(*April 3, 2019*)

09b. Compact operators

Paul Garrett garrett@math.umn.edu http://www.math.umn.edu/˜garrett/

[This document is
http://www.math.umn.edu/˜garrett/m/real/notes_2018-19/09b_compact_operators.pdf]

1. Compact operators
2. Hilbert-Schmidt operators
3. Spectral Theorem for self-adjoint compact operators
4. The Fredholm alternative
5. Simplest Rellich compactness lemma
• Appendix: topologies on finite-dimensional spaces

Among all linear operators on Hilbert spaces, the *compact* ones (defined below) are the simplest, and most closely imitate finite-dimensional operator theory. In addition, compact operators are important in practice. We prove a spectral theorem for *self-adjoint compact* operators, which does *not* use broader discussions of properties of spectra, only using the *Cauchy-Schwarz-Bunyakowsky inequality* and the *definition* of self-adjoint compact operator. The argument follows the Rayleigh-Ritz argument for finite-dimensional self-adjoint operators.

The simplest naturally occurring compact operators are the *Hilbert-Schmidt* operators, discussed below.

---

1. Compact operators

A set in a topological space is *pre-compact* if its closure is compact. [1] A linear operator $T : X \to Y$ on Hilbert spaces is *compact* when it maps the unit ball in X to a *pre-compact* set in Y. Equivalently, T is compact if and only if it maps *bounded* sequences in X to sequences in Y with *convergent subsequences*.

[1.1] Remark: The same definition makes sense for operators on *Banach* spaces, but many good features of compact operators on Hilbert spaces are not shared by compact operators on Banach spaces.

[1.2] Proposition: An operator-norm limit of compact operators is compact.

Proof: Let $T_n \to T$ in uniform operator norm, with compact T_n. Given $\varepsilon > 0$, let n be sufficiently large such that $|T_n - T| < \varepsilon/2$. Since $T_n(B)$ is pre-compact, there are finitely many $y_1, \ldots, y_t$ such that for any $x \in B$ there is i such that $|T_n x - y_i| < \varepsilon/2$. By the triangle inequality

$$|Tx - y_i| \;\le\; |Tx - T_n x| + |T_n x - y_i| \;<\; \varepsilon$$

Thus, $T(B)$ is covered by finitely many balls of radius ε. ///

A continuous linear operator is of *finite rank* if its image is finite-dimensional. A finite-rank operator is *compact*, since all balls are pre-compact in a finite-dimensional Hilbert space.

[1.3] Theorem: A compact operator $T : X \to Y$ with X, Y Hilbert spaces is an operator norm limit of *finite rank* operators.

Proof: Let B be the closed unit ball in X. Since $T(B)$ is pre-compact it is totally bounded, so for given $\varepsilon > 0$ cover $T(B)$ by open balls of radius ε centered at points $y_1, \ldots, y_n$. Let p be the orthogonal projection to the finite-dimensional subspace F spanned by the y_i and define $T_\varepsilon = p \circ T$. Note that for any $y \in Y$ and for any y_i

$$|p(y) - y_i| \le |y - y_i|$$

[1] Beware, sometimes *pre-compact* has a more restrictive meaning than having compact closure.

since $y = p(y) + y'$ with y' orthogonal to all y_i. For x in X with $|x| \le 1$, by construction there is y_i such that $|Tx - y_i| < \varepsilon$. Then

$$|Tx - T_\varepsilon x| \le |Tx - y_i| + |T_\varepsilon x - y_i| < \varepsilon + \varepsilon$$

Thus, $T_\varepsilon T$ in operator norm as $\varepsilon \to 0$. ///

[1.4] Remark: An operator that is an operator-norm limit of finite-rank operators is sometimes called *completely continuous*. Thus, we see that for operators in Hilbert spaces, the class of *compact* operators is the same as that of *completely continuous* operators.

[1.5] Remark: The equivalence of compactness and complete continuity is false in Banach spaces, although the only example known to this author (Per Enflo, *Acta Math.*, vol. 130, 1973) is complicated.

2. *Hilbert-Schmidt operators*

[2.1] Hilbert-Schmidt operators given by integral kernels

Originally *Hilbert-Schmidt* operators on function spaces $L^2(X)$ arose as operators given by *integral kernels*: for X and Y σ-finite measure spaces, and for integral kernel $K \in L^2(X \times Y)$, the associated *Hilbert-Schmidt* operator [2]

$$T : L^2(X) \longrightarrow L^2(Y)$$

is

$$Tf(y) = \int_X K(x,y)\, f(x)\, dx$$

By Fubini's theorem and the σ-finiteness, for orthonormal bases φ_α for $L^2(X)$ and ψ_β for $L^2(Y)$, the collection of functions $\varphi_\alpha(x)\psi_\beta(y)$ is an orthonormal basis for $L^2(X \times Y)$. Thus, for some scalars c_{ij},

$$K(x,y) = \sum_{ij} c_{ij}\, \overline{\varphi_i}(x)\, \psi_j(y)$$

Square-integrability is

$$\sum_{ij} |c_{ij}|^2 = |K|^2_{L^2(X\times Y)} < \infty$$

The indexing sets may as well be countable, since an uncountable sum of positive reals cannot converge. Given $f \in L^2(X)$, the image Tf is in $L^2(Y)$, since

$$Tf(y) = \sum_{ij} c_{ij} \langle f, \varphi_i\rangle\, \psi_j(y)$$

has $L^2(Y)$ norm easily estimated by

$$|Tf|^2_{L^2(Y)} \le \sum_{ij} |c_{ij}|^2 |\langle f, \varphi_i\rangle|^2\, |\psi_j|^2_{L^2(Y)} \le |f|^2_{L^2(X)} \sum_{ij} |c_{ij}|^2\, |\varphi_i|^2_{L^2(X)}\, |\psi_j|^2_{L^2(Y)}$$

$$= |f|^2_{L^2(X)} \sum_{ij} |c_{ij}|^2 = |f|^2_{L^2(X)} \cdot |K|^2_{L^2(X\times Y)}$$

The adjoint $T^* : L^2(Y) \to L^2(X)$ has kernel

$$K^*(y,x) = \overline{K(x,y)}$$

[2] The σ-finiteness is necessary to make Fubini's theorem work as expected.

by computing

$$\langle Tf,g\rangle_{L^2(Y)} \;=\; \int_Y \Big(\int_X K(x,y)f(x)\,dx\Big)\,\overline{g(y)}\;dy \;=\; \int_X f(x)\overline{\Big(\int_Y \overline{K(x,y)}\,g(y)\;dy\Big)}\;dx$$

[2.2] Intrinsic characterization of Hilbert-Schmidt operators

The *intrinsic* characterization of Hilbert-Schmidt operators $V \to W$ on Hilbert spaces V, W is as the *completion* of the space of *finite-rank* operators $V \to W$ with respect to the *Hilbert-Schmidt norm*, whose square is

$$|T|^2_{\mathrm{HS}} \;=\; \mathrm{tr}(T^*T) \qquad\qquad (\text{for } T : V \to W \text{ and } T^* : W^* \to V^*)$$

The *trace* of a finite-rank operator from a Hilbert space to itself can be described in coordinates and then proven independent of the choice of coordinates, or trace can be described *intrinsically*, obviating need for proof of coordinate-independence. First, in coordinates, for an orthonormal basis e_i of V, and finite-rank $T : V \to V$, define

$$\mathrm{tr}(T) \;=\; \sum_i \langle Te_i, e_i\rangle \qquad\qquad (\text{with reference to orthonormal basis } \{e_i\})$$

With this description, one would need to show independence of the orthonormal basis. For the intrinsic description, consider the map from $V \otimes V^*$ to finite-rank operators on V induced from the bilinear map [3]

$$v \times \lambda \;\longrightarrow\; \Big(w \to \lambda(w)\cdot v\Big) \qquad\qquad (\text{for } v \in V \text{ and } \lambda \in V^*)$$

Trace is easy to define in these terms [4]

$$\mathrm{tr}(v \otimes \lambda) \;=\; \lambda(v)$$

and

$$\mathrm{tr}\Big(\sum_{v,\lambda} v \otimes \lambda\Big) \;=\; \sum_{v,\lambda} \lambda(v) \qquad\qquad (\text{finite sums})$$

Expression of *trace* in terms of an orthonormal basis $\{e_j\}$ is easily obtained from the intrinsic form: given a finite-rank operator T and an orthonormal basis $\{e_i\}$, let $\lambda_i(v) = \langle v, e_i\rangle$. We claim that

$$T \;=\; \sum_i Te_i \otimes \lambda_i$$

Indeed,

$$\Big(\sum_i Te_i \otimes \lambda_i\Big)(v) \;=\; \sum_i Te_i \cdot \lambda_i(v) \;=\; \sum_i Te_i \cdot \langle v, e_i\rangle \;=\; T\Big(\sum_i e_i \cdot \langle v, e_i\rangle\Big) \;=\; Tv$$

---

[3] The intrinsic characterization of the tensor product $V \otimes_k W$ of two k-vectorspaces is that it is a k-vectorspace with a k-bilinear map $b : V \times W \to V \otimes_k W$ such that for any k-bilinear map $B : V \times W \to X$ there is a unique *linear* $\beta : V \otimes W \to X$ giving a commutative diagram

$$\begin{tikzcd} V \otimes_k W \arrow[dr, dotted, "\exists!"] & \\ V \times W \arrow[u, "b"] \arrow[r, dashed, "\forall B"'] & X \end{tikzcd}$$

[4] In some contexts the map $v \otimes \lambda \to \lambda(v)$ is called a *contraction*.

Then the trace is

$$\mathrm{tr}T \;=\; \mathrm{tr}\Big(\sum_i Te_i\otimes\lambda_i\Big) \;=\; \sum_i \mathrm{tr}(Te_i\otimes\lambda_i) \;=\; \sum_i \lambda_i(Te_i) \;=\; \sum_i \langle Te_i, e_i\rangle$$

Similarly, *adjoints* $T^* : W \to V$ of maps $T : V \to W$ are expressible in these terms: for $v \in V$, let $\lambda_v \in V^*$ be $\lambda_v(v') = \langle v', v\rangle$, and for $w \in W$ let $\mu_w \in W^*$ be $\mu_w(w') = \langle w', w\rangle$. Then

$$(w\otimes\lambda_v)^* \;=\; v\otimes\mu_w \qquad\qquad (\text{for } w\in W \text{ and } v\in V)$$

since

$$\langle (w\otimes\lambda_v)v', w'\rangle \;=\; \langle \lambda_v(v')w, w'\rangle \;=\; \langle v', v\rangle\langle w, w'\rangle \;=\; \langle v', \langle w', w\rangle\cdot v\rangle \;=\; \langle v', (v\otimes\mu_w)w'\rangle$$

Since it is defined as a completion, the collection of all Hilbert-Schmidt operators $T : V \to W$ is a Hilbert space, with the hermitian inner product

$$\langle S, T\rangle \;=\; \mathrm{tr}(T^*S)$$

[2.3] Proposition: The Hilbert-Schmidt norm $|\,|_{\mathrm{HS}}$ dominates the uniform operator norm $|\,|_{\mathrm{op}}$, so Hilbert-Schmidt operators are *compact*.

Proof: Given $\varepsilon > 0$, let e_1 be a vector with $|e_1| \le 1$ such that $|Tv_1| \ge |T|_{\mathrm{op}} - \varepsilon$. Extend $\{e_1\}$ to an orthonormal basis $\{e_i\}$. Then

$$|T|^2_{\mathrm{op}} \;=\; \sup_{|v|\le 1} |Tv|^2 \;\le\; |Tv_1|^2 + \varepsilon \;\le\; \varepsilon + \sum_j |Tv_j|^2 \;=\; |T|^2_{\mathrm{HS}}$$

Thus, Hilbert-Schmidt norm limits of finite-rank operators are operator-norm limits of finite-rank operators, so are compact. ///

[2.4] Integral kernels yield Hilbert-Schmidt operators

It is already nearly visible that the $L^2(X\times Y)$ norm on kernels $K(x, y)$ is the same as the Hilbert-Schmidt norm on corresponding operators $T : V \to W$, yielding

[2.5] Proposition: Operators $T : L^2(X) \to L^2(Y)$ given by integral kernels $K \in L^2(X \times Y)$ are Hilbert-Schmidt, that is, are Hilbert-Schmidt norm limits of finite-rank operators.

Proof: To prove properly that the $L^2(X\times Y)$ norm on kernels $K(x, y)$ is the same as the Hilbert-Schmidt norm on corresponding operators $T : V \to W$, T should be expressed as a limit of finite-rank operators T_n in terms of kernels $K_n(x, y)$ which are finite sums of products $\varphi(x)\otimes\psi(y)$. Thus, first claim that

$$K(x,y) \;=\; \sum_i \overline{\varphi_i}(x)\, T\varphi_i(y) \qquad\qquad (\text{in } L^2(X\times Y))$$

Indeed, the inner product in $L^2(X\times Y)$ of the right-hand side against any $\varphi_i(x)\psi_j(y)$ agrees with the inner product of the latter against $K(x, y)$, and we have assumed $K \in L^2(X\times Y)$. With $K = \sum_{ij} c_{ij}\overline{\varphi}_i\otimes\psi_j$,

$$T\varphi_i \;=\; \sum_j c_{ij}\,\psi_j$$

Since $\sum_{ij}|c_{ij}|^2$ converges,

$$\lim_i |T\varphi_i|^2 \;=\; \lim_i \sum_j |c_{ij}|^2 \;=\; 0$$

and

$$\lim_n \sum_{i>n} |T\varphi_i|^2 \;=\; \lim_n \sum_{i>n} |c_{ij}|^2 \;=\; 0$$

so the infinite sum $\sum_i \overline{\varphi}_i \otimes T\varphi_i$ converges to K in $L^2(X \times Y)$. In particular, the truncations

$$K_n(x,y) \;=\; \sum_{1 \le i \le n} \overline{\varphi}_i(x)\, T\varphi_i(y)$$

converge to $K(x,y)$ in $L^2(X \times Y)$, and give finite-rank operators

$$T_n f(y) \;=\; \int_X K_n(x,y)\, f(x)\, dx$$

We claim that $T_n \to T$ in Hilbert-Schmidt norm. It is convenient to note that by a similar argument $\overline{K(x,y)} = \sum_i T^*\psi_i(x)\overline{\psi}_i(y)$. Then

$$|T - T_n|^2_{\mathrm{HS}} \;=\; \mathrm{tr}\Big((T-T_n)^* \circ (T-T_n)\Big) \;=\; \sum_{i,j>n} \mathrm{tr}\Big(\Big(T^*\psi_i \otimes \overline{\psi}_i\Big) \circ \Big(\overline{\varphi}_j \otimes T\varphi_j\Big)\Big)$$

$$= \sum_{i,j>n} \langle T^*\psi_i, \varphi_j\rangle_{L^2(X)} \cdot \langle T\varphi_j, \psi_i\rangle_{L^2(Y)} \;=\; \sum_{i,j>n} |c_{ij}|^2 \;\longrightarrow\; 0 \qquad \text{(as } n \to \infty)$$

since $\sum_{ij} |c_{ij}|^2$ converges. Thus, $T_n \to T$ in Hilbert-Schmidt norm. ///

[2.6] **Remark:** With σ-finiteness, the argument above is correct whether K is measurable with respect to the product sigma-algebra or only with respect to the *completion*.

---

3. *Spectral theorem for self-adjoint compact operators*

The λ-eigenspace V_λ of a *self-adjoint compact* operator T on a Hilbert space T is

$$V_\lambda \;=\; \{v \in V : Tv = \lambda \cdot v\}$$

We have already shown that eigenvalues, if any, of self-adjoint T are *real*.

[3.1] **Theorem:** Let T be a self-adjoint compact operator on a non-zero Hilbert space V.
• The completion of $\oplus V_\lambda$ is all of V. In particular, there is an orthonormal basis of *eigenvectors*.
• The only possible *accumulation point* of the set of eigenvalues is 0. For infinite-dimensional V, 0 *is* an accumulation point.
• Every eigenspaces X_λ for $\lambda \neq 0$ is *finite-dimensional*. The 0-eigenspace may be $\{0\}$ or may be infinite-dimensional.
• *(Rayleigh-Ritz)* One or the other of $\pm|T|_{\mathrm{op}}$ is an eigenvalue of T.

A slightly-clever alternative expression for the operator norm is needed:

[3.2] **Lemma:** For T a *self-adjoint* continuous linear operator on a *non-zero* Hilbert space X,

$$|T|_{\mathrm{op}} \;=\; \sup_{|x| \le 1} |\langle Tx, x\rangle|$$

Proof: Let s be that supremum. By Cauchy-Schwarz-Bunyakowsky, $s \le |T|_{\mathrm{op}}$. For any $x, y \in Y$, by polarization

$$2|\langle Tx, y\rangle + \langle Ty, x\rangle| \;=\; |\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle|$$

$$\leq \; |\langle T(x+y), x+y\rangle| + |\langle T(x-y), x-y\rangle| \; \leq \; s|x+y|^2 + s|x-y|^2 = 2s(|x|^2+|y|^2)$$

With $y = t \cdot Tx$ with $t > 0$, because $T = T^*$,

$$\langle Tx, y\rangle \;=\; \langle Tx, t\cdot Tx\rangle \;=\; t\cdot |Tx|^2 \;\geq\; 0 \qquad\qquad (\text{for } y = t\cdot Tc \text{ with } t>0)$$

and

$$\langle Ty, x\rangle \;=\; \langle t\cdot T^2 x, t\cdot x\rangle \;=\; t\cdot\langle Tx, Tx\rangle \;=\; t\cdot |Tx|^2 \;\geq\; 0 \qquad\qquad (\text{for } y = t\cdot Tc \text{ with } t>0)$$

Thus,

$$|\langle Tx, y\rangle| + |\langle Ty, x\rangle| \;=\; \langle Tx, y\rangle + \langle Ty, x\rangle \;=\; |\langle Tx, y\rangle + \langle Ty, x\rangle| \qquad\qquad (\text{for } y = t\cdot Tc \text{ with } t>0)$$

From this, and from the polarization identity divided by 2,

$$|\langle Tx, y\rangle| + |\langle Ty, x\rangle| \;=\; |\langle Tx, y\rangle + \langle Ty, x\rangle| \;\leq\; s(|x|^2+|y|^2) \qquad\qquad (\text{with } y = t\cdot Tx)$$

Divide through by t to obtain

$$|\langle Tx, Tx\rangle| + |\langle T^2x, x\rangle| \;\leq\; \frac{s}{t}\cdot(|x|^2 + |Tx|^2)$$

Minimize the right-hand side by taking $t^2 = |Tx|/|x|$, and note that $\langle T^2x, x\rangle = \langle Tx, Tx\rangle$, giving

$$2|\langle Tx, Tx\rangle| \;\leq\; 2s\cdot|x|\cdot|Tx| \;\leq\; 2s\cdot|x|^2\cdot|T|_{\text{op}}$$

Thus, $|T|_{\text{op}} \leq s$. ///

Now the proof of the theorem:

Proof: The last assertion of the theorem is crucial. To prove it, use the expression

$$|T| \;=\; \sup_{|x|\leq 1} |\langle Tx, x\rangle|$$

and use the fact that any value $\langle Tx, x\rangle$ is *real*. Choose a sequence $\{x_n\}$ so that $|x_n| \leq 1$ and $|\langle Tx, x\rangle| \to |T|$. Replacing it by a subsequence if necessary, the sequence $\langle Tx, x\rangle$ of real numbers has a limit $\lambda = \pm|T|$.

Then

$$0 \leq |Tx_n - \lambda x_n|^2 \;=\; \langle Tx_n - \lambda x_n, Tx_n - \lambda x_n\rangle$$
$$=\; |Tx_n|^2 - 2\lambda\langle Tx_n, x_n\rangle + \lambda^2|x_n|^2 \;\leq\; \lambda^2 - 2\lambda\langle Tx_n, x_n\rangle + \lambda^2$$

The right-hand side goes to 0. By *compactness* of T, replace x_n by a subsequence so that Tx_n converges to some vector y. The previous inequality shows $\lambda x_n \to y$. For $\lambda = 0$, we have $|T| = 0$, so $T = 0$. For $\lambda \neq 0$, $\lambda x_n \to y$ implies

$$x_n \;\longrightarrow\; \lambda^{-1} y$$

For $x = \lambda^{-1}y$,

$$Tx \;=\; \lambda x$$

and x is the desired eigenvector with eigenvalue $\pm|T|$. ///

Now use induction. The completion Y of the sum of non-zero eigenspaces is T-stable. We claim that the orthogonal complement $Z = Y^\perp$ is T-stable, and the restriction of T to is a compact operator. Indeed, for $z \in Z$ and $y \in Y$,

$$\langle Tz, y\rangle \;=\; \langle z, Ty\rangle \;=\; 0$$

proving stability. The unit ball in Z is a subset of the unit ball B in X, so has pre-compact image $TB \cap Z$ in X. Since Z is *closed* in X, the intersection $TB \cap Z$ of Z with the pre-compact TB is pre-compact, proving T restricted to $Z = Y^\perp$ is still compact. Self-adjoint-ness is clear.

By construction, the restriction T_1 of T to Z has no eigenvalues on Z, since any such eigenvalue would also be an eigenvalue of T on Z. Unless $Z = \{0\}$ this would contradict the previous argument, which showed that $\pm|T_1|$ is an eigenvalue on a *non-zero* Hilbert space. Thus, it must be that the completion of the sum of the eigenspaces is all of X. ///

To prove that eigenspaces V_λ for $\lambda \neq 0$ are finite-dimensional, and that there are only finitely-many eigenvalues λ with $|\lambda| > \varepsilon$ for given $\varepsilon > 0$, let B be the unit ball in

$$Y = \sum_{|\lambda| > \varepsilon} X_\lambda$$

The image of B by T contains the ball of radius ε in Y. Since T is compact, this ball is *pre*-compact, so Y is finite-dimensional. Since the dimensions of the X_λ are positive integers, there can be only finitely-many of them with $|\lambda| > \varepsilon$, and each is finite-dimensional. It follows that the only possible accumulation point of the set of eigenvalues is 0, and, for X infinite-dimensional, 0 *must* be an accumulation point. ///

---

4. *The Fredholm alternative*

Here, we prove the simplest *Fredholm alternative*. An analogue exists for Banach spaces, as well. This also gives a spectral theorem for not-necessarily self-adjoint compact operators.

[4.1] Theorem: For Hilbert space X, for compact $T : X \longrightarrow X$ and $0 \neq \lambda \in \mathbb{C}$, $T - \lambda$ has closed image of codimension equal to the dimension of its kernel. *(Proof in the sequel.)*

[4.2] Corollary: For $\lambda \neq 0$, *either* $T - \lambda$ is a bijection, *or* λ is an eigenvalue. ///

[4.3] Corollary: The only non-zero spectrum of a compact operator is *point* spectrum. [5] ///

[4.4] Corollary: *Either* λ is an eigenvalue, *or* $(T - \lambda)u = v$ is solvable for u for all $v \in X$. ///

This result complements the spectral theorem for *self-adjoint* compact operators on a Hilbert space, where there is an orthonormal basis of eigenvectors. For not-necessarily-self-adjoint (or not-necessarily-normal) compact operators, it can happen that there are *no* non-zero eigenvalues. This is not a pathology: the Volterra operator

$$Vf(x) = \int_0^x f(y)\,dy \qquad\qquad (\text{for } f \in L^2[0,1])$$

is Hilbert-Schmidt, hence compact, but has no non-zero eigenvalues.

[4.5] Compact operators invertible only on finite-dimensional

For compact $T : X \to X$ with continuous inverse T^{-1}, the boundedness of T^{-1} gives a constant C such that $|T^{-1}x| \leq C \cdot |x|$ for all $x \in Y$. Invertibility implies that $TX = X$, and $|x| \leq C \cdot |Tx|$ for all $x \in X$. Thus, the image by T of the unit ball in X contains an open ball in X. Compactness implies that X is finite-dimensional.

---

[5] Recall that the *eigenvalues* or *point spectrum* of an operator T on a Hilbert space X consists of $\lambda \in \mathbb{C}$ such that $T - \lambda$ fails to be *injective*. The *continuous* spectrum consists of λ with $T - \lambda$ *injective* and with *dense* image, but *not surjective*. The *residual spectrum* consists of λ with $T - \lambda$ *injective* but $(T - \lambda)X$ not *dense*.

[4.6] Generalized eigenspaces finite-dimensional for $\lambda \neq 0$

For compact $T : X \to X$ and $\lambda \neq 0$, the kernel of $T - \lambda$ is finite-dimensional, since any restriction of T to a subspace is still compact, and T acts by a scalar on $\ker(T - \lambda)$.

By induction on n, the operator $T - \lambda$ maps $\ker(T - \lambda)^{n+1}$ to the finite-dimensional space $\ker(T - \lambda)^n$, so is finite-rank. On $\ker(T - \lambda)^{n-1}$,

$$\text{compact} \;=\; \text{finite-rank} \;=\; T - \lambda \;=\; \text{compact} - \lambda \qquad\qquad (\text{on } \ker(T - \lambda)^{n+1})$$

Thus, $\lambda \neq 0$ is compact on $\ker(T - \lambda)^{n+1}$, implying that this kernel is finite-dimensional. ///

[4.7] T compact if and only if T^* compact

Proof: First, the adjoint map $T \to T^*$ is continuous in the operator-norm topology. Indeed, $|T^*| = |T|$, because

$$|T^*|^2 \;=\; \sup_{|x| \le 1} |T^* x|^2 \;=\; \langle T^* x, T^* x\rangle \;=\; \langle TT^* x, x\rangle \;\le\; |TT^* x| \cdot |x| \;\le\; |T| \cdot |T^*| \cdot |x| \cdot |x| \;=\; |T| \cdot |T^*|$$

Dividing through by $|T^*|$ gives $|T^*| \le |T|$. Symmetrically, $|T| \le |T^*|$. Compact T is an operator-norm limit of finite-rank operators T_n. Then T^* is the operator-norm limit of the finite-rank T_n^*. ///

[4.8] $\mathrm{Im}(T - \lambda)$ is closed for $\lambda \neq 0$

Proof: Let $(T - \lambda)x_n \to y$. First consider the situation that $\{x_n\}$ is *bounded*. Compactness of T yields a convergent subsequence of Tx_n, and we replace x_n by the corresponding subsequence. Then $-\lambda x_n = y - Tx_n$ converges to $y - \lim Tx_n$, so x_n is convergent to $x_o \in X$, since $\lambda \neq 0$, and $Tx_o = y$.

To reduce the general case to the previous, first reduce to the case that $T - \lambda$ is *injective*: from above, $\ker(T - \lambda)$ is finite-dimensional. Let V be the orthogonal complement to $\ker(T - \lambda)$. Since $(T - \lambda)V = (T - \lambda)X$, to prove the image is closed it suffices to consider V, or, equivalently, that $T - \lambda$ is *injective* on X. Since $T - \lambda$ is a continuous bijection to its image, by the *open mapping theorem* it is an isomorphism to its image. Thus, there is $\delta > 0$ such that $|(T - \lambda)x| \ge \delta |x|$.

Returning to the main argument, suppose that $(T - \lambda)x_n \to y_o$. Then $(T - \lambda)(x_m - x_n) \to 0$. For $T - \lambda$ injective, $x_m - x_n \to 0$, so x_n is *bounded*, reducing to that case. ///

[4.9] $T - \lambda$ injective $\Longleftrightarrow$ surjective for $\lambda \neq 0$

Proof: Suppose $T - \lambda$ is injective. Let $V_n = (T - \lambda)^n X$. Since images under $T - \lambda$ for compact T and $\lambda \neq 0$ are closed, by induction these are *closed* subspaces of X. For $x \notin (T - \lambda)X$ and any $y \in X$,

$$(T - \lambda)^n x - (T - \lambda)^{n+1} y \;=\; (T - \lambda)^n \Big(x - (T - \lambda)y\Big)$$

Injectivity of $T - \lambda$ implies that of $(T - \lambda)^n$, so this is not 0. That is, $(T - \lambda)^n x \notin (T - \lambda)^{n+1} X$. Thus, the chain of subspaces V_n is strictly decreasing.

Take $v_n \in V_n$ such that $|v_n| = 1$ and away from V_n, say by

$$\inf_{y \in V_{n+1}} |v_n - y| \;\ge\; \tfrac{1}{2}$$

The effect of T is

$$Tv_m - Tv_{m+n} \;=\; \lambda v_m + (T - \lambda)v_m - Tv_{m+n} \;\in\; \lambda v_m + V_{n+1} \qquad\qquad (\text{integers } m \ge 1 \text{ and } n \ge 1)$$

since V_{m+1} is T-stable. Thus,

$$|Tv_m - Tv_{m+n}| \;\ge\; |\lambda| \cdot \tfrac{1}{2}$$

This is impossible, since compact T maps the bounded set $\{v_n\}$ to a pre-compact set. Thus, the chain of subspaces V_n cannot be strictly decreasing, and have surjectivity $(T-\lambda)X = X$.

On the other hand, suppose $T-\lambda$ is *surjective.* Then the adjoint $(T-\lambda)^*$ is *injective.* Since adjoints of compact operators are compact, we already know that $(T-\lambda)^*$ is *surjective.* Then $T-\lambda = (T-\lambda)^{**}$ is *injective.* ///

[4.10] $\dim\ker(T-\lambda) = \dim\mathrm{coker}(T-\lambda)$ for $\lambda \neq 0$, T compact

That is, such operators are *Fredholm operators of index* 0.

Proof: The compactness of T entails the finite-dimensionality of $\ker(T-\lambda)$ for $\lambda \neq 0$. Dually, for $y_1,\ldots,y_n \in X$ linearly independent modulo $(T-\lambda)X$, by *Hahn-Banach* there are $\eta_1,\ldots,\eta_n \in X^*$ vanishing on the image $(T-\lambda)X$ and $\eta_i(y_j) = \delta_{ij}$. Such η_i are in the kernel of the adjoint $(T-\lambda)^*$. We know T^* is compact, so $\ker(T-\lambda)^*$ is finite-dimensional.

We've proven that injectivity and surjectivity of $T-\lambda$ are equivalent, and that the kernel and cokernel are finite-dimensional. Let $x_1,\ldots,x_m$ (with $m \ge 1$) span the kernel, and let (the images of) $y_1,\ldots,y_n$ (with $n \ge 1$) span the cokernel, and show that $m = n$.

For $m \le n$, let X' be a closed complementary subspace to the kernel of $T-\lambda$, for example, its orthogonal complement. Let F be the finite-rank operator which is 0 on X' and $Fx_i = y_i$. The adjusted operator $T' = T + F$ is compact. For $(T'-\lambda)x = 0$,

$$(T-\lambda)x \;=\; Fx \;\in\; (T-\lambda)X \cap \mathrm{span}\, y_1,\ldots,y_n \;=\; \{0\}$$

That is, $T'-\lambda$ is *injective*, so is *surjective*, so $m = n$. In the opposite case $m \ge n$, let $Fx_i = y_i$ for $i \le n$, and $Fx_i = y_n$ for $i \ge n$. With $T' = T + F$ again, in this case $T'-\lambda$ is *surjective*, so is injective, and $m = n$. ///

[4.11] Discreteness of spectrum of compact operators

[4.12] Claim: For T a compact operator on a Hilbert the non-zero spectrum (if any) is *point* spectrum. The number of eigenvalues λ outside a given disk $|\lambda| \le r$ is *finite* for $r > 0$, and always 0 is in the spectrum.

Proof: For λ not an eigenvalue, we know that $T-\lambda$ is injective *and surjective*, so by the open mapping theorem it is an isomorphism. Thus, indeed, the only non-zero spectrum consists of *eigenvalues.* We also know that eigenspaces are finite-dimensional, for non-zero eigenvalues.

For infinite-dimensional Hilbert spaces, 0 inevitably lies in the spectrum, otherwise T would be *invertible.* Then $1 = T \circ T^{-1}$ is the composition of a compact operator and a continuous operator, so is compact, which is possible only in finite-dimensional spaces.

Suppose there were infinitely-many different eigenvalues $\lambda_1, \lambda_2, \ldots$ outside the closed disk $|\lambda| \le r$ with $r > 0$, with corresponding eigenvectors x_i with $|x_i| = 1$. First, the x_i are linearly independent: let $\sum_i c_i\, x_i = 0$ be a non-trivial linear dependence relation with fewest non-zero c_i's, and apply T: for an index i_o with $c_{i_o} \neq 0$, we obtain a shorter relation by suitable subtraction,

$$0 \;=\; \sum_i \lambda_i\, c_i\, x_i - \lambda_{i_o} \sum_i c_i\, x_i \;=\; \sum_{i \neq i_o} (\lambda_i - \lambda_{i_o})\, c_i\, x_i$$

Thus, there can be no non-trivial linear dependence. With V_n the span of $x_1, x_2, \ldots, x_n$, this implies that the containments $V_n \subset V_{n+1}$ are *strict.* Thus, there exist unit vectors $y_i \in V_i$ with the distance from y_i to

V_{i-1} at least $\frac{1}{2}$. Then for $i > j$

$$Ty_i - Ty_j \;=\; \lambda_i y_i + (T - \lambda_i) y_i - Ty_j \;\in\; \lambda_i y_i + V_{i+1}$$

and, thus, $|Ty_i - Ty_j| \geq |\lambda| \cdot \frac{1}{2}$. However, this contradicts the compactness of T. We conclude that there can be only finitely-many eigenvalues larger than $r > 0$. ///

---

5. *Simplest Rellich compactness lemma*

One characterization of the s^{th} Levi-Sobolev space of functions $H^s(A)$ on a product $A = (S^1)^{\times n}$ of circles $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ is as the closure of the function space of *finite* Fourier series with respect to the Levi-Sobolev norm (squared)

$$\Big|\sum_{\xi\in\mathbb{Z}^n} c_\xi\, e^{i\xi\cdot x}\Big|^2_{H^s} \;=\; \sum_{\xi\in\mathbb{Z}^n} |c_\xi|^2 \cdot (1+|\xi|^2)^s \qquad\qquad (s\in\mathbb{R}, \text{ on finite Fourier series})$$

The standard orthonormal basis for $H^s(A)$ is

$$\frac{1}{(2\pi)^{n/2}} \cdot \frac{e^{i\xi\cdot x}}{(1+|\xi|^2)^{s/2}} \qquad\qquad (\text{with } \xi\in\mathbb{Z}^n)$$

By the Plancherel theorem, the map from $L^2(\mathbb{Z}^n)$ (with counting measure) to $L^2(A)$ by

$$\{c_\xi : \xi\in\mathbb{Z}^n\} \;\longrightarrow\; \frac{1}{(2\pi)^{n/2}} \sum_{\xi\in\mathbb{Z}^n} c_\xi\, \frac{e^{i\xi\cdot x}}{(1+|\xi|^2)^{s/2}}$$

is an isometric isomorphism.

For $s > t$, there is a continuous inclusion $H^s(A) \to H^t(A)$. In terms of these orthonormal bases, there is a commutative diagram

$$\begin{tikzcd} L^2(A) \arrow[r, "T"] \arrow[d, "\approx"'] & L^2(A) \arrow[d, "\approx"] \\ H^s(A) \arrow[r, "\text{inc}"'] & H^t(A) \end{tikzcd}$$

given by

$$\begin{tikzcd} \{c_\xi\} \arrow[r, "T"] \arrow[d, "\approx"'] & \Big\{(1+|\xi|^2)^{\frac{t-s}{2}} \cdot c_\xi\Big\} \arrow[d, "\approx"] \\ \frac{1}{(2\pi)^{n/2}} \sum_\xi c_\xi\, \frac{e^{i\xi\cdot x}}{(1+|\xi|^2)^{s/2}} \arrow[r, "\text{inc}"'] & \frac{1}{(2\pi)^{n/2}} \sum_\xi (1+|\xi|^2)^{\frac{t-s}{2}} \cdot c_\xi\, \frac{e^{i\xi\cdot x}}{(1+|\xi|^2)^{t/2}} \end{tikzcd}$$

Since $s > t$, the number $\lambda_\xi = (1+|\xi|^2)^{\frac{t-s}{2}}$ are bounded by 1, and have unique limit point 0. In particular, $T : L^2(A) \to L^2(A)$ is *compact*.

Thus, we have the simplest instance of *Rellich's compactness lemma*: the inclusion $H^s(A) \to H^t(A)$ is *compact* for $s > t$.

---

6. *Appendix: topologies on finite-dimensional spaces*

In the proof that Hilbert-Schmidt operators are compact, we needed the fact that finite-dimensional subspaces of Hilbert spaces are linearly homeomorphic to $\mathbb{C}^n$ with its usual topology. In fact, it is true that *any* finite dimensional topological vector space is linearly homeomorphic to $\mathbb{C}^n$. That is, we need not assume that the space is a Hilbert space, a Banach space, a Fréchet space, locally convex, or anything else. However, the general argument is a by-product of the development of the general theory of topological vector spaces, and is best delayed. Thus, we give more proofs that apply to Hilbert and Banach spaces.

[6.1] Lemma: Let W be a finite-dimensional subspace of a pre-Hilbert space V. Let $w_1, \ldots, w_n$ be a $\mathbb{C}$-basis of W. Then the continuous linear bijection

$$\varphi : \mathbb{C}^n \to W$$

by

$$\varphi(z_1, \ldots, z_n) = \sum_i z_i \cdot w_i$$

is a homeomorphism. And W is closed.

Proof: Because vector addition and scalar multiplication are continuous, the map φ is continuous. It is obviously linear, and since the w_i are linearly independent it is an injection.

Let $v_1, \ldots, v_n$ be an *orthonormal* basis for W. Consider the continuous linear functionals

$$\lambda_i(v) = \langle v, v_i \rangle$$

As intended, we have $\lambda_i(v_j) = 0$ for $i \neq j$, and $\lambda_i(v_i) = 1$. Define continuous linear $\psi : W \to \mathbb{C}^n$ by

$$\psi(v) = (\lambda_1(v), \ldots, \lambda_n(v))$$

The inverse map to ψ is continuous, because vector addition and scalar multiplication are continuous. Thus, ψ is a linear homeomorphism $W \approx \mathbb{C}^n$.

Generally, we can use Gram-Schmidt to create an orthonormal basis v_i from a given basis w_i. Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{C}^n$. Let $f_i = \psi(w_i)$ be the inverse images in $\mathbb{C}^n$ of the w_i. Let $A : \mathbb{C}^n \to \mathbb{C}^n$ be a linear homeomorphism $\mathbb{C}^n \to \mathbb{C}^n$ sending e_i to f_i, that is, $Ae_i = f_i$. Then

$$\varphi = \psi^{-1} \circ A : \mathbb{C}^n \to W$$

since both φ and $\psi^{-1} \circ A$ send e_i to w_i. Both ψ and A are linear homeomorphisms, so the composition φ is also.

Since $\mathbb{C}^n$ is a complete metric space, so is its homeomorphic image W, so W is necessarily closed. ///

Now we give a somewhat different proof of the uniqueness of topology on finite-dimensional *normed* spaces, using the Hahn-Banach theorem. Again, invocation of Hahn-Banach is actually unnecessary, since the same conclusion will be reached (later) without local convexity. The only difference in the proof is the method of proving existence of sufficiently many linear functionals.

[6.2] Lemma: Let W be a finite-dimensional subspace of a normed space V. Let $w_1, \ldots, w_n$ be a $\mathbb{C}$-basis of W. Then the continuous linear bijection

$$\varphi : \mathbb{C}^n \to W$$

by

$$\varphi(z_1, \ldots, z_n) = \sum_i z_i \cdot w_i$$

is a homeomorphism. And W is closed.

Proof: Let v_1 be a non-zero vector in W, and from Hahn-Banach let λ_1 be a continuous linear functional on W such that $\lambda_1(v_1) = 1$. By the (algebraic) isomorphism theorem

$$\text{image } \lambda_1 \approx W/\ker \lambda_1$$

so $\dim W/\ker \lambda_1 = 1$. Take $v_2 \neq 0$ in $\ker \lambda_1$ and continuous linear functional λ_2 such that $\lambda_2(v_2) = 1$. Replace v_1 by $v_1 - \lambda_2(v_1)v_2$. Then still $\lambda_1(v_1) = 1$ and also $\lambda_2(v_1) = 0$. Thus, λ_1 and λ_2 are linearly independent, and

$$(\lambda_1, \lambda_2) : W \to \mathbb{C}^2$$

is a surjection. Choose $v_3 \neq 0$ in $\ker \lambda_1 \cap \ker \lambda_2$, and λ_3 such that $\lambda_3(v_3) = 1$. Replace v_1 by $v_1 - \lambda_3(v_1)v_3$ and v_2 by $v_2 - \lambda_3(v_2)v_3$. Continue similarly until

$$\bigcap \ker \lambda_i = \{0\}$$

Then we obtain a basis $v_1, \ldots, v_n$ for W and an continuous linear isomorphism

$$\psi = (\lambda_1, \ldots, \lambda_n) : W \to \mathbb{C}^n$$

that takes v_i to the standard basis element e_i of $\mathbb{C}^n$. On the other hand, the continuity of scalar multiplication and vector addition assures that the inverse map is continuous. Thus, ψ is a continuous isomorphism.

Now let $f_i = \psi(w_i)$, and let A be a (continuous) linear isomorphism $\mathbb{C}^n \to \mathbb{C}^n$ such that $Ae_i = f_i$. Then $\varphi = \psi^{-1} \circ A$ is a continuous linear isomorphism.

Finally, since W is linearly homeomorphic to $\mathbb{C}^n$, which is complete, any finite-dimensional subspace of a normed space is closed. ///

[6.3] Remark: The proof for normed spaces works in any topological vector space in which Hahn-Banach holds. We will see later that Hahn-Banach holds for all *locally convex* spaces. Nevertheless, as we will see, this hypothesis is unnecessary, since finite-dimensional subspaces of *arbitrary* topological vector spaces are linearly homeomorphic to $\mathbb{C}^n$.

Partial Bibliography

[Fredholm 1900] I. Fredholm, *Sur une nouvelle méthode pour la résolution du problème de Dirichlet*, Kong. Betenskaps-Akad. Förh. Stockholm (1900), 39-46.

[Fredholm 1903] I. Fredholm, *Sur une classe d'équations fonctionelles*, Acta Math. **27** (1903), 365-390.

[Riesz 1917] F. Riesz, *Über lineare Funktionalgleichungen*, Acta Math. **41** (1917), 71-98.
