BOUNDED LINEAR OPERATORS

We will be concerned with linear transformations or operators $T : X \to Y$ where X and Y are normed, indeed, generally Banach spaces. For such a transformation T define the norm of T by

$$\|T\| = \sup\{\|Tx\|_Y : \|x\|_X = 1\} = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} \,.$$

When $\|T\| < \infty$, T is called *bounded*, and this is equivalent to T being continuous as a function from X to Y. We denote by $\mathcal{B}(X, Y)$ (or $\mathcal{B}(X)$ when $X = Y$) the set of bounded linear operators from X to Y. With the norm defined above this is normed space, indeed a Banach space if Y is a Banach space. Since the composition of bounded operators is bounded, $\mathcal{B}(X)$ is in fact an algebra. If X is finite dimensional then any linear operator with domain X is bounded and conversely (requires axiom of choice). In the special case that Y is the scalar field, always either $\mathbb{R}$ or $\mathbb{C}$, $\mathcal{B}(X, Y)$ is denoted by X^*. This is the *dual* space of X. As simple examples of these notions consider the following.

For infinite-dimensional X it is not apparent that X^* contains anything but the zero functional. The (Helly)- Hahn-Banach theorem is the basic tool which overcomes this difficulty, and does much more besides. There are various equivalent forms of this result of which we consider two.

Extension Form: Given a gauge function p on a real vector space X, and a linear functional ϕ defined on a subspace Y which satisfies $|\phi| \leq p$ on Y, there is an extension $\tilde{\phi}$ of ϕ to all of X satisfying $|\tilde{\phi}| \leq p$.

This is a result for *real* vector spaces as the proof explicitly uses the order structure of $\mathbb{R}$. However, Bohnenblust and Sobczyk showed that the complex case of the extension form is an almost immediate consequence.

Separation Form: Given two disjoint convex non-empty sets in a real vector space, one of which has an internal point, there is a hyperplane which separates the two sets.

There are several other equivalent results; the import of the separation form is the highlighting of the crucial role of convexity.

A *topological vector space* is both a vector and a topological space such that the operations of addition and scalar multiplication are continuous.

PROPOSITION A topological vector space admits a non-zero continuous linear functional if and only if it has a proper, open convex subset.

For a normed space X, let X^* be the (Banach) space of continuous linear functionals on X, and denote by $\sigma(X, X^*)$ the topology on X determined by the functionals in X^*; this is the *weak topology* on X.

Consequences of Hahn-Banach for a normed space X include the following:

- The space X^* of continuous linear functionals is total, that is, it separates points of X. This is just saying $\sigma(X, X^*)$ is Hausdorff.
- For $x \in X$, $\|x\| = \sup\{|x^*(x)| : x^* \in X^*, \|x^*\| \leq 1\}$.
- A subspace Y of X is (norm) dense in X if and only if any continuous linear functional that vanishes on Y is zero (if and only if it is weakly dense).
- (Mazur) If $x \in X$ lies in the weak closure of a set Y then x lies in the norm closure of the convex hull of Y. In particular, there exists a sequence of convex combinations of elements of Y which converges in norm to x.

The topology $\sigma(X, X^*)$ for a normed space X is only given by a norm when X is finite dimensional, in which case it is equivalent to the norm topology.

Since X^* is a Banach space (even if X is not complete), repeating the process gives the space $X^{**} = (X^*)^*$. For $x \in X$, the functional Jx defined on X^* by $(Jx)(x^*) = x^*(x)$ clearly lies in X^{**}, indeed the map $J : x \mapsto Jx$ is an isometry of X into X^{**}. If J is onto then X^{**} is called *reflexive.* (NB. X and X^{**} may be isometrically isomorphic without X being reflexive.) The importance of reflexivity comes from the following compactness results. On X^* we have the weak topology

$\sigma(X^*, X^{**})$ as above, but we also have the weaker $\sigma(X^*, JX) = \sigma(X^*, X)$, the *weak*-topology.*

THEOREM (Banach-Alaoglu) The closed unit ball in X^* is $\sigma(X^*, X)$-compact.

(It follows that every Banach space can be considered as a space of continuous functions on some compact space.)

If X is reflexive then clearly $\sigma(X, X^*) = \sigma(X^{**}, X^*)$, so that the unit ball of X is weakly compact. This condition is in fact equivalent to reflexivity.

If X, Y are Banach spaces and $T \in \mathcal{B}(X, Y)$, define $T^* : Y^* \to X^*$ by $(T^*y^*)(x) = y^*(Tx)$ for $x \in X, y^* \in Y^*$. T^* is the *adjoint* of T, and Hahn-Banach gives $\|T^*\| = \|T\|$.

THE CATEGORY THEOREMS

A subset Y of a topological space X is called *meagre*, or *of the first category*, in X if Y is a countable union of nowhere dense sets in X. If Y is non-meagre it is of the *second category* (in X).

THEOREM (Baire) A complete metric space is non-meagre in itself. Equivalently, a countable intersection of dense open sets in X is dense.

There is a more general result carrying the appellation of Mittag-Leffler, but in practice the following special case often suffices.

COROLLARY If a Banach space X is a countable union of closed subsets (Y_n) then some Y_k has interior in X.

THEOREM(Uniform boundedness principle, Banach-Steinhaus theorem) Let X, Y be Banach spaces, and suppose $(T_\alpha) \subseteq \mathcal{B}(X, Y)$ is pointwise bounded. Then (T_α) is (norm) bounded. That is,

$$\left(\sup_\alpha \|T_\alpha x\| < \infty \text{ for every } x \in X\right) \Rightarrow \sup_\alpha \|T\alpha\| < \infty\,.$$

NB. That X is non-meagre in itself is critical for this result.

COROLLARIES

(i) A weakly bounded set is (strongly) bounded.

(ii) Given $(T_n) \subseteq \mathcal{B}(X, Y)$ such that $Tx = \lim_n T_n(x)$ exists for each $x \in X$, then $T \in \mathcal{B}(X, Y)$.

(iii) A Banach space cannot have countable dimension; Hamel bases in Banach spaces are topologically 'bad'.

THEOREM(Closed graph) Let X, Y be Banach spaces, $T : X \to Y$ a linear transformation. If $Gr(T)$ is closed in $X \times Y$ then T is continous.

NB. Without completeness of *both* X and Y this result may fail.

THEOREM(Bounded inverse) Let X, Y be Banach spaces, $T \in \mathcal{B}(X, Y)$ a bijection. Then $T^{-1} \in \mathcal{B}(Y, X)$.

THEOREM (Open mapping) Let X, Y be Banach spaces, $T \in \mathcal{B}(X, Y)$ a surjection. Then T is open.

Remark. OMT $\Rightarrow$ CGT is almost immediate.

Applications: Schauder bases, Hörmander-Gårding criteria for hypoellipticity, separating space of a linear operator.

HOLOMORPHIC FUNCTIONS AND THE SPECTRUM

Let X be a Banach space, $D \subseteq \mathbb{C}$ an open set. A function $f : D \to X$ is holomorphic on D if for each $z_0 \in D$, $\lim\limits_{h \to 0} \dfrac{f(z_0 + h) - f(z_0)}{h}$ exists in X.

Essentially all of classical complex variable theory holds true in this situation, with almost identical proofs. Note that the apparently weaker notion, of requiring that for each $\phi \in X^*$ the map $\phi \circ f : D \to \mathbb{C}$ be holomorphic, is in fact equivalent (use UBP on difference quotients).

If T is a linear transformation on a finite dimensional space then either

(i) T is invertible, or

(ii) zero is an eigenvalue of T.

On a general Banach space X the situation is much more complicated. Suppose $T \in \mathcal{B}(X)$ and define disjoint subsets $\mathbb{C}$ as follows :

Point Spectrum of T :

$$P\sigma(T) = \{\lambda : (\lambda I - T)x = 0 \text{ for some non} - \text{zero } x \in X\} = \{\lambda : \lambda I - T \,\text{is not}\, 1:1\}$$

Continuous Spectrum of T :

$$C\sigma(T) = \{\lambda : [(\lambda I - T)X]^{-} = X, (\lambda I - T) \text{ is } 1:1 \text{ but has no bounded inverse}\}$$

Residual Spectrum of T :

$$R_{\sigma}(T) = \{\lambda : [(\lambda I - T)X]^{-} \neq X, (\lambda I - T) \text{ is } 1:1\}$$

Resolvent Set of T :

$$\begin{aligned} \rho(T) &= \{\lambda : [(\lambda I - T)X]^{-} = X, (\lambda I - T) \text{ is } 1:1 \text{ and has a bounded inverse}\} \\ &= \{\lambda : [(\lambda I - T)] = X, (\lambda I - T) \text{ is } 1:1 \text{ and has a bounded inverse}\} \end{aligned}$$

Thus $\rho(T)$ is the complement of the *spectrum* $\sigma(T) = P\sigma(T) \cup C\sigma(T) \cup R\sigma(T)$ of T. $P\sigma(T)$ is the set of eigenvalues of T; $C\sigma(T) \cup R\sigma(T) = \emptyset$ in the finite dimensional case.

THEOREM Let X be a Banach space, $T \in \mathcal{B}(X)$. Then $\rho(T)$ is an open set, and $\lambda \mapsto (\lambda I - T)^{-1}$ is holomorphic thereon.

COROLLARY $\sigma(T)$ is a non-empty compact set, and

$$\nu(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\} = \lim \|T^n\|^{1/n} \ .$$

COMPACT OPERATORS

If X, Y are Banach spaces, a linear map $T : X \to Y$ is *compact* if it maps bounded sets into pre-compact sets. If $\mathcal{C}(X, Y)$ denotes the collection of such T, then $\mathcal{C}(X, Y)$ is a norm closed subspace of $\mathcal{B}(X, Y)$. Ascoli-Arzela shows that $T \in \mathcal{C}(X, Y)$ if and only if $T^* \in \mathcal{C}(Y^*, X^*)$. $\mathcal{C}(X, Y)$ contains the set $\mathcal{F}(X, Y)$ of finite rank operators. Here $\mathcal{F}(X, Y)$ is the span of the one-dimensional operators

$$x^* \otimes y : x \mapsto x^*(x)y \qquad (x \in X, x^* \in X^*, y \in Y)\,.$$

Thus $\mathcal{C}(X,Y)$ contains $\mathcal{F}(X,Y)^-$. Whether $\mathcal{F}(X,Y)^- = \mathcal{C}(X,Y)$ is a very deep question, related to the approximation property :

X has the *approximation property* (AP) if the identity operator on X is approximable by elements of $\mathcal{F}(X)$ uniformly on compact sets of X.

A non-obvious fact is that X has the AP if and only if $\mathcal{F}(Y,X)^- = \mathcal{C}(Y,X)$ for all Y. As a simple example of these results we show:

THEOREM Any Hilbert space, indeed any space $L^p(\mu), 1 \leq p \leq \infty$, has the approximation property.

SPECTRAL THEORY OF COMPACT OPERATORS

THEOREM (Riesz-Schauder) If $T \in \mathcal{C}(X)$ then $\sigma(T)$ is at most countable with only possible limit point 0. Further, any non-zero point of $\sigma(T)$ is an eigenvalue of finite multiplicity.

COROLLARY (Fredholm alternative) If $T \in \mathcal{C}(X)$ then either $I - T$ is invertible or $1 \in P\sigma(T)$, that is, $1 \in \rho(T) \cup P\sigma(T)$.

COMPACT SELF-ADJOINT OPERATORS ON HILBERT SPACE

In the case of $T \in \mathcal{B}(H)$ for a Hilbert space H, the definition of adjoint operator is modified slightly to use $(x, T^*y) = (Tx, y)$ as the defining relation for the adjoint T^*, and then $T^* \in \mathcal{B}(H)$. In this latter case, T is *self-adjoint* if $T = T^*$. Self-adjoint operators are very special; they are the generalization of real symmetric matrices.

THEOREM (Hilbert-Schmidt) Let T be a self-adjoint compact operator on a separable Hilbert space H. Then there is an orthonormal basis (e_n) of H such that $Te_n = \lambda_n e_n$ where $\lambda_n \to 0$.

LEMMA If T is a self-adjoint operator on a Hilbert space then $\sigma(T) \subseteq \mathbb{R}$, $R\sigma(T) = \emptyset$, and $\nu(T) = \|T\|$.

THEOREM If T is a compact operator on a Hilbert space H, then there exist orthonormal sequences (x_n), (y_n) in H, and non-negative reals (λ_n), such that

$T = \sum_n \lambda_n x_n \otimes y_n$, the series being norm convergent.

In the self-adjoint case we can take $y_n = x_n$ and $\lambda_n \in \sigma(T)$, to give $T = \sum_n \lambda_n x_n \otimes x_n$, with $Tx_n = \lambda_n x_n$.

As an indication of where this 'spectral theory' is heading, take the self-adjoint case above, and let $X_\lambda = \bigoplus_{\lambda_n \leq \lambda} \{x : Tx = \lambda_n x\}$. Set P_λ to be the projection onto X_λ. Then $\lambda \mapsto P_\lambda$ is strongly right continuous[‡] and the above series for T can be written as the Riemann-Stieltjes integral

$$T = \int_{m-}^{M} \lambda dP_\lambda \,,$$

where $m = \min\ \sigma(T), M = \max\ \sigma(T)$.

One then defines $f(T) = \int_{m-}^{M} f(\lambda) dP_\lambda$, for suitable functions f, giving rise to the functional calculus homomorphism $f \mapsto f(T)$.

More generally, for any operator T on a Hilbert space H, such that $TT^* = T^*T$ (T is *normal*), there is a unique 'spectral measure' E defined on the Borel subsets of $\mathbb{C}$ and with range in the projections on H such that for $x, y \in H$

$$(Tx, y) = \int_{\sigma(T)} \lambda d\mu_{x,y} \,.$$

where for $x, y \in H$ the measure is defined on a Borel set S by $\mu_{x,y}(S) = (E(S)x, y)$.

‡ The strong topology on $\mathcal{B}(H)$ referred to here is that determined by the seminorms $T \mapsto \|Tx\|$ for $x \in H$. There is also the weak topology, determined by the seminorms $T \mapsto |(Tx, y)|$ for $x, y \in H$.

SIGNED AND COMPLEX MEASURES

Recall that a complex measure μ on a σ-algebra Σ of subsets of a set X, is a is a function $\mu : \Sigma \to \mathbb{C}$ such that for each disjoint sequence $(E_i) \subseteq \Sigma$, $\mu(\cup_i E_i) = \sum_i \mu(E_i)$.

Given such μ, need there be a positive measure λ such that $\lambda(E) \geq |\mu(E)|$ for all $E \in \Sigma$? For any $E \in \Sigma$, such λ would have to satisfy

$$\lambda(E) \geq \sup\Big\{\sum_i |\mu(E_i)| : \cup_i E_i = E, E_i \cap E_j = \emptyset, i \neq j\Big\}\,.$$

Let us denote the right side here by $|\mu|(E)$; $|\mu|$ is the *total variation* of μ.

THEOREM $|\mu|$ is a finite positive measure.

It follows that a (signed) measure $\mu = \mu^+ - \mu^-$, where μ^+, μ^- are positive measures, and $|\mu| = \mu^+ + \mu^-$.

If λ is a positive measure on Σ, and μ is a complex measure on Σ, then μ is *absolutely continuous with respect to* λ if $\lambda(E) = 0 \Rightarrow \mu(E) = 0 (E \in \Sigma)$. One writes $\mu \ll \lambda$. The name comes from the following.

THEOREM $\mu \ll \lambda$ if an only if, given $\epsilon > 0$, there is $\delta > 0$ such that $\lambda(E) < \delta \Rightarrow |\mu(E)| < \epsilon$ for all $E \in \Sigma$.

Two measures λ, μ on Σ are *mutually singular* if there exist disjoint $A, B \in \Sigma$ such that $\lambda(E) = 0$ for any $E \in \Sigma, E \subseteq A$, and $\mu(E) = 0$ for any $E \in \Sigma, E \subseteq B$. One writes $\lambda \perp \mu$.

THEOREM Let λ be a σ-finite positive measure on Σ, μ a complex measure on Σ.

(i) (Lebesgue decomposition) There exist unique mutually singular measures μ_a, μ_s on Σ such that $\mu = \mu_s + \mu_a, \mu_s \perp \lambda, \mu_a \ll \lambda$.

(ii) (Radon-Nikodym) There is a unique $h \in L^1(\lambda)$ such that

$$\mu_a(E) = \int_E h d\lambda \qquad (E \in \Sigma)\,.$$

The function h is generally denoted by $\dfrac{d\mu}{d\lambda}$, the *Radon-Nikodym derivative* of

μ with respect to λ.

COROLLARY Let μ a complex measure on Σ. Then $|\frac{d\mu}{d|\mu|}| = 1$ off a $|\mu|$-null set.

COROLLARY Let μ a real measure on Σ. Then there exist disjoint sets $A, B \in \Sigma$ such that $A \cup B = X$, and $\mu^+(E) = \mu(E \cap A), \mu^-(E) = \mu(E \cap B)$ for $E \in \Sigma$.

THEOREM Suppose that $1 \le p < \infty$, that λ is a σ-finite positive measure on a set X, and $\phi \in L^p(\lambda)^*$. Then there is a unique $g \in L^q(\lambda)$, where $\frac{1}{p} + \frac{1}{q} = 1$, with $\|g\|_q = \|\phi\|$, and $\phi(f) = \int fg d\lambda$ $(f \in L^p(\lambda)$.

This proves the (L^p, L^q) duality for p and q conjugate indices with $1 \le p < \infty$; $(L^\infty)^* \ne L^1$ in general. The Riesz theorem that $C(\Omega)^* = M(\Omega)$, for Ω locally compact Hausdorff, follows in a similar manner from the positive case, and this latter uses the result that for $\phi \in C(\Omega)^*, \phi \ge 0$, the quantity

$$E \mapsto \inf\{\sup\{\phi(f) : 0 \le f \le \chi(U)\} : U \supset E, U \text{ open}\}$$

defines an outer measure on Ω. The resulting measure is Borel and implements ϕ.

Let us remark at this point that measures taking values in a Banach space are of major importance. The elementary results there are like those for $\mathbb{C}$, but the theory has ramifications well beyond us here.

FUBINI-TONELLI THEOREMS

Returning to the (positive) measure situation, let (X, Σ, μ) and (Y, T, ν) be two measure spaces. The problem is to define $(X \times Y, \Sigma \times T, \mu \otimes \nu)$ such that for suitable f,

$$\int_{X \times Y} f d(\mu \otimes \nu) = \int_X \Big(\int_Y f d\nu\Big) d\mu = \int_Y \Big(\int_X f d\mu\Big) d\nu \, .$$

The trick is to build up from sets we would certainly wish to be in any such $\Sigma \times T$, namely those of the form $A \times B$ with $A \in \Sigma, B \in T$. These are termed *measurable rectangles*.

LEMMA The collection of finite disjoint unions of measurable rectangles forms an algebra.

We define $\Sigma \times T = \mathcal{S}$ to be the smallest σ-algebra containing this algebra. This is not very explicit, indicative of some of the technical difficulties present. Note that some books define $\Sigma \times T$ as the completion of $\mathcal{S}$ with respect to $\mu \otimes \nu$ defined below.

EXAMPLE Lebesgue measure on $\mathbb{R}^2$: if Σ is the Lebesgue measurable sets of $\mathbb{R}$, $\Sigma \times \Sigma$ is *not* the Lebesgue measurable sets of $\mathbb{R}^2$.

To relate $\Sigma \times T$ back to Σ and T, we use

LEMMA If $E \in \Sigma \times T$, then for every $x \in X, y \in Y$,

$$E_x = \{t : (x,t) \in E\} \in T, \quad E_y = \{s : (s,y) \in E\} \in \Sigma \, .$$

Indeed, if f is $\Sigma \times T$-measurable, then for each $x \in X, y \in Y$, $t \mapsto f(x,t)$ is T-measurable and $s \mapsto f(s,y)$ is Σ-measurable.

We need another characterization of $\Sigma \times T$.

DEFINITION. A collection $\mathcal{S} \subseteq \mathcal{P}(X)$ is a monotone class if

$$\big\{(A_i), (B_i) \subseteq \mathcal{S}, A_i \uparrow, B_i \downarrow\big\} \Rightarrow \big\{\cup_i A_i \in \mathcal{S}, \cap_i B_i \in \mathcal{S}\big\} \, .$$

THEOREM Let $\mathcal{S}_0$ be an algebra of sets on X. Then the smallest σ-algebra containing $\mathcal{S}_0$ is the smallest monotone class containing $\mathcal{S}_0$.

THEOREM Let (X, Σ, μ) and (Y, T, ν) be two σ-finite measure spaces, $E \in \Sigma \times T$. Then

$x \mapsto \nu(E_x)$ is σ-measurable,

$y \mapsto \mu(E_y)$ is T-measurable, and

$$\int_X \nu(E_x) d\mu = \int_Y \mu(E_y) d\nu \ .$$

Proof. In the finite case the collection of sets with the desired property is a monotone class containing the measurable rectangles. ■

EXAMPLES.

(i) $[0,1]$, with $\mu =$ Lebesgue measure, $\nu =$ counting measure, and E the diagonal in $[0,1]^2$.

(ii) (Sierpinski) $X = Y = [0,1]$, $\Sigma = T$ the σ-algebra of countable and co-countable sets, with $\mu = \nu$ determined by $\mu(X) = 1, \mu(S) = 0$ if $|S| = 1$; and $E = \{(x,y) : x < y\}$.

DEFINITION. For $E \in \Sigma \times T$, set

$$\mu \otimes \nu(E) = \int_X \nu(E_x) d\mu = \int_Y \mu(E_y) d\nu \ .$$

We now have $(X \times Y, \Sigma \times T, \mu \otimes \nu)$ defined, with a special case of the integration result built in. The MCT now gives the following.

THEOREM Let (X, Σ, μ) and (Y, T, ν) be two σ-finite measure spaces, $f : X \times Y \to \mathbb{R}^*$ a non-negative $\Sigma \times T$-measurable function. Then

(a) $y \mapsto \int_X f(x,y) d\mu(x))$ is T-measurable,

(b) $x \mapsto \int_Y f(x,y) d\nu(y))$ is Σ-measurable,

(c)

$$\int_X \int_Y f(x,y) d\nu(y) d\mu(x) = \int_{X\times} f(x,y) d\mu \otimes \nu(x,y) = \int_Y \int_X f(x,y) d\mu(x) d\nu(y) \ .$$

THEOREM Let (X, Σ, μ) and (Y, T, ν) be two σ-finite measure spaces, $f : X \times Y \to \mathbb{R}^*$ a $\Sigma \otimes T$-measurable function. Suppose that one of

$$\int_X \int_Y |f(x,y)| d\nu(y) d\mu(x), \int_{X \times Y} |f(x,y)| d\mu \otimes \nu(x,y), \int_Y \int_X |f(x,y)| d\mu(x) d\nu(y)$$

is finite. Then

(a) $x \mapsto f(x,y) \in L^1(\mu)$ for ν-a.a. y,

(b) $y \mapsto f(x,y) \in L^1(\nu)$ for μ-a.a. x,

(c) $y \mapsto \int_X f(x,y)d\mu(x) \in L^1(\nu)$,

(d) $x \mapsto \int_Y f(x,y)d\nu(y) \in L^1(\mu)$,

(e)

$$\int_X \int_Y f(x,y)d\nu(y)d\mu(x) = \int_{X\times Y} f(x,y)d\mu\otimes\nu(x,y)) = \int_Y \int_X f(x,y)d\mu(x)d\nu(y)\,.$$

EXAMPLE The result may fail without finiteness of the integrals, as is shown by the function

$$f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2} \qquad x,y \in (0,1)\,.$$

As noted above, if Σ is the Lebesgue measurable sets of $\mathbb{R}$, $\Sigma \times \Sigma \supset \mathcal{B}(\mathbb{R}^2)$, but is not the Lebesgue measurable sets of $\mathbb{R}^2$. Since any Lebesgue measurable function on $\mathbb{R}^2$ is equal a.e. to a Borel function this creates no difficulty. Indeed most applications will be to continuous functions.

EXAMPLE. The function $f(x,y) = e^{-xy\sin x}$ on $\mathbb{R}^2$ gives, via a Fubini argument,

$$\lim_{r\to\infty} \int\limits_0^r \frac{\sin x}{x}dx = \frac{\pi}{2}\,.$$