FUNCTIONAL ANALYSIS[1]

Douglas N. Arnold[2]

References:

John B. Conway, *A Course in Functional Analysis*, 2nd Edition, Springer-Verlag, 1990.

Gert K. Pedersen, *Analysis Now*, Springer-Verlag, 1989.

Walter Rudin, *Functional Analysis*, 2nd Edition, McGraw Hill, 1991.

Robert J. Zimmer, *Essential Results of Functional Analysis*, University of Chicago Press, 1990.

CONTENTS

[1]These lecture notes were prepared for the instructor's personal use in teaching a half-semester course on functional analysis at the beginning graduate level at Penn State, in Spring 1997. They are certainly not meant to replace a good text on the subject, such as those listed on this page.

[2]Department of Mathematics, Penn State University, University Park, PA 16802.
Web: http://www.math.psu.edu/dna/.

I. Vector spaces and their topology

Basic definitions: (1) Norm and seminorm on vector spaces (real or complex). A norm defines a Hausdorff topology on a vector space in which the algebraic operations are continuous, resulting in a *normed linear space*. If it is complete it is called a Banach space.

(2) Inner product and semi-inner-product. In the real case an inner product is a positive definite, symmetric bilinear form on $X \times X \to \mathbb{R}$. In the complex case it is positive definite, Hermitian symmetric, sesquilinear form $X \times X \to \mathbb{C}$. An (semi) inner product gives rise to a (semi)norm. An inner product space is thus a special case of a normed linear space. A complete inner product space is a Hilbert space, a special case of a Banach space.

The polarization identity expresses the norm of an inner product space in terms of the inner product. For real inner product spaces it is

$$(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2).$$

For complex spaces it is

$$(x, y) = \frac{1}{4}(\|x + y\|^2 + i\|x + iy\|^2 - \|x - y\|^2 - i\|x - iy\|^2).$$

In inner product spaces we also have the parallelogram law:

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

This gives a criterion for a normed space to be an inner product space. Any norm coming from an inner product satisfies the parallelogram law and, conversely, if a norm satisfies the parallelogram law, we can show (but not so easily) that the polarization identity defines an inner product, which gives rise to the norm.

(3) A *topological vector space* is a vector space endowed with a Hausdorff topology such that the algebraic operations are continuous. Note that we can extend the notion of Cauchy sequence, and therefore of completeness, to a TVS: a sequence x_n in a TVS is Cauchy if for every neighborhood U of 0 there exists N such that $x_m - x_n \in U$ for all $m, n \geq N$.

A normed linear space is a TVS, but there is another, more general operation involving norms which endows a vector space with a topology. Let X be a vector space and suppose that a family $\{\|\cdot\|_\alpha\}_{\alpha\in\mathcal{A}}$ of seminorms on X is given which are *sufficient* in the sense that $\bigcap_\alpha\{\|x\|_\alpha = 0\} = 0$. Then the topology generated by the sets $\{\|x\|_\alpha < r\}$, $\alpha \in \mathcal{A}$, $r > 0$, makes X a TVS. A sequence (or net) x_n converges to x iff $\|x_n - x\|_\alpha \to 0$ for all α. Note that, a fortiori, $|\,\|x_n\|_\alpha - \|x\|_\alpha| \to 0$, showing that each seminorm is continuous.

If the number of seminorms is finite, we may add them to get a norm generating the same topology. If the number is countable, we may define a metric

$$d(x, y) = \sum_n 2^{-n} \frac{\|x - y\|_n}{1 + \|x - y\|_n},$$

so the topology is metrizable.

Examples: (0) On $\mathbb{R}^n$ or $\mathbb{C}^n$ we may put the l_p norm, $1 \le p \le \infty$, or the weighted l_p norm with some arbitrary positive weight. All of these norms are equivalent (indeed all norms on a finite dimensional space are equivalent), and generate the same Banach topology. Only for $p = 2$ is it a Hilbert space.

(2) If Ω is a subset of $\mathbb{R}^n$ (or, more generally, any Hausdorff space) we may define the space $C_b(\Omega)$ of bounded continuous functions with the supremum norm. It is a Banach space. If X is compact this is simply the space $C(\Omega)$ of continuous functions on Ω.

(3) For simplicity, consider the unit interval, and define $C^n([0,1])$ and $C^{n,\alpha}([0,1])$, $n \in \mathbb{N}$, $\alpha \in (0,1]$. Both are Banach spaces with the natural norms. $C^{0,1}$ is the space of Lipschitz functions. $C([0,1]) \subset C^{0,\alpha} \subset C^{0,\beta} \subset C^1([0,1])$ if $0 < \alpha \le \beta \le 1$.

(4) For $1 \le p < \infty$ and Ω an open or closed subspace of $\mathbb{R}^n$ (or, more generally, a σ-finite measure space), we have the space $L^p(\Omega)$ of equivalence classes of measurable p-th power integrable functions (with equivalence being equality off a set of measure zero), and for $p = \infty$ equivalence classes of essentially bounded functions (bounded after modification on a set of measure zero). For $1 < p < \infty$ the triangle inequality is not obvious, it is Minkowski's inequality. Since we modded out the functions with L^p-seminorm zero, this is a normed linear space, and the Riesz-Fischer theorem asserts that it is a Banach space. L^2 is a Hilbert space. If $\operatorname{meas}(\Omega) < \infty$, then $L^p(\Omega) \subset L^q(\Omega)$ if $1 \le q \le p \le \infty$.

(5) The sequence space l_p, $1 \le p \le \infty$ is an example of (4) in the case where the measure space is $\mathbb{N}$ with the counting measure. Each is a Banach space. l_2 is a Hilbert space. $l_p \subset l_q$ if $1 \le p \le q \le \infty$ (note the inequality is reversed from the previous example). The subspace c_0 of sequences tending to 0 is a closed subspace of l_∞.

(6) If Ω is an open set in $\mathbb{R}^n$ (or any Hausdorff space), we can equip $C(\Omega)$ with the norms $f \mapsto |f(x)|$ indexed by $x \in \Omega$. This makes it a TVS, with the topology being that of pointwise convergence. It is not complete (pointwise limit of continuous functions may not be continuous).

(7) If Ω is an open set in $\mathbb{R}^n$ we can equip $C(\Omega)$ with the norms $f \mapsto \|f\|_{L^\infty(K)}$ indexed by compact subsets of Ω, thus defining the topology of uniform convergence on compact subsets. We get the same toplogy by using only the countably many compact sets

$$K_n = \{x \in \Omega : |x| \le n,\ \operatorname{dist}(x, \partial\Omega) \ge 1/n\}.$$

The topology is complete.

(8) In the previous example, in the case Ω is a region in $\mathbb{C}$, and we take complex-valued functions, we may consider the subspace $H(\Omega)$ of holomorbarphic functions. By Weierstrass's theorem it is a closed subspace, hence itself a complete TVS.

(9) If $f, g \in L^1(I)$, $I = (0,1)$ and

$$\int_0^1 f(x)\phi(x)\,dx = -\int_0^1 g(x)\phi'(x)\,dx,$$

for all infinitely differentiable ϕ with support contained in I (so ϕ is identically zero near 0 and 1), then we say that f is weakly differentiable and that $f' = g$. We can then define the *Sobolev space* $W^1_p(I) = \{\, f \in L^p(I) \,:\, f' \in L^p(I)\,\}$, with the norm

$$\|f\|_{W^1_p(I)} = \left(\int_0^1 |f(x)|^p \, dx + \int_0^1 |f'(x)|^p \, dx \right)^{1/p}.$$

This is a larger space than $C^1(\bar{I})$, but still incorporates first order differentiability of f. The case $p = 2$ is particularly useful, because it allows us to deal with differentiability in a Hilbert space context. Sobolev spaces can be extended to measure any degree of differentiability (even fractional), and can be defined on arbitrary domains in $\mathbb{R}^n$.

Subspaces and quotient spaces.

If X is a vector space and S a subspace, we may define the vector space X/S of cosets. If X is normed, we may define

$$\|u\|_{X/S} = \inf_{x \in u} \|x\|_X, \text{ or equivalently } \|\bar{x}\|_{X/S} = \inf_{s \in S} \|x - s\|_X.$$

This is a seminorm, and is a norm iff S is closed.

Theorem. *If X is a Banach space and S is a closed subspace then S is a Banach space and X/S is a Banach space.*

Sketch. Suppose x_n is a sequence of elements of X for which the cosets $\bar{x}_n$ are Cauchy. We can take a subsequence with $\|\bar{x}_n - \bar{x}_{n+1}\|_{X/S} \le 2^{-n-1}$, $n = 1, 2, \ldots$. Set $s_1 = 0$, define $s_2 \in S$ such that $\|x_1 - (x_2 + s_2)\|_X \le 1/2$, define $s_3 \in S$ such that $\|(x_2 + s_2) - (x_3 + s_3)\|_X \le 1/4, \ldots$. Then $\{x_n + s_n\}$ is Cauchy in X … $\square$

A converse is true as well (and easily proved).

Theorem. *If X is a normed linear space and S is a closed subspace such that S is a Banach space and X/S is a Banach space, then X is a Banach space.*

Finite dimensional subspaces are always closed (they're complete). More generally:

Theorem. *If S is a closed subspace of a Banach space and V is a finite dimensional subspace, then $S + V$ is closed.*

Sketch. We easily pass to the case V is one-dimensional and $V \cap S = 0$. We then have that $S + V$ is algebraically a direct sum and it is enough to show that the projections $S + V \to S$ and $S + V \to V$ are continuous (since then a Cauchy sequence in $S + V$ will lead to a Cauchy sequence in each of the closed subspaces, and so to a convergent subsequence). Now the projection $\pi : X \to X/S$ restricts to a 1-1 map on V so an isomorphism of V onto its image $\bar{V}$. Let $\mu : \bar{V} \to V$ be the continuous inverse. Since $\pi(S + V) \subset \bar{V}$, we may form the composition $\mu \circ \pi|_{S+V} : S + V \to V$ and it is continuous. But it is just the projection onto V. The projection onto S is $id - \mu \circ \pi$, so it is also continuous. $\square$

Note. The sum of closed subspaces of a Banach space need not be closed. For a counterexample (in a separable Hilbert space), let S_1 be the vector space of all real sequences $(x_n)_{n=1}^\infty$ for which $x_n = 0$ if n is odd, and S_2 be the sequences for which $x_{2n} = nx_{2n-1}$, $n = 1, 2, \ldots$. Clearly $X_1 = l_2 \cap S_1$ and $X_2 = l_2 \cap S_2$ are closed subspaces of l_2, the space of square integrable sequences (they are defined as the intersection of the null spaces of continuous linear functionals). Obviously every sequence can be written in a unique way as sum of elements of S_1 and S_2:

$$(x_1, x_2, \ldots) = (0, x_2 - x_1, 0, x_4 - 2x_3, 0, x_6 - 3x_5, \ldots) + (x_1, x_1, x_3, 2x_3, x_5, 3x_5, \ldots).$$

If a sequence has all but finitely many terms zero, so do the two summands. Thus all such sequences belong to $X_1 + X_2$, showing that $X_1 + X_2$ is dense in l_2. Now consider the sequence $(1, 0, 1/2, 0, 1/3, \ldots) \in l_2$. Its only decomposition as elements of $\mathcal{S}_1$ and S_2 is

$$(1, 0, 1/2, 0, 1/3, 0, \ldots) = (0, -1, 0, -1, 0, -1, \ldots) + (1, 1, 1/2, 1, 1/3, 1, \ldots),$$

and so it does not belong to $X_1 + X_2$. Thus $X_1 + X_2$ is not closed in l_2.

Basic properties of Hilbert spaces.

An essential property of Hilbert space is that the distance of a point to a closed convex set is alway attained.

Projection Theorem. *Let X be a Hilbert space, K a closed convex subset, and $x \in X$. Then there exists a unique $\bar{x} \in K$ such that*

$$\|x - \bar{x}\| = \inf_{y \in K} \|x - y\|.$$

Proof. Translating, we may assume that $x = 0$, and so we must show that there is a unique element of K of minimal norm. Let $d = \inf_{y \in K} \|y\|$ and chose $x_n \in K$ with $\|x_n\| \to d$. Then the parallelogram law gives

$$\left\| \frac{x_n - x_m}{2} \right\|^2 = \frac{1}{2}\|x_n\|^2 + \frac{1}{2}\|x_m\|^2 - \left\| \frac{x_n + x_m}{2} \right\|^2 \le \frac{1}{2}\|x_n\|^2 + \frac{1}{2}\|x_m\|^2 - d^2,$$

where we have used convexity to infer that $(x_n + x_m)/2 \in K$. Thus x_n is a Cauchy sequence and so has a limit $\bar{x}$, which must belong to K, since K is closed. Since the norm is continuous, $\|\bar{x}\| = \lim_n \|x_n\| = d$.

For uniqueness, note that if $\|\bar{x}\| = \|\tilde{x}\| = d$, then $\|(\bar{x} + \tilde{x})/2\| = d$ and the parallelogram law gives

$$\|\bar{x} - \tilde{x}\|^2 = 2\|\bar{x}\|^2 + 2\|\tilde{x}\|^2 - \|\bar{x} + \tilde{x}\|^2 = 2d^2 + 2d^2 - 4d^2 = 0. \quad \square$$

The unique nearest element to x in K is often denoted $P_K x$, and referred to as the projection of x onto K. It satisfies $P_K \circ P_K = P_K$, the definition of a projection. This terminology is especially used when K is a closed linear subspace of X, in which case P_K is a linear projection operator.

Projection and orthogonality. If S is any subset of a Hilbert space X, let

$$S^\perp = \{\, x \in X \,:\, \langle x, s\rangle = 0 \text{ for all } s \in S \,\}.$$

Then $S^\perp$ is a closed subspace of X. We obviously have $S \cap S^\perp = 0$ and $S \subset S^{\perp\perp}$.

Claim: If S is a closed subspace of X, $x \in X$, and $P_S x$ the projection of x onto S, then $x - P_S x \in S^\perp$. Indeed, if $s \in S$ is arbitrary and $t \in \mathbb{R}$, then

$$\|x - P_S x\|^2 \le \|x - P_S x - ts\|^2 = \|x - P_S x\|^2 - 2t(x - P_S x, s) + t^2\|s\|^2,$$

so the quadratic polynomial on the right hand side has a minimum at $t = 0$. Setting the derivative there to 0 gives $(x - P_S x, s) = 0$.

Thus we can write any $x \in X$ as $s + s^\perp$ with $s \in S$ and $s^\perp \in S^\perp$ (namely $s = P_S x$, $s^\perp = x - P_S x$). Such a decomposition is certainly unique (if $\bar{s} + \bar{s}^\perp$ were another one we would have $s - \bar{s} = \bar{s}^\perp - s^\perp \in S \cap S^\perp = 0$.) We clearly have $\|x\|^2 = \|s\|^2 + \|s^\perp\|^2$.

An immediate corollary is that $S^{\perp\perp} = S$ for S a closed subspace, since if $x \in S^{\perp\perp}$ we can write it as $s + s^\perp$, whence $s^\perp \in S^\perp \cap S^{\perp\perp} = 0$, i.e., $x \in S$. We thus see that the decomposition

$$x = (I - P_S)x + P_S x$$

is the (unique) decomposition of x into elements of $S^\perp$ and $S^{\perp\perp}$. Thus $P_{S^\perp} = I - P_S$. For any subset S of X, $S^{\perp\perp}$ is the smallest closed subspace containing S.

Orthonormal sets and bases in Hilbert space.

Let $e_1, e_2, \ldots, e_N$ be orthonormal elements of a Hilbert space X, and let S be their span. Then $\sum_n \langle x, e_n\rangle e_n \in S$ and $x - \sum_n \langle x, e_n\rangle e_n \perp S$, so $\sum_n \langle x, e_n\rangle e_n = P_S x$. But $\|\sum_n \langle x, e_n\rangle e_n\|^2 = \sum_{n=1}^N \langle x, e_n\rangle^2$, so

$$\sum_{n=1}^N \langle x, e_n\rangle^2 \le \|x\|^2$$

(Bessel's inequality). Now let $\mathcal{E}$ be an orthonormal set of arbitrary cardinality. It follows from Bessel's inequality that for $\epsilon > 0$ and $x \in X$, $\{\, e \in \mathcal{E} \,:\, \langle x, e\rangle \ge \epsilon \,\}$ is finite, and hence that $\{\, e \in \mathcal{E} \,:\, \langle x, e\rangle > 0 \,\}$ is countable. We can thus extend Bessel's inequality to an arbitrary orthonormal set:

$$\sum_{e \in \mathcal{E}} \langle x, e\rangle^2 \le \|x\|^2,$$

where the sum is just a countable sum of positive terms.

It is useful to extend the notion of sums over sets of arbitrary cardinality. If $\mathcal{E}$ is an arbitary set and $f : \mathcal{E} \to X$ a function mapping into a Hilbert space (or any normed linear space or even TVS), we say

$$(\star) \qquad \sum_{e \in \mathcal{E}} f(e) = x$$

if the net $\sum_{e\in\mathcal{F}} f(e)$, indexed by the *finite* subsets $\mathcal{F}$ of $\mathcal{E}$, converges to x. In other words, $(\star)$ holds if, for any neighborhood U of the origin, there is a finite set $\mathcal{F}_0 \subset \mathcal{E}$ such that $x - \sum_{e\in\mathcal{F}} f(e) \in U$ whenever $\mathcal{F}$ is a finite subset of $\mathcal{E}$ containing $\mathcal{F}_0$. In the case $\mathcal{E} = \mathbb{N}$, this is equivalent to absolute convergence of a series. Note that if $\sum_{e\in\mathcal{E}} f(e)$ converges, then for all ϵ there is a finite $\mathcal{F}_0$ such that if $\mathcal{F}_1$ and $\mathcal{F}_2$ are finite supersets of $\mathcal{F}_0$, then $\|\sum_{e\in\mathcal{F}_1} f(e) - \sum_{e\in\mathcal{F}_2} f(e)\| \le \epsilon$. It follows easily that each of the sets $\{\, e \in \mathcal{E} \mid \|f(e)\| \ge 1/n \,\}$ is finite, and hence, $f(e) = 0$ for all but countably many $e \in \mathcal{E}$.

Lemma. *If $\mathcal{E}$ is an orthonormal subset of a Hilbert space X and $x \in X$, then*

$$\sum_{e\in\mathcal{E}} \langle x, e\rangle e$$

converges.

Proof. We may order the elements $e_1, e_2, \ldots$ of $\mathcal{E}$ for which $\langle x, e\rangle \ne 0$. Note that

$$\|\sum_{n=1}^{N} \langle x, e_n\rangle e_n\|^2 = \sum_{n=1}^{N} |\langle x, e_n\rangle|^2 \le \|x\|^2.$$

This shows that the partial sums $s_N = \sum_{n=1}^{N} \langle x, e_n\rangle e_n$ form a Cauchy sequence, and so converge to an element $\sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$ of X. As an exercise in applying the definition, we show that $\sum_{e\in\mathcal{E}} \langle x, e\rangle e = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$. Given $\epsilon > 0$ pick N large enough that $\sum_{n=N+1}^{\infty} |\langle x, e_n\rangle|^2 < \epsilon$. If $M > N$ and $\mathcal{F}$ is a finite subset of $\mathcal{E}$ containing $e_1, \ldots, e_N$, then

$$\|\sum_{n=1}^{M} \langle x, e_n\rangle e_n - \sum_{e\in\mathcal{F}} \langle x, e\rangle e\|^2 \le \epsilon.$$

Letting M tend to infinity,

$$\|\sum_{n=1}^{\infty} \langle x, e_n\rangle e_n - \sum_{e\in\mathcal{F}} \langle x, e\rangle e\|^2 \le \epsilon,$$

as required. □

Recall the proof that every vector space has a basis. We consider the set of all linearly independent subsets of the vector space ordered by inclusions, and note that if we have a totally ordered subset of this set, then the union is a linearly independent subset containing all its members. Therefore Zorn's lemma implies that there exists a maximal linearly independent set. It follows directly from the maximality that this set also spans, i.e., is a basis. In an inner product space we can use the same argument to establish the existence of an orthonormal basis.

In fact, while bases exist for all vector spaces, for infinite dimensional spaces they are difficult or impossible to construct and almost never used. Another notion of basis is much more useful, namely one that uses the topology to allow infinite linear combinations. To distinguish ordinary bases from such notions, an ordinary basis is called a Hamel basis.

Here we describe an orthonormal Hilbert space basis. By definition this is a maximal orthonormal set. By Zorn's lemma, any orthonormal set in a Hilbert space can be extended to a basis, and so orthonormal bases exist. If $\mathcal{E}$ is such an orthonormal basis, and $x \in X$, then

$$x = \sum_{e \in \mathcal{E}} \langle x, e \rangle e.$$

Indeed, we know that the sum on the right exists in X and it is easy to check that its inner product with any $e_0 \in \mathcal{E}$ is $\langle x, e_0 \rangle$. Thus $y := x - \sum_{e \in \mathcal{E}} \langle x, e \rangle e$ is orthogonal to $\mathcal{E}$, and if it weren't zero, then we could adjoin $y/\|y\|$ to $\mathcal{E}$ to get a larger orthonormal set.

Thus we've shown that any element x of X can be expressed as $\sum c_e e$ for some $c_e \in \mathbb{R}$, all but countably many of which are 0. It is easily seen that this determines the c_e uniquely, namely $c_e = \langle x, e \rangle$, and that $\|x\|^2 = \sum c_e^2$.

The notion of orthonormal basis allows us to define a Hilbert space dimension, namely the cardinality of any orthonormal basis. To know that this is well defined, we need to check that any two bases have the same cardinality. If one is finite, this is trivial. Otherwise, let $\mathcal{E}$ and $\mathcal{F}$ be two infinite orthonormal bases. For each $0 \neq x \in X$, the inner product $\langle x, e \rangle \neq 0$ for at least one $e \in \mathcal{E}$. Thus

$$\mathcal{F} \subset \bigcup_{e \in \mathcal{E}} \{\, f \in \mathcal{F} \,:\, \langle f, e \rangle \neq 0 \,\},$$

i.e., $\mathcal{F}$ is contained in the union of $\operatorname{card} \mathcal{E}$ countable sets. Therefore $\operatorname{card} \mathcal{F} \le \aleph_0 \operatorname{card} \mathcal{E} = \operatorname{card} \mathcal{E}$.

If $\mathcal{S}$ is any set, we define a particular Hilbert space $l^2(\mathcal{S})$ as the set of functions $c : \mathcal{S} \to \mathbb{R}$ which are zero off a countable set and such that $\sum_{s \in \mathcal{S}} c_s^2 < \infty$. We thus see that via a basis, any Hilbert space can be put into a norm-preserving (and so inner-product-preserving) linear bijection (or Hilbert space isomorphism) with an $l^2(\mathcal{S})$. Thus, up to isomorphism, there is just one Hilbert space for each cardinality. In particular there is only one infinite dimensional separable Hilbert space (up to isometry).

Example: The best known example of an orthonormal basis in an infinite Hilbert space is the set of functions $e_n = \exp(2\pi i n \theta)$ which form a basis for complex-valued $L^2([0,1])$. (They are obviously orthonormal, and they are a maximal orthonormal set by the Weierstrass approximation Theorem. Thus an arbitrary L^2 function has an L^2 convergent Fourier series

$$f(\theta) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{2\pi i n \theta},$$

with $\hat{f}(n) = \langle f, e_n \rangle = \int_0^1 f(\theta) e^{-2\pi i n \theta}\, d\theta$. Thus from the Hilbert space point of view, the theory of Fourier series is rather simple. More difficult analysis comes in when we consider convergence in other topologies (pointwise, uniform, almost everywhere, L^p, C^1, …).

Schauder bases. An orthonormal basis in a Hilbert space is a special example of a Schauder basis. A subset $\mathcal{E}$ of a Banach space X is called a Schauder basis if for every $x \in X$ there is a unique function $c : \mathcal{E} \to \mathbb{R}$ such that $x = \sum_{e \in \mathcal{E}} c_e e$. Schauder constructed a useful Schauder basis for $C([0,1])$, and there is useful Schauder bases in many other separable Banach spaces. In 1973 Per Enflo settled a long-standing open question by proving that there exist separable Banach spaces with no Schauder bases.

II. Linear Operators and Functionals

$B(X,Y)$ = bounded linear operators between normed linear spaces X and Y. A linear operator is bounded iff it is bounded on every ball iff it is bounded on some ball iff it is continuous at every point iff it is continuous at some point.

Theorem. *If X is a normed linear space and Y is a Banach space, then $B(X,Y)$ is a Banach space with the norm*

$$\|T\|_{B(X,Y)} = \sup_{0 \neq x \in X} \frac{\|Tx\|_Y}{\|x\|_X}.$$

Proof. It is easy to check that $B(X,Y)$ is a normed linear space, and the only issue is to show that it is complete.

Suppose that T_n is a Cauchy sequence in $B(X,Y)$. Then for each $x \in X$ $T_n x$ is Cauchy in the complete space Y, so there exists $Tx \in Y$ with $T_n x \to Tx$. Clearly $T : X \to Y$ is linear. Is it bounded? The real sequence $\|T_n\|$ is Cauchy, hence bounded, say $\|T_n\| \leq K$. It follows that $\|T\| \leq K$, and so $T \in B(X,Y)$. To conclude the proof, we need to show that $\|T_n - T\| \to 0$. We have

$$\begin{aligned}\|T_n - T\| = \sup_{\|x\|\leq 1} \|T_n x - Tx\| &= \sup_{\|x\|\leq 1} \lim_{m\to\infty} \|T_n x - T_m x\| \\ &= \sup_{\|x\|\leq 1} \limsup_{m\to\infty} \|T_n x - T_m x\| \leq \limsup_{m\to\infty} \|T_n - T_m\|.\end{aligned}$$

Thus $\limsup_{n\to\infty} \|T_n - T\| = 0$. $\square$

If $T \in B(X,Y)$ and $U \in B(Y,Z)$, then $UT = U \circ T \in B(X,Z)$ and $\|UT\|_{B(X,Z)} \leq \|U\|_{B(Y,Z)} \|T\|_{B(X,Y)}$. In particular, $B(X) := B(X,X)$ is a *Banach algebra*, i.e., it has an additional "multiplication" operation which makes it a non-commutative algebra, and the multiplication is continuous.

The dual space is $X^* := B(X,\mathbb{R})$ (or $B(X,\mathbb{C})$ for complex vector spaces). It is a Banach space (whether X is or not).

The Hahn–Banach Theorem. A key theorem for dealing with dual spaces of normed linear spaces is the Hahn-Banach Theorem. It assures us that the dual space of a nontrivial normed linear space is itself nontrivial. (Note: the norm is important for this. There exist topological vector spaces, e.g., L^p for $0 < p < 1$, with no non-zero continuous linear functionals.)

Hahn-Banach. *If f is a bounded linear functional on a subspace of a normed linear space, then f extends to the whole space with preservation of norm.*

Note that there are virtually no hypotheses beyond linearity and existence of a norm. In fact for some purposes a weaker version is useful. For X a vector space, we say that $p : X \to \mathbb{R}$ is sublinear if $p(x + y) \le p(x) + p(y)$ and $p(\alpha x) = \alpha p(x)$ for $x, y \in X$, $\alpha \ge 0$.

Generalized Hahn-Banach. *Let X be a vector space, $p : X \to \mathbb{R}$ a sublinear functional, S a subspace of X, and $f : S \to \mathbb{R}$ a linear function satisfying $f(x) \le p(x)$ for all $x \in S$, then f can be extended to X so that the same inequality holds for all $x \in X$.*

Sketch. It suffices to extend f to the space spanned by S and one element $x_0 \in X \setminus S$, preserving the inequality, since if we can do that we can complete the proof with Zorn's lemma.

We need to define $f(x_0)$ such that $f(tx_0 + s) \le p(tx_0 + s)$ for all $t \in \mathbb{R}$, $s \in S$. The case $t = 0$ is known and it is easy to use homogeneity to restrict to $t = \pm 1$. Thus we need to find a value $f(x_0) \in \mathbb{R}$ such that

$$f(s) - p(-x_0 + s) \le f(x_0) \le p(x_0 + s) - f(s) \quad \text{for all } s \in S.$$

Now it is easy to check that for any s_1, $s_2 \in S$, $f(s_1) - p(-x_0 + s_1) \le p(x_0 + s_2) - f(s_2)$, and so such an $f(x_0)$ exists. $\square$

Corollary. *If X is a normed linear space and $x \in X$, then there exists $f \in X^*$ of norm 1 such that $f(x) = \|x\|$.*

Corollary. *If X is a normed linear space, S a closed subspace, and $x \in X$, then there exists $f \in X^*$ of norm 1 such that $f(x) = \|\bar{x}\|_{X/S}$.*

Duality. If X and Y are normed linear spaces and $T : X \to Y$, then we get a natural map $T^* : Y^* \to X^*$ by $T^* f(x) = f(Tx)$ for all $f \in Y^*$, $x \in X$. In particular, if $T \in B(X, Y)$, then $T^* \in B(Y^*, X^*)$. In fact, $\|T^*\|_{B(Y^*,X^*)} = \|T\|_{B(X,Y)}$. To prove this, note that $|T^* f(x)| = |f(Tx)| \le \|f\|\|T\|\|x\|$. Therefore $\|T^* f\| \le \|f\|\|T\|$, so T^* is indeed bounded, with $\|T^*\| \le \|T\|$. Also, given any $y \in Y$, we can find $g \in Y^*$ such that $|g(y)| = \|y\|$, $\|g\| = 1$. Applying this with $y = Tx$ ($x \in X$ arbitrary), gives $\|Tx\| = |g(Tx)| = |T^* gx| \le \|T^*\|\|g\|\|x\| = \|T^*\|\|x\|$. This shows that $\|T\| \le \|T^*\|$. Note that if $T \in B(X, Y)$, $U \in B(Y, Z)$, then $(UT)^* = T^* U^*$.

If X is a Banach space and S a subset, let

$$S^a = \{\, f \in X^* \mid f(s) = 0 \quad \forall s \in S \,\}$$

denote the *annihilator* of S. If V is a subset of X^*, we similarly set

$$^aV = \{\, x \in X \mid f(x) = 0 \quad \forall f \in V \,\}.$$

Note the distinction between V^a, which is a subset of X^{**} and ^aV, which is a subset of X. All annihilators are closed subspaces.

It is easy to see that $S \subset T \subset X$ implies that $T^a \subset S^a$, and $V \subset W \subset X^*$ implies that $^aW \subset {}^aV$. Obviously $S \subset {}^a(S^a)$ if $S \subset X$ and $V \subset ({}^aV)^a$ if $V \subset X^*$. The Hahn-Banach theorem implies that $S = {}^a(S^a)$ in case S is a closed subspace of X (but it can happen that $V \subsetneq ({}^aV)^a$ for V a closed subspace of X^*. For $S \subset X$ arbitrary, $^a(S^a)$ is the smallest closed subspace of X containing the subset S, namely the closure of the span of S.

Now suppose that $T : X \to Y$ is a bounded linear operator between Banach spaces. Let $g \in Y^*$. Then $g(Tx) = 0 \quad \forall x \in X \iff T^*g(x) = 0 \quad \forall x \in X \iff T^*g = 0$. I.e.,

$$\mathcal{R}(T)^a = \mathcal{N}(T^*).$$

Similarly, for $x \in X$, $Tx = 0 \iff f(Tx) = 0 \quad \forall f \in Y^* \iff T^*f(x) = 0 \quad \forall f \in Y^*$, or

$$^a\,\mathcal{R}(T^*) = \mathcal{N}(T).$$

Taking annihilators gives two more results:

$$\overline{\mathcal{R}(T)} = {}^a\,\mathcal{N}(T^*), \qquad \overline{\mathcal{R}(T^*)} \subset \mathcal{N}(T)^a.$$

In particular we see that T^* is injective iff T has dense range; and T is injective if T^* has dense range.

Note: we will have further results in this direction once we introduce the weak*-topology on X^*. In particular, $({}^aS)^a$ is the weak* closure of a subspace S of X^* and T is injective iff T^* has weak* dense range.

Dual of a subspace. An important case is when T is the inclusion map $i : S \to X$, where S is a closed subspace of X. Then $r = i^* : X^* \to S^*$ is just the restriction map: $rf(s) = f(s)$. Hahn-Banach tells us that r is surjective. Obviously $\mathcal{N}(r) = S^a$. Thus we have a canonical isomorphism $\bar{r} : X^*/S^a \to S^*$. In fact, the Hahn-Banach theorem shows that it is an isometry. Via this isometry one often identifies X^*/S^a with S^*.

Dual of a quotient space. Next, consider the projection map $\pi : X \to X/S$ where S is a closed subspace. We then have $\pi^* : (X/S)^* \to X^*$. Since π is surjective, this map is injective. It is easy to see that the range is contained in S^a. In fact we now show that π^* maps $(X/S)^*$ onto S^a, hence provides a canonical isomorphism of S^a with $(X/S)^*$. Indeed, if $f \in S^a$, then we have a splitting $f = g \circ \pi$ with $g \in (X/S)^*$ (just define $g(c) = f(x)$ where x is any element of the coset c). Thus $f = \pi^*g$ is indeed in the range of π^*. This correspondence is again an isometry.

Dual of a Hilbert space. The identification of dual spaces can be quite tricky. The case of Hilbert spaces is easy.

Riesz Representation Theorem. *If X is a real Hilbert space, define $j : X \to X^*$ by $j_y(x) = \langle x, y\rangle$. This map is a linear isometry of X onto X^*. For a complex Hilbert space it is a conjugate linear isometry (it satisfies $j_{\alpha y} = \bar{\alpha} j_y$).*

Proof. It is easy to see that j is an isometry of X into X^* and the main issue is to show that any $f \in X^*$ can be written as j_y for some y. We may assume that $f \neq 0$, so $\mathcal{N}(f)$ is a proper closed subspace of X. Let $y_0 \in [\mathcal{N}(f)]^\perp$ be of norm 1 and set $y = (fy_0)y_0$. For all $x \in X$, we clearly have that $(fy_0)x - (fx)y_0 \in \mathcal{N}(f)$, so

$$j_y(x) = \langle x, (fy_0)y_0\rangle = \langle (fy_0)x, y_0\rangle = \langle (fx)y_0, y_0\rangle = fx. \quad \square$$

Via the map j we can define an inner product on X^*, so it is again a Hilbert space.

Note that if S is a closed subspace of X, then $x \in S^\perp \iff j_s \in S^a$. The Riesz map j is sometimes used to identify X and X^*. Under this identification there is no distinction between $S^\perp$ and S^a.

Dual of $C(\Omega)$. Note: there are two quite distinct theorems referred to as the Riesz Representation Theorem. The proceeding is the easy one. The hard one identifies the dual of $C(\Omega)$ where Ω is a compact subset of $\mathbb{R}^n$ (this can be generalized considerably). It states that there is an isometry between $C(\Omega)^*$ and the space of finite signed measures on Ω. (A finite signed measure is a set function of the form $\mu = \mu_1 - \mu_2$ where μ_i is a finite measure, and we view such as a functional on $C(X)$ by $f \mapsto \int_\Omega f\,d\mu_1 - \int_\Omega f\,d\mu_2$.) This is the real-valued case; in the complex-valued case the isometry is with complex measures $\mu + i\lambda$ where μ and λ are finite signed measures.

Dual of C^1. It is easy to deduce a representation for an arbitrary element of the dual of, e.g., $C^1([0,1])$. The map $f \mapsto (f, f')$ is an isometry of C^1 onto a closed subspace of $C \times C$. By the Hahn-Banach Theorem, every element of $(C^1)^*$ extends to a functional on $C \times C$, which is easily seen to be of the form

$$(f, g) \mapsto \int f\,d\mu + \int g\,d\nu$$

where μ and and ν are signed measures ($(X \times Y)^* = X^* \times Y^*$ with the obvious identifications). Thus any continuous linear functional on C^1 can be written

$$f \mapsto \int f\,d\mu + \int f'\,d\nu.$$

In this representation the measures μ and λ are *not* unique.

Dual of L^p. Hölder's inequality states that if $1 \leq p \leq \infty$, $q = p/(p-1)$, then

$$\int fg \leq \|f\|_{L^p}\|g\|_{L^q}$$

for all $f \in L^p$, $g \in L^q$. This shows that the map $g \mapsto \lambda_g$:

$$\lambda_g(f) = \int fg,$$

maps L^q linearly into $(L^p)^*$ with $\|\lambda_g\|_{(L^p)^*} \le \|g\|_{L^q}$. The choice $f = \operatorname{sign}(g)|g|^{q-1}$ shows that there is equality. In fact, if $p < \infty$, λ is a linear isometry of L^q onto $(L^p)^*$. For $p = \infty$ it is an isometric injection, but not in general surjective. Thus the dual of L^p is L^q for p finite. The dual of L^∞ is a very big space, much bigger than L^1 and rarely used.

Dual of c_0. The above considerations apply to the dual of the sequence spaces l_p. Let us now show that the dual of c_0 is l_1. For any $c = (c_n) \in c_0$ and $d = (d_n) \in l_1$, we define $\lambda_d(c) = \sum c_n d_n$. Clearly

$$|\lambda_d(c)| \le \sup |c_n| \sum |d_n| = \|c\|_{c_0} \|d\|_{l_1},$$

so $\|\lambda_d\|_{c_0^*} \le \|d\|_{l_1}$. Taking

$$c_n = \begin{cases} \operatorname{sign}(d_n), & n \le N \\ 0, & n > N, \end{cases}$$

we see that equality holds. Thus $\lambda : l_1 \to c_0^*$ is an isometric injection. We now show that it is onto. Given $f \in c_0^*$, define $d_n = f(e^{(n)})$ where $e^{(n)}$ is the usual unit sequence $e_m^{(n)} = \delta_{mn}$. Let $s_n = \operatorname{sign}(d_n)$. Then $|d_n| = f(s_n e^{(n)})$, so

$$\sum_{n=0}^{N} |d_n| = \sum_{n=0}^{N} f(s_n e^{(n)}) = f(\sum_{n=0}^{N} s_n e^{(n)}) \le \|f\|.$$

Letting $N \to \infty$ we conclude that $d \in l_1$. Now by construction λ_d agrees with f on all sequences with only finitely many nonzeros. But these are dense in c_0, so $f = \lambda_d$.

The bidual. If X is any normed linear space, we have a natural map $i : X \to X^{**}$ given by

$$i_x(f) = f(x), \qquad x \in X, \quad f \in X^*.$$

Clearly $\|i_x\| \le \|f\|$ and, by the Hahn-Banach theorem, equality holds. Thus X may be identified as a subspace of the Banach space X^{**}. If we define $\tilde{X}$ as the closure of $i(X)$ in X^{**}, then X is isometrically embedded as a dense subspace of the Banach space $\tilde{X}$. This determines $\tilde{X}$ up to isometry, and is what we define as the completion of X. Thus any normed linear space has a completion.

If i is onto, i.e., if X is isomorphic with X^{**} via this identification, we say that X is reflexive (which can only happen is X is complete). In particular, one can check that if X is a Hilbert space and $j : X \to X^*$ is the Riesz isomorphism, and $j^* : X^* \to X^{**}$ the Riesz isomorphism for X^*, then $i = j^* \circ j$, so X is reflexive.

Similarly, the canonical isometries of L^q onto $(L^p)^*$ and then L^p onto $(L^q)^*$ compose to give the natural map of L^p into its bidual, and we conclude that L^p (and l_p) is reflexive for $1 < p < \infty$. None of L^1, l_1, L^∞, l_∞, c_0, or $C(X)$ are reflexive.

If X is reflexive, then $i({}^aS) = S^a$ for $S \subset X^*$. In other words, if we identify X and X^{**}, the distinction between the two kinds of annihilators disappears. In particular, for reflexive Banach spaces, $\overline{\mathcal{R}(T^*)} = \mathcal{N}(T)^a$ and T is injective iff T^* has dense range.

III. Fundamental Theorems

The Open Mapping Theorem and the Uniform Boundedness Principle join the Hahn-Banach Theorem as the “big three”. These two are fairly easy consequence of the Baire Category Theorem.

Baire Category Theorem. *A complete metric space cannot be written as a countable union of nowhere dense sets.*

Sketch of proof. If the statement were false, we could write $M = \bigcup_{n\in\mathbb{N}} F_n$ with F_n a closed subset which does not contain any open set. In particular, F_0 is a proper closed set, so there exists $x_0 \in M$, $\epsilon_0 \in (0,1)$ such that $E(x_0,\epsilon_0) \subset M \setminus F_0$. Since no ball is contained in F_1, there exists $x_1 \in E(x_0,\epsilon_0/2)$ and $\epsilon_1 \in (0,\epsilon_0/2)$ such that $E(x_1,\epsilon_1) \subset M \setminus F_1$. In this way we get a nested sequence of balls such that the nth ball has radius at most 2^{-n} and is disjoint from F_n. It is then easy to check that their centers form a Cauchy sequence and its limit, which must exist by completeness, can’t belong to any F_n. □

The Open Mapping Theorem. The Open Mapping Theorem follows from the Baire Category Theorem and the following lemma.

Lemma. *Let $T : X \to Y$ be a bounded linear operator between Banach spaces. If $E(0_Y,r) \subset \overline{T(E(0_X,1))}$ for some $r > 0$, then $E(0_Y,r) \subset T(B(0_X,2))$.*

Proof. Let $U = T(E(0_X,1))$. Let $y \in Y$, $\|y\| < r$. There exists $y_0 \in U$ with $\|y-y_0\| \le r/2$. By homogeneity, there exists $y_1 \in \frac{1}{2}U$ such that $\|y-y_0-y_1\| \le r/4$, $y_2 \in \frac{1}{4}U$ such that $\|y-y_0-y_1-y_2\| \le r/8$, etc. Take $x_n \in \frac{1}{2^n}U$ such that $Tx_n = y_n$, and let $x = \sum_n x_n \in X$. Then $\|x\| \le 2$ and $Tx = \sum y_n = y$. □

Remark. The same proof works to prove the statement with 2 replaced by any number greater than 1. With a small additional argument, we can even replace it with 1 itself. However the statement above is sufficient for our purposes.

Open Mapping Theorem. *A bounded linear surjection between Banach spaces is open.*

Proof. It is enough to show that the image under T of a ball about 0 contains some ball about 0. The sets $T(E(0,n))$ cover Y, so the closure of one of them must contain an open ball. By the previous result, we can dispense with the closure. The theorem easily follows using the linearity of T. □

There are two major corollaries of the Open Mapping Theorem, each of which is equivalent to it.

Inverse Mapping Theorem or Banach’s Theorem. *The inverse of an invertible bounded linear operator between Banach spaces is continuous.*

Proof. The map is open, so its inverse is continuous. □

Closed Graph Theorem. *A linear operator between Banach spaces is continuous iff its graph is closed.*

A map between topological spaces is called closed if its graph is closed. In a general Hausdorff space, this is a weaker property than continuity, but the theorem asserts that for linear operators between Banach spaces it is equivalent. The usefulness is that a direct proof of continuity requires us to show that if x_n converges to x in X then Tx_n converges to Tx. By using the closed graph theorem, we get to assume as well that Tx_n is converging to some y in Y and we need only show that $y = Tx$.

Proof. Let $G = \{\, (x, Tx) \,|\, x \in X \,\}$ denote the graph. Then the composition $G \subset X \times Y \to X$ is a bounded linear operator between Banach spaces given by $(x, Tx) \mapsto x$. It is clearly one-to-one and onto, so the inverse is continuous by Banach's theorem. But the composition $X \to G \subset X \times Y \to Y$ is simply the T, so T is continuous. $\square$

Banach's theorem leads immediately to this useful characterization of closed imbeddings of Banach spaces.

Theorem. *Let $T : X \to Y$ be a bounded linear map between Banach spaces. Then T is one-to-one and has closed range if and only if there exists a positive number c such that*

$$\|x\| \le c\|Tx\| \qquad \forall x \in X.$$

Proof. If the inequality holds, then T is clearly one-to-one, and if Tx_n is a Cauchy sequence in $\mathcal{R}(T)$, then x_n is Cauchy, and hence x_n converges to some x, so Tx_n converges to Tx. Thus the inequality implies that $\mathcal{R}(T)$ is closed.

For the other direction, suppose that T is one-to-one with closed range and consider the map $T^{-1} : \mathcal{R}(T) \to X$. It is the inverse of a bounded isomorphism, so is itself bounded. The inequality follows immediately (with c the norm of T^{-1}). $\square$

Another useful corollary is that if a Banach space admits a second weaker or stronger norm under which it is still Banach, then the two norms are equivalent. This follows directly from Banach's theorem applied to the identity.

The Uniform Boundedness Principle. The Uniform Boundedness Principle (or the Banach-Steinhaus Theorem) also comes from the Baire Category Theorem.

Uniform Boundedness Principle. *Suppose that X and Y are Banach spaces and $\mathcal{S} \subset B(X, Y)$. If $\sup_{T\in\mathcal{S}} \|T(x)\|_Y < \infty$ for all $x \in X$, then $\sup_{T\in\mathcal{S}} \|T\| < \infty$.*

Proof. One of the closed sets $\{\, x \,|\, |f_n(x)| \le N \quad \forall n \,\}$ must contain $E(x_0, r)$ for some $x_0 \in X$, $r > 0$. Then, if $\|x\| < r$, $|f_n(x)| \le |f_n(x + x_0) - f_n(x_0)| \le N + \sup |f_n(x_0)| = M$, with M independent of n. This shows that the $\|f_n\|$ are uniformly bounded (by M/r). $\square$

In words: a set of linear operators between Banach spaces which is bounded pointwise is norm bounded.

The uniform boundedness theorem is often a way to generate counterexamples. A typical example comes from the theory of Fourier series. For $f : \mathbb{R} \to \mathbb{C}$ continuous and 1-periodic the nth partial sum of the Fourier series for f is

$$f_n(s) = \sum_{k=-n}^{n} \int_{-1}^{1} f(t)e^{-2\pi ikt}\,dt\,e^{2\pi iks} = \int_{-1}^{1} f(t)D_n(s-t)\,dt$$

where

$$D_n(x) = \sum_{k=-n}^{n} e^{2\pi ikx}.$$

Writing $z = e^{2\pi is}$, we have

$$D_n(s) = \sum_{k=-n}^{n} z^k = z^{-n}\frac{z^{2n}-1}{z-1} = \frac{z^{n+1/2}-z^{-n-1/2}}{z^{1/2}-z^{-1/2}} = \frac{\sin(2n+1)\pi x}{\sin \pi x}.$$

This is the Dirichlet kernel, a C^∞ periodic function. In particular, the value of the nth partial sum of the Fourier series of f at 0 is

$$T_nf := f_n(0) = \int_{-1}^{1} f(t)D_n(t)\,dt.$$

We think of T_n as a linear functional on the Banach space of 1-periodic continuous function endowed with the sup norm. Clearly

$$\|T_n\| \le C_n := \int_{-1}^{1} |D_n(t)|\,dt.$$

In fact this is an equality. If $g(t) = \operatorname{sign} D_n(t)$, then $\sup|g| = 1$ and $T_ng = C_n$. Actually, g is not continuous, so to make this argument correct, we approximate g by a continuous functions, and thereby prove the norm equality. Now one can calculate that $\int |D_n| \to \infty$ as $n \to \infty$. By the uniform boundedness theorem we may conclude that there exists a continuous periodic function for whose Fourier series diverges at $t = 0$.

The Closed Range Theorem. We now apply the Open Mapping Theorem to better understand the relationship between T and T^*. The property of having a closed range is significant to the structure of an operator between Banach spaces. If $T : X \to Y$ has a closed range Z (which is then itself a Banach space), then T factors as the projection $X \to X/\mathcal{N}(T)$, the isomorphism $X/\mathcal{N}(T) \to Z$, and the inclusion $Z \subset Y$. The Closed Range Theorem says that T has a closed range if and only if T^* does.

Theorem. *Let $T : X \to Y$ be a bounded linear operator between Banach spaces. Then T is invertible iff T^* is.*

Proof. If $S = T^{-1} : Y \to X$ exists, then $ST = I_X$ and $TS = I_Y$, so $T^*S^* = I_{X^*}$ and $S^*T^* = I_{Y^*}$, which shows that T^* is invertible.

Conversely, if T^* is invertible, then it is open, so there is a number $c > 0$ such that $T^* B_{Y^*}(0,1)$ contains $B_{X^*}(0,c)$. Thus, for $x \in X$

$$\begin{aligned} \|Tx\| &= \sup_{f \in B_{Y^*}(0,1)} |f(Tx)| = \sup_{f \in B_{Y^*}(0,1)} |(T^* f)x| \\ &\geq \sup_{g \in B_{X^*}(0,c)} |g(x)| = c\|x\|. \end{aligned}$$

The existence of $c > 0$ such that $\|Tx\| \geq c\|x\|$ $\forall x \in X$ is equivalent to the statement that T is injective with closed range. But since T^* is injective, T has dense range. $\square$

Lemma. *Let $T : X \to Y$ be a linear map between Banach spaces such that T^* is an injection with closed range. Then T is a surjection.*

Proof. Let E be the closed unit ball of X and $F = \overline{TE}$. It suffices to show that F contains a ball around the origin, since then, by the lemma used to prove the Open Mapping Theorem, T is onto.

There exists $c > 0$ such that $\|T^* f\| \geq c\|f\|$ for all $f \in Y^*$. We shall show that F contains the ball of radius c around the origin in Y. Otherwise there exists $y \in Y$, $\|y\| \leq c$, $y \notin F$. Since F is a closed convex set we can find a functional $f \in Y^*$ such that $|f(Tx)| \leq \alpha$ for all $x \in E$ and $f(y) > \alpha$. Thus $\|f\| > \alpha/c$, but

$$\|T^* f\| = \sup_{x \in E} |T^* f(x)| = \sup_{x \in E} |f(Tx)| \leq \alpha.$$

This is a contradiction. $\square$

Closed Range Theorem. *Let $T : X \to Y$ be a bounded linear operator between Banach spaces. Then T has closed range if and only if T^* does.*

Proof. 1) $\mathcal{R}(T)$ closed $\implies$ $\mathcal{R}(T^*)$ closed.

Let $Z = \mathcal{R}(T)$. Then $\bar{T} : X/\mathcal{N}(T) \to Z$ is an isomorphism (Inverse Mapping Theorem). The diagram

$$\begin{tikzcd} X \arrow[r, "T"] \arrow[d, "\pi"'] & Y \\ X/\mathcal{N}(T) \arrow[r, "\cong", "\bar{T}"'] & Z. \arrow[u, "\cup"'] \end{tikzcd}$$

commutes. Taking adjoints,

$$\begin{tikzcd} X^* & Y^* \arrow[l, "T^*"'] \arrow[d, "\pi"] \\ \mathcal{N}(T)^a \arrow[u, "\cup"] & Y^*/Z^a. \arrow[l, "\cong"', "\bar{T}^*"] \end{tikzcd}$$

This shows that $\mathcal{R}(T^*) = \mathcal{N}(T)^a$.

2) $\mathcal{R}(T^*)$ closed $\implies$ $\mathcal{R}(T)$ closed.

Let $Z = \overline{\mathcal{R}(T)}$ (so $Z^a = \mathcal{N}(T^*)$) and let S be the range restriction of T, $S : X \to Z$. The adjoint is $S^* : Y^*/Z^a \to X^*$, the lifting of T^* to Y^*/Z^a. Now $\mathcal{R}(S^*) = \mathcal{R}(T^*)$, is closed, and S^* is an injection. We wish to show that S is onto Z. Thus the theorem follows from the preceding lemma. $\square$

IV. Weak Topologies

The weak topology. Let X be a Banach space. For each $f \in X^*$ the map $x \mapsto |f(x)|$ is a seminorm on X, and the set of all such seminorms, as f varies over X^*, is sufficient by the Hahn-Banach Theorem. Therefore we can endow X with a new TVS structure from this family of seminorms. This is called the weak topology on X. In particular, $x_n \to x$ weakly (written $x_n \xrightarrow{w} x$) iff $f(x_n) \to f(x)$ for all $f \in X^*$. Thus the weak topology is weaker than the norm topology, but all the elements of X^* remain continuous when X is endowed with the weak topology (it is by definition the weakest topology for which all the elements of X^* are continuous).

Note that the open sets of the weak topology are rather big. If U is an weak neighborhood of 0 in an infinite dimensional Banach space then, by definition, there exists $\epsilon > 0$ and finitely many functionals $f_n \in X^*$ such that $\{\, x \,|\, |f_n(x)| < \epsilon \quad \forall N \,\}$ is contained in U. Thus U contains the infinite dimensional closed subspace $\mathcal{N}(f_1) \cap \ldots \cap \mathcal{N}(f_n)$.

If $x_n \xrightarrow{w} x$ weakly, then, viewing the x_n as linear functionals on X^* (via the canonical embedding of X into X^{**}), we see that the sequence of real numbers obtained by applying the x_n to any $f \in X^*$ is convergent and hence bounded uniformly in n. By the Uniform Boundedness Principle, it follows that the x_n are bounded.

Theorem. *If a sequence of elements of a Banach space converges weakly, then the sequence is norm bounded.*

On the other hand, if the x_n are small in norm, then their weak limit is too.

Theorem. *If $x_n \xrightarrow{w} x$ in some Banach space, then $\|x\| \le \liminf_{n\to\infty} \|x_n\|$.*

Proof. Take $f \in X^*$ of norm 1 such that $f(x) = \|x\|$. Then $f(x_n) \le \|x_n\|$, and taking the $\liminf$ gives the result. $\square$

For convex sets (in particular, for subspaces) weak closure coincides with norm closure:

Theorem. *1) The weak closure of a convex set is equal to its norm closure.*

2) A convex set is weak closed iff it is normed closed.

3) A convex set is weak dense iff it is norm dense.

Proof. The second and third statement obviously follow from the first, and the weak closure obviously contains the norm closure. So it remains to show that if x does not belong to the norm closure of a convex set E, then there is a weak neighborhood of x which doesn't intersect E. This follows immediately from the following convex separation theorem. $\square$

Theorem. *Let E be a nonempty closed convex subset of a Banach space X and x a point in the complement of E. Then there exists $f \in X^*$ such that $f(x) < \inf_{y\in E} f(y)$.*

In fact we shall prove a stronger result:

Theorem. *Let E and F be disjoint, nonempty, convex subsets of a Banach space X with F open. Then there exists $f \in X^*$ such that $f(x) < \inf_{y\in E} f(y)$ for all $x \in F$.*

(The previous result follows by taking F to be any ball about x disjoint from E.)

Proof. This is a consequence of the generalized Hahn-Banach Theorem. Pick $x_0 \in E$ and $y_0 \in F$ and set $z_0 = x_0 - y_0$ and $G = F - E + z_0$. Then G is a convex open set containing 0 but not containing z_0. (The convexity of G follows directly from that of E and F; the fact that G is open follows from the representation of $G = \bigcup_{y\in E} F - y + z_0$ as a union of open sets; obviously $0 = y_0 - x_0 + z_0 \in G$, and $z_0 \notin G$ since E and F and disjoint.)

Since G is open and convex and contains 0, for each $x \in X$, $\{\, t > 0 \,|\, t^{-1}x \in G \,\}$ is a nonempty open semi-infinite interval. Define $p(x) \in [0, \infty)$ to be the left endpoint of this interval. By definition p is positively homogeneous. Since G is convex, $t^{-1}x \in G$ and $s^{-1}y \in G$ imply that

$$(t+s)^{-1}(x+y) = \frac{t}{s+t}t^{-1}x + \frac{s}{s+t}s^{-1}y \in G,$$

whence p is subadditive. Thus p is a sublinear functional. Moreover, $G = \{\, x \in X \,|\, p(x) < 1 \,\}$.

Define a linear functional f on $X_0 := \mathbb{R}z_0$ by $f(z_0) = 1$. Then $f(tz_0) = t \le tp(z_0) = p(tz_0)$ for $t \ge 0$ and $f(tz_0) < 0 \le p(tz_0)$ for $t < 0$. Thus f is a linear functional on X_0 satisfying $f(x) \le p(x)$ there. By Hahn-Banach we can extend f to a linear functional on X satisfying the same inequality. This implies that f is bounded (by 1) on the open set G, so f belongs to X^*.

If $x \in F$, $y \in E$, then $x - y + z_0 \in G$, so $f(x) - f(y) + 1 = f(x - y + z_0) < 1$, or $f(x) < f(y)$. Therefore $\sup_{x\in F} f(x) \le \inf_{y\in E} f(y)$. Since $f(F)$ is an open interval, $f(\tilde{x}) < \sup_{x\in F} f(x)$ for all $\tilde{x} \in F$, and so we have the theorem. $\square$

The weak* topology. On the dual space X^* we have two new topologies. We may endow it with the weak topology, the weakest one such that all functionals in X^{**} are continuous, or we may endow it with the topology generated by all the seminorms $f \mapsto f(x)$, $x \in X$. (This is obviously a sufficient family of functionals.) The last is called the weak* topology and is a weaker topology than the weak topology. If X is reflexive, the weak and weak* topologies coincide.

Examples of weak and weak* convergence: 1) Consider weak convergence in $L^p(\Omega)$ where Ω is a bounded subset of $\mathbb{R}^n$. From the characterization of the dual of L^p we see that

$$f_n \xrightarrow{w*} f \text{ in } L^\infty \implies f_n \xrightarrow{w} f \text{ weak in } L^p \implies f_n \xrightarrow{w} f \text{ weak in } L^q$$

whenever $1 \le q \le p < \infty$. In particular we claim that the complex exponentials $e^{2\pi inx} \xrightarrow{w*} 0$ in $L^\infty([0,1])$ as $n \to \infty$. This is simply the statement that

$$\lim_{n\to\infty} \int_0^1 g(x) e^{2\pi inx}\,dx = 0,$$

for all $g \in L^1([0,1])$, i.e., that the Fourier coefficients of an L^1 tend to 0, which is known as the Riemann–Lebesgue Lemma. (Proof: certainly true if g is a trigonometric polynomial. The trig polynomials are dense in $C([0,1])$ by the Weierstrass Approximation Theorem, and $C([0,1])$ is dense in $L^1([0,1])$.) This is one common example of weak convergence which is not norm convergence, namely weak vanishing by oscillation.

2) Another common situation is weak vanishing to infinity. As a very simple example, it is easy to see that the unit vectors in l_p converge weakly to zero for $1 < p < \infty$ (and weak* in l_∞, but not weakly in l_1). As a more interesting example, let $f_n \in L^p(\mathbb{R})$ be a sequence of function which are uniformly bounded in L^p, and for which $f_n|_{[-n,n]} \equiv 0$. Then we claim that $f_n \to 0$ weakly in L^p if $1 < p < \infty$. Thus we have to show that

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n g\,dx = 0$$

for all $g \in L^q$. Let $S_n = \{\, x \in \mathbb{R} \mid |x| \ge n \,\}$. Then $\lim_n \int_{S_n} |g|^q\,dx = 0$ (by the dominated convergence theorem). But

$$\left|\int_{\mathbb{R}} f_n g\,dx\right| = \left|\int_{S_n} f_n g\,dx\right| \le \|f_n\|_{L^p} \|g\|_{L^q(S_n)} \le C\|g\|_{L^q(S_n)} \to 0.$$

The same proof shows that if the f_n are uniformly bounded they tend to 0 in L^∞ weak*. Note that the characteristic functions $\chi_{[n,n+1]}$ do *not* tend to zero weakly in L^1 however.

3) Consider the measure $\phi_n = 2n\chi_{[-1/n,1/n]}dx$. Formally ϕ_n tends to the delta function δ_0 as $n \to \infty$. Using the weak* topology on $C([-1,1])$ this convergence becomes precise: $\phi_n \xrightarrow{w*} \delta_0$.

Theorem (Alaoglu). *The unit ball in X^* is weak* compact.*

Proof. For $x \in X$, let $I_x = \{\, t \in \mathbb{R} : |t| \le \|x\| \,\}$, and set $\Omega = \Pi_{x\in X} I_x$. Recall that this Cartesian product is nothing but the set of all functions f on X with $f(x) \in I_x$ for all x. This set is endowed with the Cartesian product topology, namely the weakest topology such that for all $x \in X$, the functions $f \mapsto f(x)$ (from Ω to I_x) are continuous. Tychonoff's Theorem states that Ω is compact with this topology.

Now let E be the unit ball in X^*. Then $E \subset \Omega$ and the topology thereby induced on E is precisely the weak* topology. Now for each pair $x, y \in X$ and each $c \in R$, define $F_{x,y}(f) = f(x) + f(y) - f(x+y)$, $G_{x,c} = f(cx) - cf(x)$. These are continuous functions on Ω and

$$E = \bigcap_{x,y\in X} F_{x,y}^{-1}(0) \cap \bigcap_{\substack{x\in X\\ c\in\mathbb{R}}} G_{c,x}^{-1}(0).$$

Thus E is a closed subset of a compact set, and therefore compact itself. $\square$

Corollary. *If $f_n \xrightarrow{w^*} f$ in X^*, then $\|f\| \le \liminf_{n\to\infty} \|f_n\|_{X^*}$.*

Proof. Let $C = \liminf \|f_n\|$ and let $\epsilon > 0$ be arbitrary. Then there exists a subsequence (also denoted f_n) with $\|f_n\| \le C + \epsilon$. The ball of radius $C + \epsilon$ being weak* compact, and so weak* closed, $\|f\| < C + \epsilon$. Since ϵ was arbitrary, this gives the result. $\square$

On X^{**} the weak* topology is that induced by the functionals in X^*.

Theorem. *The unit ball of X is weak* dense in the unit ball of X^{**}.*

Proof. Let z belong to the unit ball of X^{**}. We need to show that for any $f_1, \dots, f_n \in X^*$ of norm 1, and any $\epsilon > 0$, the set

$$\{\, w \in X^{**} \mid |(w - z)(f_i)| < \epsilon, \quad i = 1, \dots, n \,\}$$

contains a point of the unit ball of X. (Since any neighborhood of z contains a set of this form.)

It is enough to show that there exists $y \in X$ with $\|y\| < 1 + \epsilon$ such that $(y - z)(f_i) = 0$ for each i. Because then $y/(1 + \epsilon)$ belongs to the closed unit ball of X, and

$$|((1+\epsilon)^{-1}y - z)(f_i)| = |((1+\epsilon)^{-1}y - y)(f_i)| \le \|((1+\epsilon)^{-1}y - y\| = \|y\|\frac{\epsilon}{1+\epsilon} < \epsilon.$$

Let S be the span of the f_i in X^*. Since S is finite dimensional the canonical map $X \to S^*$ is surjective. (This is equivalent to saying that if the null space of a linear functional g contains the intersection of the null spaces of a finite set of linear functionals g_i, then g is a linear combination of the g_i, which is a simple, purely algebraic result. [Proof: The nullspace of the map $(g_1, \dots, g_n) : X \to \mathbb{R}^n$ is contained in the nullspace of g, so $g = T \circ (g_1, \dots, g_n)$ for some linear $T : \mathbb{R}^n \to \mathbb{R}$.]) Consequently $X/{}^aS$ is isometrically isomorphic to S.

In particular $z|_S$ concides with $y + {}^aS$ for some $y \in X$. Since $\|z\|_{S^*} \le 1$, and we can choose the coset representative y with $\|y\| \le 1 + \epsilon$ as claimed. $\square$

Corollary. *The closed unit ball of a Banach space X is weakly compact if and only if X is reflexive.*

Proof. If the closed unit ball of X is weakly compact, then it is weak* compact when viewed as a subset of X^{**}. Thus the ball is weak* closed, and so, by the previous theorem, the embedding of the the unit ball of X contains the ball of X^{**}. It follows that the embedding of X is all of X^{**}.

The reverse direction is immediate from the Alaoglu theorem. $\square$

Hilbert–Schmidt operators.

Lemma. *Suppose that* $\{ e_i \}$ *and* $\{ \tilde{e}_i \}$ *are two orthonormal bases for a separable Hilbert space* X*, and* $T \in B(X)$*. Then*

$$\sum_{i,j} |\langle Te_i, e_j \rangle|^2 = \sum_{i,j} |\langle T\tilde{e}_i, \tilde{e}_j \rangle|^2.$$

Proof. For all $w \in X$, $\sum_j |\langle w, e_j \rangle|^2 = \|w\|^2$, so

$$\sum_{i,j} |\langle Te_i, e_j \rangle|^2 = \sum_i \|Te_i\|^2 = \sum_j \|T^* e_j\|^2.$$

But

$$\sum_i \|T^* e_i\|^2 = \sum_{i,j} |\langle T^* e_i, \tilde{e}_j \rangle|^2 = \sum_j \|T\tilde{e}_j\|^2. \quad \square$$

Definition. If $T \in B(X)$ define $\|T\|_2$ by

$$\|T\|_2^2 = \sum_{i,j} |\langle Te_i, e_j \rangle|^2 = \sum_i \|Te_i\|^2$$

where $\{ e_i \}$ is any orthonormal basis for X. T is called a Hilbert–Schmidt operator if $\|T\|_2 < \infty$, and $\|T\|_2$ is called the Hilbert–Schmidt norm of T.

We have just seen that if T is Hilbert–Schmidt, then so is T^* and their Hilbert–Schmidt norms coincide.

Proposition. $\|T\| \le \|T\|_2$.

Proof. Let $x = \sum c_i e_i$ be an arbitrary element of X. Then

$$\|Tx\|^2 = \sum_i \Big| \sum_j c_j \langle Te_j, e_i \rangle \Big|^2.$$

By Cauchy–Schwarz

$$|\sum_j c_j \langle Te_j, e_i \rangle|^2 \le \sum_j c_j^2 \cdot \sum_j |\langle Te_j, e_i \rangle|^2 = \|x\|^2 \sum_j |\langle Te_j, e_i \rangle|^2.$$

Summing on i gives the result. $\square$

Proposition. *Let Ω be an open subset of $\mathbb{R}^n$ and $K \in L^2(\Omega \times \Omega)$. Define*

$$T_K u(x) = \int_\Omega K(x,y)u(y)\,dy, \qquad \text{for all } x \in \Omega.$$

Then T_K defines a Hilbert–Schmidt operator on $L^2(\Omega)$ and $\|T_K\|_2 = \|K\|_{L^2}$.

Proof. For $x \in \Omega$, set $K_x(y) = K(x,y)$. By Fubini's theorem, $K_x \in L^2(\Omega)$ for almost all $x \in \Omega$, and

$$\|K\|_{L^2}^2 = \int \|K_x\|^2\,dx.$$

Now, $T_K u(x) = \langle K_x, u\rangle$, so, if $\{ e_i \}$ is an orthonormal basis, then

$$\begin{aligned}\|T_K\|_2^2 = \sum_i \|T_K e_i\|^2 = \sum_i \int |(T_K e_i)(x)|^2\,dx = \sum_i \int |\langle K_x, e_i\rangle|^2\,dx \\ = \int \sum_i |\langle K_x, e_i\rangle|^2\,dx = \int \|K_x\|^2\,dx = \|K\|_{L^2}^2.\end{aligned}$$

□

Compact operators.

Definition. A bounded linear operator between Banach spaces is called compact if it maps the unit ball (and therefore every bounded set) to a precompact set.

For example, if T has finite rank ($\dim \mathcal{R}(T) < \infty$), then T is compact.

Recall the following characterization of precompact sets in a metric space, which is often useful.

Proposition. *Let M be a metric space. Then the following are equivalent:*

1. *M is precompact.*
2. *For all $\epsilon > 0$ there exist finitely many sets of diameter at most ϵ which cover M.*
3. *Every sequence contains a Cauchy subsequence.*

Sketch of proof. (1) $\implies$ (2) and (3) $\implies$ (1) are easy. For (2) $\implies$ (3) use a Cantor diagonalization argument to extract a Cauchy subsequence. □

Theorem. *Let X and Y be Banach spaces and $B_c(X,Y)$ the space of compact linear operators from X to Y. Then $B_c(X,Y)$ is a closed subspace of $B(X,Y)$.*

Proof. Suppose $T_n \in B_c(X,Y)$, $T \in B(X,Y)$, $\|T_n - T\| \to 0$. We must show that T is compact. Thus we must show that $T(E)$ is precompact in Y, where E is the unit ball in X. For this, it is enough to show that for any $\epsilon > 0$ there are finitely many balls U_i of radius ϵ in Y such that

$$T(E) \subset \bigcup_i U_i.$$

Choose n large enough that $\|T - T_n\| \le \epsilon/2$, and let $V_1, V_2, \dots, V_n$ be finitely many balls of radius $\epsilon/2$ which cover $T_n E$. For each i let U_i be the ball of radius ϵ with the same center as V_i. $\square$

It follows that closure of the finite rank operators in $B(X,Y)$ is contained in $B_c(X,Y)$. In general, this may be a strict inclusion, but if Y is a Hilbert space, it is equality. To prove this, choose an orthonormal basis for Y, and consider the finite rank operators of the form PT where P is the orthogonal projection of Y onto the span of finitely many basis elements. Using the fact that TE is compact (E the unit ball of X) and that $\|P\| = 1$, we can find for any $\epsilon > 0$, an operator P of this form with $\sup_{x\in E} \|(PT - T)x\| \le \epsilon$.

The next result is obvious but useful.

Theorem. *Let X and Y be Banach spaces and $T \in B_c(X,Y)$. If Z is another Banach space and $S \in B(Y,Z)$ then ST is compact. If $S \in B(Z,X)$, then TS is compact. If $X = Y$, then $B_c(X) := B_c(X,X)$ is a two-sided ideal in $B(X)$.*

Theorem. *Let X and Y be Banach spaces and $T \in B(X,Y)$. Then T is compact if and only if T^* is compact.*

Proof. Let E be the unit ball in X and F the unit ball in Y^*. Suppose that T is compact. Given $\epsilon > 0$ we must exhibit finitely many sets of diameter at most ϵ which cover T^*F. First choose m sets of diameter at most $\epsilon/3$ which cover TE, and let Tx_i belong to the ith set. Also, let $I_1, \dots, I_n$ be n intervals of length $\epsilon/3$ which cover the interval $[-\|T\|, \|T\|]$. For any m-tuple $(j_1, \dots, j_m)$ of integers with $1 \le j_i \le n$ we define the set

$$\{\, f \in F \mid f(Tx_i) \in I_{j_i}, \quad i = 1, \dots, m \,\}.$$

These sets clearly cover F, so there images under T^* cover T^*F, so it suffices to show that the images have diameter at most ϵ. Indeed, if f and g belong to the set above, and x is any element of E, pick i such that $\|Tx - Tx_i\| \le \epsilon/3$. We know that $\|f(Tx_i) - g(Tx_i)\| \le \epsilon/3$. Thus

$$\begin{aligned} |(T^*f - T^*g)(x)| &= |(f-g)(Tx)| \\ &\le |f(Tx) - f(Tx_i)| + |g(Tx) - g(Tx_i)| + |(f-g)(Tx_i)| \le \epsilon. \end{aligned}$$

This shows that T compact $\implies$ T^* compact. Conversely, suppose that $T^* : Y^* \to X^*$ is compact. Then T^{**} maps the unit ball of X^{**} into a precompact subset of Y^{**}. But the unit ball of X may be viewed as a subset of the unit ball of its bidual, and the restriction of T^{**} to the unit ball of X coincides with T there. Thus T maps the unit ball of X to a precompact set. $\square$

Theorem. *If T is a compact operator from a Banach space to itself, then $\mathcal{N}(\mathbf{1} - T)$ is finite dimensional and $\mathcal{R}(\mathbf{1} - T)$ is closed.*

Proof. T is a compact operator that restricts to the identity on $\mathcal{N}(\mathbf{1} - T)$. Hence the closed unit ball in $\mathcal{N}(\mathbf{1} - T)$ is compact, whence the dimension of $\mathcal{N}(\mathbf{1} - T)$ is finite.

Now any finite dimensional subspace is complemented (see below), so there exists a closed subspace M of X such that $\mathcal{N}(\mathbf{1}-T)+M = X$ and $\mathcal{N}(\mathbf{1}-T)\cap M = 0$. Let $S = (\mathbf{1}-T)|_M$, so S is injective and $\mathcal{R}(S) = \mathcal{R}(\mathbf{1}-T)$. We will show that for some $c > 0$, $\|Sx\| \ge c\|x\|$ for all $x \in M$, which will imply that $\mathcal{R}(S)$ is closed. If the desired inequality doesn't hold for any $c > 0$, we can choose $x_n \in M$ of norm 1 with $Sx_n \to 0$. After passing to a subsequence we may arrange that also Tx_n converges to some $x_0 \in X$. It follows that $x_n \to x_0$, so $x_0 \in M$ and $Sx_0 = 0$. Therefore $x_0 = 0$ which is impossible (since $\|x_n\| = 1$). □

In the proof we used the first part of the following lemma. We say that a closed subspace N is complemented in a Banach space X if there is another closed subspace such that $M \oplus N = X$.

Lemma. *A finite dimensional or finite codimensional closed subspace of a Banach space is complemented.*

Proof. If M is a finite dimensional subspace, choose a basis $x_1, \ldots, x_n$ and define a linear functionals $\phi_i : M \to \mathbb{R}$ by $\phi_i(x_j) = \delta_{ij}$. Extend the ϕ_i to be bounded linear functionals on X. Then we can take $N = \mathcal{N}(\phi_1)\cap\ldots\cap\mathcal{N}(\phi_n)$.

If M is finite codimensional, we can take N to be the span of a set of nonzero coset representatives. □

A simple generalization of the theorem will be useful when we study the spectrum of compact operators.

Theorem. *If T is a compact operator from a Banach space to itself, λ a non-zero complex number, and n a positive integer, then $\mathcal{N}[(\lambda\mathbf{1}-T)^n]$ is finite dimensional and $\mathcal{R}[(\lambda\mathbf{1}-T)^n]$ is closed.*

Proof. Expanding we see that $(\lambda\mathbf{1}-T)^n = \lambda^n(\mathbf{1}-S)$ for some compact operator S, so the result reduces to the previous one. □

We close the section with a good source of examples of compact operators, which includes, for example, any matrix operator on l_2 for which the matrix entries are square-summable.

Theorem. *A Hilbert–Schmidt operator on a separable Hilbert space is compact.*

Proof. Let $\{e_i\}$ be an orthonormal basis. Let T be a given Hilbert–Schmidt operator (so $\sum_i \|Te_i\|^2 < \infty$). Define T_n by $T_n e_i = Te_i$ if $i \le n$, $T_n e_i = 0$ otherwise. Then $\|T - T_n\| \le \|T - T_n\|_2 = \sum_{i=n+1}^{\infty} \|Te_i\|^2 \to 0$. □

Spectral Theorem for compact self-adjoint operators. In this section we assume that X is a *complex* Hilbert space. If $T : X \to X$ is a bounded linear operator, we view T^* as a map from $X \to X$ via the Riesz isometry between X and X^*. That is, T^* is defined by

$$\langle T^* x, y\rangle = \langle x, Ty\rangle.$$

In the case of a finite dimensional complex Hilbert space, T can be represented by a complex square matrix, and T^* is represented by its Hermitian transpose.

Recall that a Hermitian symmetric matrix has real eigenvalues and an orthonormal basis of eigenvectors. For a self-adjoint operator on a Hilbert space, it is easy to see that any eigenvalues are real, and that eigenvectors corresponding to distinct eigenvalues are orthogonal. However there may not exist an orthonormal basis of eigenvectors, or even any nonzero eigenvectors at all. For example, let $X = L^2([0,1])$, and define $Tu(x) = x\,u(x)$ for $u \in L^2$. Then T is clearly bounded and self-adjoint. But it is easy to see that T does not have any eigenvalues.

Spectral Theorem for Compact Self-Adjoint Operators in Hilbert Space. *Let T be a compact self-adjoint operator in a Hilbert space X. Then there is an orthonormal basis consisting of eigenvectors of T.*

Before proceeding to the proof we prove one lemma.

Lemma. *If T is a self-adjoint operator on a Hilbert space, then*

$$\|T\| = \sup_{\|x\|\le 1} |\langle Tx, x\rangle|.$$

Proof. Let $\alpha = \sup_{\|x\|\le 1} |\langle Tx, x\rangle|$. It is enough to prove that

$$|\langle Tx, y\rangle| \le \alpha \|x\|\|y\|$$

for all x and y. We can obviously assume that x and y are nonzero. Moreover, we may multiply y by a complex number of modulus one, so we can assume that $\langle Tx, y\rangle \ge 0$. Then

$$\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle = 4\,\mathrm{Re}\langle Tx, y\rangle = 4|\langle Tx, y\rangle|.$$

so

$$|\langle Tx, y\rangle| \le \frac{\alpha}{4}(\|x+y\|^2 + \|x-y\|^2) = \frac{\alpha}{2}(\|x\|^2 + \|y\|^2).$$

Now apply this result with x replaced by $\sqrt{\|y\|/\|x\|}\,x$ and y replaced by $\sqrt{\|x\|/\|y\|}\,y$. $\square$

Proof of spectral theorem for compact self-adjoint operators. We first show that T has a nonzero eigenvector. If $T = 0$, this is obvious, so we assume that $T \ne 0$. Choose a sequence $x_n \in X$ with $\|x_n\| = 1$ so that $|\langle Tx_n, x_n\rangle| \to \|T\|$. Since T is self-adjoint, $\langle Tx_n, x_n\rangle \in \mathbb{R}$, so we may pass to a subsequence (still denoted x_n), for which $\langle Tx_n, x_n\rangle \to \lambda = \pm\|T\|$.

Since T is compact we may pass to a further subsequence and assume that $Tx_n \to y \in X$. Note that $\|y\| \ge |\lambda| > 0$.

Using the fact that T is self-adjoint and λ is real, we get

$$\begin{aligned}\|Tx_n - \lambda x_n\|^2 &= \|Tx_n\|^2 - 2\lambda\langle Tx_n, x_n\rangle + \lambda^2\|x_n\|^2 \\ &\le 2\|T\|^2 - 2\lambda\langle Tx_n, x_n\rangle \to 2\|T\|^2 - 2\lambda^2 = 0.\end{aligned}$$

Since $Tx_n \to y$ we infer that $\lambda x_n \to y$ as well, or $x_n \to y/\lambda \neq 0$. Applying T we have $Ty/\lambda = y$, so λ is indeed a nonzero eigenvalue.

To complete the proof, consider the set of all orthonormal subsets of X consisting of eigenvectors of T. By Zorn's lemma, it has a maximal element S. Let W be the closure of the span of S. Clearly $TW \subset W$, and it follows directly (since T is self-adjoint), that $TW^\perp \subset W^\perp$. Therefore T restricts to a self-adjoint operator on $W^\perp$ and thus, unless $W^\perp = 0$, T has an eigenvector in $W^\perp$. But this clearly contradicts the maximality of S (since we can adjoin this element to S to get a larger orthonormal set of eigenvectors). Thus $W^\perp = 0$, and S is an orthonormal basis. □

The following structure result on the set of eigenvalues is generally considered part of the spectral theorem as well.

Theorem. *If T is a compact self-adjoint operator on a Hilbert space, then the set of nonzero eigenvalues of T is either a finite set or a sequence approaching* 0 *and the corresponding eigenspaces are all finite dimensional.*

Remark. 0 may or may not be an eigenvalue, and its eigenspace may or may not be finite.

Proof. Let e_i be an orthonormal basis of eigenvectors, with $Te_i = \lambda_i e_i$. Here i ranges over some index set I. It suffices to show that $S = \{\, i \in I \mid |\lambda_i| \ge \epsilon \,\}$ is finite for all $\epsilon > 0$. Then if $i, j \in I$

$$\|Te_i - Te_j\|^2 = \|\lambda_i e_j - \lambda_j e_j\|^2 = |\lambda_i|^2 + |\lambda_j|^2,$$

so if $i, j \in S$, then $\|Te_i - Te_j\|^2 \ge 2\epsilon^2$. If S were infinite, we could then choose a sequence of unit elements in X whose image under T has no convergent subsequence, which violates the compactness of T. □

Suppose, for concreteness, that X is an infinite dimensional separable Hilbert space and that $\{\, e_n \,\}_{n\in\mathbb{N}}$ is an orthonormal basis adapted to a compact self-adjoint operator T on X. Then the map $U : X \to l_2$ given by

$$U(\sum_n c_n e_n) = (c_0, c_1, \dots),$$

is an isometric isomorphism. Moreover, when we use this map to transfer the action of T to l_2, i.e., when we consider the operator UTU^{-1} on l_2, we see that this operator is simply multiplication by the bounded sequence $(\lambda_0, \lambda_1, \dots) \in l_\infty$. Thus the spectral theorem says that every compact self-adjoint T is unitarily equivalent to a multiplication operator on l_2. (An isometric isomorphism of Hilbert spaces is also called a unitary operator. Note that it is characterized by the property $U^* = U^{-1}$.)

A useful extension is the spectral theorem for commuting self-adjoint compact operators.

Theorem. *If T and S are self-adjoint compact operators in a Hilbert space H and $TS = ST$, then there is an orthonormal basis of X whose elements are eigenvectors for both S and T.*

Proof. For an eigenvalue λ of T, let X_λ denote the corresponding eigenspace of T. If $x \in X_\lambda$, then $TSx = STx = \lambda Sx$, so $Sx \in X_\lambda$. Thus S restricts to a self-adjoint operator on X_λ, and so there is an orthonormal basis of S–eigenvectors for X_λ. These are T–eigenvectors as well. Taking the union over all the eigenvalues λ of T completes the construction. $\square$

Let T_1 and T_2 be any two self-adjoint operators and set $T = T_1 + iT_2$. Then $T_1 = (T + T^*)/2$ and $T_2 = (T - T^*)/(2i)$. Conversely, if T is any element of $B(X)$, then we can define two self-adjoint operators from these formulas and have $T = T_1 + iT_2$. Now suppose that T is compact and also *normal*, i.e., that T and T^* commute. Then T_1 and T_2 are compact and commute, and hence we have an orthonormal basis whose elements are eigenvectors for both T_1 and T_2, and hence for T. Since the real and imaginary parts of the eigenvalues are the eigenvalues of T_1 and T_2, we again see that the eigenvalues form a sequence tending to zero and all have finite dimensional eigenspaces.

We have thus shown that a compact normal operator admits an orthonormal basis of eigenvectors. Conversely, if $\{ e_i \}$ is an orthonormal basis of eigenvectors of T, then $\langle T^* e_i, e_j \rangle = 0$ if $i \neq j$, which implies that each e_i is also an eigenvector for T^*. Thus $T^*Te_i = TT^*e_i$ for all i, and it follows easily that T is normal. We have thus shown:

Spectral Theorem for compact normal operators. *Let T be a compact operator on a Hilbert space X. Then there exists an orthonormal basis for X consisting of eigenvectors of T if and only if T is normal. In this case, the set of nonzero eigenvalues form a finite set or a sequence tending to zero and the eigenspaces corresponding to the nonzero eigenvalues are finite dimensional. The eigenvalues are all real if and only if the operator is self-adjoint.*

The spectrum of a general compact operator. In this section we derive the structure of the spectrum of a compact operator (not necessarily self-adjoint or normal) on a complex Banach space X.

For any operator T on a complex Banach space, the *resolvent set* of T, $\rho(T)$ consists of those $\lambda \in \mathbb{C}$ such that $T - \lambda \mathbf{1}$ is invertible, and the spectrum $\sigma(T)$ is the complement. If $\lambda \in \sigma(T)$, then $T - \lambda\mathbf{1}$ may fail to be invertible in several ways. (1) It may be that $\mathcal{N}(T - \lambda\mathbf{1}) \neq 0$, i.e., that λ is an eigenvalue of T. In this case we say that λ belongs to the *point spectrum* of T, denoted $\sigma_p(T)$. (2) If $T - \lambda\mathbf{1}$ is injective, it may be that its range is dense but not closed in X. In this case we say that λ belongs to the *continuous spectrum* of T, $\sigma_c(T)$. Or (3) it may be that $T - \lambda\mathbf{1}$ is injective but that its range is not even dense in X. This is the *residual spectrum*, $\sigma_r(T)$. Clearly we have a decomposition of $\mathbb{C}$ into the disjoint sets $\rho(T)$, $\sigma_p(T)$, $\sigma_c(T)$, and $\sigma_r(T)$. As an example of the continuous spectrum, consider the operator $Te_n = \lambda_n e_n$ where the e_n form an orthonormal basis of a Hilbert space and the λ_n form a positive sequence tending to 0. Then $0 \in \sigma_c(T)$. If $Te_n = \lambda_n e_{n+1}$, $0 \in \sigma_r(T)$.

Now if T is compact and X is infinite dimensional, then $0 \in \sigma(T)$ (since if T were invertible, the image of the unit ball would contain an open set, and so couldn't be pre-compact). From the examples just given, we see that 0 may belong to the point spectrum, the continuous spectrum, or the residual spectrum. However, we shall show that all other elements of the spectrum are eigenvalues, i.e., that $\sigma(T) = \sigma_p(T) \cup \{0\}$, and that, as in the normal case, the point spectrum consists of a finite set or a sequence approaching zero.

The structure of the spectrum of a compact operator will be deduced from two lemmas. The first is purely algebraic. To state it we need some terminology: consider a linear operator T from a vector space X to itself, and consider the chains of subspaces

$$0 = \mathcal{N}(\mathbf{1}) \subset \mathcal{N}(T) \subset \mathcal{N}(T^2) \subset \mathcal{N}(T^3) \subset \cdots .$$

Either this chain is strictly increasing forever, or there is a least $n \geq 0$ such that $\mathcal{N}(T^n) = \mathcal{N}(T^{n+1})$, in which case only the first n spaces are distinct and all the others equal the nth one. In the latter case we say that the kernel chain for T *stabilizes* at n. In particular, the kernel chain stabilizes at 0 iff T is injective. Similarly we may consider the chain

$$X = \mathcal{R}(\mathbf{1}) \supset \mathcal{R}(T) \supset \mathcal{R}(T^2) \supset \mathcal{R}(T^3) \supset \cdots ,$$

and define what it means for the range chain to stabilize at $n > 0$. (So the range stabilizes at 0 iff T is surjective.) It could happen that neither or only one of these chains stabilizes. However:

Lemma. *Let T be a linear operator from a vector space X to itself. If the kernel chain stabilizes at m and the range chain stabilizes at n, then $m = n$ and X decomposes as the direct sum of $\mathcal{N}(T^n)$ and $\mathcal{R}(T^n)$.*

Proof. Suppose m were less than n. Since the range chain stabilizes at n, there exists x with $T^{n-1}x \notin \mathcal{R}(T^n)$, and then there exists y such that $T^{n+1}y = T^n x$. Thus $x - Ty \in \mathcal{N}(T^n)$, and, since kernel chain stabilizes at $m < n$, $\mathcal{N}(T^n) = \mathcal{N}(T^{n-1})$. Thus $T^{n-1}x = T^n y$, a contradiction. Thus $m \geq n$. A similar argument, left to the reader, establishes the reverse inequality.

Now if $T^n x \in \mathcal{N}(T^n)$, then $T^{2n}x = 0$, whence $T^n x = 0$. Thus $\mathcal{N}(T^n) \cap \mathcal{R}(T^n) = 0$. Given x, let $T^{2n}y = T^n x$, so x decomposes as $T^n y \in \mathcal{R}(T^n)$ and $x - T^n y \in \mathcal{N}(T^n)$. $\square$

The second lemma brings in the topology of compact operators.

Lemma. *Let $T : X \to X$ be a compact operator on a Banach space and $\lambda_1, \lambda_2, \ldots$ a sequence of complex numbers with $\inf |\lambda_n| > 0$. Then the following is impossible: There exists a strictly increasing chain of closed subspaces $S_1 \subset S_2 \subset \cdots$ with $(\lambda_n \mathbf{1} - T)S_n \subset S_{n-1}$ for all n.*

Proof. Suppose such a chain exists. Note that each $TS_n \subset S_n$ for each n. Since S_n/S_{n-1} contains an element of norm 1, we may choose $y_n \in S_n$ with $\|y_n\| \leq 2$, $\operatorname{dist}(y_n, S_{n-1}) = 1$. If $m < n$, then

$$z := \frac{Ty_m - (\lambda_n \mathbf{1} - T)y_n}{\lambda_n} \in S_{n-1},$$

and

$$\|Ty_m - Ty_n\| = |\lambda_n| \|y_n - z_n\| \geq |\lambda_n|.$$

This implies that the sequence (Ty_n) has no Cauchy subsequence, which contradicts the compactness of T. $\square$

We are now ready to prove the result quoted at the beginning of the subsection.

Theorem. *Let T be a compact operator on a Banach space X. Then any nonzero element of the spectrum of T is an eigenvalue. Moreover $\sigma(T)$ is either finite or a sequence approaching zero.*

Proof. Consider the subspace chains $\mathcal{N}[(\lambda\mathbf{1} - T)^n]$ and $\mathcal{R}[(\lambda\mathbf{1} - T)^n]$ (these are closed subspaces by a previous result). Clearly $\lambda\mathbf{1} - T$ maps $\mathcal{N}[(\lambda\mathbf{1} - T)^n]$ into $\mathcal{N}[(\lambda\mathbf{1} - T)^{n-1}]$, so the previous lemma implies that the kernel chain stabilizes, say at n. Now $\mathcal{R}[(\lambda\mathbf{1} - T)^n] = {}^a\mathcal{N}[(\lambda\mathbf{1} - T^*)^n]$ (since the range is closed), and since these last stabilize, the range chain stabilizes as well.

Thus we have $X = \mathcal{N}[(\lambda\mathbf{1} - T)^n] \oplus \mathcal{R}[(\lambda\mathbf{1} - T)^n]$. Thus

$$\mathcal{R}(\lambda\mathbf{1} - T) \neq X \implies \mathcal{R}(\lambda\mathbf{1} - T)^n \neq X \implies \mathcal{N}(\lambda\mathbf{1} - T)^n \neq 0 \implies \mathcal{N}(\lambda\mathbf{1} - T) \neq 0.$$

In other words $\lambda \in \sigma(T) \implies \lambda \in \sigma_p(T)$.

Finally we prove the last statement. If it were false we could find a sequence of eigenvalues λ_n with $\inf |\lambda_n| > 0$. Let $x_1, x_2, \ldots$ be corresponding nonzero eigenvectors and set $S_n = \operatorname{span}[x_1, \ldots, x_n]$. These form a strictly increasing chain of subspaces (recall that eigenvectors corresponding to distinct eigenvalues are linearly independent) and $(\lambda_n\mathbf{1} - T)S_n \subset S_{n-1}$, which contradicts the lemma. $\square$

The above reasoning also gives us the Fredholm alternative:

Theorem. *Let T be a compact operator on a Banach space X and λ a nonzero complex number. Then either (1) $\lambda\mathbf{1} - T$ is an isomorphism, or (2) it is neither injective nor surjective.*

Proof. Since the kernel chain and range chain for $S = \lambda\mathbf{1} - T$ stabilize, either they both stabilize at 0, in which case S is injective and surjective, or neither does, in which case it is neither. $\square$

We close this section with a result which is fundamental to the study of Fredholm operators.

Theorem. *Let T be a compact operator on a Banach space X and λ a nonzero complex number. Then*

$$\dim \mathcal{N}(\lambda\mathbf{1} - T) = \dim \mathcal{N}(\lambda\mathbf{1} - T^*) = \operatorname{codim} \mathcal{R}(\lambda\mathbf{1} - T) = \operatorname{codim} \mathcal{R}(\lambda\mathbf{1} - T^*).$$

Proof. Let $S = \lambda\mathbf{1} - T$. Since $\mathcal{R}(S)$ is closed

$$[X/\mathcal{R}(S)]^* \cong \mathcal{R}(S)^a = \mathcal{N}(S^*).$$

Thus $[X/\mathcal{R}(S)]^*$ is finite dimensional, so $X/\mathcal{R}(S)$ is finite dimensional, and these two spaces are of the same dimension. Thus $\operatorname{codim}\mathcal{R}(S) = \dim\mathcal{N}(S^*)$.

For a general operator S we only have $\overline{\mathcal{R}(S^*)} \subset \mathcal{N}(S)^a$, but, as we now show, when $\mathcal{R}(S)$ is closed, $\mathcal{R}(S^*) = \mathcal{N}(S)^a$. Indeed, S induces an isomorphism of $X/\mathcal{N}(S)$ onto $\mathcal{R}(S)$, and for any $f \in \mathcal{N}(S)^a$, f induces a map $X/\mathcal{N}(S)$ to $\mathbb{R}$. It follows that $f = gS$ for some bounded linear operator g on $\mathcal{R}(S)$, which can be extended to an element of X^* by Hahn-Banach. But $f = gS$ simply means that $f = S^*g$, showing that $\mathcal{N}(S)^a \subset \mathcal{R}(S^*)$ (and so equality holds) as claimed.

Thus

$$\mathcal{N}(S)^* \cong X^*/\mathcal{N}(S)^a = X^*/\mathcal{R}(S^*),$$

so $\operatorname{codim}\mathcal{R}(S^*) = \dim\mathcal{N}(S)^* = \dim\mathcal{N}(S)$.

We complete the theorem by showing that $\dim\mathcal{N}(S) \le \operatorname{codim}\mathcal{R}(S)$ and $\dim\mathcal{N}(S^*) \le \operatorname{codim}\mathcal{R}(S^*)$. Indeed, since $\mathcal{R}(S)$ is closed with finite codimension, it is complemented by a finite dimensional space M (with $\dim M = \operatorname{codim}\mathcal{R}(S)$. Since $\mathcal{N}(S)$ is finite dimensional, it is complemented by a space N. Let P denote the projection of X onto $\mathcal{N}(S)$ which is a bounded map which to the identity on $\mathcal{N}(S)$ and to zero on N. Now if $\operatorname{codim}\mathcal{R}(S) < \dim\mathcal{N}(S)$, then there is a linear map of $\mathcal{N}(S)$ onto M which is not injective. But then $T - fP$ is a compact operator and $\lambda\mathbf{1} - T + fP$ is easily seen to be surjective. By the Fredholm alternative, it is injective as well. This implies that f is injective, a contradiction. We have thus shown that $\dim\mathcal{N}(S) \le \operatorname{codim}\mathcal{R}(S)$. Since T^* is compact, the same argument shows that $\dim\mathcal{N}(S^*) \le \operatorname{codim}\mathcal{R}(S^*)$. This completes the proof. $\square$

VI. Introduction to General Spectral Theory

In this section we skim the surface of the spectral theory for a general (not necessarily compact) operator on a Banach space, before encountering a version of the Spectral Theorem for a bounded self-adjoint operator in Hilbert space. Our first results don't require the full structure of in the algebra of operators on a Banach space, but just an arbitrary Banach algebra structure, and so we start there.

The spectrum and resolvent in a Banach algebra. Let X be a Banach algebra with an identity element denoted $\mathbf{1}$. We assume that the norm in X has been normalized so that $\|\mathbf{1}\| = 1$. The two main examples to bear in mind are (1) $B(X)$, where X is some Banach space; and (2) $C(G)$ endowed with the sup norm, where G is some compact topological space, the multiplication is just pointwise multiplication of functions, and $\mathbf{1}$ is the constant function 1.

In this set up the resolvent set and spectrum may be defined as before: $\rho(x) = \{\, \lambda \in \mathbb{C} \mid x - \lambda\mathbf{1} \text{ is invertible} \,\}$, $\sigma(x) = \mathbb{C} \setminus \rho(x)$. The *spectral radius* is defined to be $r(x) = \sup|\sigma(x)|$. For $\lambda \in \rho(x)$, the resolvent is defined as $R_x(\lambda) = (x - \lambda\mathbf{1})^{-1}$.

Lemma. *If $x, y \in X$ with x invertible and $\|x^{-1}y\| < 1$, then $x - y$ is invertible,*

$$(x-y)^{-1} = \sum_{n=0}^{\infty} (x^{-1}y)^n x^{-1},$$

and $\|(x-y)^{-1}\| \le \|x^{-1}\|/(1 - \|x^{-1}y\|)$.

Proof.

$$\|\sum (x^{-1}y)^n x^{-1}\| \le \|x^{-1}\| \sum \|x^{-1}y\|^n \le \|x^{-1}\|/(1 - \|x^{-1}y\|),$$

so the sum converges absolutely and the norm bound holds. Also

$$\sum_{n=0}^{\infty} (x^{-1}y)^n x^{-1}(x-y) = \sum_{n=0}^{\infty} (x^{-1}y)^n - \sum_{n=0}^{\infty} (x^{-1}y)^{n+1} = \mathbf{1},$$

and similarly for the product in the reverse order. $\square$

As a corollary, we see that if $|\lambda| > \|x\|$, then $\lambda \mathbf{1} - x$ is invertible, i.e., $\lambda \in \rho(x)$. In other words:

Proposition. *$r(x) \le \|x\|$.*

We also see from the lemma that $\lim_{\lambda \to \infty} \|R_x(\lambda)\| = 0$. Another corollary is that if $\lambda \in \rho(x)$ and $|\mu| < \|R_x(\lambda)\|^{-1}$, then $\lambda - \mu \in \rho(x)$ and

$$R_x(\lambda - \mu) = \sum_{n=0}^{\infty} R_x(\lambda)^{n+1} \mu^n.$$

Theorem. *The resolvent $\rho(x)$ is always open and contains a neighborhood of ∞ in $\mathbb{C}$ and the spectrum is always non-empty and compact.*

Proof. The above considerations show that the resolvent is open, and so the spectrum is closed. It is also bounded, so it is compact.

To see that the spectrum is non-empty, let $f \in X^*$ be arbitrary and define $\phi(\lambda) = f[R_x(\lambda)]$. Then ϕ maps $\rho(x)$ into $\mathbb{C}$, and it is easy to see that it is holomorphic (since we have the power series expansion

$$\phi(\lambda - \mu) = \sum_{n=0}^{\infty} f[R(\lambda)^{n+1}] \mu^n$$

if μ is sufficiently small). If $\sigma(x) = \emptyset$, then ϕ is entire. It is also bounded (since it tends to 0 at infinity), so Liouville's theorem implies that it is identically zero. Thus for any $f \in X^*$, $f[(\lambda \mathbf{1} - x)^{-1}] = 0$. This implies that $(\lambda \mathbf{1} - x)^{-1} = 0$, which is clearly impossible. $\square$

Corollary (Gelfand–Mazur). *If X is a complex Banach division algebra, then X is isometrically isomorphic to $\mathbb{C}$.*

Proof. For each $0 \neq x \in X$, let $\lambda \in \sigma(x)$. Then $x - \lambda\mathbf{1}$ is not invertible, and since X is a division algebra, this means that $x = \lambda\mathbf{1}$. Thus $X = \mathbb{C}\mathbf{1}$. $\square$

Now we turn to a bit of "functional calculus." Let $x \in X$ and let f be a complex function of a complex variable which is holomorphic on the closed disk of radius $\|x\|$ about the origin. Then we make two claims: (1) plugging x into the power series expansion of f defines an element $f(x) \in X$; and (2) the complex function f maps the spectrum of x into the spectrum of $f(x)$. (In fact onto, as we shall show later in the case f is polynomial.) To prove these claims, note that, by assumption, the radius of convergence of the power series for f about the origin exceeds $\|x\|$, so we can expand $f(z) = \sum_{n=0}^{\infty} a_n z^n$ where $\sum |a_n|\|x\|^n < \infty$. Thus the series $\sum a_n x^n$ is absolutely convergent in the Banach space X; we call its limit $f(x)$. (This is the *definition* of $f(x)$. It is a suggestive abuse of notation to use f to denote the this function, which maps a subset of X into X, as well as the original complex-valued function of a complex variable.) Now suppose that $\lambda \in \sigma(x)$. Then

$$f(\lambda)\mathbf{1} - f(x) = \sum_{n=1}^{\infty} a_n(\lambda^n\mathbf{1} - x^n) = (\lambda\mathbf{1} - x)\sum_{n=1}^{\infty} a_n P_n = \sum_{n=1}^{\infty} a_n P_n(\lambda\mathbf{1} - x),$$

where

$$P_n = \sum_{k=0}^{n-1} \lambda^k x^{n-k-1}.$$

Note that $\|P_n\| \leq n\|x\|^{n-1}$, so $\sum_{n=1}^{\infty} a_n P_n$ converges to some $y \in X$. Thus

$$f(\lambda)\mathbf{1} - f(x) = (\lambda\mathbf{1} - x)y = y(\lambda\mathbf{1} - x).$$

Now $f(\lambda)\mathbf{1} - f(x)$ can't be invertible, because these formulas would then imply that $\lambda\mathbf{1} - x$ would be invertible as well, but $\lambda \in \sigma(x)$. Thus we have verified that $f(\lambda) \in \sigma\bigl(f(x)\bigr)$ for all $\lambda \in \sigma(x)$.

Theorem (Spectral Radius Formula). $r(x) = \lim_{n\to\infty} \|x^n\|^{1/n} = \inf_n \|x^n\|^{1/n}$.

Proof. If $\lambda \in \sigma(x)$, then $\lambda^n \in \sigma(x^n)$ (which is also evident algebraically), so $|\lambda^n| \leq \|x^n\|$. This shows that $r(x) \leq \inf_n \|x^n\|^{1/n}$.

Now take $f \in X^*$, and consider

$$\phi(\lambda) = f[(\lambda I - x)^{-1}] = \sum_{n=0}^{\infty} \lambda^{-n-1} f(x^n).$$

Then ϕ is clearly holomorphic for $\lambda > \|x\|$, but we know it extends holomorphically to $\lambda > r(x)$ and tends to 0 as λ tends to infinity. Let $\psi(\lambda) = \phi(1/\lambda)$. Then ψ extends analytically to zero with value zero and defines an analytic function on the open ball of radius $1/r(x)$ about zero, as does, therefore,

$$\psi(\lambda)/\lambda = \sum_{n=0}^{\infty} f(\lambda^n x^n).$$

This shows that for each $|\lambda| < 1/r(x)$ and each $f \in X^*$, $f(\lambda^n x^n)$ is bounded. By the uniform boundedness principle, the set of elements $\lambda^n x^n$ are bounded in X, say by K. Thus $\|x^n\|^{1/n} \leq K^{1/n}/|\lambda| \to 1/|\lambda|$. This is true for all $|\lambda| < 1/r(x)$, so $\limsup \|x^n\|^{1/n} \leq r(x)$. $\square$

Corollary. *If H is a Hilbert space and $T \in B(H)$ a normal operator, then $r(T) = \|T\|$.*

Proof.

$$\|T\|^2 = \sup_{\|x\|\leq 1} \langle Tx, Tx\rangle = \sup_{\|x\|\leq 1} \langle T^*Tx, x\rangle = \|T^*T\|,$$

since T^*T is self-adjoint. Using the normality of T we also get

$$\begin{aligned}\|T^*T\|^2 = \sup_{\|x\|\leq 1} \langle T^*Tx, T^*Tx\rangle &= \sup_{\|x\|\leq 1} \langle TT^*Tx, Tx\rangle = \sup_{\|x\|\leq 1} \langle T^*T^2x, Tx\rangle \\ &= \sup_{\|x\|\leq 1} \langle T^2x, T^2x\rangle = \|T^2\|^2.\end{aligned}$$

Thus $\|T\|^2 = \|T^2\|$. Replacing T with T^2 gives, $\|T\|^4 = \|T^4\|$, and similarly for all powers of 2. The result thus follows from the spectral radius formula. $\square$

As mentioned, we can now show that p maps $\sigma(x)$ *onto* $\sigma\big(p(x)\big)$ if p is a polynomial.

Spectral Mapping Theorem. *Let X be a complex Banach algebra with identity, $x \in X$, and let p a polynomial in one variable with complex coefficients. Then $p\big(\sigma(x)\big) = \sigma\big(p(x)\big)$.*

Proof. We have already shown that $p\big(\sigma(x)\big) \subset \sigma\big(p(x)\big)$. Now suppose that $\lambda \in \sigma\big(p(x)\big)$. By the Fundamental Theorem of Algebra we can factor $p - \lambda$, so

$$p(x) - \lambda\mathbf{1} = a\Pi_{i=1}^{n}(x - \lambda_i\mathbf{1}),$$

for some nonzero $a \in \mathbb{C}$ and some roots $\lambda_i \in \mathbb{C}$. Since $p(x) - \lambda\mathbf{1}$ is not invertible, it follows that $x - \lambda_i\mathbf{1}$ is not invertible for at least one i. In orther words, $\lambda_i \in \sigma(x)$, so $\lambda = p(\lambda_i) \in p\big(\sigma(x)\big)$. $\square$

Spectral Theorem for bounded self-adjoint operators in Hilbert space. We now restrict to self-adjoint operators on Hilbert space and close with a version of the Spectral Theorem for this class of operator. We follow Halmos's article "What does the Spectral Theorem Say?" (American Mathematical Monthly 70, 1963) both in the relatively elementary statement of the theorem and the outline of the proof.

First we note that self-adjoint operators have real spectra (not just real eigenvalues).

Proposition. *If H is a Hilbert space and $T \in B(H)$ is self-adjoint, then $\sigma(T) \subset \mathbb{R}$.*

Proof.

$$|\langle(\lambda\mathbf{1} - T)x, x\rangle| \geq |\operatorname{Im}\langle(\lambda\mathbf{1} - T)x, x\rangle| = |\operatorname{Im}\lambda|\|x\|^2,$$

so if $\operatorname{Im}\lambda \neq 0$, $\lambda\mathbf{1} - T$ is injective with closed range. The same reasoning shows that $(\lambda\mathbf{1} - T)^* = \bar{\lambda}\mathbf{1} - T$ is injective, so $\mathcal{R}(\lambda\mathbf{1} - T)$ is dense. Thus $\lambda \in \rho(T)$. $\square$

Spectral Theorem for self-adjoint operators in Hilbert space. *If H is a complex Hilbert space and $T \in B(H)$ is self-adjoint, then there exists a measure space Ω with measure μ, a bounded measurable function $\phi : \Omega \to \mathbb{R}$, and an isometric isomorphism $U : L^2 \to H$ such that*

$$U^{-1}TU = M_\phi$$

where $M_\phi : L^2 \to L^2$ is the operation of multiplication by ϕ. (Here L^2 means $L^2(\Omega, \mu; \mathbb{C})$, the space of complex-valued functions on Ω which are square integrable with respect to the measure μ.)

Sketch of proof. Let x be a nonzero element of H, and consider the smallest closed subspace M of H containing $T^n x$ for $n = 0, 1, \ldots$, i.e., $M = \overline{\{\, p(T)x \mid p \in \mathbb{P}_\mathbb{C} \,\}}$. Here $\mathbb{P}_\mathbb{C}$ is the space of polynomials in one variable with complex coefficients. Both M and its orthogonal complement are invariant under T (this uses the self-adjointness of T). By a straightforward application of Zorn's lemma we see that H can be written as a Hilbert space direct sum of T invariant spaces of the form of M. If we can prove the theorem for each of these subspaces, we can take direct products to get the result for all of H. Therefore we may assume from the start that $H = \overline{\{\, p(T)x \mid p \in \mathbb{P}_\mathbb{C} \,\}}$ for some x. (In other terminology, that T has a *cyclic vector* x.)

Now set $\Omega = \sigma(T)$, which is a compact subset of the real line, and consider the space $C = C(\Omega, \mathbb{R})$, the space of all continuous real-valued functions on Ω. The subspace of real-valued polynomial functions is dense in C (since any continuous function on Ω can be extended to the interval $[-r(T), r(T)]$ thanks to Tietze's extension theorem and then approximated arbitrarily closely by a polynomial thanks to the Weierstrass approximation theorem). For such a polynomial function, p, define $Lp = \langle p(T)x, x\rangle \in \mathbb{R}$. Clearly L is linear and

$$|Lp| \leq \|p(T)\|\|x\|^2 = r\big(p(T)\big)\|x\|^2,$$

by the special form of the spectral radius formula for self-adjoint operators in Hilbert space. Since $\sigma\big(p(T)\big) = p\big(\sigma(T)\big)$, we have $r\big(p(T)\big) = \|p\|_{L^\infty(\Omega)} = \|p\|_C$, and thus,

$$|Lp| \leq \|x\|^2\|p\|_C.$$

This shows that L is a bounded linear functional on a dense subspace of C and so extends uniquely to define a bounded linear functional on C.

Next we show that L is positive in the sense that $Lf \geq 0$ for all non-negative functions $f \in C$. Indeed, if $f = p^2$ for some polynomial, then

$$Lf = \langle p(T)^2 x, x \rangle = \langle p(T)x, p(T)x \rangle \geq 0.$$

For an arbitrary non-negative f, we can approximate $\sqrt{f}$ uniformly by polynomials p_n, so $f = \lim p_n^2$ and $Lf = \lim Lp_n^2 \geq 0$.

We now apply the Riesz Representation Theorem for the representation of the linear functional L on C. It state that there exists a finite measure on Ω such that $Lf = \int f \, d\mu$ for $f \in \mathbb{C}$ (it is a positive measure since L is positive). In particular, $\langle p(T)x, x \rangle = \int p \, d\mu$ for all $p \in \mathbb{P}_{\mathbb{R}}$.

We now turn to the space L^2 of complex-valued functions on Ω which are square integrable with respect to the measure μ. The subspace of complex-valued polynomial functions is dense in L^2 (since the measure is finite, the L^2 norm is dominated by the supremum norm). For such a polynomial function, q, define $Uq = q(T)x$. Then

$$\|Uq\|^2 = \|q(T)x\|^2 = \langle q(T)x, q(T)x \rangle = \langle \bar{q}(T)q(T)x, x \rangle = \int |q|^2 \, d\mu = \|q\|_{L^2}^2.$$

Thus U is an isometry of a dense subspace of L^2 into H and so extends to an isometry of L^2 onto a closed subspace of H. In fact, U is onto H itself, since, by the assumption that x is a cyclic vector for T, the range of U is dense.

Finally, define $\phi : \Omega \to \mathbb{R}$ by $\phi(\lambda) = \lambda$. If q is a complex polynomial, then $(M_\phi q)(\lambda) = \lambda q(\lambda)$, which is also a polynomial. Thus

$$U^{-1}TUq = U^{-1}Tq(T)x = U^{-1}\big((M_\phi q)(T)\big)x = M_\phi q.$$

Thus the bounded operators $U^{-1}TU$ and M_ϕ coincide on a dense subset of L^2, and hence they are equal. $\square$

For a more precise description of the measure space and the extension to normal operators, see Zimmer.