(*February 18, 2012*)

Compact Operators on Hilbert Space

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

Among all linear operators on Hilbert spaces, the *compact* ones (defined below) are the simplest, and most imitate the more familiar linear algebra of finite-dimensional operator theory. In addition, these are of considerable practical value and importance. We prove a spectral theorem for self-adjoint operators with minimal fuss. Thus, we do *not* invoke broader discussions of properties of spectra. We only need the *Cauchy-Schwarz-Bunyakowsky inequality* and the *definition* of self-adjoint compact operator. It is true that various points here admit great generalization, and receive definitive treatment only in such general setting.

---

1. *Compact operators: definition*

A set in a topological space is called **pre-compact** if its closure is compact. (Beware, sometimes this has a more restrictive meaning.) A linear operator $T : X \to Y$ from a pre-Hilbert space X to a Hilbert space Y is **compact** if 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.*

Such operators are the most amenable. Sources of such operators will be considered elsewhere. For the moment we need only concentrate on the defining property, and its use in proof of the spectral theorem below.

---

2. *Clever expression for the operator norm*

First is a little lemma, useful other places as well, which provides a necessary alternative expression for the *uniform norm*

$$|T| = |T|_{\text{unif}} = \sup_{|x|\le 1} |Tx|$$

of a continuous linear operator T from a Hilbert space X to itself. For present purposes, we say that a (continuous) linear operator $T : X \to X$ is **self-adjoint** if

$$\langle Tx, y\rangle = \langle x, Ty\rangle$$

for all $x, y \in X$.

[2.0.1] Lemma: For T a self-adjoint (continuous linear) operator on a pre-Hilbert space X

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

Proof: Let s be that supremum. By the Cauchy-Schwarz-Bunyakowsky inequality $s \le |T|$.

For any $x, y \in X$

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

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

Let $y = t \cdot Tx$ with $t > 0$. Using the self-adjointness of T,

$$|\bar{\mu}\langle Tx, y\rangle + \mu\langle Ty, x\rangle| = |\langle Tx, y\rangle| + |\langle Ty, x\rangle|$$

Dividing by 2,

$$|\langle Tx, y\rangle| + |\langle Ty, x\rangle| \leq s(|x|^2 + |y|^2)$$

Divide through by t and set $t^2 = |Tx|/|x|$ to minimize the right-hand side. This gives

$$|\langle Tx, Tx\rangle| + |\langle T^2 y, x\rangle| \leq 2s|x||Tx|$$

and

$$2|\langle Tx, Tx\rangle| \leq 2s|x||Tx| \leq 2s|x|^2|T|$$

The smallest non-negative s for which this inequality is assured to hold for all $x \in X$ is $|T|$ itself. ///

---

3. *Spectral theorem*

Let T be a *self-adjoint compact* operator on a (non-zero) Hilbert space X. For complex λ, let X_λ be the **λ-eigenspace**

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

of T on X.

[3.0.1] Theorem:

- The completion of $\oplus X_\lambda$ is all of X. That is, there is an orthonormal basis consisting of *eigenvectors*.
- The only possible *accumulation point* of the set of eigenvalues is 0, and if X is infinite-dimensional it *is* an accumulation point.
- The eigenspaces X_λ are *finite-dimensional*.
- All the eigenvalues are *real*.
- One or the other of $\pm|T|$ is an eigenvalue of T

Proof: The last assertion is the most crucial technical point. To prove it, we use the fact that for self-adjoint T we have

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

And note that because T is self-adjoint any value $\langle Tx, x\rangle$ is *real*. Then choose a sequence $\{x_n\}$ so that $|x_n| \leq 1$ and $|\langle Tx, x\rangle| \to |T|$. Then, 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. Invoking the *compactness* of T, we can replace x_n by a subsequence so as to be able to assume without loss of generality that Tx_n converges to some vector y. Then the previous inequality shows that λx_n converges to y. For $\lambda = 0$, we have $|T| = 0$, so $T = 0$. For $\lambda \neq 0$, $\lambda x_n \to y$ implies

$$x_n \to \lambda^{-1} y$$

Thus, letting $x = \lambda^{-1} y$, we have

$$Tx = \lambda x$$

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

Now we use a sort of induction. Let Y be the completion of the sum of all the eigenspaces. Then Y is T-stable. Let $Z = Y^\perp$. We claim that Z is also T-stable, and that on the Hilbert space Z the (restriction of) T is a compact operator. Indeed, for $z \in Z$ and $y \in Y$, we have

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

which proves stability *easily*. And the unit ball in Z is certainly a subset of the unit ball B in X, so is mapped by T to a pre-compact set $TB \cap Z$ in X. Since Z is *closed* in X, the intersection $TB \cap Z$ of Z with the pre-compact set TB is pre-compact. This proves that T restricted to $Z = Y^\perp$ is still compact. The self-adjoint-ness is clear.

Let T_1 be the restriction of T to Z. By construction, T_1 has no eigenvalues on Z, since any such eigenvalue would also be an eigenvalue of T on Z. But 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. ///

Before proceeding, note that in an infinite-dimensional Hilbert space Y a ball B of positive radius $r > 0$ is *not* pre-compact. Indeed, let $\{e_1, e_2, \ldots\}$ be a Hilbert space basis. Then $\{re_1, re_2, re_3 \ldots\}$ is a sequence with *no* convergent subsequence, because all these points are distance $r\sqrt{2}$ apart.

To prove that the eigenspaces 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

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

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

Finally, we prove that all eigenvalues are *real*. Indeed, let $x \in X_\lambda$. Then

$$\lambda\langle x, x\rangle = \langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}$$

which implies that λ is real if $x \neq 0$. ///

---

4. *Construction of compact operators*

[4.0.1] Proposition: Uniform-norm limits of compact operators on Banach spaces are compact.

Proof: Let $T_n \to T$ in uniform operator norm, where the T_n are compact. 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 points $y_1, \ldots, y_t$ such that for any $x \in B$ there is i such that $|T_n x - y_i| < \varepsilon/2$. By the triangle inequality

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

This proves that $T(B)$ is covered by finitely many balls of radius ε. ///

[4.0.2] Remark: The σ-finiteness hypothesis in the following theorem is necessary to make Fubini's theorem work as expected.

[4.0.3] Theorem: *(Hilbert-Schmidt)* Let X, μ and Y, ν be σ-finite measure spaces. Let $K \in L^2(X \times Y, \mu \otimes \nu)$. Then the operator

$$T : L^2(X, \mu) \to L^2(Y, \nu)$$

defined by

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

is a compact operator.

Proof: We grant ourselves that, 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 $L^2(X \times Y)$. This plausible result is non-trivial, needing Fubini's theorem and the σ-finiteness. Thus,

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

with complex c_{ij}, where we should not initially presume that the index set is countable. The square-integrability asserts that

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

In particular, this implies that the indexing sets can be taken to be countable, since an uncountable sum of positive reals cannot converge. Then, given $f \in L^2(X)$, the image Tf is in $L^2(Y)$, since

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

whose $L^2(Y)$ norm is easily estimated by

$$|Tf|_2^2 \le \sum_{ij} |c_{ij}|^2 |\langle f, \varphi_i \rangle|^2\, |\psi_j|_2^2 \le |f|_2^2 \sum_{ij} |c_{ij}|^2\, |\varphi_i|_2^2\, |\psi_j|_2^2 = |f|_2^2 \sum_{ij} |c_{ij}|^2 = |f|_2^2 \cdot |K|^2_{L^2(X \times Y)}$$

We claim that we can write

$$K(x, y) = \sum_i \overline{\varphi_i}(x)\, T\varphi_i(y)$$

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 $K(x, y)$. In particular, with the coefficients c_{ij} from above, we see that

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

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

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

In fact, for the same reason,

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

This fact is essential just below.

Truncate the kernel K by

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

These give the obvious finite-rank operators

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

Granting that any n-dimensional subspace of a Hilbert space is isomorphic to $\mathbb{C}^n$, with all open balls pre-compact, these operators are compact. We claim that they approach T in operator norm. Indeed, let $g = \sum_i c_i \varphi_i$ be in $L^2(X)$. Then

$$(T - T_n)g(y) = \sum_{i>n} b_i\, T\varphi_i(y)$$

and by the triangle inequality and Cauchy-Schwarz-Bunyakowsky inequality

$$|(T-T_n)g(y)| \le \sum_{i>n} |b_i|^2\, |T\varphi_i|_2 \le \Big(\sum_{i>n} |b_i|^2\Big)^{1/2} \Big(\sum_{i>n} |T\varphi_i|_2^2\Big)^{1/2} \le |g|_2 \cdot \Big(\sum_{i>n} |T\varphi_i|_2^2\Big)^{1/2}$$

As observed in the previous paragraph,

$$\lim_n \sum_{i>n} |T\varphi_i|_2^2 = 0$$

Thus, $|T - T_n| \to 0$. ///

[4.0.4] Remark: Given the σ-finiteness of the measure spaces, the argument above is correct whether K is measurable with respect to the product sigma-algebra or only with respect to the *completion*.

---

5. *Limits of finite-rank operators*

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

[5.0.1] Theorem: A compact operator $T : X \to Y$ with Y a Hilbert space is a uniform operator norm limit of finite rank operators, and conversely.

Proof: The converse is a special case of the previous theorem that operator norm limits of compact operators are compact (even in Banach spaces), since finite-rank operators are compact.

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 $\varepsilon \to 0$. ///

[5.0.2] Remark: The conclusion of the previous theorem is known to be false in Banach spaces, although the only example known to this author (Per Enflo, *Acta Math.*, vol. 130, 1973) is rather complicated. Certainly the previous argument using orthogonal projections cannot be employed.

[5.0.3] Remark: In the proof above 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 Frechet space, locally convex, or anything else. However, the general argument is most reasonably a by-product of the development of the general theory of topological vector spaces, and is best delayed until we do that. Thus, we give more elementary and immediate proofs that apply to Hilbert and Banach spaces, despite the fact that these hypotheses are needlessly strong.

[5.0.4] 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, we anticipate that invocation of Hahn-Banach is actually unnecessary, since the same conclusion will be reached (later) without any assumption of local convexity. The only difference in the proof is the method of proving existence of sufficiently many linear functionals.

[5.0.5] 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. ///

[5.0.6] Remark: The proof for normed spaces works in any topological vector space in which Hahn-Banach holds. We will see later that this is so 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 $\mathbb{C}^n$.
