(*February 24, 2017*)

08a. Operators on Hilbert spaces

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

[This document is http://www.math.umn.edu/˜garrett/m/real/notes_2016-17/08a-ops_on_Hsp.pdf]

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 simplest naturally occurring compact operators are the *Hilbert-Schmidt* operators.

---

1. Boundedness, continuity, operator norms

[1.1] Definition: A linear (not necessarily continuous) map $T : X \to Y$ from one Hilbert space to another is *bounded* if, for all $\varepsilon > 0$, there is $\delta > 0$ such that $|Tx|_Y < \varepsilon$ for all $x \in X$ with $|x|_X < \delta$.

[1.2] Proposition: For a linear, not necessarily continuous, map $T : X \to Y$ of Hilbert spaces, the following three conditions are equivalent:
(i) T is continuous
(ii) T is continuous at 0
(iii) T is bounded

Proof: For T continuous as 0, given $\varepsilon > 0$ and $x \in X$, there is small enough $\delta > 0$ such that $|Tx' - 0|_Y < \varepsilon$ for $|x' - 0|_X < \delta$. For $|x'' - x|_X < \delta$, using the linearity,

$$|Tx'' - Tx|_X \;=\; |T(x'' - x) - 0|_X \;<\; \delta$$

That is, continuity at 0 implies continuity.

Since $|x| = |x - 0|$, continuity at 0 is immediately equivalent to boundedness. ///

[1.3] Definition: The *kernel* $\ker T$ of a linear (not necessarily continuous) linear map $T : X \to Y$ from one Hilbert space to another is

$$\ker T \;=\; \{x \in X : Tx = 0 \in Y\}$$

[1.4] Proposition: The kernel of a continuous linear map $T : X \to Y$ is closed.

Proof: For T continuous

$$\ker T = T^{-1}\{0\} \;=\; X - T^{-1}(Y - \{0\}) \;=\; X - T^{-1}(\text{open}) \;=\; X - \text{open} \;=\; \text{closed}$$

since the inverse images of open sets by a continuous map are open. ///

[1.5] Definition: The *operator norm* $|T|$ of a linear map $T : X \to Y$ is

$$\text{operator norm } T \;=\; |T| \;=\; \sup_{x\in X\,:\,|x|_X\le 1} |Tx|_Y$$

[1.6] Corollary: A linear map $T : X \to Y$ is continuous if and only if its operator norm is finite. ///

---

2. *Adjoint maps*

[2.1] Definition: An *adjoint* T^* of a continuous linear map $T : X \to Y$ from a pre-Hilbert space X to a pre-Hilbert space Y (if T^* exists) is a continuous linear map $T^* : Y \to X$ such that

$$\langle Tx, y\rangle_Y \;=\; \langle x, T^*y\rangle_X$$

[2.2] Remark: When a pre-Hilbert space X is not complete, that is, is not a Hilbert space, an operator $T : X \to Y$ may fail to have an adjoint.

[2.3] Theorem: A continuous linear map $T : X \to Y$ from a *Hilbert* space X to a pre-Hilbert space Y has a unique adjoint T^*.

[2.4] Remark: The target space of T need not be a Hilbert space, that is, need not be complete.

Proof: For each $y \in Y$, the map

$$\lambda_y : X \longrightarrow \mathbb{C}$$

given by

$$\lambda_y(x) \;=\; \langle Tx, y\rangle$$

is a continuous linear functional on X. By Riesz-Fischer, there is a unique $x_y \in X$ so that

$$\langle Tx, y\rangle \;=\; \lambda_y(x) \;=\; \langle x, x_y\rangle$$

Try to define T^* by $T^*y = x_y$. This is a well-defined map from Y to X, and has the adjoint property $\langle Tx, y\rangle_Y = \langle x, T^*y\rangle_X$.

To prove that T^* is continuous, prove that it is bounded. From Cauchy-Schwarz-Bunyakowsky

$$|T^*y|^2 \;=\; |\langle T^*y, T^*y\rangle_X| \;=\; |\langle y, TT^*y\rangle_Y| \le |y|\cdot|TT^*y| \;\le\; |y|\cdot|T|\cdot|T^*y|$$

where $|T|$ is the operator norm. For $T^*y \neq 0$, divide by it to find

$$|T^*y| \;\le\; |y|\cdot|T|$$

Thus, $|T^*| \le |T|$. In particular, T^* is bounded. Since $(T^*)^* = T$, by symmetry $|T| = |T^*|$. Linearity of T^* is easy. ///

[2.5] Corollary: For a continuous linear map $T : X \to Y$ of Hilbert spaces, $T^{**} = T$. ///

An operator $T \in \text{End}(X)$ commuting with its adjoint is *normal*, that is,

$$TT^* \;=\; T^*T$$

This only applies to operators from a Hilbert space *to itself*. An operator T is *self-adjoint* or *hermitian* if $T = T^*$. That is, T is hermitian when

$$\langle Tx, y \rangle \;=\; \langle x, Ty \rangle \qquad\qquad (\text{for all } x, y \in X)$$

An operator T is *unitary* when

$$TT^* \;=\; \text{identity map on } Y \qquad\qquad T^*T \;=\; \text{identity map on } X$$

There are simple examples in infinite-dimensional spaces where $TT^* = 1$ does not imply $T^*T = 1$, and vice-versa. Thus, it does *not* suffice to check something like $\langle Tx, Tx \rangle = \langle x, x \rangle$ to prove unitariness. Obviously hermitian operators are normal, as are unitary operators, using the more careful definition.

---

3. *Stable subspaces and complements*

Let $T : X \to X$ be a continuous linear operator on a Hilbert space X. A vector subspace Y is T-*stable* or T-*invariant* if $Ty \in Y$ for all $y \in Y$. Often one is most interested in the case that the subspace be *closed* in addition to being *invariant*.

[3.1] Proposition: For $T : X \to X$ a continuous linear operator on a Hilbert space X, and Y a T-stable subspace, $Y^\perp$ is T^*-stable.

Proof: For $z \in Y^\perp$ and $y \in Y$,

$$\langle T^*z, y \rangle \;=\; \langle z, T^{**}y \rangle \;=\; \langle z, Ty \rangle$$

since $T^{**} = T$, from above. Since Y is T-stable, $Ty \in Y$, and this inner product is 0, and $T^*z \in Y^\perp$. ///

[3.2] Corollary: For continuous *self-adjoint* T on a Hilbert space X, and Y a T-stable subspace, both Y and $Y^\perp$ are T-stable. ///

[3.3] Remark: *Normality* of $T : X \to X$ is insufficient to assure the conclusion of the corollary, in general. For example, with the two-sided ℓ^2 space

$$X \;=\; \{\{c_n : n \in \mathbb{Z}\} : \sum_{n \in \mathbb{Z}} |c_n|^2 \;<\; \infty\}$$

the right-shift operator

$$(Tc)_n \;=\; c_{n-1} \qquad\qquad (\text{for } n \in \mathbb{Z})$$

has adjoint the *left* shift operator

$$(T^*c)_n \;=\; c_{n+1} \qquad\qquad (\text{for } n \in \mathbb{Z})$$

and

$$T^*T \;=\; TT^* \;=\; 1_X$$

So this T is not merely *normal*, but *unitary*. However, the T-stable subspace

$$Y \;=\; \{\{c_n\} \in X : c_k = 0 \text{ for } k < 0\}$$

is not T^*-stable, nor is its orthogonal complement T-stable.

On the other hand, adjoint-stable *collections* of operators have a good stability result:

[3.4] Proposition: Suppose for every T in a set A of bounded linear operators on a Hilbert space V the adjoint T^* is also in A. Then, for an A-stable subspace W of V, the orthogonal complement $W^\perp$ is also A-stable.

Proof: For y in $W^\perp$ and $T \in A$, for $x \in W$,

$$\langle x, Ty \rangle \;=\; \langle T^* x, y \rangle \in \langle W, y \rangle \;=\; \{0\}$$

since $T^* \in A$. ///

---

4. *Spectrum, eigenvalues*

For a continuous linear operator $T \in \mathrm{End}(X)$, the λ-*eigenspace* of T is

$$X_\lambda \;=\; \{x \in X : Tx = \lambda x\}$$

If this space is not simply $\{0\}$, then λ is an *eigenvalue*.

[4.1] Proposition: An eigenspace X_λ for a continuous linear operator T on X is a *closed* and T-stable subspace of X. For *normal* T the adjoint T^* acts by the scalar $\overline{\lambda}$ on X_λ.

Proof: The λ-eigenspace is the kernel of the continuous linear map $T - \lambda$, so is closed. The stability is clear, since the operator restricted to the eigenspace is a scalar operator. For $v \in X_\lambda$, using normality,

$$(T - \lambda) T^* v \;=\; T^*(T - \lambda) v \;=\; T^* \cdot 0 = 0$$

Thus, X_λ is T^*-stable. For $x, y \in X_\lambda$,

$$\langle (T^* - \overline{\lambda}) x, y \rangle \;=\; \langle x, (T - \lambda) y \rangle \;=\; \langle x, 0 \rangle$$

Thus, $(T^* - \overline{\lambda}) x = 0$. ///

[4.2] Proposition: For T *normal*, for $\lambda \neq \mu$, and for $x \in X_\lambda, y \in X_\mu$, always $\langle x, y \rangle = 0$. For T *self-adjoint*, if $X_\lambda \neq 0$ then $\lambda \in \mathbb{R}$. For T *unitary*, if $X_\lambda \neq 0$ then $|\lambda| = 1$.

Proof: Let $x \in X_\lambda$, $y \in X_\mu$, with $\mu \neq \lambda$. Then

$$\lambda \langle x, y \rangle \;=\; \langle Tx, y \rangle \;=\; \langle x, T^* y \rangle \;=\; \langle x, \overline{\mu} y \rangle \;=\; \mu \langle x, y \rangle$$

invoking the previous result. Thus,

$$(\lambda - \mu) \langle x, y \rangle \;=\; 0$$

giving the result. For T self-adjoint and x a non-zero λ-eigenvector,

$$\lambda \langle x, x \rangle \;=\; \langle x, T^* x \rangle \;=\; \langle x, Tx \rangle \;=\; \langle x, \lambda x \rangle \;=\; \overline{\lambda} \langle x, x \rangle$$

Thus, $(\lambda - \overline{\lambda}) \langle x, x \rangle = 0$. Since x is non-zero, the result follows. For T unitary and x a non-zero λ-eigenvector,

$$\langle x, x \rangle \;=\; \langle T^* T x, x \rangle \;=\; \langle Tx, Tx \rangle \;=\; |\lambda|^2 \cdot \langle x, x \rangle$$

///

In what follows, for a complex scalar λ write simply λ for scalar multiplication by λ on a Hilbert space X.

[4.3] Definition: The *spectrum* $\sigma(T)$ of a continuous linear operator $T : X \to X$ on a Hilbert space X is the collection of complex numbers λ such that $T - \lambda$ *does not have* a continuous linear inverse.

[4.4] Definition: The *discrete spectrum* $\sigma_{\text{disc}}(T)$ is the collection of complex numbers λ such that $T - \lambda$ fails to be *injective*. In other words, the discrete spectrum is the collection of eigenvalues.

[4.5] Definition: The *continuous spectrum* $\sigma_{\text{cont}}(T)$ is the collection of complex numbers λ such that $T - \lambda \cdot 1_X$ *is* injective, *does* have dense image, but fails to be *surjective*.

[4.6] Definition: The *residual spectrum* $\sigma_{\text{res}}(T)$ is everything else: neither discrete nor continuous spectrum. That is, the residual spectrum of T is the collection of complex numbers λ such that $T - \lambda \cdot 1_X$ *is* injective, and *fails* to have dense image (so is certainly not surjective).

[4.7] Remark: To see that there are no *other* possibilities, note that the *Closed Graph Theorem* implies that a bijective, continuous, linear map $T : X \to Y$ of Banach spaces has continuous inverse. Indeed, granting that the inverse exists as a linear map, its graph is

$$\text{graph of } T^{-1} \;=\; \{(y, x) \in Y \times X \;:\; (x, y) \text{ in the graph of } T \;\subset X \times Y\}$$

Since the graph of T is closed, the graph of T^{-1} is closed, and by the Closed Graph Theorem T^{-1} is continuous.

The potential confusion of *residual* spectrum does not occur in many situations of interest"

[4.8] Proposition: A normal operator $T : X \to X$ has *empty* residual spectrum.

Proof: The adjoint of $T - \lambda$ is $T^* - \overline{\lambda}$, so consider $\lambda = 0$ to lighten the notation. Suppose that T does *not* have dense image. Then there is non-zero z such that

$$0 \;=\; \langle z, Tx\rangle \;=\; \langle T^*z, x\rangle \qquad\qquad \text{(for every } x \in X)$$

Therefore $T^*z = 0$, and the 0-eigenspace Z of T^* is non-zero. Since $T^*(Tz) = T(T^*z) = T(0) = 0$ for $z \in Z$, T^* stabilizes Z. That is, Z is both T and T^*-stable. Therefore, $T = (T^*)*$ acts on Z by (the complex conjugate of) 0, and T has non-trivial 0-eigenvectors, contradiction. ///

---

5. *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*.

[5.1] 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.

[5.2] 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

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

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

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.

[5.3] Theorem: A compact operator $T : X \to Y$ with Y a Hilbert space 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

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

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 $\varepsilon \to 0$. ///

[5.4] 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.

[5.5] Remark: The theorem is false in Banach spaces, although the only example known to this author (Per Enflo, *Acta Math.*, vol. 130, 1973) is complicated.

---

6. *Hilbert-Schmidt operators*

[6.1] 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

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

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

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

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

by computing

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

[6.2] 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]

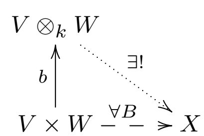
$$v \times \lambda \;\longrightarrow\; \Big(w \to \lambda(w)\cdot v\Big) \qquad\qquad (\text{for } v \in V \text{ and } \lambda \in V^*)$$

Trace is easy to define in these terms [4]

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

---

[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



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

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,

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

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

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

[6.3] 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. ///

[6.4] 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

[6.5] 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

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

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

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

[6.6] 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*.

---

7. 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*.

[7.1] 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 $\pm|T|_{\text{op}}$ is an eigenvalue of T.

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

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

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

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

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

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

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

and

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

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,

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

Divide through by t to obtain

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

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

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

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

Now the proof of the theorem:

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

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

and use the fact that any value $\langle Tx, x\rangle$ is *real*. Choose a sequence $\{x_n\}$ so that $|x_n| \le 1$ 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

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

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 $\lambda \ne 0$, $\lambda x_n \to y$ 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. ///

---

8. *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.

[8.1] **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.

[8.2] 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. ///

[8.3] 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 $\mathbb{C}^n$.