Contemporary Mathematics
Volume **720**, 2018
https://doi.org/10.1090/conm/720/14520

Fundamentals of spectral theory

Thomas Ransford

ABSTRACT. The mini-course developed the elements of spectral theory needed for the other lectures at the conference. No prior knowledge of functional analysis was assumed. The notes below are a copy of what was presented during the lectures, together with some exercises.

Introduction

Two fundamental theorems from spectral theory of matrices:

- Every $n \times n$ matrix has eigenvalues (at least 1 and at most n).
- Every normal $n \times n$ matrix has an orthonormal basis of eigenvectors.

Our aims:

- Establish analogues of these theorems in infinite dimensions.
- Apply these new results to the theory of differential equations.

Plan of the mini-course:

(1) Normed spaces and operators
(2) Invertible operators
(3) The spectrum
(4) Hilbert spaces
(5) Operators on Hilbert spaces
(6) Compact operators
(7) Spectral theorem
(8) Sturm–Liouville equation

1. Normed spaces and operators

1.1. Normed spaces.

DEFINITION. A *normed space* is a vector space X equipped with a *norm*, namely a function $\|\cdot\| : X \to \mathbb{R}$ satisfying:

- $\|x\| \geq 0$ for all $x \in X$, with equality iff $x = 0$;
- $\|\lambda x\| = |\lambda|\|x\|$ for all $\lambda \in \mathbb{C}$ and all $x \in X$;
- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

2010 *Mathematics Subject Classification.* Primary 47A10; Secondary 34B05, 47B05, 47B15.

©2018 American Mathematical Society

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

Note that X is a metric space, with metric $d(x,y) := \|x-y\|$.

X is *complete* if every Cauchy sequence in X converges to a limit in X. A complete normed space is called a *Banach space*.

1.2. Examples of normed spaces.

EXAMPLE 1. $X = \mathbb{C}^n$ with the Euclidean norm:

$$\|(\xi_1,\dots,\xi_n)\|_2 := \Bigl(\sum_{k=1}^n |\xi_k|^2\Bigr)^{1/2}.$$

EXAMPLE 2. $X = C[a,b]$ with the sup norm:

$$\|f\|_\infty := \sup_{t\in[a,b]} |f(t)|.$$

EXAMPLE 3. $X = \{\text{polynomials}\}$ with the coefficient norm

$$\Bigl\|\sum_{k=0}^n a_k t^k\Bigr\| := \sum_{k=0}^n |a_k|.$$

Examples 1 and 2 are Banach spaces, Example 3 is not.

1.3. Absolute convergence implies convergence.

THEOREM. *Let (x_k) be a sequence in a Banach space X. If $\sum_k \|x_k\| < \infty$, then $\sum_k x_k$ converges in X.*

PROOF. Set $s_n := \sum_{k=1}^n x_k$. For $n > m$, we have

$$\|s_n - s_m\| = \Bigl\|\sum_{k=m+1}^n x_k\Bigr\| \le \sum_{k=m+1}^n \|x_k\|.$$

As $\sum_k \|x_k\| < \infty$, the last term is small if m,n are large enough. Thus (s_n) is a Cauchy sequence, so it converges in X. $\square$

1.4. Operators.

THEOREM. *Let X, Y be normed spaces and let $T : X \to Y$ be a linear map. The following are equivalent:*

(1) *T is continuous on X;*
(2) *T is bounded on the unit ball of X;*
(3) *there is a constant C such that $\|Tx\| \le C\|x\|$ for all $x \in X$.*

DEFINITION. A map T satisfying (1)–(3) above is called a *(bounded) operator*. The smallest constant C such that (3) holds is called the *operator norm* of T, denoted $\|T\|$.

1.5. Examples of operators.

EXAMPLE 1. Let $A = (a_{jk})$ be an $n \times n$ matrix, and define $T : \mathbb{C}^n \to \mathbb{C}^n$ by $Tx := Ax$. Then, writing $x = (\xi_1,\dots,\xi_n)$,

$$\|Tx\|_2^2 = \sum_{j=1}^n \Bigl|\sum_{k=1}^n a_{jk}\xi_k\Bigr|^2 \le \sum_{k=1}^n\sum_{j=1}^n |a_{jk}|^2 \|x\|_2^2.$$

So T is a bounded operator with $\|T\| \le (\sum_{j,k} |a_{jk}|^2)^{1/2}$.

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

EXAMPLE 2. Define $T : C[0,1] \to C[0,1]$ by $Tf(t) := tf(t)$.

$$\|Tf\|_\infty = \sup_{t\in[0,1]} |tf(t)| \leq \sup_{t\in[0,1]} |f(t)| = \|f\|_\infty.$$

So T is bounded with $\|T\| \leq 1$. In fact $\|T\| = 1$ (take $f \equiv 1$).

EXAMPLE 3. Define $T : C[0,1] \to C[0,1]$ by $Tf(t) := \int_0^t f$.

$$\|Tf\|_\infty = \sup_{t\in[0,1]} \left|\int_0^t f\right| \leq \int_0^1 |f| \leq \|f\|_\infty.$$

So T is bounded with $\|T\| \leq 1$. In fact $\|T\| = 1$ (take $f \equiv 1$).

EXAMPLE 4. Let $X = \{\text{polynomials}\}$ with the coefficient norm, and define $T : X \to X$ by $Tp := p'$. If $p_n(t) = t^n$, then

$$\|p_n\| = 1 \quad \text{and} \quad \|Tp_n\| = n.$$

Conclusion: this T is unbounded.

1.6. The space of bounded operators. We write $B(X,Y)$ for the set of bounded operators $T : X \to Y$.

THEOREM. *$B(X,Y)$ is a normed space with respect to the operator norm. Further, if Y is complete, then so is $B(X,Y)$.*

Important special cases:

- $Y = \mathbb{C}$. We write $X^* := B(X,\mathbb{C})$, the *dual space* of X. Elements of the dual are called (continuous) *linear functionals*.
- $Y = X$. We write $B(X) := B(X,X)$, which is now an algebra. If $S,T \in B(X)$, then also $ST \in B(X)$ and

$$\|ST\| \leq \|S\|\|T\|.$$

Exercises

1A Justify the completeness or incompleteness each of the spaces X in §1.2.

1B Let X be a Banach space and let (x_k) be vectors in X such that $\sum_{k\geq1} \|x_k\| < \infty$. Prove that

$$\Big\|\sum_{k\geq1} x_k\Big\| \leq \sum_{k\geq1} \|x_k\|.$$

1C Prove the theorem in §1.4.

1D Prove the theorem in §1.6.

1E Prove that, if $S,T \in B(X)$, then $\|ST\| \leq \|S\|\|T\|$.

2. Invertible operators

2.1. Invertible operators. Let X be a normed space. As usual, $B(X)$ denotes the space of bounded operators on X.

We say that $T \in B(X)$ is *invertible* if there exists $S \in B(X)$ such that $ST = TS = I$. This S is unique, and is denoted by T^{-1}.

The set of all invertible operators on X is a group, denoted $GL(X)$.

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

2.2. The fundamental lemma.

LEMMA. *Assume that X is a Banach space. If $T \in B(X)$ and $\|T\| < 1$, then $(I-T)$ is invertible, and $(I-T)^{-1} = \sum_{k\geq 0} T^k$.*

PROOF. Since $\sum_{k\geq 0}\|T^k\| \leq \sum_{k\geq 0}\|T\|^k < \infty$, the series $\sum_{k\geq 0} T^k$ is convergent in the Banach space $B(X)$. Further

$$(I-T)\sum_{k\geq 0} T^k = \lim_{n\to\infty}(I-T)\sum_{k=0}^{n} T^k = \lim_{n\to\infty}(I-T^{n+1}) = I.$$

Likewise $\sum_{k\geq 0} T^k(I-T) = I$. □

2.3. The group of invertible operators is open. Recall: $GL(X)$ denotes the group of invertible operators on X.

THEOREM. *Assume that X is a Banach space. Then $GL(X)$ is open in $B(X)$.*

PROOF. By the fundamental lemma, I lies in the interior of $GL(X)$. Given $S \in GL(X)$, the map $T \mapsto ST$ is a homeomorphism of $B(X)$ onto itself taking $GL(X)$ onto itself and mapping I to S. Hence S lies in the interior of $GL(X)$ as well. Thus $GL(X)$ is open in $B(X)$. □

2.4. Continuity of inversion.

THEOREM. *Assume that X is a Banach space. Then the map $T \mapsto T^{-1}$ is continuous on $GL(X)$.*

PROOF. By the fundamental lemma, if $S_n \to I$, then $S_n^{-1} \to I$. Indeed, writing $S_n = I - T_n$, we have $\|T_n\| \to 0$, so for n large

$$\|(I-T_n)^{-1} - I\| = \Big\|\sum_{k\geq 1} T_n^k\Big\| \leq \sum_{k\geq 1}\|T_n\|^k \to 0.$$

Hence, for a general $S \in GL(X)$, we have

$$\begin{aligned} S_n \to S &\Rightarrow S_nS^{-1} \to I \\ &\Rightarrow (S_nS^{-1})^{-1} \to I \\ &\Rightarrow SS_n^{-1} \to I \\ &\Rightarrow S_n^{-1} \to S^{-1}. \end{aligned}$$

□

2.5. A generalization of the fundamental lemma.

LEMMA. *Assume that X is a Banach space. If $T \in B(X)$ and $\|T^k\| < 1$ for some $k \geq 1$, then $(I-T)$ is invertible.*

PROOF. By the fundamental lemma, $(I-T^k)$ is invertible. Also,

$$I - T^k = (I-T)\Big(\sum_{j=0}^{k-1} T^j\Big) = \Big(\sum_{j=0}^{k-1} T^j\Big)(I-T),$$

whence it follows that $(I-T)$ is invertible. □

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

2.6. Application to an initial-value problem.

THEOREM. *Let $b > 0$, let $n \geq 1$, let $a_0, \ldots, a_{n-1} \in C[0,b]$, let $g \in C[0,b]$ and let $x_0, \ldots, x_{n-1} \in \mathbb{C}$. Then there exists a unique solution f to*

$$(*) \begin{cases} f^{(n)}(t) + a_{n-1}(t) f^{(n-1)}(t) + \cdots + a_0(t) f(t) = g(t), \\ f^{(j)}(0) = x_j \quad (0 \leq j < n). \end{cases}$$

Idea for the proof. Convert into a system of first-order ODEs. For example, if $n = 2$, then $(*)$ is equivalent to

$$(**) \begin{cases} F'(t) + A(t)F(t) = G(t), \\ F(0) = X_0, \end{cases}$$

where

$$A := \begin{pmatrix} 0 & -1 \\ a_0 & a_1 \end{pmatrix}, \quad F = \begin{pmatrix} f_0 \\ f_1 \end{pmatrix}, \quad G = \begin{pmatrix} 0 \\ g \end{pmatrix}, \quad X_0 = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}.$$

So it is enough to prove

THEOREM. *Let $A : [0,b] \to B(\mathbb{C}^n)$ and $G : [0,b] \to \mathbb{C}^n$ be continuous maps, and let $X_0 \in \mathbb{C}^n$. Then there exists a unique solution F to*

$$(**) \begin{cases} F'(t) + A(t)F(t) = G(t), \\ F(0) = X_0. \end{cases}$$

PROOF. F solves $(**)$ iff $(I+V)F = H$, where

$$(VF)(t) := \int_0^t A(s)F(s)\,ds \quad \text{and} \quad H(t) := \int_0^t G(s)\,ds + X_0.$$

Here $V : Y \to Y$, where $Y := C([0,b], \mathbb{C}^n)$, a Banach space.
So it is enough to show that $V \in B(Y)$ and $(I+V)$ is invertible.

(1) Proof that V is a bounded operator on Y: Let $F \in C([0,b], \mathbb{C}^n)$. For each $t \in [0,b]$,

$$\begin{aligned} \|VF(t)\|_{\mathbb{C}^n} &= \left\| \int_0^t A(s)F(s)\,ds \right\|_{\mathbb{C}^n} \\ &\leq \int_0^t \|A(s)F(s)\|_{\mathbb{C}^n}\,ds \\ &\leq \int_0^t \|A(s)\|_{B(\mathbb{C}^n)} \|F(s)\|_{\mathbb{C}^n}\,ds \\ &\leq M \int_0^t \|F(s)\|_{\mathbb{C}^n}\,ds, \end{aligned}$$

where $M := \sup_{s \in [0,b]} \|A(s)\|_{B(\mathbb{C}^n)} < \infty$. Hence

$$\|VF\|_Y = \sup_{t \in [0,b]} \|VF(t)\|_{\mathbb{C}^n} \leq Mb \sup_{s \in [0,b]} \|F(s)\| = Mb\|F\|_Y.$$

Conclusion: V is a bounded operator on Y and $\|V\| \leq Mb$.

(2) Proof that $(I+V)$ is invertible: Repeat the previous computation with F replaced by $V^{k-1}F$ to get

$$\|V^k F(t)\|_{\mathbb{C}^n} \leq M \int_0^t \|V^{k-1}F(s)\|_{\mathbb{C}^n}\,ds.$$

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

Hence, by induction on k,

$$\|V^k F(t)\|_{\mathbb{C}^n} \leq \frac{M^k t^k}{k!} \|F\|_Y.$$

Hence

$$\|V^k F\|_Y \leq \frac{M^k b^k}{k!} \|F\|_Y.$$

Hence

$$\|V^k\| \leq \frac{M^k b^k}{k!}.$$

If k is large enough, then $M^k b^k / k! < 1$, so $\|V^k\| < 1$. By the (generalized) fundamental lemma, $(I + V)$ is invertible. Done! $\square$

Exercises

2A Let (T_n) be a sequence of invertible operators in $B(X)$ converging to a non-invertible operator T. Show that $\|T_n^{-1}\| \to \infty$. [*Hint*: $\|I - T_n^{-1}T\| \leq \|T_n^{-1}\| \|T_n - T\|$.]

2B Let $n \geq 2$. Determine A, F, G, X_0 so that the systems $(*)$ and $(**)$ in §2.6 are equivalent.

2C Show that the solution F to $(**)$ in §2.6 satisfies

$$\|F\|_\infty \leq B\Big(\|X_0\| + \int_0^b \|G(s)\|\, ds\Big),$$

where B is a constant depending on A, but not on X_0 or G.

2D Combine the two preceding exercises to show that the solution f to $(*)$ in §2.6 satisfies

$$\max_{0 \leq j \leq n-1} \|f^{(j)}\|_\infty \leq C\Big(\max_{0 \leq j \leq n-1} |x_j| + \int_0^b |g(s)|\, ds\Big),$$

where C is a constant depending on $a_0, \ldots, a_{n-1}$, but not on $x_0, \ldots, x_{n-1}$ or g.

3. The spectrum

Throughout this section, X is a Banach space and $T \in B(X)$.

3.1. Eigenvalues and eigenvectors. $\lambda \in \mathbb{C}$ is an *eigenvalue* of T if there exists a non-zero $x \in X$ such that $Tx = \lambda x$. The vector x is a called an *eigenvector*.

Warning: If $\dim X = \infty$ then T may have no eigenvalues!

Example. $X := C[0,1]$ and $Tf(t) := tf(t)$ (see Exercise 3A).

Two ways round this problem:

- impose extra conditions of T (shall explore this later);
- 'broaden' the definition of eigenvalue (subject for this section).

3.2. Spectrum.

Definition. The *spectrum* of T is defined by

$$\sigma(T) := \{\lambda \in \mathbb{C} : T - \lambda I \text{ not invertible}\}.$$

If $\dim X < \infty$, then $\sigma(T) = \{\text{eigenvalues of } T\}$.
If $\dim X = \infty$, then $\{\text{eigenvalues of } T\} \subset \sigma(T)$, but inclusion may be strict.

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

Example. Let $X := C[0,1]$ and $Tf(t) := tf(t)$.

- If $(T - \lambda I)$ is invertible, then there exists $f \in C[0,1]$ such that $(t-\lambda)f(t) = 1$ for all $t \in [0,1]$, so $\lambda \notin [0,1]$.
- Conversely, if $\lambda \notin [0,1]$, then $t \mapsto 1/(t-\lambda) \in C[0,1]$, and multiplication by this function is an inverse to $(T - \lambda I)$.
- Conclusion: $\sigma(T) = [0,1]$.

3.3. Spectral radius formula.

Theorem. *The spectrum $\sigma(T)$ is a non-empty compact set, and*

$$\max\{|\lambda| : \lambda \in \sigma(T)\} = \lim_{n\to\infty} \|T^n\|^{1/n} = \inf_{n\geq 1} \|T^n\|^{1/n}.$$

Proof. In three steps:

(1) $\rho(T) := \lim_{n\to\infty} \|T^n\|^{1/n}$ exists and equals $\inf_{n\geq 1} \|T^n\|^{1/n}$;
(2) $\sigma(T)$ is a closed subset of the disk $|z| \leq \rho(T)$;
(3) there exists $\lambda \in \sigma(T)$ with $|\lambda| = \rho(T)$.

Note: $\rho(T)$ is called the *spectral radius* of T.

Step 1: $\lim_{n\to\infty} \|T^n\|^{1/n} = \inf_{n\geq 1} \|T^n\|^{1/n}$.

Set $\eta := \inf_{n\geq 1} \|T^n\|^{1/n}$. It is enough to show that

$$\limsup_{n\to\infty} \|T^n\|^{1/n} \leq \eta. \qquad (*)$$

Let $\epsilon > 0$ and choose $m \geq 1$ such that $\|T^m\|^{1/m} < \eta + \epsilon$.

Given $n \geq 1$, write $n = qm + r$, where $0 \leq r < m$. Then

$$\|T^n\| \leq \|T^m\|^q \|T^r\| \leq (\eta+\epsilon)^{mq} \|T^r\| = (\eta+\epsilon)^{n-r} \|T^r\| \leq (\eta+\epsilon)^n M,$$

where M is a constant independent of n. Hence

$$\limsup_{n\to\infty} \|T^n\|^{1/n} \leq \eta + \epsilon.$$

Letting $\epsilon \to 0$, we get $(*)$.

Step 2: $\sigma(T)$ is a closed subset of the disk $|z| \leq \rho(T)$.

Define $F : \mathbb{C} \to B(X)$ by $F(\lambda) := T - \lambda I$. Clearly F is continuous. Also, recalling that $GL(X)$ denotes the invertible operators on X,

$$F^{-1}(GL(X)) = \mathbb{C} \setminus \sigma(T).$$

As $GL(X)$ is open in $B(X)$, we have $\mathbb{C} \setminus \sigma(T)$ is open in $\mathbb{C}$, whence $\sigma(T)$ is closed in $\mathbb{C}$.

Now suppose that $|\lambda| > \rho(T) = \inf_{n\geq 1} \|T^n\|^{1/n}$. Then there exists $n \geq 1$ such that $\|(T/\lambda)^n\| < 1$. By the generalized form of the fundamental lemma, it follows that $(I - T/\lambda)$ is invertible, and therefore $\lambda \notin \sigma(T)$. Conclusion:

$$\lambda \in \sigma(T) \implies |\lambda| \leq \rho(T).$$

Step 3: There exists $\lambda \in \sigma(T)$ with $|\lambda| = \rho(T)$.

Two cases to consider:

Case 1: $\rho(T) = 0$.

We need to show that T is not invertible. Suppose, on the contrary, that T is invertible. Then, for all $n \geq 1$, we have

$$1 = \|I\| = \|T^n (T^{-1})^n\| \leq \|T^n\| \|T^{-1}\|^n.$$

Then $\|T^n\|^{1/n} \geq 1/\|T^{-1}\| > 0$ for all n, contradicting $\rho(T) = 0$.

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

Case 2: $\rho(T) > 0$.

Without loss of generality $\rho(T) = 1$. Then, by Step 2, $\sigma(T) \subset \overline{\mathbb{D}}$. We need to show that $\sigma(T) \cap \partial\mathbb{D} \neq \emptyset$. Suppose, on the contrary, that $\sigma(T) \subset \mathbb{D}$. Then $(I - zT)$ is invertible for all $z \in \overline{\mathbb{D}}$.

As $z \mapsto (I - zT)^{-1} : \overline{\mathbb{D}} \to B(X)$ is continuous on a compact set, it is uniformly continuous. So, given $\epsilon > 0$, there is $\delta > 0$ such that

$$z, w \in \overline{\mathbb{D}},\ |z - w| \leq \delta \quad \Rightarrow \quad \|(I - zT)^{-1} - (I - wT)^{-1}\| \leq \epsilon.$$

Using the identity $(I - S^2)^{-1} = ((I - S)^{-1} + (I + S)^{-1})/2$ and induction on n, it follows that, for all $n \geq 0$,

$$|z - w| \leq \delta \quad \Rightarrow \quad \|(I - (zT)^{2^n})^{-1} - (I - (wT)^{2^n})^{-1}\| \leq \epsilon.$$

In particular, taking $z = 1$ and $w = (1 - \delta)$, we get

$$\|(I - T^{2^n})^{-1} - (I - ((1 - \delta)T)^{2^n})^{-1}\| \leq \epsilon.$$

To summarize, $\forall \epsilon > 0\ \exists \delta > 0\ \forall n \geq 0$

$$\|(I - T^{2^n})^{-1} - (I - ((1 - \delta)T)^{2^n})^{-1}\| \leq \epsilon. \tag{$*$}$$

Now $\rho((1 - \delta)T) = (1 - \delta) < 1$, so $((1 - \delta)T)^{2^n} \to 0$ and hence

$$(I - ((1 - \delta)T)^{2^n})^{-1} \to I.$$

Together with $(*)$, this shows that, for all sufficiently large n,

$$\|(I - T^{2^n})^{-1} - I\| \leq 2\epsilon.$$

Thus $(I - T^{2^n})^{-1} \to I$, whence $(I - T^{2^n}) \to I$ and $T^{2^n} \to 0$.
This contradicts the supposition that $\rho(T) = 1$. Done! □

3.4. Application: Fundamental theorem of algebra.

THEOREM. *Every polynomial* $p(z) = z^n + a_1 z^{n-1} + \cdots + a_n$ *has a root.*

PROOF. We have $p(z) = \det(zI - A)$, where

$$A := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 \end{pmatrix}.$$

Take $\lambda \in \sigma(A)$. Then $p(\lambda) = \det(\lambda I - A) = 0$. □

3.5. Some further developments.

- λ is an *approximate eigenvalue* of T if there exists a sequence (x_n) of unit vectors in X such that $\|(T - \lambda I)x_n\| \to 0$. One can show every $\lambda \in \partial\sigma(T)$ is an approximate eigenvalue. Hence every $T \in B(X)$ has approximate eigenvalues.
- *Spectral mapping theorem*: If p is a polynomial then $\sigma(p(T)) = p(\sigma(T))$.
- *Holomorphic functional calculus.* One can define an operator $f(T)$ for every function f holomorphic on a neighborhood of $\sigma(T)$. Idea: use Cauchy integral formula

$$f(T) = \frac{1}{2\pi i} \int_\Gamma f(z)(zI - T)^{-1}\, dz.$$

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

Exercises

3A Define $T : C[0,1] \to C[0,1]$ by $Tf(t) := tf(t)$. Prove that T has no eigenvalues.

3B Let $S \in B(X)$ and suppose that $I \pm S$ are both invertible. Show that $I - S^2$ is invertible and

$$(I - S^2)^{-1} = \frac{(I-S)^{-1} + (I+S)^{-1}}{2}.$$

3C (Spectral mapping theorem) Let $T \in B(X)$ and let p be a polynomial. Show that

$$\sigma(p(T)) = p(\sigma(T)).$$

[*Hint:* factorize $p(z) - \lambda = c(z - z_1) \dots (z - z_n)$.]

3D Let $T \in B(X)$ and let $\lambda \in \partial\sigma(T)$. Show that λ is an approximate eigenvalue, i.e. there exist unit vectors $x_n \in X$ such that $\|(T - \lambda I)x_n\| \to 0$. [*Hint:* Use Exercise 2A.]

4. Hilbert spaces

4.1. Inner products.

Definition. A map $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{C}$ is an *inner product* if:

- $x \mapsto \langle x, y \rangle$ is linear, for each $y \in X$;
- $\langle y, x \rangle = \overline{\langle x, y \rangle}$, for all $x, y \in X$;
- $\langle x, x \rangle \geq 0$ for all $x \in X$, with equality iff $x = 0$.

Properties of inner products:

- $\|x\| := \langle x, x \rangle^{1/2}$ is a norm on X.
- $|\langle x, y \rangle| \leq \|x\|\|y\|$ (Cauchy–Schwarz inequality).
- If $\|x_n - x\| \to 0$ and $\|y_n - y\| \to 0$, then $\langle x_n, y_n \rangle \to \langle x, y \rangle$.
- $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ (parallelogram identity).
- $\langle x, y \rangle = (1/4) \sum_{k=0}^{3} i^k \|x + i^k y\|^2$ (polarization identity).

4.2. Hilbert spaces. A *Hilbert space* is a vector space H, equipped with an inner product, such that the resulting normed space is complete.

Examples:

- $H = \mathbb{C}^n$
  $\langle x, y \rangle := \sum_{k=1}^{n} x_k \overline{y}_k$
- $H = \ell^2 := \{x = (x_1, x_2, \dots) : \sum_1^\infty |x_k|^2 < \infty\}$
  $\langle x, y \rangle := \sum_{k=1}^{\infty} x_k \overline{y}_k$.
- $H = L^2[a, b] := \{f : [a, b] \to \mathbb{C} : \int_a^b |f(t)|^2 \, dt < \infty\}$
  $\langle f, g \rangle := \int_a^b f(t)\overline{g(t)} \, dt$
- $H = L^2(S, \mu) := \{f : S \to \mathbb{C} : \int_S |f|^2 \, d\mu < \infty\}$
  $\langle f, g \rangle := \int_S f\overline{g} \, d\mu$.

Henceforth, H designates a Hilbert space.

4.3. Orthonormal sequences.

Definition. A (finite or infinite) sequence (e_k) in H is *orthonormal (o.n.)* if

$$\langle e_j, e_k \rangle = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$$

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

Examples:

- $H = \mathbb{C}^n$; $\quad e_1 = (1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$
- $H = \ell^2$; $\quad e_1 = (1, 0, 0, \ldots), e_2 = (0, 1, 0, \ldots), \ldots$
- $H = L^2[0, 1]$; $\quad e_k(t) = e^{2\pi ikt}$ $(k \in \mathbb{Z})$

4.4. Bessel's inequality.

THEOREM (Bessel's inequality). *Let (e_k) be an orthonormal sequence in H. Then, for all $x \in H$,*

$$\sum_k |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$

PROOF. For $n \geq 1$, we have

$$0 \leq \Big\| x - \sum_{k=1}^n \langle x, e_k \rangle e_k \Big\|^2 = \|x\|^2 - \sum_{k=1}^n |\langle x, e_k \rangle|^2.$$

Let $n \to \infty$. $\square$

4.5. Riesz–Fischer theorem.

THEOREM (Riesz–Fischer). *Let (e_k) be an o.n. sequence in H and let (c_k) be complex scalars. Then $\sum_{k=1}^\infty c_k e_k$ converges in H iff $\sum_{k=1}^\infty |c_k|^2 < \infty$.*

PROOF. Set $s_n := \sum_{k=1}^n c_k e_k$, and note that, for $n > m$,

$$\|s_n - s_m\|^2 = \Big\| \sum_{k=m+1}^n c_k e_k \Big\|^2 = \sum_{k=m+1}^n |c_k|^2.$$

Now use completeness of H. $\square$

REMARKS. (1) $\sum_k c_k e_k$ is independent of the order of the terms.
(2) By Bessel and Riesz–Fischer, the series $\sum_k \langle x, e_k \rangle e_k$ converges.

4.6. Parseval's identity.

THEOREM. *Let (e_k) be an orthonormal sequence in H and let $x \in H$. TFAE:*

- *x lies in the closure of the span of the (e_k);*
- *$x = \sum_k \langle x, e_k \rangle e_k$;*
- *$\|x\|^2 = \sum_k |\langle x, e_k \rangle|^2$ (Parseval's identity).*

PROOF. The equivalence of the first two statements follows from

$$\Big\| x - \sum_{k=1}^n c_k e_k \Big\|^2 = \Big\| x - \sum_{k=1}^n \langle x, e_k \rangle e_k \Big\|^2 + \Big\| \sum_{k=1}^n (\langle x, e_k \rangle - c_k) e_k \Big\|^2.$$

The equivalence of the last two statements follows from

$$\Big\| x - \sum_{k=1}^n \langle x, e_k \rangle e_k \Big\|^2 = \|x\|^2 - \sum_{k=1}^n |\langle x, e_k \rangle|^2. \qquad \square$$

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

4.7. Orthonormal bases.

DEFINITION. An *orthonormal basis* of H is an o.n. sequence (e_k) that spans a dense subspace of H. By the preceding theorem, for all $x \in H$,

$$x = \sum_k \langle x, e_k \rangle e_k \quad \text{and} \quad \|x\|^2 = \sum_k |\langle x, e_k \rangle|^2.$$

A metric space is *separable* if it has a countable dense subset. Separable H include: $\mathbb{C}^n$, ℓ^2, $L^2[a,b]$.

THEOREM. *A Hilbert space has an orthonormal basis iff it is separable.*

PROOF. 'Only if' : consider $\{\sum_{k=1}^n c_k e_k : c_k \in \mathbb{Q} + i\mathbb{Q}, n \geq 1\}$.
'If': Choose (v_k) with dense span, then use Gram–Schmidt. □

4.8. Orthogonal complements.

Given $S \subset H$, we define

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

Note that $S^\perp$ is a closed subspace of H (whether S is or not).

THEOREM. *Let M be a closed subspace of H. Then $H = M \oplus M^\perp$.*

PROOF. (For H separable.) Clearly $M \cap M^\perp = \{0\}$. We must show that $M + M^\perp = H$. Let $x \in H$. Let (e_k) be an o.n. basis of M and set

$$y := \sum_k \langle x, e_k \rangle e_k.$$

Then $y \in M$. Also $\langle x - y, e_k \rangle = 0$ for all k, so $x - y \in M^\perp$. Thus

$$x = y + (x - y) \in M + M^\perp.$$ □

4.9. Linear functionals on a Hilbert space.

THEOREM (Riesz representation theorem). *Let $\phi : H \to \mathbb{C}$ be a continuous linear functional. Then there exists a unique $y \in H$ such that*

$$\phi(x) = \langle x, y \rangle \quad (x \in H).$$

PROOF. *Existence*: If $\phi = 0$, then we can take $y = 0$. Otherwise, $\ker \phi$ is a proper closed subspace of H. Since $H = \ker \phi \oplus (\ker \phi)^\perp$, there exists non-zero vector $y \in (\ker \phi)^\perp$. Then ϕ and $x \mapsto \langle x, y \rangle$ are linear functionals with the same kernel, so one is a multiple of the other. We can ensure that this multiple is 1 by replacing y by a suitable multiple of itself.

Uniqueness: If y_1, y_2 are two candidates, then $\langle x, y_1 - y_2 \rangle = 0$ for all x, in particular for $x = y_1 - y_2$. This implies $y_1 - y_2 = 0$. □

Exercises

4A Prove the properties of inner products listed in §4.1.

4B Prove that ℓ^2 is complete.

4C Let (e_k) be an o.n. basis of H and let $(c_k) \in \ell^2$. Show that $\sum_k c_k e_k$ is independent of the order of the sum. [*Hint:* Show that $\sum_k \langle c_k e_k, y \rangle$ is absolutely convergent for each $y \in H$.]

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

4D (Gram–Schmidt algorithm) Let $(v_1, v_2, v_3, \dots)$ be a (finite or infinite) sequence of linearly independent vectors in H. Define inductively

$$f_k := v_k - \sum_{j=1}^{k-1} \langle v_k, e_j \rangle e_j \quad \text{and} \quad e_k := f_k / \|f_k\|.$$

Show that (e_k) is an orthonormal sequence and that $\operatorname{span}\{e_1, \dots, e_n\} = \operatorname{span}\{v_1, \dots, v_n\}$ for each n.

4E Show that every subset of a separable metric space is itself separable.

5. Operators on Hilbert spaces

5.1. Adjoint of an operator.

THEOREM. *Each $T \in B(H)$ has a unique* adjoint $T^* \in B(H)$ *such that*

$$\langle Tx, y \rangle = \langle x, T^* y \rangle \quad (x, y \in H).$$

PROOF. Fix $y \in H$. The map $x \mapsto \langle Tx, y \rangle$ is a continuous linear functional on H, so by the Riesz representation theorem, there exists a unique $z \in H$ such that

$$\langle Tx, y \rangle = \langle x, z \rangle \quad (x \in H).$$

Define $T^* y := z$. It is routine to check that T^* is linear. Lastly,

$$\|T^* x\|^2 = \langle T^* x, T^* x \rangle = \langle T T^* x, x \rangle \le \|T\| \|T^* x\| \|x\|,$$

whence $\|T^* x\| \le \|T\| \|x\|$. Thus $T^* \in B(H)$ and $\|T^*\| \le \|T\|$. □

5.2. Calculation of adjoints.

THEOREM. *Let (e_k) be an o.n. basis and suppose $Te_k = \sum_j a_{jk} e_j$ for all k. Then $T^* e_k = \sum_j \overline{a_{kj}} e_j$ for all k.*

PROOF. We have $T^* e_k = \sum_j \langle T^* e_k, e_j \rangle e_j$, and

$$\langle T^* e_k, e_j \rangle = \overline{\langle e_j, T^* e_k \rangle} = \overline{\langle T e_j, e_k \rangle} = \overline{\langle \sum_i a_{ij} e_i, e_k \rangle} = \overline{a_{kj}}. \qquad \square$$

Examples:

- If $T \in B(\ell^2)$ is the right shift, then T^* is the left shift.
- (Diagonal operators) If (e_k) is an o.n. basis for H and $Te_k = \lambda_k e_k$ for all k, then $T^* e_k = \overline{\lambda}_k e_k$ for all k.

5.3. Properties of adjoints.

Algebraic properties: $(S+T)^* = S^* + T^*$, $(\lambda T)^* = \overline{\lambda} T^*$, $(ST)^* = T^* S^*$, $T^{**} = T$.

Norm properties: $\|T^*\| = \|T\|$ and $\|T^* T\| = \|T\|^2$.

PROOF OF THE NORM PROPERTIES. We already know that $\|T^*\| \le \|T\|$. Also $\|T\| = \|T^{**}\| \le \|T^*\|$. Further $\|T^* T\| \le \|T^*\| \|T\| = \|T\|^2$. Finally,

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, T^* T x \rangle \le \|T^* T\| \|x\|^2,$$

whence $\|T\|^2 \le \|T^* T\|$. □

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

5.4. Self-adjoint operators.

DEFINITION. $T \in B(H)$ is *self-adjoint* if $T^* = T$.

Examples:

- For $T \in B(H)$, let $A := (T + T^*)/2$ and $B := (T - T^*)/2i$. Then A, B are self-adjoint and $T = A + iB$.
- If $T \in B(H)$, then T^*T is self-adjoint.
- Let $P : H \to M$ be the orthogonal projection onto a closed subspace M. Then P is self-adjoint.

5.5. Spectrum of self-adjoint operators.

THEOREM. *If T is a self-adjoint operator, then $\sigma(T) \subset \mathbb{R}$.*

PROOF. Let $\alpha + i\beta \in \sigma(T)$. Then $\alpha + i(\beta + t) \in \sigma(T + itI)$, so

$$|\alpha + i(\beta + t)|^2 \le \|T + itI\|^2 = \|(T + itI)^*(T + itI)\| = \|T^2 + t^2 I\|.$$

Hence

$$\alpha^2 + \beta^2 + 2\beta t \le \|T^2\|.$$

As this holds for all $t \in \mathbb{R}$, we must have $\beta = 0$. □

5.6. Normal operators.

DEFINITION. $T \in B(H)$ is *normal* if $TT^* = T^*T$.

Examples:

- Every self-adjoint operator is normal.
- Every diagonal operator is normal.
- Let A, B be self adjoint. Then $A + iB$ is normal iff $AB = BA$.
- The right shift S on ℓ^2 is *not* normal: $S^*S = I \ne SS^*$.

5.7. Properties of normal operators. Let $T \in B(H)$ be normal.

- $\|T^*x\| = \|Tx\|$ for all $x \in H$.
- If $Te = \lambda e$, then $T^*e = \overline{\lambda} e$.
- If $Te = \lambda e$ and $Tf = \mu f$ where $\lambda \ne \mu$, then $\langle e, f\rangle = 0$.
- $\|T^2\| = \|T\|^2$. Indeed:
$$\|T^2\|^2 = \|(T^2)^*(T^2)\| = \|(T^*T)^*(T^*T)\| = \|T^*T\|^2 = \|T\|^4.$$
- $\rho(T) = \|T\|$. Indeed: $\|T^{2^n}\| = \|T\|^{2^n}$ for all n, so
$$\rho(T) = \lim_{n\to\infty} \|T^{2^n}\|^{1/2^n} = \|T\|.$$

5.8. Spectral theorem in finite dimensions.

THEOREM. *Assume* $\dim H < \infty$. *Then every normal $T \in B(H)$ has an o.n. basis of eigenvectors.*

PROOF.

- Pick an o.n. basis of each eigenspace of T.
- Put them together to get a (finite) o.n. sequence (e_n).
- Let M be the (closed) subspace spanned by (e_n).
- Since T is normal, $T^*(M) \subset M$, whence $T(M^\perp) \subset M^\perp$.
- If $M^\perp \ne \{0\}$, then $T|M^\perp$ has an eigenvector e; then $e \in M \cap M^\perp$, so $e = 0$, contradiction.
- Conclusion: $M = H$ and (e_n) is an o.n. basis of H. □

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

5.9. Spectral theorem in infinite dimensions? If $\dim H = \infty$, then it is no longer true that every normal T has an o.n. basis of eigenvectors. Indeed, it may happen that T has no eigenvectors at all, even if T is self-adjoint.

EXAMPLE. $H = L^2[0,1]$ and $T : f(t) \mapsto tf(t)$.

We need a further ingredient...

Exercises

5A Prove that the adjoint T^* is a linear map.
5B Prove the algebraic properties of adjoints listed in §5.3.
5C Let M be a closed subspace of H and let $P : H \to M$ be the orthogonal projection of H onto M. Show that P is self-adjoint.
5D Let $T \in B(H)$. Show that T is a normal operator if and only if $\|Tx\| = \|T^*x\|$ for all $x \in H$. [*Hint for 'if':* Use the polarization identity (see §4.1).]
5E Let $T \in B(H)$. Show that $\|T\| = \rho(T^*T)^{1/2}$. Hence calculate $\left\|\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right\|$.

6. Compact operators

6.1. Compact operators.

DEFINITION. $T \in B(H)$ is *compact* if, whenever (x_n) is a bounded sequence, (Tx_n) contains a convergent subsequence.

Examples:

- If $T \in B(H)$ and has finite rank (i.e. $\dim(T(H)) < \infty$), then it is compact.
- If $\dim H = \infty$, then I_H is not compact. [*Proof:* Let (e_n) be an o.n. sequence. Then (e_n) is bounded. But it contains no convergent subsequence since $\|e_n - e_m\| = \sqrt{2}$ for all $m \neq n$.]

The set of compact operators on H is denoted by $K(H)$.

6.2. Structure of $K(H)$.

THEOREM. *$K(H)$ is an ideal in $B(H)$.*

PROOF. Show:

- $S, T \in K(H) \Rightarrow (S + T) \in K(H)$.
- $S \in B(H), T \in K(H) \Rightarrow ST, TS \in K(H)$. □

COROLLARY. *If $\dim H = \infty$, then no invertible operator on H is compact.*

THEOREM. *$K(H)$ is closed in $B(H)$.*

PROOF. Let $(T_k) \in K(H)$ and let $T_k \to T$. Let (x_n) be bounded.

- There exists $N_1 \subset \mathbb{N}$ such that $(T_1 x_n)_{n \in N_1}$ converges.
- There exists $N_2 \subset N_1$ such that $(T_2 x_n)_{n \in N_2}$ converges. Etc.
- Take N to be the 1st element of N_1, the 2nd element of N_2, etc. Then $(T_k x_n)_{n \in N}$ converges for each k.
- As $\|T - T_k\| \to 0$, get $(Tx_n)_{n \in N}$ is Cauchy, so convergent.

Conclusion: T is compact. □

COROLLARY. *If T is a limit of bounded, finite-rank operators, then it is compact.*

THEOREM. *If $T \in K(H)$ then $T^* \in K(H)$.*

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

PROOF. Let (x_n) be a bounded sequence. As TT^* is compact, there exists $N \subset \mathbb{N}$ such that $(TT^*x_n)_{n\in N}$ converges. Since

$$\|T^*x_n - T^*x_m\|^2 \leq \|x_n - x_m\|\|TT^*(x_n - x_m)\|,$$

it follows that $(T^*x_n)_{n\in N}$ is Cauchy, and hence convergent. $\square$

REMARKS. (1) To summarize: $K(H)$ is a closed $*$-ideal in $B(H)$.
(2) It can be shown that if H is infinite-dimensional and separable, then $K(H)$ is the unique non-trivial closed ideal in $B(H)$.

6.3. Compact diagonal operators.

THEOREM. *Let $T \in B(H)$, let (e_n) be an o.n. basis of H, and suppose that $Te_n = \lambda_n e_n$ for all n. Then T is compact iff $\lambda_n \to 0$.*

PROOF. Suppose that $\lambda_n \not\to 0$. Then there is a sequence $N \subset \mathbb{N}$ and $\delta > 0$ such that $|\lambda_n| \geq \delta$ for all $n \in N$. For $m, n \in N$ we then have

$$\|Te_n - Te_m\|^2 = |\lambda_n|^2 + |\lambda_m|^2 \geq 2\delta^2,$$

so $(Te_n)_{n\in N}$ has no convergent subsequence. So T is not compact.

Suppose conversely that $\lambda_n \to 0$. Let T_n be the diagonal operator with eigenvalues
$(\lambda_1, \ldots, \lambda_n, 0, 0, \ldots)$. Then T_n is finite rank and

$$\|T - T_n\| = \sup_{k>n} |\lambda_k| \to 0 \quad (n \to \infty).$$

So T is compact. $\square$

6.4. Hilbert–Schmidt operators.

LEMMA. *Let $T \in B(H)$, and let (e_n) and (f_m) be o.n. bases of H. Then*

$$\sum_n \|Te_n\|^2 = \sum_m \|T^*f_m\|^2.$$

PROOF. By Parseval,

$$\sum_n \|Te_n\|^2 = \sum_n \sum_m |\langle Te_n, f_m\rangle|^2 = \sum_m \sum_n |\langle e_n, T^*f_m\rangle|^2 = \sum_m \|T^*f_m\|^2. \quad \square$$

DEFINITION. $T \in B(H)$ is a *Hilbert–Schmidt* operator if $\sum_n \|Te_n\|^2 < \infty$ for some (and hence for every) o.n. basis (e_n).

THEOREM. *Every Hilbert–Schmidt operator T is compact.*

PROOF. Let (e_n) be an o.n. basis. For $x \in H$, we have

$$Tx = T\Big(\sum_{k=1}^{\infty} \langle x, e_k\rangle e_k\Big) = \sum_{k=1}^{\infty} \langle x, e_k\rangle Te_k.$$

Set $T_nx := \sum_{k=1}^n \langle x, e_k\rangle Te_k$. By Cauchy–Schwarz and Parseval,

$$\|(T - T_n)x\| \leq \sum_{k=n+1}^{\infty} |\langle x, e_k\rangle| \|Te_k\| \leq \|x\| \Big(\sum_{k=n+1}^{\infty} \|Te_k\|^2\Big)^{1/2},$$

so $\|T - T_n\| \leq (\sum_{k=n+1}^{\infty} \|Te_k\|^2)^{1/2} \to 0$ as $n \to \infty$. As each T_n is of finite rank, it follows that T is compact. $\square$

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

6.5. Integral operators.

THEOREM. *Let $K \in L^2([a,b] \times [a,b])$. For $f \in L^2[a,b]$, define*

$$T_K f(x) := \int_a^b K(x,y) f(y)\,dy.$$

(1) *$T_K \in B(L^2[a,b])$ with $\|T_K\| \leq \|K\|_2$*
(2) *T_K is a Hilbert–Schmidt operator, hence compact.*

PROOF. (1) Write $K_x(y) := K(x,y)$.

- Since $K \in L^2([a,b]^2)$, we have $K_x \in L^2[a,b]$ for a.e. $x \in [a,b]$.
- For all such x, we have $T_K f(x) = \langle f, \overline{K_x} \rangle$.
- By Cauchy–Schwarz, $|T_K f(x)| \leq \|K_x\|_2 \|f\|_2$ a.e.
- Hence $T_K f \in L^2([a,b])$ with $\|T_K f\|_2 \leq \|K\|_2 \|f\|_2$.

(2) Let (e_n) be an o.n. basis of $L^2[a,b]$. Then

$$\begin{aligned}\sum_n \|T_K e_n\|^2 &= \sum_n \int_a^b |(T_K e_n)(x)|^2\,dx = \sum_n \int_a^b |\langle e_n, \overline{K_x} \rangle|^2\,dx \\ &= \int_a^b \sum_n |\langle e_n, \overline{K_x} \rangle|^2\,dx = \int_a^b \|\overline{K_x}\|_2^2\,dx = \|K\|_2^2 < \infty. \qquad \square\end{aligned}$$

REMARK. All this works equally well in any countably generated, σ-finite measure space.

6.6. A compact operator with no eigenvalues. Define $T \in B(\ell^2)$ by $T := SD$, where S is the right shift and D is the diagonal operator with $De_n := e_n/n$ for all n.

- As S is bounded and D is compact, T is also compact.
- A computation gives $\|T^n\| = 1/n!$. By the spectral radius formula,

  $$\rho(T) = \lim_{n\to\infty} \|T^n\|^{1/n} = \lim_{n\to\infty} (1/n!)^{1/n} = 0.$$

  So the only possible eigenvalue of T is 0. But T is injective, so 0 is not an eigenvalue either.

Exercises

6A Prove that, if $S, T \in K(H)$, then $(S+T) \in K(H)$.
6B Prove that, if $S \in B(H)$ and $T \in K(H)$, then both $ST, TS \in K(H)$.
6C In the notation of §6.5, show that $(T_K)^* = T_L$, where $L(x,y) := \overline{K(y,x)}$.
6D In the example in §6.6, prove that $\|T^n\| = 1/n!$.

7. The spectral theorem

7.1. Spectral theorem.

THEOREM. *Let H be a separable Hilbert space and let $T \in B(H)$ be a compact normal operator. Then H has an o.n. basis of eigenvectors of T.*

The proof hinges on the following key lemma:

LEMMA. *A compact normal operator $T \in B(H)$ has an eigenvector.*

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

7.2. Proof of key lemma. Suppose first T is compact and *self-adjoint.* WLOG $\|T\| = 1$.

- Choose $x_n \in H$ with $\|x_n\| = 1$ such that $\|Tx_n\| \to 1$.
- $\|(T^2 - I)x_n\|^2 = \|T^2 x_n\|^2 + \|x_n\|^2 - 2\|Tx_n\|^2 \to 0$.
- By compactness of T, a subsequence $T^2 x_{n_j} \to$ some $y \in H$.
- As $(T^2 - I)x_{n_j} \to 0$, it follows that $x_{n_j} \to y$.
- Hence $\|y\| = 1$ and $(T^2 - I)y = 0$.
- If $(T - I)y = 0$, then y is an eigenvector of T.
- If $(T - I)y = z \neq 0$, then $(T + I)z = (T^2 - I)y = 0$, and so z is an eigenvector of T.

Conclusion: T has an eigenvector in this case.

Now suppose T is compact and *normal.*

- Set $A := (T + T^*)/2$ and $B := (T - T^*)/2i$.
- As A is compact and self-adjoint, it has an eigenvalue α.
- Set $K := \ker(A - \alpha I)$.
- As T is normal, $AB = BA$. It follows that $B(K) \subset K$.
- As $B|_K$ is compact and self-adjoint, it has an eigenvector e.
- So e is an eigenvector of both A and B.
- As $T = A + iB$, it follows that e is an eigenvector of T.

Conclusion: T has an eigenvector in this case too. □

7.3. Proof of spectral theorem. Let T be a compact normal operator on a separable H.

- Pick an o.n. basis of each eigenspace of T.
- Put them together to get a (countable) o.n. sequence (e_n).
- Let M be the closed subspace spanned by (e_n).
- Since T is normal, $T^*(M) \subset M$, whence $T(M^\perp) \subset M^\perp$.
- If $M^\perp \neq \{0\}$, then by the lemma $T|_{M^\perp}$ has an eigenvector e; then $e \in M \cap M^\perp$, so $e = 0$, contradiction.
- Conclusion: $M = H$ and (e_n) is an o.n. basis of H. □

7.4. Eigenvalue decay.

THEOREM. *Let $T \in B(H)$ with an o.n. basis (e_n) of eigenvectors, $Te_n = \lambda_n e_n$.*

(1) *If T is compact, then $\lambda_n \to 0$.*
(2) *If T is Hilbert–Schmidt, then $\sum_n |\lambda_n|^2 < \infty$.*

PROOF. (1) This is just the Theorem in §6.3.
(2) If T is Hilbert–Schmidt, then

$$\sum_n |\lambda_n|^2 = \sum_n \|Te_n\|^2 < \infty. \qquad \square$$

COROLLARY. *Every compact operator on H is the limit of a sequence of finite-rank operators.*

PROOF. If T is compact and self-adjoint, then it is a diagonal operator whose eigenvalues tend to zero, so it is the limit of finite-rank operators (see §6.3).

If T is compact but not self-adjoint, we write it as $T = A + iB$, where $A := (T + T^*)/2$ and $B := (T - T^*)/2i$ are both compact and self-adjoint. There exist finite-rank operators $A_n \to A$ and $B_n \to B$. Set $T_n := A_n + iB_n$. Then T_n is finite-rank and $T_n \to T$. □

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

7.5. Min-max principle. Let $\mathcal{M}_n$ denote the set of (closed) subspaces of H of dimension n.

THEOREM. *Let $T \in B(H)$ be a compact self-adjoint operator with positive eigenvalues, ordered so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$. Then, for each n,*

$$\lambda_n = \max_{M \in \mathcal{M}_n} \min_{\substack{x \in M \\ \|x\|=1}} \langle Tx, x\rangle = \min_{M \in \mathcal{M}_{n-1}} \max_{\substack{x \in M^\perp \\ \|x\|=1}} \langle Tx, x\rangle.$$

PROOF. We need two facts:

- For all x, we have $\langle Tx, x\rangle = \sum_k \lambda_k |\langle x, e_k\rangle|^2$.
- If $\dim M > \dim N$, then $M \cap N^\perp \neq \{0\}$ (as $P_N|_M : M \to N$ has a non-trivial kernel).

Let us write $M_n := \operatorname{span}\{e_1, \dots, e_n\}$.

- If $\dim M = n$, then $\exists x_0 \in M \cap M_{n-1}^\perp$ with $\|x_0\| = 1$, so

$$\min_{\substack{x \in M \\ \|x\|=1}} \langle Tx, x\rangle \leq \langle Tx_0, x_0\rangle = \sum_{k=n}^{\infty} \lambda_k |\langle x_0, e_k\rangle|^2 \leq \lambda_n.$$

- If $\dim M = n-1$, then then $\exists x_0 \in M_n \cap M^\perp$ with $\|x_0\| = 1$, so

$$\max_{\substack{x \in M^\perp \\ \|x\|=1}} \langle Tx, x\rangle \geq \langle Tx_0, x_0\rangle = \sum_{k=1}^{n} \lambda_k |\langle x_0, e_k\rangle|^2 \geq \lambda_n.$$

- Equality is attained in the first case if $M = M_n$ and in the second if $M = M_{n-1}$. □

7.6. Further developments. Let $T \in K(H)$. The *singular values* σ_k of T are the square roots of the eigenvalues of T^*T, listed in decreasing order.

- *Singular-value decomposition:* Given $T \in K(H)$, there are o.n. sequences (e_k) and (f_k) such that $Tx = \sum_k \sigma_k \langle x, e_k\rangle f_k$.
- *Distance formula:* If $T \in K(H)$ then, for each $n \geq 1$,

  $$\sigma_n(T) = \min\{\|T - R\| : R \in B(H),\ \operatorname{rank} R < n\}.$$

  Hence T is the limit of a sequence of finite-rank operators.
- *Weyl's inequality*: Let $T \in K(H)$ and let $\lambda_1, \dots, \lambda_n$ be n eigenvalues of T. Then $|\lambda_1 \lambda_2 \dots \lambda_n| \leq \sigma_1 \sigma_2 \dots \sigma_n$.
- *Schatten classes:* $S_p(H) := \{T : \sum_k |\sigma_k(T)|^p < \infty|\}$.
  $S_2(H) =$ Hilbert–Schmidt operators.
  $S_1(H) =$ trace-class operators.
- *Spectral theorem for general normal operators:* Every bounded normal operator on H is unitarily equivalent to multiplication by a bounded function on some L^2-space.

Exercises

7A (Singular-value decomposition) Let $T \in B(H)$ be a compact operator.
- Show that T^*T is a compact self-adjoint operator.
- Let (e_k) be an eigenbasis of T^*T. Show that the vectors (Te_k) are mutually orthogonal.
- Show that the eigenvalues of T^*T are positive, σ_k^2 say. Order the non-zero ones so that $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots$.

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

- Set $f_k := Te_k/\sigma_k$ if $\sigma_k \neq 0$. Show that the resulting (e_k) and (f_k) are o.n. sequences and

$$Tx = \sum_k \sigma_k \langle x, e_k\rangle f_k \quad (x \in H).$$

7B Let $T \in B(H)$ be a compact operator, with singular-value decomposition $Tx = \sum_k \sigma_k \langle x, e_k\rangle f_k$.

- Show that, if $R \in B(H)$ and $\operatorname{rank}(R) < n$, then $\ker(R) \cap \operatorname{span}\{e_1, \ldots, e_n\} \neq \{0\}$, and deduce that $\|T - R\| \geq \sigma_n$.
- Show that, if $Rx := \sum_{k=1}^{n-1} \sigma_k \langle x, e_k\rangle f_k$, then $\operatorname{rank}(R) < n$ and $\|T - R\| \leq \sigma_n$.
- Deduce the distance formula:

$$\sigma_n(T) = \min\{\|T - R\| : R \in B(H),\ \operatorname{rank}(R) < n\}.$$

7C Let H be a separable Hilbert space and E be a subset of H such that $\langle e, f\rangle = 0$ for all $e, f \in E$. Prove that E is at most countable. [This was used in proving the spectral theorem in §7.3.] [Hint: Show that the open balls $\{x \in H : \|x - e\| < \|e\|/2\}$ $(e \in E, e \neq 0)$ are disjoint.]

8. Sturm–Liouville equation

8.1. Formulation of the problem. Sturm–Liouville problem (SL): Find $y \in C^2[a, b]$ and $\lambda \in \mathbb{C}$ satisfying:

- **(SL1) differential equation:**

$$-(py')' + qy = \lambda y,$$

  where $p \in C^1[a, b]$ and $p > 0$, and $q \in C[a, b]$ is real-valued, and
- **(SL2) boundary conditions:**

$$\begin{cases} \alpha_0 y(a) + \alpha_1 y'(a) = 0 \\ \beta_0 y(b) + \beta_1 y'(b) = 0, \end{cases}$$

  where $(\alpha_0, \alpha_1), (\beta_0, \beta_1) \in \mathbb{R}^2 \setminus \{(0, 0)\}$.

8.2. Examples. The existence of non-trivial solutions y depends on the parameter λ.

Example. Consider the Sturm–Liouville problem:

$$\begin{cases} -y'' = \lambda y \text{ on } [0, 1] \\ y(0) = 0,\ y(1) = 0. \end{cases}$$

This has a non-trivial solution y iff $\lambda = k^2\pi^2$ $(k = 1, 2, 3, \ldots)$.
In this case $y(t) = c \sin(k\pi t)$.

Remark. This particular SL-problem arises in the wave equation for a stretched string. Other applications include Schrödinger's equation and Legendre's equation.

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

8.3. Notation. In what follows:

- $Y := \{y \in C^2[a,b] : y \text{ satisfies (SL2)}\}$.
- $L : C^2[a,b] \to C[a,b]$ is defined by

$$Ly := -(py')' + qy.$$

Thus y is a solution to (SL) iff $y \in Y$ and $Ly = \lambda y$.
If $y \not\equiv 0$, we call it an *eigenfunction* of (SL) and λ its *eigenvalue.*

REMARK. This is an abuse of terminology, since $L(Y) \not\subset Y$.

8.4. The main theorem.

THEOREM. *With the above notation:*

- *The eigenvalues of (SL) form a real sequence* (λ_k) *such that*

$$\lambda_k \to \infty \quad \text{and} \quad \sum_k \lambda_k^{-2} < \infty.$$

- *The corresponding normalized eigenfunctions* (y_k) *form an o.n. basis of* $L^2[a,b]$.

Basic idea of the proof:

- Construct an integral operator T that is an inverse to the differential operator L.
- Apply the spectral theorem to T, and deduce result for L.

And now for the details ...

8.5. Proof of main theorem.

LEMMA 1 (Lagrange's identity). *If* $y_1, y_2 \in Y$, *then* $\langle Ly_1, y_2\rangle = \langle y_1, Ly_2\rangle$.

PROOF. Integrating by parts, we have

$$(*) \qquad \langle Ly_1, y_2\rangle = \int_a^b py_1'\overline{y_2'}\,dt + \int_a^b qy_1\overline{y_2}\,dt - \Big[py_1'\overline{y_2}\Big]_a^b.$$

Thus $\langle Ly_1, y_2\rangle - \langle y_1, Ly_2\rangle = -\Big[p.(y_1'\overline{y_2} - \overline{y_2'}y_1)\Big]_a^b = 0$, as $y_j \in Y$. □

Consequences:

- Every eigenvalue λ of (SL) is real.
- Eigenfunctions having distinct eigenvalues are orthogonal.

LEMMA 2. *There exists a constant* C *such that each eigenvalue* $\lambda > C$.

PROOF. Suppose not.

- Then there exist at least five eigenvalues $\lambda_1, \dots, \lambda_5 < \min q$.
- Let $y_1, \dots, y_5$ be the associated normalized eigenfunctions.
- By linear algebra, there exist $\gamma_1, \dots, \gamma_5 \in \mathbb{C}$, not all zero, such that $y := \sum_1^5 \gamma_j y_j$ satisfies $y(a) = y'(a) = y(b) = y'(b) = 0$.
- By $(*)$, $\langle Ly, y\rangle = \int_a^b p|y'|^2 + \int_a^b q|y|^2 \ge (\min q)\sum_1^5 |\gamma_j|^2$.
- Also $\langle Ly, y\rangle = \sum_1^5 \lambda_j|\gamma_j|^2 < (\min q)\sum_1^5 |\gamma_j|^2$. Contradiction. □

Consequence: Replacing q by $q + C$ if necessary, WLOG we may suppose that all eigenvalues $\lambda > 0$. In particular, $L|_Y$ is injective.

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

LEMMA 3. *There exist real-valued* $u, v \in C^2[a,b]$ *such that*

- $Lu = Lv = 0$;
- $\alpha_0 u(a) + \alpha_1 u'(a) = 0$;
- $\beta_0 v(b) + \beta_1 v'(b) = 0$;
- $p.(u'v - v'u) \equiv 1$.

PROOF. We use the existence/uniqueness theorem for the initial-value problem (section 2).

- $\exists u \in C^2, u \not\equiv 0$ such that $Lu = 0$ and $\alpha_0 u(a) + \alpha_1 u'(a) = 0$,
  $\exists v \in C^2, v \not\equiv 0$ such that $Lv = 0$ and $\beta_0 v(b) + \beta_1 v'(b) = 0$.
- $p.(u'v - v'u)$ is a constant because
$$[p.(u'v - v'u)]' = (Lu)v + pu'v' - (Lv)u - pu'v' = 0.$$
- If the constant is 0, then $v \in Y$, contradicting $L|_Y$ injective.
- Replace u by a multiple of itself to make the constant 1. □

LEMMA 4. *With* u, v *as above, define* $G : [a,b] \times [a,b] \to \mathbb{R}$ *by*
$$G(s,t) := \begin{cases} u(s)v(t), & s \le t, \\ u(t)v(s), & t \le s, \end{cases}$$
and $T : L^2[a,b] \to L^2[a,b]$ *by*
$$Tf(s) := \int_a^b G(s,t) f(t)\, dt.$$
If $f \in C[a,b]$, *then* $Tf \in Y$ *and* $L(Tf) = f$. *Consequently*
$$T : C[a,b] \to Y$$
$$L : Y \to C[a,b]$$
are mutually inverse mappings.

PROOF. Fix $f \in C[a,b]$.

- We have
$$Tf(s) = v(s) \int_a^s uf + u(s) \int_s^b vf,$$
$$(Tf)'(s) = v'(s) \int_a^s uf + u'(s) \int_s^b vf.$$
  It follows that $Tf \in C^2[a,b]$.
- Further
$$\alpha_0 (Tf)(a) + \alpha_1 (Tf)'(a) = (\alpha_0 u(a) + \alpha_1 u'(a)) \int_a^b vf = 0.$$
  Likewise $\beta_0 (Tf)(b) + \beta_1 (Tf)'(b) = 0$. So $Tf \in Y$.
- Lastly, writing $U(s) := \int_a^s uf$ and $V := \int_s^b vf$, we have
$$\begin{aligned} L(Tf) &= -(p(v'U + u'V))' + q(vU + uV) \\ &= (Lv)U + (Lu)V + p.(u'v - v'u)f = f. \end{aligned}$$
□

CONCLUSION OF PROOF OF MAIN THEOREM.

- As $G \in L^2([a,b] \times [a,b])$, the operator T is Hilbert–Schmidt. Also, as G is real-valued and symmetric, T is self-adjoint.

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms

- By the spectral theorem, $L^2[a,b]$ has an orthonormal basis of eigenvectors (y_k) of T and the eigenvalues satisfy $\sum_k |\mu_k|^2 < \infty$.
- The (y_k) automatically belong to Y. Indeed, since T maps $L^2[a,b]$ into $C[a,b]$ and $y_k = (1/\mu_k)Ty_k$, we have $y_k \in C[a,b]$. Likewise, since T maps $C[a,b]$ into Y, similar reasoning gives $y_k \in Y$.
- As T, L are mutually inverse, they have the same eigenvectors, and the eigenvalues of L are the reciprocals of those of T.
- Hence the (normalized) eigenfunctions of L form an o.n. basis of $L^2[a,b]$. Also the eigenvalues of L satisfy $\sum_k |\lambda_k|^{-2} < \infty$.
- Lastly, as already observed, the eigenvalues of L are real and bounded below. Consequently $\lambda_k \to \infty$ as $k \to \infty$. □

8.6. Further developments.

- The eigenvalues of L are simple (i.e. $\dim\ker(L - \lambda_k I) = 1$).
- The eigenfunctions y_k can be chosen real-valued.
- *Oscillation theorem:* y_k has exactly $k-1$ zeros in (a,b).
- SL with weights: $-(py')' + qy = \lambda r y$, where $r(t) > 0$.
- Singular Sturm–Liouville problems.
- Boundary-value problems in higher dimensions ...

Exercises

8A Let y_1, y_2 be two eigenfunctions of (SL) corresponding to the same eigenvalue λ. Show that there exist $\gamma_1, \gamma_2 \in \mathbb{C}$, not both zero, such that $z := \gamma_1 y_1 + \gamma_2 y_2$ satisfies $z(a) = z'(a) = 0$, and deduce that $z \equiv 0$. [*Hint for the last part:* Use the uniqueness of the solution to $(*)$ in §2.6.]

8B Let y be an eigenfunction of (SL). Show that $\overline{y}$ is also an eigenfunction with the same eigenvalue. Deduce that there exists a real-valued eigenfunction with the same eigenvalue.

Département de mathématiques et de statistique, Université Laval, Québec (QC), G1V0A6, Canada

Email address: `ransford@mat.ulaval.ca`

Licensed to Cornell Univ. Prepared on Sat Jan 2 09:40:19 EST 2021for download from IP 132.174.252.179.
License or copyright restrictions may apply to redistribution; see https://www.ams.org/publications/ebooks/terms