BANACH AND HILBERT SPACE REVIEW

CHRISTOPHER HEIL

These notes will briefly review some basic concepts related to the theory of Banach and Hilbert spaces. We are not trying to give a complete development, but rather review the basic definitions and theorems, mostly without proof.

1. Banach Spaces

Definition 1.1 (Norms and Normed Spaces)**.** Let X be a vector space (= linear space) over the field $\mathbb{C}$ of complex scalars. Then X is a *normed linear space* if for every $f \in X$ there is a real number $\|f\|$, called the *norm* of f, such that:

(a) $\|f\| \geq 0$,

(b) $\|f\| = 0$ if and only if $f = 0$,

(c) $\|cf\| = |c|\,\|f\|$ for every scalar c, and

(d) $\|f + g\| \leq \|f\| + \|g\|$.

Definition 1.2 (Convergent and Cauchy sequences)**.** Let X be a normed space, and let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of elements of X.

(a) $\{f_n\}_{n\in\mathbb{N}}$ *converges* to $f \in X$ if $\lim_{n\to\infty} \|f - f_n\| = 0$, i.e., if

$$\forall\, \varepsilon > 0, \quad \exists\, N > 0, \quad \forall\, n \geq N, \quad \|f - f_n\| < \varepsilon.$$

In this case, we write $\lim_{n\to\infty} f_n = f$ or $f_n \to f$.

(b) $\{f_n\}_{n\in\mathbb{N}}$ is *Cauchy* if

$$\forall\, \varepsilon > 0, \quad \exists\, N > 0, \quad \forall\, m, n \geq N, \quad \|f_m - f_n\| < \varepsilon.$$

Definition 1.3 (Banach Spaces)**.** It is easy to show that any convergent sequence in a normed linear space is a Cauchy sequence. However, it may or may not be true in an arbitrary normed linear space that all Cauchy sequences are convergent. A normed linear space X which does have the property that all Cauchy sequences are convergent is said to be *complete*. A complete normed linear space is called a *Banach space*.

Date: Revised September 7, 2006.

Example 1.4. The following are all Banach spaces under the given norms. Here p can be in the range $1 \le p < \infty$.

$$
\begin{aligned}
L^p(\mathbb{R}) &= \Big\{f \colon \mathbb{R} \to \mathbb{C} \;:\; \int_{\mathbb{R}} |f(x)|^p \, dx < \infty\Big\}, & \|f\|_p &= \Big(\int_{\mathbb{R}} |f(x)|^p \, dx\Big)^{1/p}. \\
L^\infty(\mathbb{R}) &= \big\{f \colon \mathbb{R} \to \mathbb{C} \;:\; f \text{ is essentially bounded}\big\}, & \|f\|_\infty &= \operatorname*{ess\,sup}_{x\in\mathbb{R}} |f(x)|. \\
C_b(\mathbb{R}) &= \big\{f \in L^\infty(\mathbb{R}) : \; f \text{ is bounded and continuous}\big\}, & \|f\|_\infty &= \sup_{x\in\mathbb{R}} |f(x)|, \\
C_0(\mathbb{R}) &= \big\{f \in C_b(\mathbb{R}) \;:\; \lim_{|x|\to\infty} f(x) = 0\big\}, & \|f\|_\infty &= \sup_{x\in\mathbb{R}} |f(x)|.
\end{aligned}
$$

Remark 1.5. (a) To be more precise, the elements of $L^p(\mathbb{R})$ are equivalence classes of functions that are equal almost everywhere, i.e., if $f = g$ a.e. then we identify f and g as elements of $L^p(\mathbb{R})$.

(b) If $f \in C_b(\mathbb{R})$ then $\operatorname{ess\,sup}_{x\in\mathbb{R}} |f(x)| = \sup_{x\in\mathbb{R}} |f(x)|$. Thus we regard $C_b(\mathbb{R})$ as being a subspace of $L^\infty(\mathbb{R})$; more precisely, each element $f \in C_b(\mathbb{R})$ determines an equivalence class of functions that are equal to it a.e., and it is this equivalence class that belongs to $L^\infty(\mathbb{R})$.

(c) The L^∞ norm on $C_b(\mathbb{R})$ is called the *uniform norm*. Convergence with respect to this norm coincides with the definition of uniform convergence.

(d) Sometimes the notation $C(\mathbb{R})$ is used to denote the space that we call $C_b(\mathbb{R})$, or $C(\mathbb{R})$ may denote the space of all possible continuous functions on $\mathbb{R}$. In this latter case, $\|\cdot\|_\infty$ would not be a norm on $C(\mathbb{R})$, since $\|f\|_\infty$ would not be finite for all $f \in C(\mathbb{R})$.

Definition 1.6 (Closed Subspaces)**.** Let X be a Banach space.

(a) A vector $f \in X$ is a *limit point* of a set $S \subset X$ if there exist vectors $g_n \in S$ that converge to f. In particular, every element of S is a limit point of S.

(b) A subset S of a Banach space X is *closed* if it contains all its limit points. In other words, S is closed if whenever $\{g_n\}_{n\in\mathbb{N}}$ is a sequence of elements of S and $g_n \to f \in X$, then f must be an element of S.

If Y is a closed sub*space* of a Banach space X, then it is itself a Banach space under the norm of X. Conversely, if Y is a subspace of X and Y is a Banach space under the norm of X, then Y is a closed subspace of X.

Example 1.7. $C_b(\mathbb{R})$ and $C_0(\mathbb{R})$ are closed subspaces of $L^\infty(\mathbb{R})$ under the L^∞ norm.

The following are also subspaces of $L^\infty(\mathbb{R})$, but they are not a Banach spaces under the L^∞ norm because they are not closed subspaces of $L^\infty(\mathbb{R})$:

$$
\begin{aligned}
C_c(\mathbb{R}) &= \{f \in C_b(\mathbb{R}) : f \text{ has compact support}\}, \\
C_b^\infty(R) &= \{f \in C_b(\mathbb{R}) : f \text{ is infinitely differentiable and } f^{(m)} \in L^\infty(\mathbb{R}) \text{ for every } m \ge 0\}, \\
\mathcal{S}(\mathbb{R}) &= \{f \in C_b^\infty(\mathbb{R}) : x^m f^{(n)}(x) \in L^\infty(\mathbb{R}) \text{ for every } m, n \ge 0\}.
\end{aligned}
$$

The space $\mathcal{S}(\mathbb{R})$ is called the *Schwartz space*.

Example 1.8 (Dense Subspaces)**.** Suppose that S is a subspace of a Banach space X. The *closure* of S is the smallest closed subset $\bar{S}$ of X that contains S. It can be shown that $\bar{S}$ is the union of S and all all of the limit points of S. If $\bar{S} = T$ then S is said to be *dense* in T.

For example, the space $C_c(\mathbb{R})$ is dense in $C_0(\mathbb{R})$ in the $L^\infty(\mathbb{R})$ norm. That is, every function in $C_0(\mathbb{R})$ can be written as a limit (in the L^∞ norm) of functions from $C_c(\mathbb{R})$. Similarly, $\mathcal{S}(\mathbb{R})$ is dense in $C_0(\mathbb{R})$ in L^∞ norm. However, neither $C_c(\mathbb{R})$ nor $\mathcal{S}(\mathbb{R})$ nor $C_0(\mathbb{R})$ is dense in $L^\infty(\mathbb{R})$ using the L^∞ norm. This is because the L^∞-norm limit of a sequence of continuous functions is continuous, so a bounded but discontinuous function could not be written as a *uniform* limit of functions from these spaces.

On the other hand, $C_c(\mathbb{R})$, $C_0(\mathbb{R})$, and $\mathcal{S}(\mathbb{R})$ are all dense in $L^p(\mathbb{R})$ (in L^p norm, not L^∞ norm) for each $1 \le p < \infty$.

Definition 1.9 (Sequence spaces)**.** There are also many useful Banach spaces whose elements are sequences of complex numbers. Be careful to distinguish between an element of such a space, which is a sequence of numbers, and a sequence of elements of such a space, which would be a sequence of sequences of numbers.

The elements of a sequence space may be indexed by any countable index set I, typically the natural numbers $\mathbb{N} = \{1, 2, \dots\}$ or the integers $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$. If a sequence is indexed by the natural numbers then a typical sequence a has the form $a = (a_n)_{n\in\mathbb{N}} = (a_1, a_2, \dots)$, while if they are indexed by the integers then a typical sequence has the form $a = (a_n)_{n\in\mathbb{Z}} = (\dots, a_{-2}, a_{-1}, a_0, a_1, a_2, \dots)$. In either case, the a_n are complex numbers.

The following are Banach spaces whose elements are sequences of complex numbers:

$$\begin{aligned}
\ell^p(I) &= \Big\{(a_n)_{n\in I} \;:\; \sum_{n\in I} |a_n|^p < \infty\Big\}, & \|(a_n)\|_p &= \Big(\sum_{n\in I} |a_n|^p\Big)^{1/p}, \quad 1 \le p < \infty,\\
\ell^\infty(I) &= \{(a_n)_{n\in I} \;:\; \sup_{n\in I} |a_n| < \infty\}, & \|(a_n)\|_\infty &= \Big(\sup_{n\in I} |a_n|\Big).
\end{aligned}$$

Example 1.10 (Finite dimensional Banach spaces)**.** The finite-dimensional complex Euclidean space $\mathbb{C}^n$ is a Banach space. There are infinitely many possible choices of norms for $\mathbb{C}^n$. In particular, each of the following determines a norm for $\mathbb{C}^n$:

$$\begin{aligned}
|u|_p &= \big(|u_1|^p + \cdots + |u_n|^p\big)^{1/p}, \quad 1 \le p < \infty,\\
|u|_\infty &= \max\{|u_1|, \cdots, |u_n|\}.
\end{aligned}$$

The norm $|\cdot|_2$ is the usual Euclidean norm on $\mathbb{C}^n$.

2. Hilbert Spaces

Definition 2.1 (Inner Products)**.** Let H be a vector space. Then H is an *inner product space* if for every $f, g \in X$ there exists a complex number $\langle f, g\rangle$, called the *inner product* of f and g, such that:

(a) $\langle f, f\rangle$ is real and $\langle f, f\rangle \geq 0$.
(b) $\langle f, f\rangle = 0$ if and only if $f = 0$.
(c) $\langle g, f\rangle = \overline{\langle f, g\rangle}$,
(d) $\langle af_1 + bf_2, g\rangle = a\langle f_1, g\rangle + b\langle f_2, g\rangle$.

Each inner product determines a norm by the formula $\|f\| = \langle f, f\rangle^{1/2}$. Hence every inner product space is a normed linear space. The *Cauchy–Schwarz inequality* states that $|\langle f, g\rangle| \leq \|f\|\,\|g\|$ for every $f, g \in H$.

If an inner product space H is complete, then it is called a *Hilbert space.* In other words, a Hilbert space is a Banach space whose norm is determined by an inner product.

Example 2.2. $L^2(\mathbb{R})$ is a Hilbert space when the inner product is defined by

$$\langle f, g\rangle \;=\; \int_{\mathbb{R}} f(x)\,\overline{g(x)}\,dx.$$

$\ell^2(I)$ is a Hilbert space when the inner product is defined by

$$\big\langle (a_n),\, (b_n)\big\rangle \;=\; \sum_{n\in I} a_n \bar{b}_n.$$

However, neither $L^p(\mathbb{R})$ nor ℓ^p is a Hilbert space when $p \neq 2$.

Example 2.3 (Finite dimensional Hilbert spaces)**.** The space $\mathbb{C}^n$, finite-dimensional complex Euclidean space, is a Hilbert space. The inner product is just the usual dot product of vectors:

$$\langle u, v\rangle \;=\; u \cdot v \;=\; u_1\bar{v}_1 + \cdots + u_n\bar{v}_n.$$

This inner product determines the usual Euclidean norm $|\cdot|_2$ defined by

$$|u|_2 \;=\; (|u_1|^2 + \cdots + |u_n|^2)^{1/2}.$$

However, none of the norms $|\cdot|_p$ with $p \neq 2$ are determined by an inner product. There are many other possible inner products for $\mathbb{C}^n$.

3. Complete Sequences in Hilbert Spaces

Warning: the term "complete" has several entirely distinct mathematical meanings. We have already said that "a Banach space is complete" if every Cauchy sequence in the space converges. The term "complete sequences" defined in this section is a completely separate definition that applies to sets of vectors in a Hilbert or Banach space (although we will only define it for Hilbert spaces).

Definition 3.1 (Complete Sequences)**.** Let H be a Hilbert space, and let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of elements of H. The *finite linear span* of $\{f_n\}_{n\in\mathbb{N}}$ is the set of all finite linear combinations of the f_n, i.e.,

$$\operatorname{span}\{f_n\}_{n\in\mathbb{N}} \;=\; \biggl\{\sum_{n=1}^{N} c_n f_n : \text{ all } N>0 \text{ and all } c_1,\dots,c_N \in \mathbb{C}\biggr\}.$$

The finite span $\operatorname{span}\{f_n\}_{n\in\mathbb{N}}$ is a subspace of H. If $\operatorname{span}\{f_n\}_{n\in\mathbb{N}}$ is dense in H, then the sequence $\{f_n\}_{n\in\mathbb{N}}$ is said to be *complete* (or *total* or *fundamental*) in H.

By definition, the finite linear span is dense in H if the closure of the finite linear span is all of H. The closure of the finite span is the *closed span*, consisting of all limits of elements of the finite span, i.e.,

$$\overline{\operatorname{span}}\{f_n\}_{n\in\mathbb{N}} \;=\; \Bigl\{f \in H : f = \lim_{k\to\infty} h_k \text{ for some } h_k \in \operatorname{span}\{f_n\}_{n\in\mathbb{N}}\Bigr\},$$

Both $\operatorname{span}\{f_n\}_{n\in\mathbb{N}}$ and $\overline{\operatorname{span}}\{f_n\}_{n\in\mathbb{N}}$ are subspaces of H, but generally only the closed span is a closed subspace of H.

By definition, $\{f_n\}_{n\in\mathbb{N}}$ is complete if every element $f \in H$ is a limit point of $\operatorname{span}\{f_n\}_{n\in\mathbb{N}}$ and therefore can be written as the limit of a sequence of finite linear combinations of the elements f_n. In other words, given any $\varepsilon > 0$ there must exist some $N > 0$ and some scalars $c_1(f,\varepsilon),\dots,c_N(f,\varepsilon)$ such that

$$\biggl\|f - \sum_{n=1}^{N} c_n(f,\varepsilon)\, f_n\biggr\| \;<\; \varepsilon. \tag{3.1}$$

Note that N and the coefficients $c_n(f,\varepsilon)$ depend on *both* ε and f.

Example 3.2. A sequence $\{f_n\}_{n\in\mathbb{N}}$ is a *Schauder basis* (or simply a *basis*) for H if each element $f \in H$ can be written as a unique "infinite linear combination" of the f_n. That is, given $f \in H$ there must exist unique coefficients $c_n(f)$ so that

$$f \;=\; \sum_{n=1}^{\infty} c_n(f)\, f_n. \tag{3.2}$$

What the series in (3.2) really means is that the partial sums are converging to f, i.e.,

$$\lim_{N\to\infty} \biggl\|f - \sum_{n=1}^{N} c_n(f)\, f_n\biggr\| \;=\; 0.$$

Thus, $\{f_n\}_{n\in\mathbb{N}}$ is a basis if for each $f \in H$ there exist unique scalars $\{c_n(f)\}_{n\in\mathbb{N}}$ (depending *only* on f) such that if $\varepsilon > 0$ then there exists an $N > 0$ so that

$$\biggl\|f - \sum_{n=1}^{N} c_n(f)\, f_n\biggr\| \;<\; \varepsilon. \tag{3.3}$$

Comparing (3.3) with (3.1) we see that being a basis is a more stringent condition than being a complete set, i.e., all bases are complete, but not all complete sets need be bases. In particular, when $\{f_n\}_{n\in\mathbb{N}}$ is a basis, the coefficients $c_n(f)$ do *not* depend on ε, while for a complete set they can depend on both f *and* ε. The number N depends on both ε and f in both cases, the difference being that for a basis, if you make ε smaller then you simply have to take *more* of the coefficients $c_n(f)$ to make equation (3.3) valid, while for a complete set, if you make ε smaller then you may have to take both a larger N and a possibly *completely different* choice of scalars $c_n(f, \varepsilon)$ to make equation (3.1) valid.

The following result gives a characterization of complete sets, namely, a sequence $\{f_n\}_{n\in\mathbb{N}}$ is complete if and only if there is no element other than 0 that is orthogonal to every f_n.

Theorem 3.3. Let $\{f_n\}_{n\in\mathbb{N}}$ be a set of elements from a Hilbert space H. Then the following two statements are equivalent, i.e., each statement implies the other.

(a) $\{f_n\}_{n\in\mathbb{N}}$ is complete in H.

(b) The only element $f \in H$ which satisfies $\langle f, f_n\rangle = 0$ for every n is $f = 0$.

Proof. (a) $\Rightarrow$ (b). Suppose that $\{f_n\}_{n\in\mathbb{N}}$ is complete, and assume $f \in H$ satisfies $\langle f, f_n\rangle = 0$ for every n. Choose any $\varepsilon > 0$. Then by definition of completeness, we can find an $N > 0$ and coefficients $c_1, \dots, c_N$ so that the element $h = \sum_{n=1}^{N} c_n f_n$ lies within ε of f, i.e., $\|f - h\| < \varepsilon$. Note that

$$\langle f, h\rangle \;=\; \Big\langle f, \sum_{n=1}^{N} c_n f_n \Big\rangle \;=\; \sum_{n=1}^{N} \bar{c}_n \,\langle f, f_n\rangle \;=\; 0.$$

Therefore, by Cauchy–Schwarz,

$$\|f\|^2 \;=\; \langle f, f\rangle \;=\; \langle f, f-h\rangle \;+\; \langle f, h\rangle \;=\; \langle f, f-h\rangle \;\le\; \|f\|\,\|f-h\| \;\le\; \varepsilon\,\|f\|.$$

Hence $\|f\| \le \varepsilon$. Since this is true for every $\varepsilon > 0$, it follows that $\|f\| = 0$, and therefore that $f = 0$.

Here is another wording of the same proof. Begin as before with a complete set $\{f_n\}_{n\in\mathbb{N}}$ and an element $f \in H$ satisfying $\langle f, f_n\rangle = 0$ for every n. Then $\langle f, h\rangle = 0$ for every finite linear combination $h = \sum_{n=1}^{N} c_n f_n$, i.e., for every $h \in \operatorname{span}\{f_n\}_{n\in\mathbb{N}}$. Since f is an element of H, for each $\varepsilon = 1/k$ we can find an element $h_k \in \operatorname{span}\{f_n\}_{n\in\mathbb{N}}$ such that $\|f - h_k\| < 1/k$. Then $f = \lim_{k\to\infty} h_k$, so

$$\|f\|^2 \;=\; \langle f, f\rangle \;=\; \lim_{k\to\infty} \langle f, h_k\rangle \;=\; \lim_{k\to\infty} 0 \;=\; 0.$$

(b) $\Rightarrow$ (a). We will prove the contrapositive statement. Suppose that statement (a) is false, i.e., that $\{f_n\}_{n\in\mathbb{N}}$ is not complete. Then the closed span $\overline{\operatorname{span}}\{f_n\}_{n\in\mathbb{N}}$ cannot be all of H. Therefore, there must be some element $f \in H$ that is *not* in $\overline{\operatorname{span}}\{f_n\}_{n\in\mathbb{N}}$.

A basic Hilbert space fact is that we can define the *orthogonal projection* of H onto a closed subspace of H. Here, $S = \overline{\operatorname{span}}\{f_n\}_{n\in\mathbb{N}}$ is the closed subspace, and the orthogonal projection onto S is the unique mapping $P \colon H \to H$ satisfying

(a) $Ph \in S$ for every $h \in H$, and $Ph = h$ if $h \in S$ (and therefore $P^2 = P$), and

(b) $\langle Pg, h\rangle = \langle g, Ph\rangle$ for every $g, h \in H$ (i.e., P is *self-adjoint*).

Now let $g = f - Pf$. If $g = 0$ then $f = Pf \in S$, contradicting the fact that $f \notin S$. Hence $g \ne 0$. Since $f_n \in S$, we have $Pf_n = f_n$. Therefore,

$$\langle g, f_n\rangle \;=\; \langle f, f_n\rangle - \langle Pf, f_n\rangle \;=\; \langle f, f_n\rangle - \langle f, Pf_n\rangle \;=\; \langle f, f_n\rangle - \langle f, f_n\rangle \;=\; 0.$$

Thus the nonzero element g is orthogonal to every element f_n. Hence statement (b) is false. □

4. Operators, Functionals, and the Dual Space

An *operator* is a function T that maps one vector space X into another vector space Y (often, an "operator" is required to be a linear mapping). We write $T : X \to Y$ to mean that T is a function with domain X and range contained in Y.

Definition 4.1 (Notation for Operators)**.** Let X, Y be normed linear spaces, and suppose that $T\colon X \to Y$. We write either $T(x)$ or Tx to denote the image of an element $x \in X$.

(a) T is *linear* if $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ for every $x, y \in X$ and $\alpha, \beta \in \mathbb{C}$.

(b) T is *injective* if $T(x) = T(y)$ implies $x = y$.

(c) The *kernel* or *nullspace* of T is $\ker(T) = \{x \in X : T(x) = 0\}$.

(c) The *range* of T is $\operatorname{range}(T) = \{T(x) : x \in X\}$.

(d) The *rank* of T is the vector space dimension of its range, i.e., $\operatorname{rank}(T) = \dim(\operatorname{range}(T))$. In particular, T is *finite-rank* if $\operatorname{range}(T)$ is finite-dimensional.

(d) T is *surjective* if $\operatorname{range}(T) = Y$.

(e) T is a *bijection* if it is both injective and surjective.

Definition 4.2 (Continuous and Bounded Operators)**.** Assume that X and Y are normed linear spaces, and that $T : X \to Y$ is an operator. We use $\|\cdot\|_X$ to denote the norm on the space X, and $\|\cdot\|_Y$ to denote the norm on Y.

(a) T is *continuous* if $f_n \to f$ implies $T(f_n) \to T(f)$, i.e., if

$$\lim_{n\to\infty} \|f - f_n\|_X = 0 \qquad \Longrightarrow \qquad \lim_{n\to\infty} \|T(f) - T(f_n)\|_Y = 0.$$

(b) T is *bounded* if there is a finite real number C so that $\|T(f)\|_Y \le C\,\|f\|_X$ for every $f \in X$. The smallest such C is called the *operator norm* of T and is denoted by $\|T\|$. That is, $\|T\|$ is the smallest number such that

$$\forall\, f \in X, \quad \|Tf\|_Y \;\le\; \|T\|\,\|f\|_X.$$

We have the equalities

$$\|T\| \;=\; \sup_{f\neq 0} \frac{\|T(f)\|_Y}{\|f\|_X} \;=\; \sup_{\|f\|_X\le 1} \|T(f)\|_Y \;=\; \sup_{\|f\|_X=1} \|T(f)\|_Y.$$

(c) For linear operators T we have the fundamental fact that continuity and boundedness are equivalent, i.e.,

$$T \text{ is continuous} \iff T \text{ is bounded.}$$

Definition 4.3 (Linear Functionals)**.** A *functional* is an operator whose range is the space of scalars $\mathbb{C}$. In other words, a functional is a mapping $T\colon X \to \mathbb{C}$, so $T(f)$ is a *number* for every $f \in X$. If T is a *linear* functional, then we know that it is continuous if and only if it is bounded. Hence the terms "continuous linear functional" and "bounded linear functional" are interchangeable. Also note that since $T(f)$ is a scalar, the norm of $T(f)$ is just its absolute value $|T(f)|$. Hence the norm of a linear functional T is given by the formulas

$$\|T\| \;=\; \sup_{f\neq 0} \frac{|T(f)|}{\|f\|_X} \;=\; \sup_{\|f\|_X\leq 1} |T(f)| \;=\; \sup_{\|f\|_X=1} |T(f)|,$$

and $\|T\|$ is the smallest number such that

$$\forall\, f \in X, \quad |T(f)| \;\leq\; \|T\|\,\|f\|_X.$$

Example 4.4 (The Delta Functional)**.** We can define a linear functional δ on $C_0(\mathbb{R})$ by the formula

$$\delta(f) \;=\; f(0).$$

It is a functional because $\delta(f) = f(0)$ is a number for every f. It is easy to see that δ is linear.

Let's show that δ is a continuous linear functional. To do this, we show that δ is bounded. Recall that $C_0(\mathbb{R})$ is a Banach space under the L^∞ norm. So, suppose that $f \in C_0(\mathbb{R})$ and that $\|f\|_\infty = 1$. Because f is a *continuous* function its essential supremum is an actual supremum, i.e., $\|f\|_\infty = \sup |f(x)|$. Hence,

$$|f(0)| \;\leq\; \sup_{x\in\mathbb{R}} |f(x)| \;=\; \|f\|_\infty \;=\; 1.$$

Therefore,

$$\|\delta\| \;=\; \sup_{\|f\|_\infty=1} |\delta(f)| \;=\; \sup_{\|f\|_\infty=1} |f(0)| \;\leq\; \sup_{\|f\|_\infty=1} 1 \;=\; 1.$$

Thus δ is a bounded linear functional on $C_0(\mathbb{R})$, and its norm is at most 1. In fact we have $\|\delta\| = 1$, since we can easily find a function $f \in C_0(\mathbb{R})$ such that $f(0) = 1$ and $\|f\|_\infty = 1$.

Example 4.5 (Linear Functionals by Integration)**.** Consider the case $X = L^p(\mathbb{R})$. Recall *Hölder's Inequality*: if $f \in L^p(\mathbb{R})$ and $g \in L^{p'}(\mathbb{R})$ where p' is such that $\frac{1}{p} + \frac{1}{p'} = 1$, then $fg \in L^1(\mathbb{R})$, and

$$\|fg\|_1 \;\leq\; \|f\|_p\,\|g\|_{p'}.$$

For example, if $p = 1$ then $p' = \infty$, if $p = 2$ then $p' = 2$, and if $p = \infty$ then $p' = 1$.

Now let $g \in L^{p'}(\mathbb{R})$ be a fixed function. Then because of Hölder's inequality, we can define a functional $T_g : L^p(\mathbb{R}) \to \mathbb{C}$ by the formula

$$T_g(f) \;=\; \int_{\mathbb{R}} f(x)\,\overline{g(x)}\,dx.$$

Hölder's inequality tells us that

$$|T_g(f)| \;\leq\; \int_{\mathbb{R}} |f(x)|\,|g(x)|\,dx \;=\; \|fg\|_1 \;\leq\; \|f\|_p\,\|g\|_{p'}.$$

Therefore T_g is bounded, with

$$\|T_g\| \;=\; \sup_{\|f\|_p=1} |T_g(f)| \;\le\; \|g\|_{p'} \;<\; \infty.$$

In fact, we actually have $\|T_g\| = \|g\|_{p'}$. *Each function $g \in L^{p'}(\mathbb{R})$ determines a continuous linear functional T_g on $L^p(\mathbb{R})$.*

Because the formula for T_g looks just like the definition of the inner product on $L^2(\mathbb{R})$, we often use the same notation. That is, we write

$$T_g(f) \;=\; \langle f, g\rangle \;=\; \int_{\mathbb{R}} f(x)\,\overline{g(x)}\,dx. \tag{4.1}$$

However, you should remember that, unlike the L^2 case, the functions f and g are coming from different places: $f \in L^p(\mathbb{R})$ while $g \in L^{p'}(\mathbb{R})$. Thus, the notation $\langle f, g\rangle$ in (4.1) is not a true inner product except in the special case $p = p' = 2$.

Because of this example, we often abuse the inner product notation and use it represent the action of arbitrary linear functionals $T : X \to \mathbb{C}$. That is, we often write $\langle f, T\rangle$ instead of $T(f)$.

Definition 4.6 (Dual Spaces)**.** If X is a normed linear space then its *dual space* is the set of all continuous linear functionals with domain X. The dual space is usually denoted by either X' or X^*. In other words,

$$X^* \;=\; \{T\colon X \to \mathbb{C} \,:\, T \text{ is continuous and linear}\}.$$

The dual space is a Banach space, even if X is only a normed space.

Example 4.7 (Dual of $L^p(\mathbb{R})$)**.** We've already seen that every function $g \in L^{p'}(\mathbb{R})$ determines a continuous linear functional T_g on $L^p(\mathbb{R})$. We usually "identify" g with T_g. In other words, because each g determines a unique T_g, we say that a function $g \in L^{p'}(\mathbb{R})$ "is" a continuous linear functional on $L^p(\mathbb{R})$, and so each $g \in L^{p'}(\mathbb{R})$ "is" an element of the dual space of $L^p(\mathbb{R})$. In this sense, we write

$$L^{p'}(\mathbb{R}) \;\subset\; (L^p(\mathbb{R}))^*, \quad \text{for } 1 \le p \le \infty.$$

In fact, there is more to the story. When $1 \le p < \infty$, Hölder's inequality can be refined to show that *every bounded linear functional on $L^p(\mathbb{R})$ comes from integration against a unique function in $L^{p'}(\mathbb{R})$*. That is, every continuous linear functional $T\colon L^p(\mathbb{R}) \to \mathbb{C}$ is equal to T_g for some unique $g \in L^{p'}(\mathbb{R})$. Thus, we actually have

$$(L^p(\mathbb{R}))^* \;=\; L^{p'}(\mathbb{R}), \quad \text{for } 1 \le p < \infty.$$

This equality fails for $p = \infty$, i.e., $(L^\infty(\mathbb{R}))^* \ne L^1(\mathbb{R})$, because there exist continuous linear functionals on $L^\infty(\mathbb{R})$ that are not given by integration against an L^1 function.

Example 4.8 (Hilbert space duals)**.** Consider the Hilbert space case of the preceding example, i.e., the case $p = 2$. In this case we have $p = p' = 2$, and therefore $(L^2(\mathbb{R}))^* = L^2(\mathbb{R})$. This fact is true for every Hilbert space: i.e., if H is a Hilbert space then $H^* = H$. Only Hilbert spaces are self-dual.

Example 4.9 (Reflexivity)**.** Note the following: since $(p')' = p$, we have

$$1 < p < \infty \qquad \Longrightarrow \qquad (L^p(\mathbb{R}))^* = L^{p'}(\mathbb{R}) \text{ and } (L^{p'}(\mathbb{R}))^* = L^p(\mathbb{R}).$$

However,

$$p = 1 \qquad \Longrightarrow \qquad (L^1(\mathbb{R}))^* = L^\infty(\mathbb{R}) \text{ and } (L^\infty(\mathbb{R}))^* \neq L^1(\mathbb{R}).$$

Therefore $L^p(\mathbb{R})$ for $1 < p < \infty$ is said to be *reflexive*, while $L^1(\mathbb{R})$ is not. A general Banach space X is reflexive if $(X^*)^* = X$.

5. Adjoints

Definition 5.1 (Adjoints)**.** Let X and Y be Banach spaces, and let $S\colon X \to Y$ be a bounded linear operator. Fix an element $g^* \in Y^*$, and define a functional $f^*\colon X \to \mathbb{F}$ by

$$\langle f, f^* \rangle \;=\; \langle Sf, g^* \rangle, \qquad f \in X.$$

Then f^* is linear since S and g^* are linear. Further,

$$|\langle f, f^* \rangle| \;=\; |\langle Sf, g^* \rangle| \;\le\; \|Sf\|_Y \, \|g^*\|_{Y^*},$$

so

$$\|f^*\| \;=\; \sup_{\|f\|_X = 1} |\langle f, f^* \rangle| \;\le\; \|g^*\|_{Y^*} \sup_{\|f\|_X=1} \|Sf\|_Y \;=\; \|g^*\|_{Y^*} \, \|S\| \;<\; \infty. \tag{5.1}$$

Hence f^* is bounded, so $f^* \in X^*$. Thus, for each $g^* \in Y^*$ we have defined a functional $f^* \in X^*$. Therefore, we can define an operator $S^*\colon Y^* \to X^*$ by setting $S^*(g^*) = f^*$. This mapping S^* is linear, and by (5.1) we have

$$\|S^*\| \;=\; \sup_{\|g^*\|_{Y^*}=1} \|S^*(g^*)\|_{X^*} \;=\; \sup_{\|g^*\|_{Y^*}=1} \|f^*\|_{X^*} \;\le\; \sup_{\|g^*\|_{Y^*}=1} \|g^*\|_{Y^*} \, \|S\| \;=\; \|S\|.$$

In fact, it is true that $\|S^*\| = \|S\|$. This operator S^* is called the *adjoint* of S.

The fundamental property of the adjoint can be restated as follows: $S^* : Y^* \to X^*$ is the unique mapping which satisfies

$$\forall\, f \in X, \quad \forall\, g^* \in Y^*, \quad \langle Sf, g^* \rangle \;=\; \langle f, S^*(g^*) \rangle. \tag{5.2}$$

Remark 5.2. Assume that $X = H$ and $Y = K$ are Hilbert spaces. Then $H = H^*$ and $K = K^*$. Therefore, if $S\colon H \to K$ then its adjoint S^* maps K back to H. Moreover, by (5.2), the adjoint $S^*\colon K \to H$ is the unique mapping which satisfies

$$\forall\, f \in H, \quad \forall\, g \in K, \quad \langle Sf, g \rangle \;=\; \langle f, S^* g \rangle. \tag{5.3}$$

We make the following further definitions specifically for operators $S\colon H \to H$ which map a Hilbert space H into itself.

Definition 5.3. Let H be a Hilbert space.

(a) $S\colon H \to H$ is *self-adjoint* if $S = S^*$. By (5.3),

$$S \text{ is self-adjoint} \quad \iff \quad \forall\, f, g \in H, \quad \langle Sf, g\rangle \;=\; \langle f, Sg\rangle.$$

It can be shown that if S is self-adjoint, then $\langle Sf, f\rangle$ is real for every f, and

$$\|S\| \;=\; \sup_{\|f\|=1} |\langle Sf, f\rangle|.$$

(b) $S\colon H \to H$ is *positive*, denoted $S \geq 0$, if $\langle Sf, f\rangle$ is real and $\langle Sf, f\rangle \geq 0$ for every $f \in H$. It can be shown that a positive operator on a complex Hilbert space is self-adjoint.

(c) $S\colon H \to H$ is *positive definite*, denoted $S > 0$, if $\langle Sf, f\rangle$ is real and $\langle Sf, f\rangle > 0$ for every $f \neq 0$.

(d) If S, $T\colon H \to H$, then we write $S \geq T$ if $S - T \geq 0$. Similarly, $S > T$ if $S - T > 0$.

As an example, consider the finite-dimensional Hilbert spaces $H = \mathbb{C}^n$ and $K = \mathbb{C}^m$. A linear operator $S\colon \mathbb{C}^n \to \mathbb{C}^m$ is simply an $m \times n$ matrix with complex entries, and its adjoint $S^*\colon \mathbb{C}^m \to \mathbb{C}^n$ is simply the $n \times m$ matrix given by the conjugate transpose of S. In this case, the matrix S^* is often called the *Hermitian* of the matrix S, i.e., $S^* = S^{\mathrm{H}} = \overline{S}^{\mathrm{T}}$.

6. Compact Operators

In this section we consider compact operators. We will focus on operators on Hilbert spaces, although compact operators on Banach spaces are equally important.

Definition 6.1 (Compact Operator)**.** Let H, K be Hilbert spaces. A linear operator $T\colon H \to K$ is *compact* if for every sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ in the unit ball of H (i.e., $\|f_n\|_2 \leq 1$ for all n), there is a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ such that $\{Tf_{n_k}\}_{k\in\mathbb{N}}$ converges in K.

Theorem 6.2 (Properties of Compact Operators)**.** Let H, K be Hilbert spaces, and let $T\colon H \to K$ be linear.

(a) If T is compact, then it is bounded.

(b) If T is bounded and has finite rank, then T is compact.

(c) If $T_n\colon H \to K$ are compact and $T_n \to T$ in operator norm, i.e., $\lim_{n\to\infty} \|T - T_n\| = 0$, then T is compact.

(d) T is compact if and only if there exist bounded, finite-rank operators $T_n\colon H \to K$ such that $T_n \to T$ in operator norm.

Theorem 6.3 (Compositions and Compact Operators)**.** Let H_1, H_2, H_3 be Hilbert spaces.

(a) If $A\colon H_1 \to H_2$ is bounded and $T\colon H_2 \to H_3$ is compact, then $TA\colon H_1 \to H_3$ is compact.

(b) If $T\colon H_1 \to H_2$ is compact and $A\colon H_2 \to H_3$ is bounded, then $AT\colon H_1 \to H_3$ is compact.

The following is a fundamental result for operators that are both compact and self-adjoint.

Theorem 6.4 (Spectral Theorem for Compact Self-Adjoint Operators)**.** Let H be a Hilbert space, and let $T\colon H \to H$ be compact and self-adjoint. Then there exist nonzero real numbers $\{\lambda_n\}_{n\in J}$, either finitely many or $\lambda_n \to 0$ if infinitely many, and an orthonormal basis $\{e_n\}_{n\in J}$ of $\overline{\mathrm{range}(T)}$, such that

$$Tf \;=\; \sum_{n\in J} \lambda_n \, \langle f, e_n\rangle \, e_n, \qquad f \in H.$$

Each λ_n is an eigenvalue of T, and each e_n is a corresponding eigenvector. If T is a positive operator, then $\lambda_n > 0$ for each $n \in J$.

Definition 6.5 (Singular Numbers)**.** Let $T\colon H \to H$ be an arbitrary compact operator on a Hilbert space H. Then T^*T and TT^* are both compact and self-adjoint, and in fact are positive operators. In particular, by the Spectral Theorem, there exists an orthonormal sequence $\{e_n\}_{n\in J}$ and corresponding positive real numbers $\{\mu_n\}_{n\in J}$ such that

$$T^*Tf \;=\; \sum_{n\in J} \mu_n \, \langle f, e_n\rangle \, e_n, \qquad f \in H.$$

(a) The *singular numbers* of T are $s_n = \mu_n^{1/2}$. The vectors e_n are the *singular vectors* of T.

(b) Given $1 \le p < \infty$, the *Schatten class* $\mathcal{I}_p$ consists of all compact operators $T\colon H \to H$ such that

$$\|T\|_{\mathcal{I}_p} \;=\; \biggl(\sum_{n\in J} s_n^p\biggr)^{1/p} \;<\; \infty.$$

(c) We say that T is a *Hilbert–Schmidt operator* if $T \in \mathcal{I}_2$, i.e., if $\sum_{n\in J} s_n^2 < \infty$.

(d) We say that T is a *trace-class operator* if $T \in \mathcal{I}_1$, i.e., if $\sum_{n\in J} s_n < \infty$.

In particular, if T is compact and self-adjoint, then the singular numbers are simply the absolute values of the eigenvalues, i.e., $s_n = |\lambda_n|$. By the Spectral Theorem, we must have $s_n \to 0$, and the definition of the Schatten classes simply quantifies the rate of convergence to zero.

As an example, consider an *integral operator* on $L^2(\mathbb{R})$, formally defined by the rule

$$Lf(x) \;=\; \int_{\mathbb{R}} k(x,y)\, f(y)\, dy, \tag{6.1}$$

where k is a fixed measurable function on $\mathbb{R}^2$, called the *kernel* of L (not to be confused with nullspace of L). The following result summarizes some properties of integral operators

Theorem 6.6 (Integral Operators)**.** Let $k\colon \mathbb{R}^2 \to \mathbb{R}$ be measurable, and let L be formally defined by (6.1).

(a) If $k \in L^2(\mathbb{R}^2)$ then L is a bounded and compact mapping of $L^2(\mathbb{R})$ into itself, and $\|L\| \le \|k\|_2$.

(b) L is a Hilbert–Schmidt operator if and only if $k \in L^2(\mathbb{R}^2)$. In this case,

$$\|k\|_2 \;=\; \|L\|_{\mathcal{I}_2} \;=\; \left(\sum_{n\in J} s_n^2\right)^{1/2}.$$

(c) The adjoint L^* is the integral operator whose kernel is $\overline{k(y,x)}$. In particular, L is self-adjoint if and only if $k(x,y) = \overline{k(y,x)}$.