best known results.

Riesz-Thorin Theorem. Let (Ω, μ) be a measure space and let $B \subset L_{p_1}(\Omega, \mu) \cap L_{p_2}(\Omega, \mu)$ be a linear set which is dense in both $L_{p_i}(\Omega, \mu)$, $i = 1, 2$, with $1 \leq p_1 < p_2 \leq \infty$. Assume that we have a linear map T defined on B with values in the measurable functions on a measure space (Ω_1, μ_1). Assume also that for every $f \in B$ we have $\|Tf\|_{p_i} \leq C\|f\|_{p_i}$ for $i = 1, 2$. Then for every p, $p_1 \leq p \leq p_2$ and $f \in B$ we have

$$\|Tf\|_p \leq C\|f\|_p$$

so T extends to a continuous linear operator from $L_p(\Omega, \mu)$ to $L_p(\Omega_1, \mu_1)$.

The proof of this theorem can be found in Katznelson [1968], Zygmund [1968], Dunford-Schwartz [1958], or in books devoted to interpolation of operators like Bennet-Sharpley [1988] or Krein-Petunin-Semenow [1978].

7. We say that a function f on a measure space (Ω, μ) is of *weak type* p, $0 < p < \infty$ and write $f \in L_{p,\infty}(\Omega, \mu)$ if there exists a constant K such that for every $\lambda \in \mathbb{R}_+$ we have

$$\mu\{\omega \in \Omega \colon |f(\omega)| \geq \lambda\} \leq \left(\frac{K}{\lambda}\right)^p.$$

We denote the space of all functions of weak type p by $L_{p,\infty}$ and the inf of all constants K which can appear in the above inequality for a given f by $\|f\|_{p,\infty}$. This quantity is not a norm but it is a convenient notation.

If T is a linear map defined on a dense, linear subset $B \subset L_p(\Omega, \mu)$ with values in the measurable functions on (Ω_1, μ_1), then we say that T is of *weak type (p-p)* if there exists a constant K such that

$$\mu_1\{\omega_1 \in \Omega_1 \colon |Tf(\omega_1)| \geq \lambda\} \leq K\left(\frac{\|f\|_p}{\lambda}\right)^p$$

for every $\lambda \in \mathbb{R}_+$ and $f \in B$. If $p = \infty$ we interpret this condition to mean $\|Tf\|_\infty \leq K\|f\|_\infty$.

Marcinkiewicz Theorem. Suppose that B is a linear set dense in both $L_{p_1}(\Omega, \mu)$ and $L_{p_2}(\Omega, \mu)$, $1 \leq p_1 < p_2 \leq \infty$ and suppose that T is a linear map defined on B with values in the measurable functions on a measure space (Ω_1, μ_1). If T is of weak type $(p_1 - p_1)$ and of weak type $(p_2 - p_2)$ then for every $p, p_1 < p < p_2$ there exists a constant C_p such that $\|Tf\|_p \leq C_p\|f\|_p$ for all $f \in B$.

So we get that T extends to a continuous operator from $L_p(\Omega, \mu)$ into $L_p(\Omega_1, \mu_1)$ for $p_1 < p < p_2$. For the constant C_p we have the estimate $C_p \leq K \max(\frac{1}{(p-p_1)}, \frac{1}{(p_2-p)})$. The proof can be found in Stein [1970], Zygmund [1968] or in any of the books on interpolation of operators mentioned in **6**.

8. The *Rademacher functions* $(r_n(t))_{n=1}^\infty$ are defined on [0,1] as $r_n(t) = \operatorname{sgn} \sin 2^n t\pi$. The alternative description is that $r_1(t) = 1$ if $0 \leq t < \frac{1}{2}$ and $r_1(t) = -1$ if $\frac{1}{2} < t \leq 1$ and $r_{n+1}(t)$ takes value 1 on the left hand half of each interval where $r_n(t)$ is constant and takes value -1 on the right hand half of each such interval. Using probabilistic language we can say that $(r_n(t))_{n=1}^\infty$ is a sequence of independent random variables, each taking value 1 with probability $\frac{1}{2}$ and value -1 with probability $\frac{1}{2}$.

Whatever the description, we easily see that $(r_n(t))_{n=1}^\infty$ is an orthonormal system in $L_2[0,1]$. Clearly this system is not complete. The main fact about Rademacher functions, which will be used repeatedly in this book is

Khintchine's inequality. There exist constants A_p, B_p, $0 < p < \infty$ such that for all (finite) sequences of scalars $(a_n)_{n=1}^\infty$ and every p, $0 < p < \infty$ we have

$$A_p\Big\|\sum_{n\geq 1} a_n r_n\Big\|_p \leq \Big\|\sum_{n\geq 1} a_n r_n\Big\|_2 = \Big(\sum_{n\geq 1} |a_n|^2\Big)^{\frac{1}{2}} \leq B_p\Big\|\sum_{n\geq 1} a_n r_n\Big\|_p.$$

There is also a similar inequality for lacunary sequences of exponents. More precisely we have the following: If $(n_k)_{k=1}^\infty$ is a sequence of natural numbers such that $\inf_k(n_{k+1}/n_k) = \lambda > 1$ then there exist constants $A_p = A(p, \lambda)$ and $B_p = B(p, \lambda)$ for $0 < p < \infty$ such that for every sequence of scalars $(a_k)_{k=1}^\infty$ and every p, $0 < p < \infty$ we have

$$A_p\Big\|\sum_{k=1}^\infty a_k e^{in_k\theta}\Big\|_p \leq \Big\|\sum_{k=1}^\infty a_k e^{in_k\theta}\Big\|_2 \leq B_p\Big\|\sum_{k=1}^\infty a_k e^{in_k\theta}\Big\|_p.$$