II. Compact Self-Adjoint Operators, 34-53

DOI: 10.3792/euclid/9781429799911-2

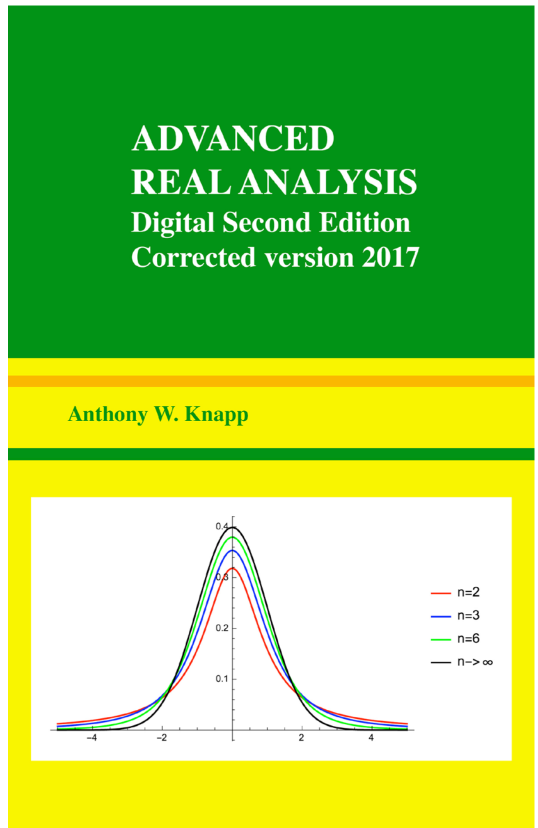
from

Advanced Real Analysis

Digital Second Edition
Corrected version 2017

Anthony W. Knapp

Full Book DOI: 10.3792/euclid/9781429799911
ISBN: 978-1-4297-9991-1



Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, N.Y. 11733–1729, U.S.A.
Email to: aknapp@math.stonybrook.edu
Homepage: www.math.stonybrook.edu/∼aknapp

Title: Advanced Real Analysis
Cover: Normal distribution as a limit of Gosset's t distribution; see page 421.

Mathematics Subject Classification (2010): 46–01, 42–01, 43–01, 35–01, 34–01, 47–01, 58–01, 60A99, 60F05, 28C10, 42C40, 65T60.

First Edition, ISBN-13 978-0-8176-4382-9

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

Published by the Author

CHAPTER II

Compact Self-Adjoint Operators

Abstract. This chapter proves a first version of the Spectral Theorem and shows how it applies to complete the analysis in Sturm's Theorem of Section I.3.

Section 1 introduces compact linear operators from a Hilbert space into itself and characterizes them as the limits in the operator norm topology of the linear operators of finite rank. The adjoint of a compact operator is compact.

Section 2 proves the Spectral Theorem for compact self-adjoint operators on a Hilbert space, showing that such operators have orthonormal bases of eigenvectors with eigenvalues tending to 0.

Section 3 establishes two versions of the Hilbert–Schmidt Theorem concerning self-adjoint integral operators with a square-integrable kernel. The abstract version gives an L^2 expansion of the members of the image of the operator in terms of eigenfunctions, and the concrete version, valid when the kernel is continuous and the space is compact metric, proves that the eigenfunctions are continuous and the expansion in terms of eigenfunctions is uniformly convergent.

Section 4 introduces unitary operators on a Hilbert space, establishing the equivalence of three conditions that may be used to define them.

Section 5 studies compact linear operators on an abstract Hilbert space, with special attention to two kinds—the Hilbert–Schmidt operators and the operators of trace class. All three sets of operators—compact, Hilbert–Schmidt, and trace-class—are ideals in the algebra of all bounded linear operators and are closed under the operation of adjoint. Trace-class implies Hilbert–Schmidt, which implies compact. The product of two Hilbert–Schmidt operators is of trace class.

1. Compact Operators

Let H be a real or complex Hilbert space with inner product[1] $(\cdot\,,\,\cdot)$ and norm $\|\cdot\|$. A bounded linear operator $L : H \to H$ is said to be **compact** if L carries the closed unit ball of H to a subset of H that has compact closure, i.e., if each bounded sequence $\{u_n\}$ in H has the property that $\{L(u_n)\}$ has a convergent subsequence.[2] The first three conclusions of the next proposition together give a characterization of the compact operators on H.

[1]This book follows the convention that inner products are linear in the first variable and conjugate linear in the second variable.

[2]Some books use the words "completely continuous" in place of "compact" for this kind of operator.

Proposition 2.1. Let $L : H \to H$ be a bounded linear operator on a Hilbert space H. Then

(a) L is compact if the image of L is finite dimensional,
(b) L is compact if L is the limit, in the operator norm, of a sequence of compact operators,
(c) L compact implies that there exist bounded linear operators $L_n : H \to H$ such that $L = \lim L_n$ in the operator norm and the image of each L_n is finite dimensional,
(d) L compact implies L^* compact.

PROOF. For (a), let M be the image of L. Being finite dimensional, M is closed and is hence a Hilbert space. Let $\{v_1, \dots, v_k\}$ be an orthonormal basis. The linear mapping that carries each v_j to the j^{th} standard basis vector e_j in the space of column vectors is then a linear isometry of M onto $\mathbb{R}^k$ or $\mathbb{C}^k$. In $\mathbb{R}^k$ and $\mathbb{C}^k$, the closed ball about 0 of radius $\|L\|$ is compact, and hence the closed ball about 0 of radius $\|L\|$ in M is compact. The latter closed ball contains the image of the closed unit ball of H under L, and hence L is compact.

For (b), let B be the closed unit ball of H. Write $L = \lim L_n$ in the operator norm, each L_n being compact. Since the subsets of a complete metric space having compact closure are exactly the totally bounded subsets, it is enough to prove that $L(B)$ is totally bounded. Let $\epsilon > 0$ be given, and choose n large enough so that $\|L_n - L\| < \epsilon/2$. With n fixed, $L_n(B)$ is totally bounded since $L_n(B)$ is assumed to have compact closure. Thus we can find finitely many points $v_1, \dots, v_k$ such that the open balls of radius $\epsilon/2$ about the v_j's together cover $L_n(B)$. We shall prove that the open balls of radius ϵ about the v_j's together cover $L(B)$. In fact, if u is given with $\|u\| \le 1$, choose j with $\|L_n(u) - v_j\| < \epsilon/2$. Then $\|L(u) - v_j\| \le \|L(u) - L_n(u)\| + \|L_n(u) - v_j\| < \|L_n - L\|\|u\| + \frac{\epsilon}{2} \le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$, as required.

For (c), we may assume that H is infinite dimensional. Since L is compact, there exists a compact subset K of H containing the image of the closed unit ball. As a compact metric space, K is separable. Let $\{w_n\}$ be a countable dense set, and let M be the smallest closed vector subspace of H containing all w_n. Since the closure of $\{w_n\}$ contains K, M contains K. The subspace M is separable: in fact, if the scalars are real, then the set of all rational linear combinations of the w_n's is a countable dense set; if the scalars are complex, then we obtain a countable dense set by allowing the scalars to be of the form $a + bi$ with a and b rational.

Since M is a closed vector subspace, it is a Hilbert space and has an orthonormal basis S. The set S must be countable since the open balls of radius $1/2$ centered at the members of S are disjoint and would otherwise contradict the fact that every topological subspace of a separable topological space is Lindelöf. Thus let us list the members of S as $v_1, v_2, \dots$. For each n, let M_n be the (closed) linear span of $\{v_1, \dots, v_n\}$, and let E_n be the orthogonal projection on M_n. The linear operator $E_n L$ is bounded, being a composition of bounded linear operators, and its image is contained in the finite-dimensional space M_n. Hence it is enough to show for each $\epsilon > 0$ that there is some n with $\|(1 - E_n)L\| < \epsilon$. If this condition were to fail, we could find some $\epsilon > 0$ such that $\|(1 - E_n)L\| \geq \epsilon$ for every n. With ϵ fixed in this way, choose for each n some vector u_n of norm 1 such that $\|(1 - E_n)L(u_n)\| \geq \epsilon/2$. The sequence $\{L(u_n)\}$ lies in the compact set K. Choose a convergent subsequence $\{L(u_{n_k})\}$, and let $v = \lim L(u_{n_k})$. For n_k sufficiently large, we have $\|v - L(u_{n_k})\| \leq \epsilon/4$. In this case,

$$\|(1 - E_{n_k})v\| \geq \|(1 - E_{n_k})L(u_{n_k})\| - \|(1 - E_{n_k})(v - L(u_{n_k}))\| \geq \tfrac{\epsilon}{2} - \tfrac{\epsilon}{4} = \tfrac{\epsilon}{4}.$$

On the other hand, v is in M, and v is of the form $v = \sum_{j=1}^{\infty} (v, v_j)v_j$. In this expression we have $E_n(v) = \sum_{j=1}^{n} (v, v_j)v_j$, and these partial sums converge to v in H. In short, $\lim_n E_n v = v$. Then $\|(1 - E_n)v\|$ tends to 0, and this contradicts our estimate $\|(1 - E_{n_k})v\| \geq \frac{\epsilon}{4}$.

For (d), first suppose that the image of L is finite dimensional, and choose an orthonormal basis $\{u_1, \dots, u_n\}$ of the image. Then $L(u) = \sum_{j=1}^{n} (L(u), u_j)u_j = \sum_{j=1}^{n} (u, L^*(u_j))u_j$. Taking the inner product with v gives $(u, L^*(v)) = (L(u), v) = \sum_{j=1}^{n} (u, L^*(u_j))(u_j, v)$. This equality shows that $L^*(v)$ and $\sum_{j=1}^{n} (v, u_j)L^*(u_j)$ have the same inner product with every u. Thus they must be equal, and we conclude that the image of L^* is finite dimensional.

Now suppose that L is any compact operator on H. Given $\epsilon > 0$, use (c) to choose a bounded linear operator L_n with finite-dimensional image such that $\|L - L_n\| < \epsilon$. Since a bounded linear operator and its adjoint have the same norm, $\|L^* - L_n^*\| < \epsilon$. Since L_n^* has finite-dimensional image, according to what we have just seen, and since we can obtain such an approximation for any $\epsilon > 0$, (b) shows that L^* is compact. □

2. Spectral Theorem for Compact Self-Adjoint Operators

Let $L : H \to H$ be a bounded linear operator on the real or complex Hilbert space H. One says that a nonzero vector v is an **eigenvector** of L if $L(v) = cv$ for some constant c; the constant c is called the corresponding **eigenvalue**. The set of all u for which $L(u) = cu$ is a closed vector subspace; under the assumption that this subspace is not 0, it is called the **eigenspace** for the eigenvalue c.

In the finite-dimensional case, the self-adjointness condition $L^* = L$ means that L corresponds to a **Hermitian matrix** A, i.e., a matrix equal to its conjugate transpose, once one fixes an ordered orthonormal basis. In this case it is shown in linear algebra that the members of an orthonormal basis can be chosen to be eigenvectors of L, the eigenvalues all being real. In terms of matrices, the corresponding matrix A is conjugate via a **unitary matrix**, i.e., a matrix whose conjugate transpose is its inverse, to a diagonal matrix with real entries. This result is called the Spectral Theorem for such linear operators or matrices. A quick proof goes as follows: An eigenvector v of L with eigenvalue c has $(L-cI)(v)=0$, and this implies that the matrix A of L has the property that $A-cI$ has a nonzero null space. Hence $\det(A-cI)=0$ if and only if c is an eigenvalue of L. One readily sees from the self-adjointness of L that all complex roots of $\det(A-cI)$ have to be real. Moreover, if L carries a vector subspace M into itself, then L carries $M^\perp$ into itself as well. Finite-dimensionality forces A to have a complex eigenvalue, and this must be real. Hence there is a nonzero vector u with $L(u)=cu$ for some real c. Normalizing, we may assume that u has norm 1. If M consists of the scalar multiples of u, then L carries $M^\perp$ to itself, and the restriction of L to $M^\perp$ is self adjoint. Proceeding inductively, we obtain a system of orthogonal eigenvectors for L, each of norm 1.

A certain amount of this argument works in the infinite-dimensional case. In fact, suppose that L is self adjoint. Then any u in H has

$$(L(u),u)=(u,L^*(u))=(u,L(u))=\overline{(L(u),u)},$$

and hence the function $u\mapsto(L(u),u)$ is real-valued. If u is an eigenvector in H with eigenvalue c, i.e., if $L(u)=cu$, then $c(u,u)=(L(u),u)$ is real; since (u,u) is real and nonzero, c is real. If u_1 and u_2 are eigenvectors for distinct eigenvalues c_1 and c_2, then u_1 and u_2 are orthogonal because

$$(c_1-c_2)(u_1,u_2)=(c_1u_1,u_2)-(u_1,c_2u_2)=(L(u_1),u_2)-(u_1,L(u_2))=0.$$

If M is a vector subspace of H with $L(M)\subseteq M$, then also $L(M^\perp)\subseteq M^\perp$ because $m\in M$ and $m^\perp\in M^\perp$ together imply

$$0=(L(m),m^\perp)=(m,L(m^\perp)).$$

These observations prove everything in the following proposition except the last statement.

Proposition 2.2. If $L:H\to H$ is a bounded self-adjoint linear operator on a Hilbert space H, then $u\mapsto(L(u),u)$ is real-valued, every eigenvalue of L is real, eigenvectors under L for distinct eigenvalues are orthogonal, and every vector subspace M with $L(M)\subseteq M$ has $L(M^\perp)\subseteq M^\perp$. In addition,

$$\|L\|=\sup_{\|u\|\le 1}|(L(u),u)|.$$

PROOF. We are left with proving the displayed formula. Inequality in one direction is easy: we have

$$\sup_{\|u\|\le 1} |(L(u),u))| \le \sup_{\substack{\|u\|\le 1,\\ \|v\|\le 1}} |(L(u),v)| = \|L\|.$$

With $C = \sup_{\|u\|\le 1}(L(u),u))$, we are therefore to prove that $\|L\| \le C$, hence that $\|L(u)\| \le C\|u\|$ for all u. In doing so, we may assume that $u \neq 0$ and $L(u) \neq 0$. Let t be a positive real number. Since $(L^2(u),u) = (L(u),L(u))$, we have

$$\begin{aligned}
&\|L(u)\|^2 \\
&\quad = \tfrac{1}{4}\Big[\big(L(tu+t^{-1}L(u)),\ tu+t^{-1}L(u)\big) - \big(L(tu-t^{-1}L(u)),\ tu-t^{-1}L(u)\big)\Big] \\
&\quad \le \tfrac{1}{4}\Big[C\|tu+t^{-1}L(u)\|^2 + C\|tu-t^{-1}L(u)\|^2\Big] \\
&\quad = \tfrac{1}{2}\,C\Big[\|tu\|^2 + \|t^{-1}L(u)\|^2\Big],
\end{aligned}$$

the last step following from the parallelogram law. By differential calculus the minimum of an expression $a^2t^2 + b^2t^{-2}$, in which a and b are positive constants, is attained when $t^2 = b/a$. Here $a = \|u\|$ and $b = \|L(u)\|$, and thus $\|L(u)\|^2 \le \frac{C}{2}\big[\|L(u)\|\|u\| + \|L(u)\|\|u\|\big] = C\|L(u)\|\|u\|$. Dividing by $\|L(u)\|$ gives $\|L(u)\| \le C\|u\|$ and completes the proof. □

In the infinite-dimensional case, in which we work with the operator L but no matrix, consider what is needed to imitate the proof of the finite-dimensional Spectral Theorem and thereby find an orthonormal basis of vectors carried by L to multiples of themselves. In the formula of Proposition 2.2, if we can find some u with $\|u\| = 1$ such that $\|L\| = |(L(u),u)|$, then this u satisfies $\|L\|\|u\|^2 = |(L(u),u)| \le \|L(u)\|\|u\| \le \|L\|\|u\|^2$, and we conclude that $|(L(u),u)| = \|L(u)\|\|u\|$, i.e., that equality holds in the Schwarz inequality. Reviewing the proof of the Schwarz inequality, we see that $L(u)$ and u are proportional. Thus u is an eigenvector of L, and we can at least get started with the proof.

Unfortunately an orthonormal basis of eigenvectors need not exist for a self-adjoint L without an extra hypothesis. In fact, take $H = L^2([0,1])$ with $(f,g) = \int_0^1 f\bar{g}\,dx$, and define $L(f)(x) = xf(x)$. This linear operator L has norm 1, and the equality $(f, L(g)) = \int_0^1 xf(x)\overline{g(x)}\,dx = (L(f),g)$ shows that L is self adjoint. On the other hand, the only function f with $xf = cf$ a.e. for some constant c is the 0 function. Thus we get no eigenvectors at all, and the supremum in the formula of Proposition 2.2 need not be attained.

The hypothesis that we shall add to obtain an orthonormal basis of eigenvectors is that L is compact in the sense of the previous section. Each compact self-adjoint operator has an orthonormal basis of eigenvectors, according to the following theorem.

Theorem 2.3 (Spectral Theorem for compact self-adjoint operators). Let $L : H \to H$ be a compact self-adjoint linear operator on a real or complex Hilbert space H. Then H has an orthonormal basis of eigenvectors of L. In addition, for each scalar λ, let

$$H_\lambda = \{u \in H \mid L(u) = \lambda u\},$$

so that $H_\lambda - \{0\}$ consists exactly of the eigenvectors of L with eigenvalue λ. Then the number of eigenvalues of L is countable, the eigenvalues are all real, the spaces H_λ are mutually orthogonal, each H_λ for $\lambda \neq 0$ is finite dimensional, any orthonormal basis of H of eigenvectors under L is the union of orthonormal bases of the H_λ's, and for any $\epsilon > 0$, there are only finitely many λ with $H_\lambda \neq 0$ and $|\lambda| \geq \epsilon$. Moreover, either or both of $\|L\|$ and $-\|L\|$ are eigenvalues, and these are the eigenvalues with the largest absolute value.

PROOF. We know from Proposition 2.2 that the eigenvalues of L are all real and that the spaces H_λ are mutually orthogonal. In addition, the formula $\|L\| = \sup_{\|u\|\leq 1} \|L(u)\|$ shows that no eigenvalue can be greater than $\|L\|$ in absolute value.

The theorem certainly holds if $L = 0$ since every nonzero vector is an eigenvector. Thus we may assume that $\|L\| > 0$.

The main step is to produce an eigenvector with one of $\|L\|$ and $-\|L\|$ as eigenvalue. Taking the equality $\|L\| = \sup_{\|u\|\leq 1} |(L(u), u))|$ of Proposition 2.2 into account, choose a sequence $\{u_n\}$ with $\|u_n\| = 1$ such that $\lim_n |(L(u_n), u_n)| = \|L\|$. Since the proposition shows that $(L(u_n), u_n)$ has to be real, we may assume that this sequence is chosen so that $\lambda = \lim_n (L(u_n), u_n)$ exists. Then $\lambda = \pm\|L\|$. Using the compactness of L and passing to a subsequence if necessary, we may assume that $L(u_n)$ converges to some limit v_0. Meanwhile,

$$\begin{aligned} 0 \leq \|L(u_n) - \lambda u_n\|^2 &= \|L(u_n)\|^2 - 2\lambda \operatorname{Re}(L(u_n), u_n) + \lambda^2 \|u_n\|^2 \\ &\leq \|L\|^2 - 2\lambda \operatorname{Re}(L(u_n), u_n) + \lambda^2. \end{aligned}$$

The equalities $\lambda^2 = \|L\|^2$ and $\lim_n (L(u_n), u_n) = \lambda$ show that the right side tends to 0, and thus $\lim_n \|L(u_n) - \lambda u_n\| = 0$. Since $\lim_n \|L(u_n) - v_0\| = 0$ also, the triangle inequality shows that $\lim \lambda u_n$ exists and equals v_0. Since $\lambda \neq 0$, $\lim u_n$ exists and $v_0 = \lambda \lim u_n$. Consequently $\|v_0\| = |\lambda| \lim \|u_n\| = |\lambda| = \|L\| \neq 0$. Applying L to the equation $v_0 = \lambda \lim u_n$ and taking into account that L is continuous and that $\lim L(u_n) = v_0$, we see that $L(v_0) = \lambda v_0$. Thus v_0 is an eigenvector with eigenvalue λ, and the main step is complete.

Now consider the collection of all orthonormal systems of eigenvectors for L, and order it by inclusion upward. A chain consists of nested such systems, and the union of the members of a chain is again such an orthonormal system. By Zorn's Lemma the collection contains a maximal element S. Let M be the smallest closed vector subspace containing this maximal orthonormal system S of eigenvectors. Since the collection of all finite linear combinations of members of S is dense in M, the continuity of L shows that $L(M) \subseteq M$. By Proposition 2.2, $L(M^\perp) \subseteq M^\perp$. The equality $(L(u), v) = (u, L(v))$ for any two members u and v of $M^\perp$ shows that the restriction of L to $M^\perp$ is self adjoint, and this restriction is certainly bounded and compact. Arguing by contradiction, suppose $M^\perp \neq 0$. Then either $L = 0$ or else $L \neq 0$ and the main step above shows that L has an eigenvector in $M^\perp$. Thus L has an eigenvector v_0 of norm 1 in $M^\perp$ in either case. But then $S \cup \{v_0\}$ would be an orthonormal system of eigenvectors properly containing S, in contradiction to the maximality. We conclude that $M^\perp = 0$. Since M is a closed vector subspace of H, it satisfies $M^{\perp\perp} = M$. Therefore $M = (M^\perp)^\perp = 0^\perp = H$, and H has an orthonormal basis of eigenvectors.

With the orthonormal basis $S = \{v_\alpha\}$ of eigenvectors fixed, consider all v_α's for which the corresponding eigenvalue λ_α has $|\lambda_\alpha| \geq \epsilon$. If α_1 and α_2 are two distinct such indices, we have

$$\begin{aligned}
\|L(v_{\alpha_1}) - L(v_{\alpha_2})\|^2 &= \|\lambda_{\alpha_1} v_{\alpha_1} - \lambda_{\alpha_2} v_{\alpha_2}\|^2 \\
&= \|\lambda_{\alpha_1} v_{\alpha_1}\|^2 + \|\lambda_{\alpha_2} v_{\alpha_2}\|^2 \quad \text{by the Pythagorean theorem} \\
&= |\lambda_{\alpha_1}|^2 + |\lambda_{\alpha_2}|^2 \\
&\geq 2\epsilon^2.
\end{aligned}$$

If there were infinitely many such eigenvectors v_{α_n}, the bounded sequence $\{L(v_{\alpha_n})\}$ could not have a convergent subsequence, in contradiction to compactness. Thus only finitely many members of S have eigenvalue with absolute value $\geq \epsilon$.

Fix $\lambda \neq 0$, let S_λ be the finite set of members of S with eigenvalue λ, and let H_λ be the linear span of S_λ. If v is an eigenvector of L for the eigenvalue λ beyond the vectors in H_λ, then the expansion

$$v = \sum_{v_\alpha \in S_\lambda} (v, v_\alpha) v_\alpha + \sum_{v_\alpha \in S - S_\lambda} (v, v_\alpha) v_\alpha$$

shows that $(v, v_\alpha) \neq 0$ for some v_α in $S - S_\lambda$. This v_α must have eigenvalue λ' different from λ, and then Proposition 2.2 gives the contradiction $(v, v_\alpha) = 0$. We conclude that H_λ is the entire eigenspace for eigenvalue λ and that the eigenvalues of the members of S are the only eigenvalues of L.

For each positive integer n, we know that only finitely many eigenvalues λ corresponding to members of S have $|\lambda| \geq 1/n$. Since every eigenvalue of L is the eigenvalue for some member of S, the number of eigenvalues λ of L with $|\lambda| \geq 1/n$ is finite. Taking the union of these sets as n varies, we see that the number of eigenvalues of L is countable. This completes the proof. □

3. Hilbert–Schmidt Theorem

The Hilbert–Schmidt Theorem was postponed from Section I.3, where it was used in connection with Sturm–Liouville theory. The nub of the matter is the Spectral Theorem for compact self-adjoint operators on a Hilbert space, Theorem 2.3. But the actual result quoted in Section I.3 contains an overlay of measure theory and continuity. Correspondingly there is an abstract Hilbert–Schmidt Theorem, which combines the Spectral Theorem with the measure theory, and then there is a concrete form, which adds the hypothesis of continuity and obtains extra conclusions from it.

The abstract theorem works with an **integral operator** on L^2 of a σ-finite measure space (X, μ), the operator being of the form

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

where $K(x, y)$ is measurable on $X \times X$. The function K is called the **kernel** of the operator.[3] If f is in $L^2(X, \mu)$, then the Schwarz inequality gives $|Tf(x)| \leq \|K(x, \cdot\,)\|_2 \|f\|_2$ for each x in X. Squaring both sides, integrating, and taking the square root yields $\|Tf\|_2 \leq \big(\int_{X\times X} |K|^2\, d(\mu\times\mu)\big)^{1/2} \|f\|_2$. As a linear operator on $L^2(X, \mu)$, T therefore has operator norm satisfying

$$\|T\| \leq \Big(\int_X \int_X |K(x, y)|^2\, d\mu(x)\, d\mu(y)\Big)^{1/2} = \|K\|_2.$$

In particular, T is bounded if K is square-integrable on $X \times X$. In this case the adjoint of T is given by

$$T^* g(x) = \int_X \overline{K(y, x)} g(y)\, d\mu(y)$$

because $(Tf, g) = \int_X \int_X K(x, y) f(y) \overline{g(x)}\, d\mu(y)\, d\mu(x)$ and because the asserted form of T^* has

$$\begin{aligned}(f, T^*g) &= \textstyle\int_X f(x) \overline{\big(\int_X \overline{K(y, x)} g(y)\, d\mu(y)\big)}\, d\mu(x)\\ &= \textstyle\int_X \int_X f(x) K(y, x) \overline{g(y)}\, d\mu(y)\, d\mu(x).\end{aligned}$$

[3] Not to be confused with the abstract-algebra notion of "kernel" as the set mapped to 0.

Theorem 2.4 (Hilbert–Schmidt Theorem, abstract form). Let (X, μ) be a σ-finite measure space, and let $K(\cdot, \cdot)$ be a complex-valued L^2 function on $X \times X$ such that $K(x, y) = \overline{K(y, x)}$ for all x and y in X. Then the linear operator T defined by

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

is a self-adjoint compact operator on the Hilbert space $L^2(X, \mu)$ with $\|T\| \le \|K\|_2$. Consequently if for each complex $\lambda \ne 0$, a vector subspace V_λ of $L^2(X, \mu)$ is defined by

$$V_\lambda = \big\{ f \in L^2(X, \mu) \;\big|\; Tf = \lambda f \big\},$$

then each V_λ is finite dimensional, the space V_λ is nonzero for only countably many λ, the spaces V_λ are mutually orthogonal with respect to the inner product on $L^2(X, \mu)$, the λ's with $V_\lambda \ne 0$ are all real, and for any $\epsilon > 0$, there are only finitely many λ with $V_\lambda \ne 0$ and $|\lambda| \ge \epsilon$. The largest value of $|\lambda|$ for which $V_\lambda \ne 0$ is $\|T\|$. Moreover, the vector subspace of L^2 orthogonal to all V_λ is the kernel of T, so that if $v_1, v_2, \ldots$ is an enumeration of the union of orthonormal bases of the spaces V_λ with $\lambda \ne 0$, then for any f in $L^2(X, \mu)$,

$$Tf = \sum_{n=1}^{\infty} (Tf, v_n) v_n,$$

the series on the right side being convergent in $L^2(X, \mu)$.

PROOF. Theorem 2.3 shows that it is enough to prove that the self-adjoint bounded linear operator T is compact. Choose a sequence of simple functions K_n square integrable on $X \times X$ such that $\lim_n \|K - K_n\|_2 = 0$, and define $T_n f(x) = \int_X K_n(x, y) f(y)\, d\mu(y)$. The linear operator T_n is bounded with $\|T_n\| \le \|K_n\|_2$, and it has finite-dimensional image since K_n is simple. By Proposition 2.1a, T_n is compact. Since $\|T - T_n\| \le \|K - K_n\|_2$ and since the right side tends to 0, T is exhibited as the limit of T_n in the operator norm and is compact by Proposition 2.1b. □

Now we include the overlay of continuity. The additional assumptions are that X is a compact metric space, μ is a Borel measure on X that assigns positive measure to every nonempty open set, and K is continuous on $X \times X$. The additional conclusions are that the eigenfunctions for the nonzero eigenvalues are continuous and that the series expansion actually converges absolutely uniformly as well as in L^2. The result used in Section I.3 was the special case of this result with $X = [a, b]$ and μ equal to Lebesgue measure.

Theorem 2.5 (Hilbert–Schmidt Theorem, concrete form). Let X be a compact metric space, let μ be a Borel measure on X that assigns positive measure to every nonempty open set, and let $K(\,\cdot\,,\,\cdot\,)$ be a complex-valued continuous function on $X \times X$ such that $K(x, y) = \overline{K(y, x)}$ for all x and y in X. Then the linear operator T defined by

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

is a self-adjoint compact operator on the Hilbert space $L^2(X, \mu)$ with $\|T\| \le \|K\|_2$, and its image lies in $C(X)$. Consequently the vector subspace V_λ of $L^2(X, \mu)$ defined for any complex $\lambda \neq 0$ by

$$V_\lambda = \big\{ f \in L^2(X, \mu) \,\big|\, Tf = \lambda f \big\}$$

consists of continuous functions, each V_λ is finite dimensional, the space V_λ is nonzero for only countably many λ, the spaces V_λ are mutually orthogonal with respect to the inner product on $L^2(X, \mu)$, the λ's with $V_\lambda \neq 0$ are all real, and for any $\epsilon > 0$, there are only finitely many λ with $V_\lambda \neq 0$ and $|\lambda| \ge \epsilon$. The largest value of $|\lambda|$ for which $V_\lambda \neq 0$ is $\|T\|$. If $v_1, v_2, \dots$ is an enumeration of the union of orthonormal bases of the spaces V_λ with $\lambda \neq 0$, then for any f in $L^2(X, \mu)$,

$$Tf(x) = \sum_{n=1}^{\infty} (Tf, v_n) v_n(x),$$

the series on the right side being absolutely uniformly convergent for x in X.

REMARK. The hypothesis that μ assigns positive measure to every nonempty open set is used only to identify $\sum_{n=1}^{\infty} (Tf, v_n) v_n(x)$ with $Tf(x)$ at every point. Without this particular hypothesis on μ, the series is still absolutely uniformly convergent, but its sum is shown to equal $Tf(x)$ only almost everywhere with respect to μ.

PROOF. Given $\epsilon > 0$, choose $\delta > 0$ by uniform continuity of K such that $|K(x, y) - K(x_0, y_0)| \le \epsilon$ whenever (x, y) and (x_0, y_0) are at distance $\le \delta$. If f is in $L^2(X, \mu)$ and the points x and x_0 are at distance $\le \delta$, then (x, y) and (x_0, y) are at distance $\le \delta$ and hence

$$\begin{aligned} |Tf(x) - Tf(x_0)| &\le \textstyle\int_X |K(x, y) - K(x_0, y)| |f(y)|\, d\mu(y) \\ &\le \epsilon \textstyle\int_X |f(y)|\, d\mu(y) \le \epsilon \|f\|_2 (\mu(X))^{1/2}, \end{aligned}$$

the last step following from the Schwarz inequality. This proves that Tf is continuous for each f in $L^2(X, \mu)$. In particular, if $Tf = \lambda f$ with $\lambda \neq 0$, then $f = T(\lambda^{-1} f)$ exhibits f as in the image of T and therefore as continuous.

Everything in the theorem now follows from Theorem 2.4 except for the absolute uniform convergence to $Tf(x)$ in the last sentence of the theorem.

For the absolute uniform convergence, let $(\cdot\,,\,\cdot)$ denote the inner product in $L^2(X,\mu)$. We begin by considering the function $\overline{K(x,\,\cdot)}$ for fixed x. It satisfies

$$(\overline{K(x,\,\cdot)}, v_n) = \textstyle\int_X \overline{K(x,y)}\,\overline{v_n(y)}\,d\mu(y) = \overline{(Tv_n)(x)} = \overline{\lambda_n}\,\overline{v_n(x)}$$

if v_n is in V_{λ_n}, and Bessel's inequality gives

$$\sum_{n=1}^{N} |\lambda_n|^2 |v_n(x)|^2 \le \int_X |K(x,y)|^2\,d\mu(y) \le \|K\|_{\sup}^2 \mu(X) \qquad (*)$$

for all N and x. Since the v_n form an orthonormal basis of $V_0^{\perp}$,

$$\lim_{N\to\infty} \Big\| Tg - \textstyle\sum_{n=1}^{N} (Tg, v_n) v_n \Big\|_2 = 0 \qquad (**)$$

for all g in $L^2(X,\mu)$. Meanwhile, we have

$$(Tg, v_n)v_n(x) = (g, Tv_n)v_n(x) = \lambda_n (g, v_n) v_n(x).$$

Application of the Schwarz inequality and $(*)$ gives

$$\begin{aligned}
\sum_{n=M}^{N} |(Tg, v_n)v_n(x)| &= \sum_{n=M}^{N} |\lambda_n (g, v_n) v_n(x)| \\
&\le \Big(\sum_{n=M}^{N} |\lambda_n|^2 |v_n(x)|^2 \Big)^{1/2} \Big(\sum_{n=M}^{N} |(g, v_n)|^2 \Big)^{1/2} \\
&\le \|K\|_{\sup} \mu(X)^{1/2} \Big(\sum_{n=M}^{N} |(g, v_n)|^2 \Big)^{1/2}.
\end{aligned}$$

Bessel's inequality shows that the series $\sum_{n=1}^{\infty} |(g, v_n)|^2$ converges and has sum $\le \|g\|_2^2$. Therefore $\sum_{n=M}^{N} |(g, v_n)|^2$ tends to 0 as M and N tend to infinity, and the rate is independent of x. Consequently the series $\sum_{n=1}^{\infty} |(Tg, v_n)v_n(x)|$ is uniformly Cauchy, and it follows that the series $\sum_{n=1}^{\infty} (Tg, v_n)v_n(x)$ is absolutely uniformly convergent for x in X. Since the uniform limit of continuous functions is continuous, the sum has to be a continuous function. Since $(**)$ shows that $\sum_{n=1}^{N} (Tg, v_n)v_n$ converges in $L^2(X,\mu)$ to Tg, a subsequence of $\sum_{n=1}^{N} (Tg, v_n)v_n(x)$ converges almost everywhere to $Tg(x)$. Since Tg is continuous, the set where $\sum_{n=1}^{\infty} (Tg, v_n)v_n(x) \ne Tg(x)$ is an open set. The fact that this set has measure 0 implies, in view of the hypothesis on μ, that this set is empty. Thus $\sum_{n=1}^{N} (Tg, v_n)v_n(x)$ converges absolutely uniformly to $Tg(x)$. □

4. Unitary Operators

In $\mathbb{C}^N$, a unitary matrix corresponds in the standard basis to a **unitary** linear transformation U, i.e., one with $U^* = U^{-1}$. Such a transformation preserves inner products and therefore carries any orthonormal basis to another orthonormal basis. Conversely any linear transformation from $\mathbb{C}^N$ to itself that carries some orthonormal basis to another orthonormal basis is unitary. For the infinite-dimensional case we define a linear operator to be **unitary** if it satisfies the equivalent conditions in the following proposition.[4]

Proposition 2.6. If V is a real or complex Hilbert space, then the following conditions on a linear operator $U : V \to V$ are equivalent:

(a) $UU^* = U^*U = 1$,

(b) U is onto V, and $(Uv, Uv') = (v, v')$ for all v and v' in V,

(c) U is onto V, and $\|Uv\| = \|v\|$ for all v in V.

A unitary operator carries any orthonormal basis to an orthonormal basis. Conversely if $\{u_i\}$ and $\{v_i\}$ are orthonormal bases, then there exists a unique bounded linear operator U such that $Uu_i = v_i$ for all i, and U is unitary.

REMARKS. In the finite-dimensional case the condition "$UU^* = 1$" in (a) and the condition "U is onto V" in (b) and (c) follow from the rest, but that implication fails in the infinite-dimensional case. Any two orthonormal bases have the same cardinality, by Proposition 12.11 of *Basic*, and hence the index sets for $\{u_i\}$ and $\{v_i\}$ in the statement of the proposition may be taken to be the same.

PROOF. If (a) holds, then $UU^* = 1$ proves that U is onto, and $U^*U = 1$ proves that $(Uv, Uv') = (U^*Uv, v') = (v, v')$. Thus (b) holds. In the reverse direction, suppose that (b) holds. From $(U^*Uv, v') = (Uv, Uv') = (v, v')$ for all v and v', we see that $U^*U = 1$. Thus U is one-one. Since U is assumed onto, it has a two-sided inverse, which must then equal U^* since any left inverse equals any right inverse. Thus (a) holds, and (a) and (b) are equivalent. Conditions (b) and (c) are equivalent by polarization.

If $\{u_i\}$ is an orthonormal basis and U is unitary, then $(Uu_i, Uu_j) = (u_i, u_j) = \delta_{ij}$ by (b), and hence $\{Uu_i\}$ is an orthonormal set. If $(v, Uu_i) = 0$ for all i, then $(U^*v, u_i) = 0$ for all i, $U^*v = 0$, and $v = U(U^*v) = U0 = 0$. So $\{Uu_i\}$ is an orthonormal basis.

If $\{u_i\}$ and $\{v_i\}$ are orthonormal bases, define U on finite linear combinations of the u_i by $U\big(\sum_i c_i u_i\big) = \sum_i c_i v_i$. Then $\big\|U\big(\sum_i c_i u_i\big)\big\|^2 = \big\|\sum_i c_i v_i\big\|^2 =$

[4] This book uses the term "unitary" for both real and complex Hilbert spaces. A unitary linear operator from a *real* Hilbert space into itself is traditionally said to be **orthogonal**, but there is no need to reject the word "unitary" for real Hilbert spaces.

$\sum_i |c_i|^2 = \| \sum_i c_i u_i \|^2$. Hence U extends to a bounded linear operator on V, necessarily preserving norms. It must be onto V since it preserves norms and its image contains the dense set of finite linear combinations $\sum_i c_i v_i$. Thus (c) holds, and U is unitary. □

Since unitary operators are exactly the invertible linear operators that preserve inner products, they are the ones that serve as isomorphisms of a Hilbert space with itself. Theorem 2.3 and Proposition 2.6 together give us a criterion for deciding whether two compact self-adjoint operators on a Hilbert space are related to each other by an underlying isomorphism of the Hilbert space: the criterion is that the two operators have the same eigenvalues, that the dimension of the eigenspace for each nonzero eigenvalue of one operator match the dimension of the eigenspace for that eigenvalue of the other operator, and that the Hilbert-space dimension of the zero eigenspaces of the two operators match.

5. Classes of Compact Operators

In this section we bring together various threads concerning compact operators, integral operators, the Hilbert–Schmidt Theorem, the Hilbert–Schmidt norm of a square matrix, and traces of matrices. The end product is to consist of some relationships among these notions, together with the handy notion of the trace of an operator. Once we have multiple Fourier series available as a tool in the next chapter, we will be able to supplement the results of the present section and obtain a formula for computing the trace of certain kinds of integral operators. Let us start with various notions about bounded linear operators from an abstract real or complex Hilbert space V to itself, touching base with familiar notions when $V = \mathbb{C}^n$.

Compact linear operators were discussed in Section 1. Compactness means that the image of the closed unit ball has compact closure in V. We know from Proposition 2.1 that the compact linear operators are exactly those that can be approximated in the operator norm topology by linear operators with finite-dimensional image. The adjoint of a compact linear operator is compact. Being the members of the closure of a vector subspace, the compact linear operators form a vector subspace. When $V = \mathbb{C}^n$, every linear operator is of course compact.

If L is a compact linear operator, then LA and AL are compact whenever A is a bounded linear operator. In fact, if L_n is a sequence of linear operators with finite-dimensional image such that $\|L - L_n\| \to 0$, then $\|LA - L_n A\| \le \|L - L_n\| \|A\| \to 0$; since $L_n A$ has finite-dimensional image, LA is compact. To see that AL is compact, we take the adjoint: L^* is compact, and hence $L^* A^* = (AL)^*$ is compact; since $(AL)^*$ is compact, so is AL. In algebraic terminology the compact linear operators form a two-sided ideal in the algebra of all bounded linear operators.

Next we introduce Hilbert–Schmidt operators. If L is a bounded linear operator on V and if $\{u_i\}$ and $\{v_j\}$ are orthonormal bases of V, then Parseval's equality gives

$$\begin{aligned}\textstyle\sum_i \|Lu_i\|^2 &= \textstyle\sum_{i,j} |(Lu_i, v_j)|^2 = \sum_{i,j} |(u_i, L^* v_j)|^2 \\ &= \textstyle\sum_{i,j} |\overline{(L^* v_j, u_i)}|^2 = \sum_{i,j} |(L^* v_j, u_i)|^2 = \sum_j \|L^* v_j\|^2.\end{aligned}$$

Application of this formula twice shows that if we replace $\{u_i\}$ by a different orthonormal basis $\{u_i'\}$, we get $\sum_i \|Lu_i\|^2 = \sum_i \|Lu_i'\|^2$. The expression

$$\|L\|_{\mathrm{HS}}^2 = \sum_i \|Lu_i\|^2 = \sum_{i,j} |(Lu_i, v_j)|^2,$$

which we therefore know to be independent of both orthonormal bases $\{u_i\}$ and $\{v_j\}$, is the square of what is called the **Hilbert–Schmidt norm** $\|L\|_{\mathrm{HS}}$ of L.

For the finite-dimensional situation in which the underlying Hilbert space is $\mathbb{R}^n$ or $\mathbb{C}^n$, we can take $\{u_i\}$ and $\{v_j\}$ both to be the standard orthonormal basis, and then the Hilbert–Schmidt norm of the linear function corresponding to a matrix A is just $\big(\sum_{i,j} |A_{ij}|^2\big)^{1/2}$.

Our computation with $\|L\|_{\mathrm{HS}}$ above shows that

$$\|L\|_{\mathrm{HS}} = \|L^*\|_{\mathrm{HS}}.$$

The bounded linear operators that have finite Hilbert–Schmidt norm are called **Hilbert–Schmidt operators**. The name results from the following proposition.

Proposition 2.7. Let (X, μ) be a σ-finite measure space such that $L^2(X, \mu)$ is separable, and let $K(\,\cdot\,,\,\cdot\,)$ be a complex-valued L^2 function on $X \times X$. Then the linear operator T defined by

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

is a compact operator on the Hilbert space $L^2(X, \mu)$ with $\|T\|_{\mathrm{HS}} = \|K\|_2$.

REMARK. No self-adjointness is assumed in this proposition.

PROOF. If $\{u_i\}$ is an orthonormal basis of $L^2(X, \mu)$, then the functions $(u_j \otimes \bar{u}_i)(x, y) = u_j(x)\overline{u_i(y)}$ form an orthonormal basis of $L^2(X \times X, \mu \times \mu)$ as a consequence of Proposition 12.9 of *Basic*. Hence

$$(Tu_i, u_j) = \textstyle\int_X \int_X K(x, y) u_i(y) \overline{u_j(x)}\, d\mu(x)\, d\mu(y) = (K, (u_j \otimes \bar{u}_i)).$$

Taking the square of the absolute value of both sides and summing on i and j, we obtain $\|T\|_{\mathrm{HS}}^2 = \|K\|_2^2$. $\square$

Returning to an abstract Hilbert space V and the bounded linear operators on it, let us observe for any L that

$$\|L\| \le \|L\|_{\text{HS}}.$$

In fact, if u in V has $\|u\| = 1$, then the singleton set $\{u\}$ can be extended to an orthonormal basis $\{u_i\}$, and we obtain $\|Lu\|^2 \le \sum_i \|Lu_i\|^2 = \|L\|_{\text{HS}}^2$. Taking the supremum over u with $\|u\| = 1$, we see that $\|L\|^2 \le \|L\|_{\text{HS}}^2$. Two easier but related inequalities are that

$$\|AL\|_{\text{HS}} \le \|A\|\|L\|_{\text{HS}} \qquad \text{and} \qquad \|LA\|_{\text{HS}} \le \|A\|\|L\|_{\text{HS}}.$$

The first of these follows from the inequality $\|ALu_i\|^2 \le \|A\|^2\|Lu_i\|^2$ by summing over an orthonormal basis. The second follows from the first because $\|LA\|_{\text{HS}} = \|(LA)^*\|_{\text{HS}} = \|A^*L^*\|_{\text{HS}} \le \|A^*\|\|L^*\|_{\text{HS}} = \|A\|\|L\|_{\text{HS}}$.

Any Hilbert–Schmidt operator is compact. In fact, if L is Hilbert–Schmidt, let $\{u_i\}$ be an orthonormal basis, let $\epsilon > 0$ be given, and choose a finite set F of indices i such that $\sum_{i \notin F} \|Lu_i\|^2 < \epsilon$. If E is the orthogonal projection on the span of the u_i for i in F, then we obtain $\|L^* - EL^*\|^2 = \|L - LE\|^2 \le \|L - LE\|_{\text{HS}}^2 = \sum_i \|(L - LE)u_i\|^2 < \epsilon$. Hence L^* can be approximated in the operator norm topology by operators with finite-dimensional image and is compact; since L^* is compact, L is compact.

The sum of two Hilbert–Schmidt operators is Hilbert–Schmidt. In fact, we have $\|(L+M)u_i\| \le \|Lu_i\| + \|Mu_i\| \le 2\max\{\|Lu_i\|, \|Mu_i\|\}$. Squaring gives $\|(L+M)u_i\|^2 \le 4\max\{\|Lu_i\|^2, \|Mu_i\|^2\} \le 4(\|Lu_i\|^2 + \|Mu_i\|^2)$, and the result follows when we sum on i. Thus the Hilbert–Schmidt operators form a vector subspace of the bounded linear operators on V, in fact a vector subspace of the compact operators on V. As is true of the compact operators, the Hilbert–Schmidt operators form a two-sided ideal in the algebra of all bounded linear operators; this fact follows from the inequalities $\|AL\|_{\text{HS}} \le \|A\|\|L\|_{\text{HS}}$ and $\|LA\|_{\text{HS}} \le \|A\|\|L\|_{\text{HS}}$.

The vector space of Hilbert–Schmidt operators becomes a normed linear space under the Hilbert–Schmidt norm. Even more, it is an inner-product space. To see this, let L and M be Hilbert–Schmidt operators, and let $\{u_i\}$ be an orthonormal basis. We define $\langle L, M\rangle = \sum_i (Lu_i, Mu_i)$. This sum is absolutely convergent as we see from two applications of the Schwarz inequality: $\sum_i |(Lu_i, Mu_i)| \le \sum_i \|Lu_i\|\|Mu_i\| \le \big(\sum_i \|Lu_i\|^2\big)^{1/2}\big(\sum_i \|Mu_i\|^2\big)^{1/2} = \|L\|_{\text{HS}}\|M\|_{\text{HS}} < \infty$. Substituting from the definitions, we readily check that

$$\langle L, M\rangle = \begin{cases} \displaystyle\sum_{k\in\{0,2\}} \tfrac{i^k}{4}\|L + i^k M\|_{\text{HS}}^2 & \text{if } V \text{ is real,} \\ \displaystyle\sum_{k=0}^{3} \tfrac{i^k}{4}\|L + i^k M\|_{\text{HS}}^2 & \text{if } V \text{ is complex.} \end{cases}$$

Hence the definition of $\langle L, M\rangle$ is independent of the orthonormal basis. It is immediate from the definition and the above convergence that the form $\langle\,\cdot\,,\,\cdot\,\rangle$ makes the vector space of Hilbert–Schmidt operators into an inner-product space with associated norm $\|\,\cdot\,\|_{\text{HS}}$.

If L has finite-dimensional image, then L is a Hilbert–Schmidt operator. In fact, let E be the orthogonal projection on image L, take an orthonormal basis $\{u_i \mid i \in F\}$ of image L, and extend to an orthonormal basis $\{u_i \mid i \in S\}$ of V; here F is a finite subset of S. Then $\sum_{i\in S}\|Lu_i\|^2 = \sum_{i\in S}\|ELu_i\|^2 = \sum_{i\in S}\|L^*Eu_i\|^2 = \sum_{i\in F}\|L^*u_i\|^2 < \infty$. Thus the Hilbert–Schmidt operators form an ideal between the ideal of compact operators and the ideal of operators with finite-dimensional image.

Now we turn to a generalization of the trace $\operatorname{Tr} A = \sum_i A_{ii}$ of a square matrix A. This generalization plays a basic role in distribution theory, in index theory for partial differential equations, and in representation theory. In this section we shall describe the operators, and at the end of Chapter III we shall show how traces can be computed for simple integral operators. Realistic applications tend to be beyond the scope of this book.

Although the trace of a linear operator on $\mathbb{C}^n$ may be computed as the sum of the diagonal entries of the matrix of the operator in any basis, we shall continue to use orthonormal bases. Thus the expression we seek to extend to any Hilbert space V is $\sum_i (Lu_i, u_i)$. The operators of "trace class" are to be a subset of the Hilbert–Schmidt operators. It might at first appear that the condition to impose for the definition of trace class is that $\sum_i (Lu_i, u_i)$ be absolutely convergent for some orthonormal basis, but this condition is not enough. In fact, if a bounded linear operator L is defined on a Hilbert space with orthonormal basis $u_1, u_2, \ldots$ by $Lu_i = u_{i+1}$ for all i, then $(Lu_i, u_i) = 0$ for all i; on the other hand, $\|Lu_i\|^2 = 1$ for all i, and L is not Hilbert–Schmidt.

We say that a bounded linear operator L on V is of **trace class** if it is a compact operator[5] such that $\sum_i |(Lu_i, v_i)| < \infty$ for all orthonormal bases $\{u_i\}$ and $\{v_i\}$. Since compact operators are closed under addition and under passage to adjoints, we see directly from the definition that the sum of two trace-class operators is of trace class and that the adjoint of a trace-class operator is of trace class. The operator $L = B^*A$ with A and B Hilbert–Schmidt is an example of a trace-class operator. In fact, the operator L is compact as the product of two compact operators; also, $(Lu_i, v_i) = (B^*Au_i, v_i) = (Au_i, Bv_i)$, and we therefore have $\sum_i |(Lu_i, v_i)| = \sum_i |(Au_i, Bv_i)| \le \sum_i \|Au_i\|\|Bv_i\| \le$

[5]This condition is redundant; it is enough to assume boundedness. However, to proceed without using compactness of L, we would have to know that L^*L has a "positive semidefinite" square root, which requires having the full Spectral Theorem for bounded self-adjoint operators. This theorem is not available until the end of Chapter IV. The development here instead gets by with the Spectral Theorem for compact self-adjoint operators (Theorem 2.3).

$\big(\sum_i \|Au_i\|^2\big)^{1/2}\big(\sum_i \|Bv_i\|^2\big)^{1/2} = \|A\|_{\text{HS}}\|B\|_{\text{HS}}$. The following proposition shows that there are no other examples.

Proposition 2.8. If $L : V \to V$ is a trace-class operator on the Hilbert space V, then L factors as $L = B^*A$ with A and B Hilbert–Schmidt. Moreover, the supremum of $\sum_i |(Lu_i, v_i)|$ over all orthonormal bases $\{u_i\}$ and $\{v_i\}$ equals the infimum, over all Hilbert–Schmidt A and B such that $L = B^*A$, of the product $\|A\|_{\text{HS}}\|B\|_{\text{HS}}$.

PROOF. First we produce a factorization. Since L is a compact operator, L^*L is a compact self-adjoint operator, and Theorem 2.3 shows that L^*L has an orthonormal basis of eigenvectors w_i with real eigenvalues λ_i tending to 0. Since $\lambda_i(w_i, w_i) = (L^*Lw_i, w_i) = (Lw_i, Lw_i)$, we see that all λ_i are ≥ 0. Define a bounded linear operator T by $Tw_i = \sqrt{\lambda_i}\, w_i$ for all i. The operator T is self adjoint, it has $(Tv, v) \geq 0$ for all v, its kernel N is the smallest closed vector subspace containing all the w_i with $\lambda_i = 0$, and its image is dense in $N^\perp$. Since $N \cap N^\perp = 0$, T is one-one from $N^\perp$ into $N^\perp$. Thus $Tv \mapsto Lv$ is a well-defined linear function from a dense vector subspace of $N^\perp$ into V. The map $Tv \mapsto Lv$ has the property that $\|Lv\|^2 = (Lv, Lv) = (L^*Lv, v) = (T^2v, v) = (Tv, Tv) = \|Tv\|^2$. Thus $Tv \mapsto Lv$ is a linear isometry from a dense vector subspace of $N^\perp$ into V. Since V is complete, $Tv \mapsto Lv$ extends to a linear isometry $U : N^\perp \to V$. This U satisfies $L = UT$.

Let I be the set of indices i for the orthonormal basis $\{w_i\}$, and let P be the subset with $\lambda_i > 0$. By polarization, U preserves inner products in carrying $N^\perp$ into V. Extend U to all of V by setting it equal to 0 on N, so that U^* is well defined. The system $\{w_i\}_{i\in P}$ is an orthonormal basis of $N^\perp$, and hence the system $\{f_i\}_{i\in P}$ with $f_i = Uw_i$ for $i \in P$ is an orthonormal set in V. Since $U : N^\perp \to V$ is isometric, we have $(w_i, U^*f_i) = (Uw_i, f_i) = (Uw_i, Uw_i) = (w_i, w_i)$. Since Tw_i is a multiple of w_i, we obtain $(Tw_i, U^*f_i) = (Tw_i, w_i)$. Therefore

$$\begin{aligned}\sum_{i\in P} |(Lw_i, f_i)| &= \sum_{i\in P} |(UTw_i, f_i)| = \sum_{i\in P} |(Tw_i, U^*f_i)| \\ &= \sum_{i\in P} |(Tw_i, w_i)| = \sum_{i\in P} (Tw_i, w_i).\end{aligned}$$

Extend $\{f_i\}_{i\in P}$ to an orthonormal basis $\{f_i\}$ of V; since any two orthonormal bases of a Hilbert space have the same cardinality, we can index the new vectors of this set by $I - P$. The operators L and T have the same kernel, and thus the sums for $i \in P$ can be extended over all i in I to give

$$\sum_{i\in I} |(Lw_i, f_i)| = \sum_{i\in I} (Tw_i, w_i).$$

Define a bounded linear operator S on V by $Sw_i = \sqrt[4]{\lambda_i}\, w_i$ for all i. Then $|(Sw_i, w_j)|^2 = \delta_{ij}(S^2w_i, w_i) = \delta_{ij}(Tw_i, w_i)$, and hence S is a Hilbert–Schmidt operator with $\|S\|^2_{\text{HS}} = \sum_{i \in I} (Tw_i, w_i)$. Take $A = S$ and $B^* = US$; each of these is Hilbert–Schmidt since $\|US\|_{\text{HS}} \le \|U\|\|S\|_{\text{HS}}$, and we have $B^*A = USS = UT = L$. This proves the existence of a decomposition $B^*A = L$.

For the bases $\{w_i\}$ and $\{f_i\}$, we have just seen that

$$\|A\|_{\text{HS}}\|B\|_{\text{HS}} \le \|S\|_{\text{HS}}\|U\|\|S\|_{\text{HS}} \le \|S\|^2_{\text{HS}} = \sum_{i \in I} (Tw_i, w_i) = \sum_{i \in I} |(Lw_i, f_i)|.$$

But if $L = B'^*A'$ is any decomposition of L as the product of Hilbert–Schmidt operators and if $\{u_i\}$ and $\{v_i\}$ are any two orthonormal bases, we have

$$\begin{aligned}\sum_i |(Lu_i, v_i)| &= \sum_i |(B'^*A'u_i, v_i)| = \sum_i |(A'u_i, B'v_i)| \\ &\le \sum_i \|A'u_i\|\|B'v_i\| \le \|A'\|_{\text{HS}}\|B'\|_{\text{HS}}.\end{aligned}$$

Therefore
$$\sup \sum_i |(Lu_i, v_i)| \le \inf \|A'\|_{\text{HS}}\|B'\|_{\text{HS}},$$

as asserted. □

If $\{u_i\}$ is an orthonormal basis of V and L is of trace class, we can thus write $L = B^*A$ with A and B Hilbert–Schmidt. We define the **trace** of L to be

$$\operatorname{Tr} L = \textstyle\sum_i (Lu_i, u_i) = \sum_i (B^*Au_i, u_i) = \sum_i (Au_i, Bu_i) = \langle A, B \rangle.$$

The series $\sum_i (Lu_i, u_i)$ is absolutely convergent by definition of trace class. The trace of L is independent of the orthonormal basis since it equals $\langle A, B \rangle$, and it is independent of A and B since it equals $\sum_i (Lu_i, u_i)$.

In practice it is not so easy to check from the definition that L is of trace class. But there is a simple sufficient condition.

Proposition 2.9. If $L : V \to V$ is a bounded linear operator on the Hilbert space V and if $\sum_{i,j} |(Lu_i, v_j)| < \infty$ for some orthonormal bases $\{u_i\}$ and $\{v_j\}$, then L is of trace class.

PROOF. Since $|(Lu_i, v_i)| \le \|L\|$, we have $|(Lu_i, v_j)|^2 \le \|L\||(Lu_i, v_j)|$ for all i and j, and it follows from the finiteness of $\sum_{i,j} |(Lu_i, v_j)|$ that $\|L\|^2_{\text{HS}} = \sum_{i,j} |(Lu_i, v_j)|^2$ is finite. Thus L is a Hilbert–Schmidt operator and has to be compact.

If $\{e_k\}$ and $\{f_l\}$ are orthonormal bases, we expand $e_k = \sum_i (e_k, u_i)u_i$ and $f_k = \sum_j (f_k, v_j)v_j$ and substitute to obtain $(Le_k, f_k) = \sum_{i,j} (e_k, u_i)(Lu_i, v_j)\overline{(f_k, v_j)}$. Taking the absolute value and summing on k gives

$$\sum_k |(Le_k, f_k)| \le \sum_{i,j} |(Lu_i, v_j)| \sum_k |(e_k, u_i)\overline{(f_k, v_j)}|.$$

Application of the Schwarz inequality to the sum on k and then Bessel's inequality to each factor of the result yields

$$\begin{aligned}\sum_k |(Le_k, f_k)| &\le \sum_{i,j} |(Lu_i, v_j)| \Big(\sum_k |(e_k, u_i)|^2\Big)^{1/2} \Big(\sum_k |(f_k, v_j)|^2\Big)^{1/2} \\ &\le \sum_{i,j} |(Lu_i, v_j)| \|u_i\| \|v_j\| = \sum_{i,j} |(Lu_i, v_j)| < \infty,\end{aligned}$$

and therefore L is of trace class. □

6. Problems

1. Let (S, μ) be a σ-finite measure space, let f be in $L^\infty(S, \mu)$, and let M_f be the bounded linear operator on $L^2(S, \mu)$ given by $M_f(g) = fg$.
   (a) Find a necessary and sufficient condition for M_f to have an eigenvector.
   (b) Find a necessary and sufficient condition for M_f to be compact.
2. Let L be a compact operator on a Hilbert space, and let λ be a nonzero complex number. Prove that if $\lambda I - L$ is one-one, then the image of $\lambda I - L$ is closed.
3. Prove for a Hilbert space V that the normed linear space of Hilbert–Schmidt operators with the norm $\|\cdot\|_{\text{HS}}$ is a Banach space.
4. If L is a trace-class operator on a Hilbert space V, let $\|L\|_{\text{TC}}$ equal the supremum of $\sum_i |(Lu_i, v_i)|$ over all orthonormal bases $\{u_i\}$ and $\{v_i\}$. By Proposition 2.8 this equals the infimum, over all Hilbert–Schmidt A and B such that $L = B^*A$, of the product $\|A\|_{\text{HS}} \|B\|_{\text{HS}}$. Prove that the vector space of trace-class operators is a normed linear space under $\|\cdot\|_{\text{TC}}$ as norm.
5. If L is a trace-class operator on a complex Hilbert space V and A is a bounded linear operator, prove that $\operatorname{Tr} AL = \operatorname{Tr} LA$ and conclude that $\operatorname{Tr}(BLB^{-1}) = \operatorname{Tr} L$ for any bounded linear operator B.

Problems 6–8 deal with some extensions of Theorem 2.3 to situations involving several operators. A bounded linear operator L is said to be **normal** if $LL^* = L^*L$.

6. Suppose that $\{L_\alpha\}$ is a finite commuting family of compact self-adjoint operators on a Hilbert space. Prove that there exists an orthonormal basis consisting of simultaneous eigenvectors for all L_α.
7. Fix a *complex* Hilbert space V.
   (a) Prove that the decomposition $L = \frac{1}{2}(L + L^*) + i\,\frac{1}{2i}(L - L^*)$ exhibits any normal operator $L : V \to V$ as a linear combination of commuting self-adjoint operators.
   (b) Prove that the operators in (a) are compact if L is compact.
   (c) State an extension of Theorem 2.3 that concerns compact normal operators on a complex Hilbert space.

8. Fix a Hilbert space V.
   (a) Prove that a unitary operator from V to itself is always normal.
   (b) Under what circumstances is a unitary operator compact?

Problems 9–13 indicate an approach to second-order ordinary differential equations by integral equations in a way that predates the use of the Hilbert–Schmidt Theorem.

9. For $\omega \neq 0$, show that the unique solution $u(t)$ on $[a, b]$ of the equation $u'' + \omega^2 u = g(t)$ and the initial conditions $u(a) = 1$ and $u'(a) = 0$ is

$$u(t) = \cos \omega(t - a) + \omega^{-1} \int_a^t g(s) \sin \omega(t - s)\, ds.$$

10. Let $\rho(t)$ be a continuous function on $[a, b]$, and let $u(t)$ be the unique solution of the equation $u'' + [\omega^2 - \rho(t)]u = 0$ and the initial conditions $u(a) = 1$ and $u'(a) = 0$. Show that u satisfies the integral equation

$$u(t) - \omega^{-1} \int_a^t \rho(s) \sin \omega(t - s) u(s)\, ds = \cos \omega(t - a),$$

which is of the form $u(t) - \int_a^t K(t, s)u(s)\, ds = f(t)$, where $K(t, s)$ is continuous on the triangle $a \leq s \leq t \leq b$.

11. Let $K(t, s)$ be continuous on the triangle $a \leq s \leq t \leq b$. For f continuous on $[a, b]$, define $(Tf)(t) = \int_a^t K(t, s) f(s)\, ds$.
   (a) Prove that f continuous implies Tf continuous.
   (b) Put $M = \max |K(t, s)|$. If f has $C = \int_a^b |f(t)|\, dt$, prove inductively that $|(T^n f)(t)| \leq \frac{CM^n}{(n-1)!} (t - a)^{n-1}$ for $n \geq 1$.
   (c) Deduce that the series $f + Tf + T^2 f + \cdots$ converges uniformly on $[a, b]$.

12. Set $u = f + Tf + T^2 f + \cdots$ in the previous problem, and prove that u satisfies $u - Tu = f$.

13. In the previous problem prove that $u = f + Tf + T^2 f + \cdots$ is the only solution of $u - Tu = f$.