Spectral Theory Notes

Zhang Ruiyang

Contents

Preface

This is the notes taken while taking MATH0071 Spectral Theory by Prof Leonid Parnovski at UCL in Term 1, 2022.

In Chapter 1, we recall some basic results from linear algebra and functional analysis to motivate the study of spectral theory. In Chapter 2, we start by introducing the two key objects - the spectrum and the resolvent - and discuss various fundamental properties and results about them. We then move on to two specific types of operators - projections and compact operators - in Chapter 3. Note that we only require the space to be Banach in the first three chapters. From there onwards, we move to Hilbert spaces and study some theories in Chapter 4. The final chapter, Chapter 5, studies some additional topics of spectral theory.

The sections with asterisks are not examinable, mostly those involved with unbounded operators. However, it would be beneficial to still read them as many interesting spectra come from differential operators, which are unbounded.

Note that in this course, $\mathbb{N} = \{1, 2, 3, \cdots\}$.

Chapter 1

Introduction

Spectral theory studies the spectrum of operators. Operator is a very broad term, and different kinds of operators (self-adjoint, bounded, etc.) have very different properties, and naturally very different properties of their spectrum. In this course, we will develop the spectral theory of bounded operators rigorously, and that of unbounded operators in a hand-wavy manner due to time constraints. Differential operators are unbounded, and they are of great significance and provide a lot of motivation for the development of spectral theory, which is why we believe it would be helpful to mention the spectrum of unbounded operators every now and then.

This course will start with discussing spectral theory in Banach spaces, and we will prove as much as we can in it. Then we will move on to Hilbert spaces. Some people teach spectral theory by diving right into Hilbert spaces, but since many results will remain valid in Banach spaces we might as well start with those.

So far, we have used the term 'spectrum' a lot. This is not too new of a concept. Consider a $n \times n$ matrix A with $\mathbb{R}$ entries, which is a linear operator that maps from $\mathbb{R}^n$ to $\mathbb{R}^n$. The set of eigenvalues of A is then the spectrum of A. In the following section, we will start with a recap of the results for these finite-dimensional operators.

In this course, all mappings/operators are going to be **linear**, unless stated otherwise. Also, we can only discuss the spectrum for operators that map from a space to itself, so we will only consider those operators. Finally, we might be working with vector spaces, and those vector spaces will be on field $\mathbb{F}$, which would be either $\mathbb{R}$ or $\mathbb{C}$, and we should assume they are on $\mathbb{C}$ unless stated otherwise.

1.1 Finite-Dimensional Recap

Consider an $n \times n$ matrix A. Now that we know the spectrum of A is just the set of its eigenvalues, how should we find those?

By definition, if we have

$$Av = \lambda v$$

for a non-zero vector v, then we say λ is an eigenvalue of A with v being its corresponding eigenvector (or the other way around). Equivalently, we the above equation is just $(A - \lambda I)v = 0$. Since v is non-zero, this equation implies that $A - \lambda I$ has *non-trivial kernel* (i.e. not $\{0\}$), which also means $A - \lambda I$ is *not injective*. So far, everything we have derived *does not* rely on the fact that A is a finite-dimensional operator.

In the finite-dimensional setting, if an operator is not injective, it is also not surjective, and not bijective. So, $A - \lambda I$ is not bijective, meaning that it is also not invertible, which is to say

$$\det(A - \lambda I) = 0.$$

The determinant of $A - \lambda I$ is called the **characteristic polynomial** $\chi_A(\lambda)$ of A, which is a polynomial of degree n. Thus, the eigenvalues of A are just the set of roots of the characteristic polynomial. By the Fundamental Theorem of Algebra, the characteristic polynomial is of degree n, so it has n roots in $\mathbb{C}$, although some of them might be repeating. This means there will always be n eigenvalues for A.

If all eigenvalues are distinct, A is **diagonalisable**, meaning that there exists matrix T such that $T^{-1}AT = D$ for a diagonal matrix D. Notice that this T should also have columns that are orthogonal to each other and are unit vectors, so T is an orthonormal matrix.

If a matrix has repeated eigenvalues, it is still possible to be diagonalisable, as long as the algebraic multiplicity and geometric multiplicity of every eigenvalue of the matrix coincide. Let us define these two multiplicities. For matrix A with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$, we can factorise the characteristic function as

$$\chi_A(\lambda) = (\lambda_1 - \lambda)^{a_1} \cdots (\lambda_k - \lambda)^{a_k}.$$

In this case, a_k is the **algebraic multiplicity** of the eigenvalue λ_k. Next, for an eigenvalue λ, we will consider its corresponding eigenspace $V_\lambda = \{v \ : \ Av = \lambda v\}$, which is the space of all possible eigenvectors with eigenvalue λ. The dimension of eigenspace V_λ is then defined to be the **geometric multiplicity** of λ. We claim that for a fixed eigenvalue, its algebraic multiplicity is greater than or equal to its geometric multiplicity. A consequence of this is that a matrix with distinct eigenvalues will have algebraic multiplicity 1 for all its eigenvalues and thus geometric multiplicity 1 as well, so these two quantities coincide for every eigenvalue, resulting in this matrix being diagonalisable.

It is not always the case that the two multiplicities coincide, of course. Consider matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

It can be computed that A has repeated eigenvalues 0, so it has algebraic multiplicity 2, while its geometric multiplicity is only 1.

It is going to be a bit of an effort to check for the coincidence of two multiplicities of all eigenvalues of a matrix. For some matrices, it would be obvious after first glance to realise it is diagonalisable. A matrix A that satisfies $A^* = A$, with A^* being the adjoint of A, is called **self-adjoint** (or **Hermitian**), and such matrices would be diagonalisable. The adjoint of a matrix is just the transpose of the complex conjugate of A, while if the matrix is on $\mathbb{R}$ it is simply the transpose. So, if we are only on $\mathbb{R}$, a self-adjoint matrix is called **symmetric**.

Self-adjoint is a very useful property to have. A lot of horrible things might happen if the matrix does not have this property. To illustrate one potential catastrophe, consider the following two matrices. Matrix A is 31×31, with entries 2 on $(1, 2), (2, 3), \ldots, (30, 31)$ and entries 0 everywhere else. Matrix B is almost the same as A, but the bottom left entry of B is 2^{-30} instead of 0. Notice that B is a slightly perturbed version of A, and the magnitude of this perturbation is almost negligible, especially if it is stored in a computer (almost all computers will view 0 and 2^{-30} as the same). However, if we consider the characteristic polynomials of A and B, we would have

$$\chi_A(\lambda) = -\lambda^{31}$$

and

$$\chi_B(\lambda) = -\lambda^{31} + 1.$$

This means, the spectrum of A, denoted by σ_A, is $\{0\}$, while the spectrum of B is 31-th roots of unity. Graphically speaking, on a complex plane, σ_A is simply the origin while σ_B is the equally spread out points on the unit circle. Those are very different, yet for a computer, this difference cannot be noticed, which is why eigenvalues computed by a computer might not be accurate. This issue, however, does not arise for self-adjoint matrices, which is why they are so great.

Earlier on, we demonstrated a sequence of equivalent formulations of eigenvalues and eigenvectors. We started with a few alternative formulations that do not rely on the operator being finite-dimensional. We only incorporate the fact that the operator is finite-dimensional into the derivation when we claim that a finite-dimensional operator that is not injective is not surjective and not bijective either. This equivalence does not hold for infinite dimensional operators. Let us consider an example when this does not hold.

Consider the linear mapping $T : S \to S$ for some infinite dimensional set S. If $S = \mathbb{N}$ being the natural number starting from 1 (note that in this course $\mathbb{N}$ starts from 1), the mapping $T : n \mapsto n + 1$ is injective but not surjective, and the mapping $P : n \mapsto n - 1$ for $n \neq 1$ and $1 \mapsto 1$ is surjective but not injective.

Alternatively, consider $S = l_2$ where l_2 is the set of infinite dimensional vectors with a finite sum of squares of its entries. The mapping

$$A(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$$

is injective but not surjective, and the mapping

$$B(x_1, x_2, x_3, \ldots) = (x_2, x_3, \ldots)$$

is surjective but not injective. Also, notice that

$$AB(x_1, x_2, \ldots) = (0, x_2, x_3, \ldots)$$

and

$$BA(x_1, x_2, \ldots) = (x_1, x_2, x_3, \ldots)$$

which is the identity. This hints at the idea of left inverse and right inverse, which will be discussed in detail later.

1.2 Spaces Basics

Let X be a vector space over $\mathbb{F}$.

Definition 1.1. *A **norm** is a function $\|\cdot\| : X \to [0, \infty)$ such that*

1. $\|x\| = 0 \iff x = 0$
2. $\|\lambda x\| = |\lambda|\|x\|$ *for all* $\lambda \in \mathbb{F}$ *and* $x \in X$
3. $\|x + y\| \leq \|x\| + \|y\|$ *for all* $x, y \in X$.

*Given this norm, we can then induce a **metric** d if we let $d(x, y) := \|x - y\|$ for all $x, y \in X$.*

An important property that we would like to have is completeness. A normed space (i.e. a vector space equipped with a norm) is complete if every Cauchy sequence converges.

Definition 1.2. *A **Banach space** $(X, \|\cdot\|)$ is a normed space that becomes complete with respect to the induced metric d.*

So far we have metric and norm, and a metric can be induced by a norm. Next, we would like to go one step further and define the inner product.

Definition 1.3. *A vector space X is an **inner product space** if there exists a function, known as the **inner product**, $(\cdot, \cdot) : X \times X \to \mathbb{F}$ such that*

1. $(x, x) \geq 0$ *for all* $x\ inX$
2. $(x, x) = 0 \iff x = 0$
3. $(\lambda x + \mu y, z) = \lambda(x, z) + \mu(y, z)$ *for all* $\lambda, \mu \in \mathbb{F}$ *and* $x, y, z \in X$.
4. $(x, y) = \overline{(y, x)}$ *for all* $x, y \in X$.

Given an inner product $(\cdot, \cdot)$, we can induce a norm $\|\cdot\|$ by setting $\|x\| = \sqrt{(x, x)}$. To verify this function is indeed a norm, we would need to use the **Cauchy-Schwarz-Buniakovski inequality** (more commonly known as Cauchy-Schwarz inequality), which states that $|(x, y)| \leq \|x\| \cdot \|y\|$ for all $x, y \in X$.

A normed space that is complete is known as the Banach space. Now that we have an inner product space, we would have the following definition.

Definition 1.4. *An inner product space $(X, (\cdot, \cdot))$ is a **Hilbert space** if it is complete in the induced norm $\|\cdot\|$.*

1.3 Spaces Examples

Here, we will give a few examples of spaces that we will be working with.

Example. We will start with two infinite-dimensional vector spaces, c_{00} and c_0.

$$c_{00} = \{x = (x_1, x_2, \ldots, x_n, 0, \ldots), x_j \in \mathbb{C}\}$$

and

$$c_0 = \{x = (x_1, x_2, \ldots), x_j \in \mathbb{C}, x_n \to 0 \text{ as } n \to \infty\}.$$

We can equip these two spaces with the sup norm, which is $\|x\| = \sup_j |x_j|$. The first space is not Banach, while the second space is. To see why the first space is not, consider the vector $x^n = (1, 1/2, \ldots, 1/n, 0, \ldots)$ which is in c_{00}. It is clear that $\{x^n\}$ is Cauchy, but it does not converge in c_{00}, although it does converge in c_0.

Example. We will consider another infinite-dimensional vector space l_p.

$$l_p = \left\{ x = (x_1, x_2, \ldots), x_j \in \mathbb{C} \;\middle|\; \sum_{j=1}^{\infty} |x_j|^p < \infty \right\}.$$

The space can be equipped with the l_p norm, defined by $\|x\|_p := \left[\sum_{j=1}^{\infty} |x_j|^p\right]^{1/p}$. This space would be a normed space and Banach space for $1 \leq p < \infty$. When $p = 2$, this space would be a Hilbert space, with the inner product $(x, y) := \sum_{j=1}^{\infty} x_j \overline{y_j}$.

Example. The last vector space that we will consider here is the l_∞ space, which is defined by

$$l_\infty = \{x = (x_1, x_2, \ldots), x_j \in \mathbb{C}\}$$

and the norm of this space is $\|x\|_\infty = \sup_j |x_j| < \infty$.

Example. Moving on to spaces of functions. Set $X = C[a,b]$ be the space of continuous functions on $[a,b]$, with $a < b$ and $a, b \in \mathbb{R}$. For every function $f : X \to X$ in $C[a,b]$, we will define the norm $\|f\| = \sup_{t\in[a,b]} |f(t)|$. This space is a Banach space.

Example. We then consider the function space $C_p[a,b]$ of continuous functions on $[a,b]$ with norm defined to be $\|f\|_p = \left[\int_a^b |f(t)|^p dt\right]^{1/p}$ for $f \in C_p[a,b]$. This is a normed space for $1 \leq p < \infty$. This space is, however, not complete, and the completion of it is the $L_p[a,b]$ space. An example to illustrate the incompleteness of C_p is the sequence of functions f_n that takes 0 from a to $a + (b-a)/2 - 1/n$, 1 from $a + (b-a)/2 + 1/n$ to b, and linear in between. This function will converge to a function that takes 0 in the first half and 1 in the second half, with a discontinuity in between - which is not in C_p.

1.4 Operators Basics

At this stage, we have established spaces for the operators to be a mapping from. Next, we will develop some results about such operators.

Theorem 1.5. *Let X, Y be normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ respectively, and $A : X \to Y$. The following are equivalent:*

1. *A is continuous*
2. *A is continuous at any point*
3. *A is continuous at 0*
4. *there exists constant $c > 0$ such that $\|Ax\|_Y \leq c\|x\|_X$ for all $x \in X$.*

*We would also say A is a **bounded** operator.*

The last formulation in the above theorem helps us to define the norm of an operator.

Definition 1.6. *For a bounded operator A, we can define $\|A\|$ in the following equivalent ways:*

1. $\inf\{c > 0 \mid \|Ax\|_Y \leq c\|x\|_X$ *for all* $x \in X\}$
2. $\inf\{c > 0 \mid \|Ax\|_Y \leq c$ *for all possible* $\|x\| \leq 1\}$
3. $\sup_{x\in X, x\neq 0} \frac{\|Ax\|_Y}{\|x\|_X}$
4. $\sup_{x\in X, x\neq 0, \|x\|\leq 1} \|Ax\|_Y$.

In particular, we have $\|Ax\| \leq \|A\| \cdot \|x\|$ for all $x \in X$.

If the operator norms satisfy $\|AB\| \leq \|A\| \cdot \|B\|$ for any operators A, B of a vector space of operators, the space will be called a **Banach algebra**.

Definition 1.7. *Let X, Y be Banach spaces. We say that (A, D_A) is an **unbounded** operator from X to Y if D_A, the domain of A, is a linear subspace of X and $A : D_A \to Y$.*

Now we will discuss some properties of operators. If A is bounded, it is also continuous, meaning that if we have $\{x_n\}$ with $x_n \to x$, then $Ax_n \to Ax$. If A is closed, then it means for a sequence $\{x_n\}$ with $x_n \to x$ and $Ax_n \to y$, we would have $Ax = y$.

Theorem 1.8. *Let X, Y be normed spaces. Then, the space $B(X, Y)$ of all bounded operators $A : X \to Y$ is a normed space with operator norm. The space $B(X, Y)$ is Banach if and only if Y is Banach.*

Next, we will define the dual space.

Definition 1.9. *Let X be a normed space. A* ***functional*** *$f : X \to \mathbb{F}$ is a bounded mapping. The collection of all functionals is called the* **dual space** *X^* of X, and a dual space is always a Banach space.*

Some examples of dual spaces are shown next. $l_p^* \cong l_q$ where $1/p + 1/q = 1$ for $1 < p, q < \infty$. $l_1^* \cong l_\infty$. $c_0^* \cong l_1$. A space will be called **reflexive** if $X^{**} \cong X$, i.e. the dual of the dual is, in a rough sense, itself.

Given that we have defined the dual space, we can then define the two types of convergences in a normed space.

Definition 1.10. *Let X be a normed space, and consider $\{x_n\}$ with $x_n \in X$ and $x \in X$. We say that*

1. *x_n converges to x strongly if $\|x_n - x\| \to 0$. We will denote it as $x_n \to x$ or s-lim $x_n = x$ or $\lim x_n = x$*
2. *x_n converges to x weakly if for all functionals $f \in X^*$, we have $f(x_n) \to f(x)$. We will denote it as w-lim $x_n = x$ or $x_n \rightharpoonup x$.*

From the notation, it is obvious that the first type of convergence is stronger than the second. Let us consider the following an example of weak convergence that is not strong convergence.

Example. Let $X = c_0$, and we define $x_n = e_n = (0, \ldots, 0, 1, 0, \ldots)$ with 1 being the n-th entry. We will also consider $x = (0, 0, \ldots)$. First, we claim that $x_n \rightharpoonup x$. To show this, for any $f \in c_0^*$, we would want to have $f(x_n) \to f(x)$. We know that $f = (f_1, f_2, \ldots) \in l_1$, so $\sum |f_j| < \infty$ and with $f(y) = \sum f_j y_j$. This means, $f(x_n) = f_n \to 0 = f(x)$ as $n \to \infty$ since $f \in l_1$. Next, it is trivial that $\|x_n - x\| \neq 0$, meaning that x_n does not converge to x strongly.

1.5 Convergence of Operators

Suppose we have a sequence of operators A_n and A in $B(X)$, the set of bounded operators from X to X. There are three types of convergence of operators.

1. We say A_n converges to A **uniformly** (in norm) if $\|A_n - A\| \to 0$. We denote it by $\lim A_n = A$.
2. We say A_n converges to A **strongly** if for all $x \in X$, we have $A_n x \to Ax$, i.e. $\|A_n x - Ax\| \to 0$. We denote it by s-lim $A_n = A$.
3. We say A_n converges to A **weakly** if for all $x \in X$, we have $A_n x \rightharpoonup Ax$, i.e. for all $f \in A^*$, we have $f(A_n x) \to f(Ax)$. We denote it by w-lim $A_n = A$.

It is clear that uniform convergence $\implies$ strong convergence $\implies$ weak convergence. We will show one example of A_n converging strongly to A while it does not converge uniformly.

Example. Consider $X = l^2$. We define $A_n x = (x_1, \ldots, x_n, 0, \ldots)$. For $n > m$, we have

$$(A_n - A_m)x = (0, \ldots, 0, x_{m+1}, \ldots, x_n, 0, \ldots),$$

which implies $\|(A_n - A_m)x\| \leq 1 \cdot \|x\|$. Also, consider $e_n = (0, \ldots, 0, 1, 0, \ldots) \in l^2$. We have

$$(A_n - A_m)e_n = 1 \cdot \|e_n\|$$

which means the inequality is sharp, and the operator norm of $\|A_n - A_m\|$ is 1, meaning that $\{A_n\}$ is not Cauchy and thus $\{A_n\}$ does not converge uniformly. However, A_n does converge strongly to identity I. Notice that

$$\|A_n x - Ix\| = \|(0, \ldots, x_{n+1}, x_{n+2}, \ldots)\| = \left|\sum_{j=n+1}^{\infty} x_j^2\right|^{1/2} \to 0$$

as $n \to \infty$ since $x \in l^2$.

Next, we will consider two types of sequence convergence.

Definition 1.11. *Let X be a normed space. Consider a sequence x_j, $x \in X$. We say that the series $\sum_{j=1}^{\infty} x_j$* ***converges*** *to x if $s\text{-}\lim S_n = x$ where the partial sum $S_n = \sum_{j=1}^{n} x_j$. We say the series $\sum_{j=1}^{\infty} x_j$* ***converges absolutely*** *if $\sum_{j=1}^{\infty} \|x_j\|$ converges.*

Theorem 1.12. *When X is a Banach space, a series that is converging absolutely is convergent.*

Proof. Since X is a Banach space, it is complete for every Cauchy sequence converges. Assuming $\{x_j\}$ in X converges absolutely to $x \in X$, we have, for any $\varepsilon > 0$, there exists some N such that for $n > m > N$, we have

$$\sum_{j=1}^{n} \|x_j\| - \sum_{j=1}^{m} \|x_j\| = \sum_{j=m+1}^{n} \|x_j\| \leq \varepsilon.$$

Using triangular inequality, we have

$$\left\|\sum_{j=1}^{n} x_j - \sum_{j=1}^{m} x_j\right\| = \left\|\sum_{j=m+1}^{n} x_j\right\| \leq \sum_{j=m+1}^{n} \|x_j\| \leq \varepsilon$$

where the first inequality above involves the triangular inequality. Thus, the sequence of partial sums is Cauchy, so it converges, as desired. □

1.6 Inverse and Perturbation of Operators

We first recall some definitions from Linear Algebra that are used in Functional Analysis.

Let $A \in B(X, Y)$ be a bounded operator between normed spaces X and Y. We denote the **kernel** of A as

$$\mathrm{Ker}(A) := \{x \in X \mid Ax = 0\}$$

and the **range** of A as

$$\mathrm{Ran}(A) := \{y \in Y \mid \exists x \in X \text{ with } Ax = y\} = \{Ax \mid x \in X\}.$$

Notice that $\mathrm{Ker}(A)$ is a closed subspace of X.

Saying A is invertible (or bijective) is equivalent to say that (1) $\mathrm{Ker}(A) = \{0\}$, i.e. A is injective (or one-to-one), and (2) $\mathrm{Ran}(A) = Y$, i.e. A is surjective (or onto). Notice if A is invertible, we will denote its inverse as A^{-1} and we will have $AA^{-1} = A^{-1}A = I$, where the identity is in its corresponding spaces. Next, we have the following result.

Theorem 1.13 (Banach Inverse Mapping). *If X and Y are Banach spaces, and if $A \in B(X, Y)$ and A^{-1} exists, then $A^{-1} \in B(Y, X)$.*

This theorem is a direct result of the open mapping theorem, which we state below, since a bounded operator is a continuous operator.

Theorem 1.14 (Open Mapping Theorem). *If X and Y are Banach spaces, and if $A \in B(X, Y)$ and A is surjective, then A is an open map (i.e. for any G that is open in X, $A(G)$ is open in Y).*

Inverse might not always exist, but sometimes we could have partial inverses known as the **left inverse** and the **right inverse**.

Definition 1.15. *Let $A \in B(X, Y)$ for normed spaces X and Y. A **left inverse** A_l^{-1} of A is an operator from Y to X such that $A_l^{-1}A = I_X$. A **right inverse** A_r^{-1} of A is an operator from Y to X such that $AA_r^{-1} = I_Y$.*

It is clear that a left inverse will exist if $\mathrm{Ker}(A) = \{0\}$, and a right inverse will exist if $\mathrm{Ran}(A) = Y$. This then implies the following theorem.

Theorem 1.16. *Let $A \in B(X, Y)$ for normed spaces X and Y. If A has both left inverse A_l^{-1} and right inverse A_r^{-1}, then A is invertible, and we have $A^{-1} = A_l^{-1} = A_r^{-1}$.*

This theorem should be clear from the definitions of left and right inverses. Next, we will have several properties about inverses, which are standard Linear Algebra.

Lemma 1.17. *Assuming the operators and product of operators in the following are well defined. We have:*

1. *For $A, B \in B(X, Y)$, if both A^{-1} and B^{-1} exist, then (AB) is invertible, and $(AB)^{-1} = B^{-1}A^{-1}$.*
2. *For $A, B \in B(X, X)$, if $AB = BA$ and (AB) is invertible, then A and B are invertible.*
3. *For $A, B \in B(X, X)$, if $AB = BA$ and A^{-1} exists, then $A^{-1}B = BA^{-1}$.*

Proof. (1) Notice that A^{-1} and B^{-1} exist implies $A^{-1}A = AA^{-1} = I$ and $B^{-1}B = BB^{-1} = I$. So, we have $B^{-1}A^{-1}AB = B^{-1}B = I$ and $ABB^{-1}A^{-1} = AA^{-1} = I$, thus (AB) is invertible with its inverse being $(B^{-1}A^{-1})$.

(2) Since (AB) is invertible, so there exist some S such that $SAB = ABS = I$. This means (BS) is a right inverse of A. Also, $SAB = SBA$ since $AB = BA$, and this implies that (SB) is a left inverse of A. So, A has both left and right inverse, meaning that it is invertible. A similar argument can be established for B to show that B is invertible.

(3) We have $A^{-1}B = A^{-1}BAA^{-1} = A^{-1}ABA^{-1} = BA^{-1}$, as desired. □

Notice that the second part of the above lemma is some form of the converse of the first part, but with an additional condition on the commutativity of A and B. We will establish an example in the following to show that the second part of the lemma will not hold without the commutativity.

Example. Consider $X = l^2$, so each element $x \in X$ will be of the form $x = (x_1, x_2, \ldots)$. Define two operators A and B as

$$A(x_1, x_2, \ldots) := (0, x_1, x_2, \ldots)$$

and

$$B(x_1, x_2, x_3, \ldots) := (x_2, x_3, \ldots).$$

Notice that $BA = I$ and is invertible, while $AB \neq I$, as

$$AB(x_1, x_2, \ldots) = (0, x_2, x_3, \ldots).$$

Clearly, A and B are not invertible. So the second part of the above lemma will not hold without $AB = BA$.

We will end this section by discussing two perturbation lemmas. A perturbation is the concept of making a change of small magnitude to some object, and we would like to see how stable that object is to the perturbation.

Lemma 1.18 (First Perturbation Lemma)**.** *Let X be a Banach space and $A \in B(X)$. If $\|A\| < 1$, then $(I - A)$ is invertible, and*

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n = I + A + A^2 + \ldots.$$

Also, $\|(I - A)^{-1}\| \leq \frac{1}{1-\|A\|}$.

Proof. Define $S = \sum_{n=0}^{\infty} A^n$. This series converges uniformly, which can be established by showing that the partial sum converges absolutely. Recall that $\|A^k\| \leq \|A\|^k$ since $\|AB\| \leq \|A\|\|B\|$. We have,

$$\sum_{k=0}^{\infty} \|A^k\| \leq \sum_{k=0}^{\infty} \|A\|^k < \infty$$

since $\|A\| < 1$. We have shown in the previous section that absolute convergence implies convergence in Banach spaces.

Next, we would like to show $S(I - A) = (I - A)S = I$. We have

$$\sum_{k=0}^{n} A^k (I - A) = I - A^n \to I$$

as $n \to \infty$, and similar for $(I - A)S$. So $(I - A)^{-1}$ exists.

To show the norm bound of this inverse, we have

$$\|(I - A)^{-1}\| = \left\| \sum_{n=0}^{\infty} A^n \right\| \leq \sum_{n=0}^{\infty} \|A^n\| \leq \sum_{n=0}^{\infty} \|A\|^n = \frac{1}{1 - \|A\|}$$

since $\|A\| < 1$. $\square$

Lemma 1.19 (Second Perturbation Lemma)**.** *Let X, Y be Banach spaces and $A, B \in B(X, Y)$. Let A be invertible, and $\|B\| < 1/\|A^{-1}\|$. Then, $(A + B)$ is invertible,*

$$(A + B)^{-1} = A^{-1} \sum_{j=0}^{\infty} (-1)^j (BA^{-1})^j,$$

and we have

$$\|(A + B)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|B\| \cdot \|A^{-1}\|}.$$

Proof. Notice that $A+B = A(I-(-A^{-1}B))$. From the conditions, we know that A is invertible. Next, we have

$$\|-A^{-1}B\| \leq \|A^{-1}\| \cdot \|B\| < 1.$$

So, from the first perturbation lemma, $I-(-A^{-1}B)$ is invertible as well. So, from a previous lemma, if X and Y are invertible then XY is invertible too, so here, $(A+B)$ is invertible since A and $I - A^{-1}B$ are invertible.

To compute the explicit form of the inverse, we have

$$(A+B)^{-1} = [A(I-(-A^{-1}B))]^{-1} = \left[\sum_{j=0}^{\infty}(-1)^j(A^{-1}B)^j\right]A^{-1} = A^{-1}\sum_{j=0}^{\infty}(-1)^j(BA^{-1})^j,$$

using the expression from the first perturbation lemma.

Similarly, for the norm of the inverse, we have

$$\begin{aligned}\|(A+B)^{-1}\| &= \|[A(I-(-A^{-1}B))]^{-1}\| \\ &\leq \|A^{-1}\| \cdot \|(I-(-A^{-1}B))^{-1}\| \\ &\leq \frac{\|A^{-1}\|}{1-\|A^{-1}B\|} \leq \frac{\|A^{-1}\|}{1-\|A^{-1}\|\cdot\|B\|},\end{aligned}$$

as desired. $\square$

Chapter 2

Spectrum and Resolvent

2.1 Basic Definitions

Spectral theory studies the spectrum of various kinds of operators, and finally, we will define in this section what the spectrum of an operator is.

Definition 2.1. *Let X be a Banach space, and $A \in B(X)$.*

- *The **resolvent set** $\rho(A)$ is the set of all $\lambda \in \mathbb{C}$ such that $(A - \lambda I)$ is invertible, and $(A - \lambda I)^{-1}$ is bounded.*
- *The **spectrum** $\sigma(A)$ is the complement of $\rho(A)$, so we have $\sigma(A) = \mathbb{C} \backslash \rho(A)$.*
- *We say $\lambda \in \mathbb{C}$ is an **eigenvalue** if there exists some **eigenvector** (or **eigenfunction**) $x \in X \backslash \{0\}$ such that $Ax = \lambda x$.*
- *The set of all eigenvalues of A is called a **point spectrum** $\sigma_p(A)$.*

Remark. The boundedness condition of the inverse in the first definition above holds automatically for bounded operators due to the existence of inverse and the inverse mapping theorem. For an unbounded operator A, its resolvent and spectrum and such are defined in the say way, but notice that for λ to be in the resolvent set, we need to make sure $(A - \lambda I)^{-1}$ is bounded, which always holds for bounded A but not for unbounded A.

Theorem 2.2. *For a Banach space X and operator $A \in B(X)$, we have $\sigma_p(A) \subset \sigma(A)$.*

Proof. Let $\lambda \in \sigma_p(A)$. Then, we have $Ax = \lambda x$ for some $x \neq 0$. So, clearly $(A - \lambda I)x = 0$ with non-zero x, which makes $x \in \operatorname{Ker}(A - \lambda I)$. So, $\operatorname{Ker}(A - \lambda I) \neq \{0\}$, thus $\lambda \in \sigma(A)$. □

Let us consider some examples of the spectrum and the point spectrum.

Example. If X is a finite-dimensional operator, then $\sigma_p(X) = \sigma(X)$. However, this relationship rarely holds when the operators are infinite-dimensional.

Example. Consider $X = C[0, 1]$ with $\| \cdot \|_{\sup}$ norm. We define operator A by $(Af)(t) := tf(t)$ for any $f \in C[0, 1]$.

For some $\lambda \in \mathbb{C}$, we have

$$(A - \lambda I)f(t) = (t - \lambda)f(t)$$

and we are interested in the λ values that make $A - \lambda I$ invertible. Consider some $g \in C[0,1]$, we have

$$(A - \lambda I)^{-1} g(t) = \frac{g(t)}{t - \lambda}.$$

This will be a well-defined operator for $\lambda \notin [0,1]$, so such λ are not in the spectrum.

If $\lambda \in [0,1]$, we have

$$\mathrm{Ker}(A - \lambda I) = \{f \mid tf(t) = \lambda f(t)\} = \{f \mid (t - \lambda) f(t) = 0\} = \{0\}$$

since there will always be an undefined point when $t = \lambda$. So, there is no eigenvalues, and thus $\sigma_p(A) = \phi$. Next, we have

$$\mathrm{Ran}(A - \lambda I) = \{(t - \lambda) f(t) \mid f \in C[0,1]\} = \{g \in C[0,1] \mid g(\lambda) = 0\} \neq X.$$

Thus, all such λ will make $A - \lambda I$ not invertible, so $\sigma(A) = [0,1]$. Clearly, $\sigma_p(A) \neq \sigma(A)$ in this case.

Now, we will show some properties of the spectrum of a bounded operator.

Theorem 2.3. *Let X be a Banach space, and $A \in B(X)$. Then, we have*

1. *$\sigma(A) \subseteq B_C(0, \|A\|) = \{\lambda \in \mathbb{C} \mid |\lambda| \leq \|A\|\}$.*
2. *Suppose $\lambda_0 \in \rho(A)$ and $|\lambda - \lambda_0| < 1/\|(A - \lambda_0 I)^{-1}\|$, then $\lambda \in \rho(A)$.*

Moreover, $\sigma(A)$ is compact in $\mathbb{C}$.

Proof. The compactness of $\sigma(A)$ is a direct consequence of the two parts. The first part indicates that $\sigma(A)$ is bounded, while the second part indicates that the complement of $\sigma(A)$, the resolvent set $\rho(A)$, is open, which implies that $\sigma(A)$ is closed. So, $\sigma(A)$ is closed and bounded, which means it is compact.

(1) Suppose $|\lambda| > \|A\|$. Then, $A - \lambda I = (-\lambda)(I - A/\lambda)$. Notice that $\|A/\lambda\| = \|A\|/|\lambda| < 1$, we can apply the first perturbation lemma and say that $(A - \lambda I)^{-1}$ exists. So, $\lambda \in \rho(A)$, and this implies $\sigma(A) \subseteq B_C(0, \|A\|)$, as desired.

(2) $A - \lambda I = (A - \lambda_0 I) + (\lambda_0 - \lambda) I$. Notice that the first term on the right is invertible since $\lambda_0 \in \rho(A)$, and the second term is a small perturbation. We have,

$$\|(\lambda_0 - \lambda) I\| = |(\lambda_0 - \lambda)| < \frac{1}{\|(A - \lambda_0 I)^{-1}\|}$$

from the condition. So by the second perturbation lemma, we know that $(A - \lambda_0 I) + (\lambda_0 - \lambda) I = A - \lambda I$ is invertible, meaning that $\lambda \in \rho(A)$. This means, for every point λ_0 in $\rho(A)$, we can always find an open ball with radius $1/\|(A - \lambda_0 I)^{-1}\|$ and center λ_0, such that it is contained in $\rho(A)$, thus $\rho(A)$ is open in $\mathbb{C}$. $\square$

Remark. As one might have noticed in the proof of this theorem, we would commonly seek help from the resolvent set when we want to know something about its complement, the spectrum. This approach will reoccur many times.

Remark. Suppose $\lambda_0 \in \rho(A)$ and $|\lambda - \lambda_0| < 1/[2\|(A - \lambda_0 I)^{-1}\|]$, then we have

$$\|(A - \lambda I)^{-1}\| < \frac{\|(A - \lambda_0 I)^{-1}\|}{1 - \|(A - \lambda_0 I)^{-1}\| \cdot |\lambda - \lambda_0|} < 2\|(A - \lambda_0 I)^{-1}\|$$

using the second perturbation lemma. This bound does not depend on λ, so it is uniformly bounded. We will use this result later on.

2.2 Resolvent

Definition 2.4. *The operator-valued function of complex variable $R(A;\lambda)$, defined by $\lambda \mapsto (A-\lambda I)^{-1}$ for every $\lambda \in \rho(A)$, is called the* ***resolvent*** *of A.*

We have two identities about the resolvent, both of which are easy to show.

Lemma 2.5 (First Resolvent Identity)**.**

$$R(A;\lambda) - R(A;\lambda_0) = (\lambda - \lambda_0)R(A;\lambda)R(A;\lambda_0)$$

for all $\lambda, \lambda_0 \in \rho(A)$.

Proof.

$$\begin{aligned}
R(A;\lambda) - R(A;\lambda_0) &= (A-\lambda I)^{-1} - (A-\lambda_0 I)^{-1} \\
&= (A-\lambda I)^{-1}[(A-\lambda_0 I) - (A-\lambda I)](A-\lambda_0 I)^{-1} \\
&= (A-\lambda I)^{-1}(\lambda-\lambda_0)(A-\lambda_0 I)^{-1} \\
&= (\lambda-\lambda_0)(A-\lambda I)^{-1}(A-\lambda_0 I)^{-1}.
\end{aligned}$$

□

Remark. The order of $R(A;\lambda)$ and $R(A;\lambda_0)$ can be switched around.

Lemma 2.6 (Second Resolvent Identity)**.**

$$R(A;\lambda) - R(B;\lambda) = R(A;\lambda)(B-A)R(B;\lambda)$$

for all $A, B \in B(X)$ and $\lambda \in \rho(A) \cap \rho(B)$.

Proof.

$$\begin{aligned}
R(A;\lambda) - R(B;\lambda) &= (A-\lambda I)^{-1} - (B-\lambda I)^{-1} \\
&= (A-\lambda I)^{-1}[(B-\lambda I) - (A-\lambda I)](B-\lambda I)^{-1} \\
&= (A-\lambda I)^{-1}(B-A)(A-\lambda_0 I)^{-1}.
\end{aligned}$$

□

Remark. The order of $R(A;\lambda)$ and $R(B;\lambda)$ *cannot* be switched around.

We will first state and not prove the following theorem-definition, where we study what it means for an operator-valued function to be holomorphic.

Theorem 2.7. *Let Z be a complex Banach space, the domain $\Omega \subset \mathbb{C}$ be an open set, and $F : \Omega \to Z$ is a Z-valued function. Then, the following are equivalent:*

1. *For all $\lambda_0 \in \Omega$, there exists*

$$F'(\lambda_0) = \frac{dF}{d\lambda}\bigg|_{\lambda=\lambda_0} = \lim_{\lambda\to\lambda_0} \frac{F(\lambda)-F(\lambda_0)}{\lambda-\lambda_0} \in Z,$$

and this limit exists meaning that

$$\left\| F'(\lambda_0) - \frac{F(\lambda)-F(\lambda_0)}{\lambda-\lambda_0} \right\| \to 0$$

as $\lambda \to \lambda_0$.

2. For all $\lambda_0 \in \Omega$, there exists a neighbourhood of λ_0 where F has the form

$$F(\lambda) = \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n F_n(\lambda_0)$$

where $F_n(\lambda_0) \in Z$ are coefficients of this decomposition.

3. For all $G \in Z^*$, there is a complex-valued function

$$\Omega \ni \lambda \mapsto G(F(\lambda)) \in \mathbb{C}$$

is holomorphic in Ω.

4. If $Z = B(X, Y)$ with Banach spaces X and Y, then for all $x \in X$ and $g \in Y^*$, the complex-valued function

$$\Omega \ni \lambda \mapsto g(F(\lambda)x) \in \mathbb{C}$$

is holomorphic.

If these conditions are satisfied, we say that F is ***holomorphic*** *in Ω.*

Next, we will show a few more identities about the resolvent.

Theorem 2.8. *Let X be a complex Banach space and $A \in B(X)$. Then, the resolvent $R(A; \cdot)$ is a holomorphic $B(X)$-valued function defined on the resolvent set $\rho(A)$. Moreover,*

1. we have

$$\frac{d}{d\lambda} R(A; \lambda)\Big|_{\lambda=\lambda_0} = R(A; \lambda_0)^2$$

for all $\lambda_0 \in \rho(A)$.

2. we have

$$-\lambda R(A; \lambda) \to I$$

as $\lambda \to \infty$.

3. we have

$$\|R(A; \lambda)\| \geq \frac{1}{d(\lambda, \sigma(A))}$$

where $d(\lambda, \sigma(A)) = \inf_{\mu \in \sigma(A)} |\lambda - \mu|$.

Proof. (1) Pick some $\lambda_0 \in \rho(A)$. We know that $R(A; \lambda)$ is bounded in some neighbourhood of λ_0, by part 2 of Theorem 2.3. We also know that, by the first resolvent identity, we have

$$R(A; \lambda) - R(A; \lambda_0) = (\lambda - \lambda_0) R(A; \lambda) R(A; \lambda_0),$$

which is equivalent to

$$R(A; \lambda) = R(A; \lambda_0) + (\lambda - \lambda_0) R(A; \lambda) R(A; \lambda_0),$$

and the second term of the above equation goes to zero as $\lambda \to \lambda_0$ since $R(A; \lambda)$ is bounded. Thus, taking the limit on both sides of the above equation gives us

$$\lim_{\lambda \to \lambda_0} R(A; \lambda) = R(A; \lambda_0).$$

Moreover, the first resolvent identity provides us with

$$\frac{R(A; \lambda) - R(A; \lambda_0)}{\lambda - \lambda_0} = R(A; \lambda) R(A; \lambda_0),$$

and taking limit with $\lambda \to \lambda_0$ will thus yield the desired identity.

(2) We have, consider λ with $|\lambda| > \|A\|$, then

$$\begin{aligned} \|-\lambda R(A;\lambda) - I\| &= \|-\lambda(A-\lambda I)^{-1} - I\| = \|(I - A/\lambda)^{-1} - I\| \\ &= \left\| \sum_{n=0}^{\infty} \left(\frac{A}{\lambda}\right)^n - I \right\| \quad \text{using the first perturbation lemma} \\ &= \left\| \sum_{n=1}^{\infty} \left(\frac{A}{\lambda}\right)^n \right\| \le \sum_{n=1}^{\infty} \left\| \left(\frac{A}{\lambda}\right)^n \right\| \le \sum_{n=1}^{\infty} \left\| \frac{A}{\lambda} \right\|^n = \sum_{n=1}^{\infty} \frac{\|A\|^n}{|\lambda|^n} \\ &= \frac{1}{1 - \|A\|/|\lambda|} \cdot \frac{\|A\|}{|\lambda|} = \frac{\|A\|}{|\lambda| - \|A\|} \to 0 \end{aligned}$$

as $\lambda \to \infty$. Also, since $\|-\lambda R(A;\lambda) - I\| \ge 0$, we have our desired identity.

(3) Let $\lambda_0 \in \rho(A)$. Then, by part 2 of Theorem 2.3, we have $B(\lambda_0, 1/\|R(A;\lambda_0)\|) \subset \rho(A)$. Thus, since $\sigma(A)$ is the complement of $\rho(A)$, we have

$$d(\lambda_0, \sigma(A)) \ge \frac{1}{\|R(A;\lambda_0)\|},$$

which yields the desired identity after rearrangement. □

2.3 Spectrum is Non-Empty

Before showing that the spectrum is not empty, we need an auxiliary lemma.

Lemma 2.9. *Let X be a normed space and $x \in X$. Then, there exists a functional $f \in X^*$ such that $f(x) = \|x\|$ and $\|f\| = 1$.*

This result is a corollary of the Hahn-Banach Theorem.

Theorem 2.10. *For a Banach space X that is not $\{0\}$ and $A \in B(X)$, we have $\sigma(A) \neq \emptyset$.*

Proof. Suppose $\sigma(A) = \emptyset$, then $\rho(A) = \mathbb{C}$, and $R(A;\lambda)$ is defined on the whole $\mathbb{C}$ and thus is holomorphic. Take some $x \in X$ and $g \in X^*$, and we shall consider the function f that maps

$$\mathbb{C} \ni \lambda \mapsto g(R(A;\lambda)x) =: f(\lambda) \in \mathbb{C}.$$

This function is holomorphic on $\mathbb{C}$, and thus is a constant function due to Liouville's theorem.

Since $f(\lambda) \to 0$ as $\lambda \to \infty$, as $\|f(\lambda)\| \le \|g\|\|R(A;\lambda)\|\|x\|$ where $\|R(A;\lambda)\| \to 0$ as $\lambda \to \infty$ by the second identity of Theorem 2.8, we have $f(z) = 0$ for all z.

This means, we have $f(17) = g(R(A;17)x) = 0$. Using Lemma 2.9, we can pick a good functional g such that $\|g\| = 1$ and $g(R(A;17)x) = \|R(A;17)x\|$. This means we have

$$g(R(A;17)x) = \|R(A;17)x\| = 0.$$

Thus, $R(A;17)x = 0$, which implies $(A - 17I)(A - 17I)^{-1}x = 0$, and $x = 0$. This x is arbitrary, yet $x = 0$ for all $x \in X$, which is a contradiction. □

Let us compute the spectrum for a few examples.

Example. Consider a set $L = \{\lambda_1, \lambda_2, \cdots\}$ where $\lambda_j \in \mathbb{C}$ for all j, and $\|\lambda_j\| \leq M$ for some constant M. Let $X = l_1$, and $A : X \to X$ with $Ae_n = \lambda_n e_n$ for all n, which means A is a diagonal matrix with L being its diagonal. So,

$$A(x_1, x_2, \ldots) = (\lambda_1 x_1, \lambda_2 x_2, \ldots)$$

and $\|A\| = \sup |\lambda_j|$. It is obvious from the definition that $\sigma_p(A) = \{\lambda_1, \lambda_2, \ldots\}$. Next, since the spectrum is compact, as proved in Theorem 2.3, and the point spectrum is contained in the spectrum, we know that the spectrum contains the closure of $\{\lambda_1, \lambda_2, \ldots\}$. The natural question to ask at this point is, are there any other points in the spectrum? Consider $\lambda \notin \overline{\{\lambda_1, \lambda_2, \ldots\}}$ and $\lambda \in \sigma(A)$. We then would have

$$(A - \lambda I)^{-1}(x_1, x_2, \ldots) = \big((\lambda_1 - \lambda)^{-1} x_1, (\lambda_2 - \lambda)^{-1} x_2, \ldots\big)$$

which should not be well-defined as $\lambda \in \sigma(A)$, but it is, given that $\lambda \notin \overline{\{\lambda_1, \lambda_2, \ldots\}}$. So, we have a contradiction, and thus $\sigma(A) = \overline{\{\lambda_1, \lambda_2, \ldots\}}$.

Example. Once again, we are working in $X = l_1$. Consider the left shift operator $A(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$. We know from before that $\|A\| = 1$. In matrix form, this is an infinite-dimensional Jordan block with zeroes on the diagonal and ones right above the diagonal. First, let us find out the point spectrum of A. If A is finite-dimensional, the point spectrum will be zero. However, since we are in infinite-dimensional, things might be different. Suppose for some λ and $x \neq 0$, we have $Ax = \lambda x$. This means we have

$$A(x_1, x_2, \ldots) = (x_2, x_3, \ldots) = (\lambda x_1, \lambda x_2, \ldots),$$

which implies

$$x_2 = \lambda x_1, x_3 = \lambda x_2 = \lambda^2 x_1, \ldots, x_n = \lambda^{n-1} x_1, \ldots,$$

so $x = x_1(1, \lambda, \lambda^2, \ldots)$. Since $x \neq 0$, we have $x_1 \neq 0$. For this x to be in l_1, we must have

$$\sum_{n=0}^{\infty} |\lambda^n| < \infty,$$

meaning that $|\lambda| < 1$. So, we have $\sigma_p(A) = \{|\lambda| < 1\} = B(0, 1)$. Next, similar to above, we know that $\sigma(A) \supseteq B_c(0, 1) = \{|\lambda| \leq 1\}$. However, we also know that $\sigma(A) \subseteq B_c(0, \|A\|) = B_c(0, 1)$, as shown in Theorem 2.3, thus $\sigma(A) = B_c(0, 1) = \{|\lambda| \leq 1\}$.

To summarise the properties we have established so far about the spectrum $\sigma(A)$ for $A \in B(X)$ with X being a Banach space, we have

1. $\overline{\sigma_p(A)} \subseteq \sigma(A)$
2. $\sigma(A)$ is compact
3. $\sigma(A) \neq \emptyset$.

2.4 Spectral Radius

Definition 2.11 (Spectral Radius)**.** *The* ***spectral radius*** *of $A \in B(X)$ for a Banach space X is defined as $r(A) := \sup\{|\lambda| : \lambda \in \sigma(A)\} \in \mathbb{R}_+$. Notice that $r(A) \leq \|A\|$ and $\sigma(A) \subseteq B_c(0, r(A))$.*

Theorem 2.12. *Let X be a complex Banach space, and $A \in B(X)$. Then, $r(A) = \lim_{n\to\infty} \|A^n\|^{1/n}$.*

Proof. In order to prove the desired result, we will prove two partial results instead:

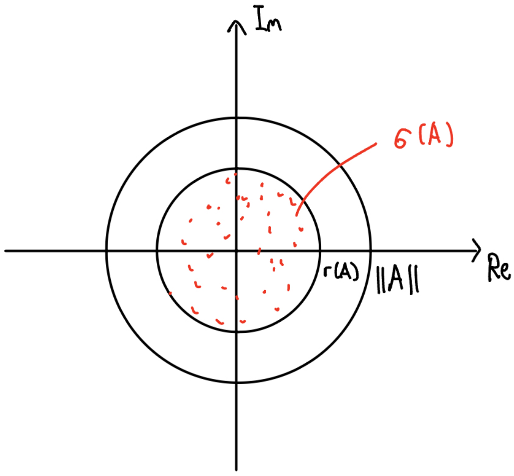

Figure 2.1: Spectral Radius

1. $r(A) \leq \liminf \|A^n\|^{1/n}$
2. $r(A) \geq \limsup \|A^n\|^{1/n}$.

Since $\liminf \leq \limsup$, these two partial results will yield

$$r(A) = \liminf \|A^n\|^{1/n} = \limsup \|A^n\|^{1/n} = \lim \|A^n\|^{1/n}.$$

Somehow, it is usually easier to establish an upper bound than to establish a lower bound. This is the case here as well.

(1) Suppose $\lambda \in \sigma(A)$. Since

$$A^n - \lambda^n I = (A - \lambda I)(A^{n-1} + A^{n-2}\lambda + \cdots + \lambda^{n-1} I)$$

and operators on the right-hand side of the above equation commute, knowing that $(A - \lambda I)$ is not invertible implies that $(A^n - \lambda^n I)$ is not invertible either, which means $\lambda^n \in \sigma(A^n)$. Therefore, given

$$r(A) = \sup\{|\lambda| : \lambda \in \sigma(A)\},$$

we have

$$\begin{aligned} r(A)^n &= [\sup\{|\lambda| : \lambda \in \sigma(A)\}]^n \\ &= \sup\{|\lambda|^n : \lambda \in \sigma(A)\} \quad \text{since } x^n \text{ is monotone for } x > 0 \\ &= \sup\{|\lambda^n| : \lambda \in \sigma(A)\} \leq \sup\{|\mu| : \mu \in \sigma(A^n)\} \\ &= r(A^n) \leq \|A^n\|. \end{aligned}$$

Thus, $r(A) \leq \|A^n\|^{1/n}$, and taking the $\liminf$ on both sides yields the desired partial result.

(2) Suppose we pick some λ outside the $\|A\|$-radius open ball, i.e. $|\lambda| > \|A\|$. Then, $\lambda \in \rho(A)$, and we have

$$R(A; \lambda) = (A - \lambda I)^{-1} = (-\lambda)^{-1}(I - A/\lambda)^{-1} = -\sum_{n=0}^{\infty} \frac{A^n}{\lambda^{n+1}}$$

where the last equality is due to the first perturbation lemma, Lemma 1.18, as $\|A/\lambda\| < 1$ by construction.

Let $x \in X$ and $g \in X^*$, and we consider the scalar valued function $f(\lambda) = g(R(A;\lambda)x)$. We know that $f(\lambda)$ is holomorphic for $|\lambda| > r(A)$ as $R(A;\lambda)$ is defined for $\lambda \in \rho(A)$. If $\|\lambda\| \geq \|A\|$, we know from the derivation above that

$$f(\lambda) = g(R(A;\lambda)x) = g\left(-\sum_{n=0}^{\infty} \frac{A^n x}{\lambda^{n+1}}\right) = -\sum_{n=0}^{\infty} \lambda^{-n-1} g(A^n x).$$

By Laurent's theorem, since f is holomorphic for $|\lambda| > r(A)$, the above decomposition thus holds for all $|\lambda| > r(A)$.

Now, take $\lambda = ae^{i\theta}$ where a is just slightly greater than $r(A)$. We have

$$f(ae^{i\theta}) = -\sum_{n=0}^{\infty} a^{-n-1} e^{-i\theta(n+1)} g(A^n x).$$

We then multiply the above equation by $a^{m+1}e^{i\theta(m+1)}$ for some fixed $m \in \mathbb{N}$, and we get

$$a^{m+1}e^{i\theta(m+1)} f(ae^{i\theta}) = -\sum_{n=0}^{\infty} a^{m-n} e^{i\theta(m-n)} g(A^n x).$$

Integrate the above equation from 0 to 2π with regards to θ yields

$$\begin{aligned}
\int_0^{2\pi} a^{m+1}e^{i\theta(m+1)} f(ae^{i\theta})\ d\theta &= \int_0^{2\pi} -\sum_{n=0}^{\infty} a^{m-n} e^{i\theta(m-n)} g(A^n x)\ d\theta \\
&= -\sum_{n=0}^{\infty} a^{m-n} g(A^n x) \int_0^{2\pi} e^{i\theta(m-n)}\ d\theta \\
&= -2\pi \sum_{n=0}^{\infty} g(A^n x),
\end{aligned}$$

since the integral $\int_0^{2\pi} e^{i\theta(m-n)}\ d\theta$ is only non-zero, and takes 2π instead, when $m = n$, so we just set our fixed m to be n.

Thus, we have

$$\begin{aligned}
|g(A^m x)| &= \left| -\frac{1}{2\pi} a^{m+1} \int_0^{2\pi} e^{i\theta(m+1)} f(ae^{i\theta}) \right| \\
&= \left| -\frac{1}{2\pi} a^{m+1} \int_0^{2\pi} e^{i\theta(m+1)} g\left(R(A;ae^{i\theta})x\right)\ d\theta \right| \\
&\leq \frac{a^{m+1}}{2\pi} \left| \int_0^{2\pi} e^{i\theta(m+1)} g\left(R(A;ae^{i\theta})x\right)\ d\theta \right| \\
&\leq \frac{a^{m+1}}{2\pi} \int_0^{2\pi} \left| e^{i\theta(m+1)} g\left(R(A;ae^{i\theta})x\right) \right|\ d\theta \\
&= \frac{a^{m+1}}{2\pi} \int_0^{2\pi} \left| g\left(R(A;ae^{i\theta})x\right) \right|\ d\theta \\
&\leq \frac{a^{m+1}}{2\pi} \|g\| \cdot \|x\| \int_0^{2\pi} \left\| R(A;ae^{i\theta}) \right\|\ d\theta.
\end{aligned}$$

We denote $M(a) := \sup_{0\leq\theta\leq 2\pi} \|R(A; ae^{i\theta})\|$, and that means

$$|g(A^m x)| \leq a^{m+1}\|g\| \cdot \|x\| M(a).$$

We can pick, using Lemma 2.9, a good functional $g \in X^*$ such that $\|g\| = 1$ and $g(A^m x) = \|A^m x\|$. So, we have

$$\|A^m x\| \leq a^{m+1}\|x\|M(a) \quad \iff \quad \sup_{x\neq 0} \frac{\|A^m x\|}{\|x\|} \leq a^{m+1}M(a),$$

and this means $\|A^m\| \leq a^{m+1}M(a)$ and $\|A^m\|^{1/m} \leq a \cdot a^{1/m} M(a)^{1/m}$. Taking the lim sup for m on both sides of the inequality will yield

$$\limsup \|A^m\|^{1/m} \leq \limsup a \cdot a^{1/m} M(a)^{1/m} = a.$$

Since the argument above is valid for any $a > r(A)$, we would have

$$\limsup \|A^m\|^{1/m} \leq r(A),$$

as desired. □

2.5 Spectral Mapping Theorem

We would like to know if the spectrum of a function of an operator is the function of the spectrum of the operator, i.e. for some operator A, we would want to know if $\sigma(f(A)) = f(\sigma(A))$ holds. Turns out, this holds when f is a polynomial. We do not really have similar results for other classes of functions, and it becomes more case-specific for those scenarios.

Let $p(\zeta) = \sum_{k=0}^{N} a_k \zeta^k$ be a polynomial with the non-zero leading term $a_N \neq 0$. Let A be a bounded operator $A \in B(X)$ with X being a Banach space. Then, we can define a polynomial of this operator as

$$p(A) = \sum_{k=0}^{N} a_k A^k$$

where $A^k = A \circ A \cdots \circ A$ and $A^0 = I$.

Theorem 2.13 (Spectral Mapping Theorem)**.** *For $A \in B(X)$ and $p(\zeta)$ a polynomial , we have*

$$\sigma(p(A)) = p(\sigma(A)) = \{p(\zeta) \mid \zeta \in \sigma(A)\}.$$

Proof. Take $\mu \in \mathbb{C}$ and solve for $p(\zeta) = \mu$. This has N roots as p is a degree N polynomial. Let $\lambda_1, \lambda_2, \cdots, \lambda_N \in \mathbb{C}$ be the roots of $p(\zeta) = \mu$, counting multiplicity. Then, we have

$$p(\zeta) - \mu = (\zeta - \lambda_1)(\zeta - \lambda_2)\cdots(\zeta - \lambda_N)$$

and

$$p(A) - \mu = (A - \lambda_1 I)(A - \lambda_2 I)\cdots(a - \lambda_N I).$$

We should note that the terms on the right of the above equation commute.

Notice that we then have the following equivalent results: $\mu \neq \sigma(A) \iff p(A) - \mu I$ is not invertible $\iff (A - \lambda_k I)$ is invertible for all $k \iff \lambda_k \notin \sigma(A)$ for all $k \iff \mu \notin p(\sigma(A))$. □

This marks the end of the abstract theory of the spectrum of operators.

Chapter 3

Projections and Compact Operators

3.1 Projection

We will study two specific types of operators. In this section, we will study projections, and in the next section, we will study compact operators. Let us start by defining what a projection is.

Definition 3.1. *Let X be a normed space. An operator $P \in B(X)$ is a* ***projection*** *if $P^2 = P$.*

Lemma 3.2. *Let $P \in B(X)$ be a projection. Then, $Q = I - P$ is also a projection, $PQ = 0 = QP$, and $Ran(P) = Ker(Q)$ and $Ran(Q) = Ker(P)$.*

Proof. Notice that $Q^2 = (I - P)^2 = I - 2P + P^2 = I - P = Q$, so Q is a projection. Also, $PQ = P(I - P) = P - P^2 = 0$, and also $QP = (I - P)P = P - P^2 = 0$.

Finally, since $QP = 0$, Q maps the range of P into 0, so the kernel of Q contains the range of P. Next, for some $x \in \mathrm{Ker}(Q)$, $0 = Qx = (I - P)x = x - Px$, meaning that $Px = x$ so x is in the range of P. Thus, $\mathrm{Ran}(P) = \mathrm{Ker}(Q)$. For the other property, it is similar. From $PQ = 0$, we know that $\mathrm{Ran}(Q) \subseteq \mathrm{Ker}(P)$. For $x \in \mathrm{Ker}(P)$, we have $0 = -Px = (I - P)x - x = Qx - x$, so $Qx = x$ and $x \in \mathrm{Ran}(Q)$. Thus, $\mathrm{Ran}(Q) = \mathrm{Ker}(P)$. $\square$

Lemma 3.3. *Let $P \in B(X)$ be a projection. Then, $Ran(P)$ is closed and $X = Ran(P) \bigoplus Ker(P)$.*

Remark. One way to define projection in linear algebra is to say that we can express the space X as a direct sum of V_1 and V_2, and the projection will map any $x \in X$ to its corresponding component in either V_1 or V_2 (same subspace for the same projection). This lemma aims to show that the present definition of projection satisfies this linear algebra definition as well.

Proof. We set $Q = I - P$, and from the previous lemma, we know that $\mathrm{Ran}(P) = \mathrm{Ker}(Q)$. The kernel of a bounded operator is closed, so $\mathrm{Ran}(P)$ is closed. Next, we know that $I = (I - P) + P$, so for any $x \in X$, we have $x = Qx + Px$, with $Qx \in \mathrm{Ker}(P)$ and $Px \in \mathrm{Ran}(P)$. This means we can decompose any element of X into the sum of elements from $\mathrm{Ran}(P)$ and $\mathrm{Ker}(P)$. So, $X = \mathrm{Ran}(P) + \mathrm{Ker}(P)$. To show this decomposition is unique, we just need to show in addition that the intersection of these two subspaces only contains 0.

Suppose $x \in \mathrm{Ran}(P)$. Then, it means that there is some $y \in X$ such that $Py = x$, which means that $P^2y = Px = Py = x$. Furthermore, if we also have $x \in \mathrm{Ker}(P)$, $Px = 0 = x$, which means the intersection $\mathrm{Ran}(P) \cap \mathrm{Ker}(P) = \{0\}$, as desired. $\square$

Next, we would like to know about the spectrum of a projection. There are two trivial projections - zero operator and identity operator. It is easy to note that the spectrum of the zero operator is 0 and that of the identity operator is 1. We would like to study the spectrum of, thus, non-trivial projections.

Theorem 3.4. *Let $P \in B(X)$ be a projection, and P is a non-trivial projection (so neither zero nor identity). Then, $\sigma(P) = \{0, 1\}$.*

Proof. For projection P, we know that $P^2 - P = 0$. So, using the spectral mapping theorem, we have $\sigma(P^2 - P) = \sigma(0) = 0 = \sigma(P)^2 - \sigma(P) = \{\zeta^2 - \zeta \mid \zeta \in \sigma(P)\}$. So, if $\zeta \in \sigma(P)$, $\zeta = 0$ or 1. This means the spectrum of P could only contain 0 and 1. Next, we would like to show if these two are really contained in the spectrum.

Since $P \neq 0$, we have $\mathrm{Ran}(P) = \mathrm{Ker}(I - P) \neq \{0\}$, meaning that $I - P$ is not invertible, so $1 \in \sigma(P)$.

Similarly, since $P \neq I$ which means $I - P \neq 0$, we have $\mathrm{Ran}(I - P) = \mathrm{Ker}(P) \neq \{0\}$, meaning that P is not invertible, so $0 \in \sigma(P)$.

Thus, $\sigma(P) = \{0, 1\}$, as desired. $\square$

3.2 Compact Operator

Definition 3.5. *Let X be a normed space, and $K \subseteq X$.*

1. *K is **relatively compact** if each sequence in K has a Cauchy subsequence.*
2. *K is **compact** if every sequence in K has a converging subsequence.*

Proposition 3.6. *If a subset $K \subset X$ is compact, then it is closed and bounded. If a subset $K \subset X$ is relatively compact, then it is bounded. If X is finite-dimensional, then the converse of the previous two statements is also true.*

The converse of the two statements does not hold for infinite-dimensional X.

Example. Consider $X = l_1$ and a sequence $\{x_n\}$ in X with $x_n = e_n = (0, 0, \ldots, 1, 0, \ldots)$ with 1 being the n-th element. It is easy to notice that $\{x_n\} \subset B_c(0, 1)$ so it is closed and bounded, yet this sequence is not Cauchy as $d(x_n, x_m) = 2$ for all $n \neq m$. This means $B_c(0, 1)$ is not relatively compact in l_1, although being closed and bounded.

Proposition 3.7. *Let X be finite-dimensional, and $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on X. Then, these two norms are equivalent, i.e. there exists constants $c_1, c_2 > 0$ such that $c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1$ for all $x \in X$.*

Lemma 3.8. *Let X be a normed space and X_0 be a linear finite-dimensional subspace of X. Then, X_0 is closed.*

Proof. Let $\{e_1, e_2, \ldots, e_n\}$ be a basis of X_0. Define a norm $\|\cdot\|_0$ on X_0 such that if $x = \sum_{j=1}^{n} a_j e_j$, then $\|x\|_0 = \sup_j |a_j|$. By the previous proposition, we know that $\|\cdot\|_0$ is equivalent to the norm $\|\cdot\|_X$, which is the norm of X restricted to X_0.

Suppose we have a sequence $\{x_m\}$ in X_0 with $x_m = \sum_{j=1}^n a_j^m e_j$ for $m = 1, 2, \ldots$ that is converging in $\|\cdot\|_X$ to some $x \in X$. Then, $\{x_m\}$ is a Cauchy sequence in $\|\cdot\|_X$, so it is also Cauchy in $\|\cdot\|_0$ as these two norms are equivalent in X_0. Thus, it means that, for each $j = 1, 2, \ldots, n$, $\{a_j^m\}_{m=1}^\infty$ is Cauchy, and it converges as $a_j^m \in \mathbb{F}$ and $\mathbb{F}$ is complete. So, $a_j^m \to a_j$ as $m \to \infty$ for some a_j.

We define $y := \sum_{j=1}^n a_j e_j \in X_0$. Notice that $\|x_m - y\|_0 \to 0$ as $m \to \infty$ by the construct of y. Since the two norms $\|\cdot\|_0$ and $\|\cdot\|_X$ are equivalent, we have $\|x_m - y\|_X \to 0$ as $m \to \infty$ as well, so $x = y \in X_0$. Thus, X_0 is closed, as desired. $\square$

Now we are ready to define what a compact operator is.

Definition 3.9. *Let X and Y be normed space. A linear operator $T : X \to Y$ is a **compact operator** if it maps bounded sets of X to relatively compact sets of Y. We denote the collection of all such operators by $Com(X, Y)$, and $Com(X) := Com(X, X)$. Clearly, since every relatively compact set is bounded, we have $Com(X, Y) \subseteq B(X, Y)$.*

Lemma 3.10. *An operator $T : X \to Y$ is compact if and only if $T(B_c(0, 1))$ is relatively compact in Y.*

Proof. The forward direction is trivially true by the definition of a compact operator. For the backward direction, we first notice that the scaling of values of a sequence does not affect its Cauchy-ness. So, if $T(B_c(0, 1))$ is relatively compact, $T(B_c(0, r))$ is relatively compact as well for any $r > 0$. Also, a translation will not affect the Cauchy-ness as well. Furthermore, we notice that every bounded set is the subset of a large enough closed ball, so it will be mapped to a relatively compact set after T, as desired. $\square$

Theorem 3.11. *Let X, Y, Z be normed spaces.*

1. *If $T_1, T_2 \in Com(X, Y)$, $\alpha, \beta \in \mathbb{F}$, then $\alpha T_1 + \beta T_2 \in Com(X, Y)$.*
2. *If A is bounded, T is compact, then both AT and TA are compact when the products make sense.*
3. *If $T_n \in Com(X, Y)$ for all n, and $\|T_n - T\| \to 0$ as $n \to \infty$, then $T \in Com(X, Y)$.*

Proof. (1) Consider a sequence $\{x_n\}$ with $x_n \in B_c(0, 1) \subset X$. Since T_1 is compact, there exists a subsequence $\{x_n^{(1)}\}$ of $\{x_n\}$ such that $\{T_1 x_n^{(1)}\}$ is Cauchy. Since T_2 is compact, there exists a (sub)subsequence $\{x_n^{(2)}\}$ of $\{x_n^{(1)}\}$ such that $\{T_2 x_n^{(2)}\}$ is Cauchy. This then means $\{\alpha T_1 x_n^{(2)} + \beta T_2 x_n^{(2)}\}$ is Cauchy, so $\alpha T_1 + \beta T_2 \in \mathrm{Com}(X, Y)$.

(2) We will use the closed ball version of the compact operator definition. $(TA)(B_c(0, 1)) = T(A(B_c(0, 1)))$. Since A is bounded, $A(B_c(0, 1))$ is bounded, so $T(A(B_c(0, 1)))$ is relatively compact. Thus, TA is compact. Next, $(AT)(B_c(0, 1)) = A(T(B_c(0, 1)))$. $T(B_c(0, 1))$ is relatively compact, and A is bounded so every Cauchy sequence will still be Cauchy after the map, so $A(T(B_c(0, 1)))$ is relatively compact. Thus, AT is compact as well.

(3) Let $\{x_n\}$ be a sequence with $x_n \in B_c(0, 1)$, i.e. $\|x_n\| \leq 1$. Since T_1 is compact, there exists subsequence $\{x_n^{(1)}\}$ of $\{x_n\}$ such that $\{T_1 x_n^{(1)}\}$ is Cauchy. Since T_2 is compact, there exists subsequence $\{x_n^{(2)}\}$ of $\{x_n^{(1)}\}$ such that $\{T_2 x_n^{(2)}\}$ is Cauchy. $\ldots$ Since T_m is compact, there exists subsequence $\{x_n^{(m)}\}$ of $\{x_n^{(m-1)}\}$ such that $\{T_m x_n^{(m)}\}$ is Cauchy.

Let $y_n := x_n^{(n)}$. Then, $\{y_n\}$ is a subsequence of $\{x_n\}$.

For each m, $\{T_m y_n\}_{n=1}^\infty$ is Cauchy. To see this, we notice that $\{y_n\}_{n=m}^\infty = \{x_n^{(n)}\}_{n=m}^\infty$ is a subsequence of $\{x_n^{(m)}\}_{n=1}^\infty$, so $\{T_m y_n\}_{n=m}^\infty$ is Cauchy and thus $\{T_m y_n\}_{n=1}^\infty$ is Cauchy as well.

We also know that $\{Ty_n\}_{n=1}^\infty$ is Cauchy. To see this, we first fix some $\varepsilon > 0$. Since $\|T - T_m\| \to 0$ as $m \to \infty$, there exists some $M > 0$ such that $\|T - T_M\| < \varepsilon/3$. Since $\{T_M y_n\}$ is Cauchy as shown above, there exists some $N > 0$ such that for all $n_1, n_2 > N$, we have $\|T_M y_{n_1} - T_M y_{n_2}\| < \varepsilon/3$. Then, we have, for $n_1, n_2 > N$, there is

$$\begin{aligned}\|Ty_{n_1} - Ty_{n_2}\| &\leq \|Ty_{n_1} - T_m y_{n_1}\| + \| - Ty_{n_2} + T_m y_{n_2}\| + \|T_m y_{n_1} - T_m y_{n_2}\| \\ &< \|T - T_m\| \cdot \|y_{n_1}\| + \|T - T_m\| \cdot \|y_{n_2}\| + \frac{\varepsilon}{3} \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.\end{aligned}$$

So $\{Ty_n\}_{n=1}^\infty$ is Cauchy.

Thus, we have shown that T is compact. □

Remark. Part 1 of the theorem indicates that Com(X) is a linear subspace of $B(X)$. Part 2 of the theorem indicates that Com(X) is an ideal in $B(X)$. Part 3 of the theorem indicates that Com(X) is closed. Thus, Com(X) is a closed ideal of $B(X)$. In fact, any closed ideal of $B(X)$ is a subset of Com(X).

Definition 3.12. *We say that $T \in B(X, Y)$ is a* ***finite dimensional****, or* ***finite rank****, operator if Ran(T) is finite-dimensional. Then, $T \in$ Com(X, Y), since $T(B_c(0, 1))$ is closed so it is also relatively compact as it is finite-dimensional.*

Let us consider an example.

Example. For $X = C([0, 1])$, and $k = k(s, t) \in C([0, 1] \times [0, 1])$. Define $T : X \to X$ by

$$(Tf)(t) := \int_0^1 k(s, t)f(s)\ ds.$$

This T is compact. To see this, we first notice that it is bounded, as $\|T\| \leq \sup_{s,t} \|k(s, t)\|$. Next, we construct

$$(T_n f)(t) := \int_0^1 k_n(s, t)f(s)\ ds.$$

If $T_n f$ is of finite rank which makes T_n compact as it is bounded, and $k_n \to k$ so $\|T_n - T\| \to 0$, then by part 3 of Theorem 3.11, T is compact.

We will set $k_n(s, t) := \sum_{j=0}^{N_1} \sum_{l=0}^{N_2} a_{jl} s^j t^l$ for some N_1, N_2, a_{jl} that depend on n. By Weierstrass approximation theorem, this will converge to $k(s, t)$. Using this k_n, we have

$$(T_n f)(t) := \int_0^1 \sum_{j=0}^{N_1} \sum_{l=0}^{N_2} a_{jl} s^j t^l f(s)\ ds = \sum_{l=0}^{N_2} t^l \left[\int_0^1 \sum_{j=0}^{N_1} a_{jl} s^j f(s)\ ds \right]$$

which is a polynomial of degree N_2, and it is in the span of $\{1, t, t^2, \ldots, t^{N_2}\}$. So, $T_n f$ is a finite-dimensional operator.

Thus, T is compact, as desired.

Lemma 3.13 (Almost Orthogonality Lemma)**.** *Let X be a normed space and $X_0 \subset X$ be a closed linear subspace of X with $X_0 \neq X$. Then, for each $\varepsilon > 0$, there exists $z \in X \backslash X_0$ such that $\|z\| = 1$ and $d(z, X_0) \geq 1 - \varepsilon$, i.e. $\|z - x\| \geq 1 - \varepsilon$ for all $x \in X_0$.*

Proof. The result is trivial for $\varepsilon \geq 1$ as the metric is non-negative. Consider $0 < \varepsilon < 1$. Since $X_0 \neq X$, there exists $x_1 \in X \backslash X_0$. Since X_0 is closed, we let $d := d(x_1, X_0) = \inf_{x \in X_0} \|x_1 - x\| > 0$. Notice that for the given ε, we have $d/(1-\varepsilon) > d$, so there exists some $y \in X_0$ such that $\|x_1 - y\| \leq d/(1-\varepsilon)$ (if this claim is not true then there would be a contradiction with d being the infimum). Let $z := (x_1 - y)/\|x_1 - y\|$, so $\|z\| = 1$. Let $x \in X_0$, then

$$\begin{aligned}
\|z - x\| &= \left\| \frac{x_1 - y}{\|x_1 - y\|} - x \right\| \\
&= \frac{1}{\|x_1 - y\|} \cdot \|x_1 - y - x \cdot \|x_1 - y\|\| \\
&\geq \frac{1-\varepsilon}{d} \|x_1 - (y + x \cdot \|x_1 - y\|)\| \\
&\geq \frac{1-\varepsilon}{d} \cdot d = 1 - \varepsilon
\end{aligned}$$

where the last inequality is due to the face that $y + x \cdot \|x_1 - y\|$ is in X_0 as both x and y are. □

Remark. If we are in Hilbert space, we would be able to have the concept of full orthogonality. Here, since we are only in Banach space, this approximate result is the best of what we can obtain, which is already very helpful.

Lemma 3.14. *Let X be infinite-dimensional. Then $B_C(0,1) \subset X$ is not relatively compact.*

Proof. Take $x_1 \in X$ such that $\|x_1\| = 1$, and we put $X_1 = \text{span}\{x_1\}$. Then, $\text{dim}X_1 = 1$ so $X_1 \neq X$. Also, X_1 is closed since it is a finite-dimensional span. By almost orthogonality lemma, there exists some $x_2 \in X \backslash X_1$ such that $\|x_2\| = 1$ and $d(x_2, X_1) \geq 1 - 1/2 = 1/2$ (take $\varepsilon = 1/2$), i.e. $d(x_2, x) \geq 1/2$ for all $x \in X_1$. In particular, we have $d(x_2, x_1) \geq 1/2$. Put $X_2 = \text{span}\{x_1, x_2\}$, and repeat this procedure to obtain x_3 and X_3, x_4 and X_4, etc. This gives us a sequence $\{x_n\}$ such that $x_n \in B_C(0,1)$ for all n yet $d(x_n, x_m) \geq 1/2$ for all $n \neq m$, so this sequence has no Cauchy subsequence, and thus $B_C(0,1)$ is not relatively compact. □

Corollary 3.15. *The identity operator in X is compact $\iff$ X is finite-dimensional.*

Theorem 3.16. *Let X be an infinite-dimensional Banach space and $T \in Com(X)$. T is not invertible.*

Proof. If T is invertible, then T^{-1} exists and T^{-1} is bounded by Banach inverse mapping theorem. $T \cdot T^{-1}$ is compact as T is compact and T^{-1} is bounded. However, since identity is not compact for infinite-dimensional X, we have a contradiction, so T is not invertible. □

Corollary 3.17. *Let X be an infinite-dimensional Banach space and $T \in Com(X)$. Then $0 \in \sigma(T)$.*

In the following, we will always let X be an infinite-dimensional Banach space and $T \in \text{Com}(X)$.

Theorem 3.18. *Let $T \in Com(X)$ and $\lambda \neq 0$ is an eigenvalue of T. Then, the geometric multiplicity of λ is finite, i.e. $\text{dim}X_\lambda < \infty$ where $X_\lambda := \{x \in X | Tx = \lambda x\}$.*

Proof. $T|_{X_\lambda} = \lambda I|_{X_\lambda}$ is compact as it is the restriction of a compact operator to a closed subspace. So $\dim X_\lambda$ is finite as it would not be compact if it is infinite-dimensional, since infinite-dimensional identity is not compact. $\square$

Lemma 3.19. *Let $T \in Com(X)$, $\lambda \neq 0$. If λ is not an eigenvalue, then there exists $c > 0$ such that $\|(T - \lambda I)x\| \geq c\|x\|$ for all $x \in X$.*

Proof. Suppose not. Then, for $c = 1/k$, there exists $x_k \in X$ such that $\|(T - \lambda I)x_k\| < \|x_k\|/k$. Put $z_k = x_k/\|x_k\|$, then $\|z_k\| = 1$ and $\|(T - \lambda I)z_k\| < 1/k$, so $\lim_{k\to\infty} |(T - \lambda I)z_k\| = 0$. Since T is compact, there is a subsequence $\{z_{k_j}\}$ of $\{z_k\}$ such that Tz_{k_j} is Cauchy, so it converges to $z := \lim Tz_{k_j}$ as X is Banach. Then,

$$z_{k_j} = \frac{1}{\lambda}[Tz_{k_j} - (T - \lambda I)z_{k_j}] \to \frac{z}{\lambda} =: z'.$$

Then, $\|z'\| = 1$. Also,

$$\|(T - \lambda I)z'\| = \|(T - \lambda I)\lim z_{k_j}\| = \|\lim(T - \lambda I)z_{k_j}\| = 0,$$

so λ is an eigenvalue with eigenvector z'. We have a contradiction. $\square$

Theorem 3.20. *Let X be a Banach space and $A \in B(X)$ such that there exists $c > 0$ with $\|Ax\| \geq c\|x\|$ for all $x \in X$. Then Ker$(A) = \{0\}$ and Ran(A) is closed.*

Proof. If $\mathrm{Ker}(A) \neq \{0\}$, then there exists some $x \neq 0$ and $x \in \mathrm{Ker}(A)$ such that $\|Ax\| = 0$ yet we also have $\|Ax\| \geq c\|x\| > 0$. Contradiction. So $\mathrm{Ker}(A) = \{0\}$.

Suppose $y \in \overline{\mathrm{Ran}(A)}$, i.e. there exists a sequence $\{y_k\}$ in $\mathrm{Ran}(A)$ with $y_k \to y$ and $y \in \overline{\mathrm{Ran}(A)}$. So, $y_k \in \mathrm{Ran}(A)$ implies there exists $x_k \in X$ such that $y_k = Ax_k$. As $\|Ax\| \geq c\|x\|$, we have

$$\|x_n - x_m\| \leq \|A(x_n - x_m)\|/c = \|y_n - y_m\|/c \to 0$$

as $n, m \to \infty$, so $\{x_n\}$ is Cauchy. So, $Ax = A(\lim x_n) = \lim(Ax_n) = \lim y_n = y$, so $y \in \mathrm{Ran}(A)$, and $\mathrm{Ran}(A)$ is thus closed. $\square$

Corollary 3.21. *Let X be a Banach space and $A \in B(X)$ such that there exists $c > 0$ with $\|Ax\| \geq c\|x\|$ for all $x \in X$. Then, for all $n \in \mathbb{N}$, Ker$(A^n) = \{0\}$ and RanA^n is closed.*

Proof. $\|A^n x\| \geq c\|A^{n-1}x\| \geq c^n\|x\|$ for all $x \in X$. c^n is just some constant, so we would obtain the desired result by applying the previous theorem. $\square$

Theorem 3.22. *Let X be a Banach space, $T \in Com(X)$, $\lambda \neq 0$, and λ is not an eigenvalue. Then $(T - \lambda I)$ is invertible, and $\lambda \notin \sigma(T)$.*

Proof. Put $X_0 := X$, and $X_n := \mathrm{Ran}((T - \lambda I)^n)$ for $n \in \mathbb{N}$. X_n is closed for all n from lemma and corollary above. Also,

$$\begin{aligned} X_{n+1} = (T - \lambda I)^{n+1}X &= (T - \lambda I)[(T - \lambda I)^n x] = (T - \lambda I)X_n \\ &= (T - \lambda I)^n[(T - \lambda I)x] = (T - \lambda I)^n X_1. \end{aligned}$$

Thus, $X_{n+1} \subseteq X_n$.

The desired results will follow trivially if we have the following two claims: (1) there exists n such that $X_{n+1} = X_n$, (2) the minimum $k \in \mathbb{N}$ such that $X_k = X_{k+1}$ is 0. With these two results, we would have $X_0 = X_1 = \mathrm{Ran}(T - \lambda I)$ and $(T - \lambda I)$ has a trivial kernel, so it is invertible and $\lambda \notin \sigma(T)$. Let us prove these two claims one by one.

(1) Suppose not, i.e. $X_{n+1} \neq X_n$ for all possible n. Then, by almost orthogonality lemma, there exist some $x_n \in X_n \backslash X_{n+1}$ such that $\|x_n\| = 1$ and $\|x_n - z\| \geq 1/2$ for all $z \in X_{n+1}$. If $m > n$, then

$$\begin{aligned}
\|Tx_n - Tx_m\| &= \|(T - \lambda I)(x_m - x_n) + \lambda I(x_m - x_n)\| \\
&= |\lambda| \cdot \left\| \frac{T - \lambda I}{\lambda}(x_m - x_n) + x_m - x_n \right\| \\
&= |\lambda| \cdot \|(\text{something in } X_{n+1}) - x_n\| \\
&\geq |\lambda|/2.
\end{aligned}$$

Notice that the last equality is because the first term in the norm is $T - \lambda I$ applied to some element of X_n, so it is in X_{n+1}, and x_m is in X_{n+1} as $m > n$. So, $\{Tx_n\}$ has no Cauchy subsequence, so T is not compact which is a contradiction. Thus, this claim is shown.

As a consequence of this claim, we should note that if $X_{n+1} = X_n$ for some n, then $X_n = X_{n+1} = X_{n+2} = \cdots$, as $X_{n+2} = (T - \lambda I)X_{n+1} = (T - \lambda I)X_n = X_{n+1} = X_n$.

(2) Suppose not. Then, X_{k-1} exists, and we have $X_{k-1} \neq X_k = X_{k+1}$. Take some $x \in X_{k-1} \backslash X_k$. Then $(T - \lambda I)x \in X_k = X_{k+1}$, so there exists $y \in X_k$ such that $(T - \lambda I)x = (T - \lambda I)y \implies (T - \lambda I)(x - y) = 0$. Since $x \notin X_k$ yet $y \in X_k$, we have $x - y \neq 0$, so λ is an eigenvalue, which is a contradiction. $\square$

Corollary 3.23. *Every non-zero point of $\sigma(T)$ with $T \in Com(X)$ is an eigenvalue of T.*

Theorem 3.24. *Let X be a Banach space and $T \in Com(X)$. Then, $\sigma(T)$ is at most countable, and the only possible accumulation point is the origin.*

Proof. We will prove that $\sigma(T) \cap \{\lambda \in \mathbb{C} \mid |\lambda| > \delta\}$ is finite for any $\delta > 0$. This is equivalent to the desired result.

Suppose not. Fix some δ, so there exists countably many distinct $\lambda_1, \lambda_2, \ldots \in \sigma(T)$ such that $|\lambda_j| > \delta$. Since $\lambda_j \neq 0$, by Theorem 3.22, we know that $\lambda_j \in \sigma_p(T)$, so there exists some corresponding eigenvector x_j such that $Tx_j = \lambda_j x_j$.

Put $X_n = \mathrm{span}\{x_1, x_2, \ldots, x_n\}$ and $\dim X_n = n$. So, we have $X_1 \subset X_2 \subset \cdots$, and they are proper subsets. Notice that by our definition, we have $TX_n \subseteq X_n$ and $(T - \lambda_n I)X_n \subseteq X_{n-1}$. In fact, we have equalities for these two statements.

By almost orthogonality lemma, there exists some $y_n \in X_n$ such that $\|y_n\| = 1$ and $\|y_n - x\| \geq 1/2$ for all $x \in X_{n-1}$. Let $n > m$, and we have

$$\begin{aligned}
\|Ty_n - Ty_m\| &= \|\lambda_n y_n + (T - \lambda_n)y_n - Ty_m\| \\
&= \left\| \lambda_n \left[y_n - \frac{1}{\lambda_n}[-(T - \lambda_n)y_n + Ty_m] \right] \right\| \\
&\geq \frac{1}{2}\|\lambda_n\| > \frac{1}{2}\delta.
\end{aligned}$$

The first inequality above uses almost orthogonal lemma, as $(T-\lambda_n)y_n \in X_{n-1}$ and $Ty_m \in X_m \subset X_{n-1}$. So, $\{Ty_n\}$ has no Cauchy subsequence and $y_n \in B_C(0,1)$, so $T \notin \text{Com}(X)$. A contradiction. $\square$

Lemma 3.25. *Let $T \in Com(X,Y)$, then $T^* \in Com(Y^*, X^*)$.*

Chapter 4

Spectrum Theory in Hilbert Spaces

4.1 Hilbert Space Basics

We will cover basic definitions and results of Hilbert space theory that we will use later without proof. We will usually use H to denote some Hilbert space.

Definition 4.1. *A system $\{x_\alpha\}_{\alpha \in J}$ is called **orthogonal** if $x_\alpha \perp x_\beta$ for $\alpha \neq \beta$, where $x_\alpha \perp x_\beta \iff (x_\alpha, x_\beta) = 0$. It is **orthonormal** if $\|x_\alpha\| = (x_\alpha, x_\alpha) = 1$ for all $\alpha \in J$.*

Theorem 4.2 (Pythagoras Theorem)**.** *If $\{x_j\}_{j=1}^n$ is orthogonal, then*

$$\left\| \sum_{j=1}^{n} x_j \right\|^2 = \sum_{j=1}^{n} \|x_j\|^2 .$$

Theorem 4.3 (Polarisation Identity)**.** *For any $x, y \in H$ with Hilbert space H, we have*

$$\begin{aligned} 4(x,y) &= \|x+y\|^2 - \|x-y\|^2 && \text{if } \mathbb{F} = \mathbb{R}, \\ 4(x,y) &= \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2 && \text{if } \mathbb{F} = \mathbb{C}. \end{aligned}$$

Theorem 4.4 (Parallelogram Law)**.** *For any $x, y \in H$ with Hilbert space H, we have*

$$2(\|x\|^2 + \|y\|^2) = \|x+y\|^2 + \|x-y\|^2.$$

Theorem 4.5 (Jordon-von Neumann Theorem)**.** *If a Banach space satisfies the parallelogram law, then it is a Hilbert space and the inner product is defined via polarisation identity.*

Theorem 4.6. *Let L be a closed linear subspace of Hilbert space H, and $x \in H$. Then, there exists unique $y \in L$ such that $\|x - y\| = \inf_{z \in L} \|x - z\| = d(x, L)$ and $(x - y, z) = 0$ for all $z \in L$.*

Definition 4.7. *Let $M \subseteq H$ be a set. The **orthogonal complement** to M is the set $M^\perp := \{x \in H \mid (x, y) = 0 \ \forall y \in M\}$.*

Theorem 4.8. *For a Hilbert space H,*

1. *$M^{\perp}$ is a closed linear subspace of H.*
2. *$M_1 \subseteq M_2 \implies M_2^{\perp} \subseteq M_1^{\perp}$.*
3. *$M \subseteq M^{\perp\perp}$. Moreover, $M^{\perp\perp} = \overline{span\{M\}}$.*
4. *Let M be a linear subspace of H, then $M^{\perp} = \{0\} \iff M$ is dense in H.*

Theorem 4.9. *If M is a closed linear subspace of H, then $H = M \dot{+} M^{\perp}$.*

Remark. Notice that we will use $\dot{+}$ to denote the direct sum and $\bigoplus$ to denote the orthogonal sum.

Let $\{e_\alpha\}_{\alpha \in J}$ be an orthonormal set. For $x \in H$, we call (x, e_α) a **Fourier coefficient** of x.

Theorem 4.10 (Bessell Inequality)**.**

$$\|x\|^2 \geq \sum_{\alpha \in J} (x, e_\alpha)^2.$$

Theorem 4.11. *The following are equivalent:*

1. *(Parseval Identity) For all $x \in H$, $\|x\|^2 = \sum (x, e_\alpha)^2$.*
2. *(Fourier Expansion) For all $x \in H$, $\|x\|^2 = \sum (x, e_\alpha) e_\alpha$.*
3. *$x = 0 \iff (x, e_\alpha) = 0$ for all $\alpha \in J$.*
4. *$span(\{e_\alpha\})$ is dense in H.*

*Then, we say the orthonormal system $\{e_\alpha\}_{\alpha \in J}$ is **complete** in H.*

Theorem 4.12 (Riesz Representation Theorem)**.** *Let $f \in H^*$ for some Hilbert space H. Then, there exists a unique $z \in H$ such that $f(x) = (x, z)$ for all $x \in H$, and $\|f\|_{H^*} = \|z\|$.*

Definition 4.13. *Let $A \in B(H)$. Then there exists unique operator $A^* \in B(H)$ such that $(Ax, y) = (x, A^* y)$ for all $x, y \in H$. We will call A^* the **adjoint** of A.*

Now, we will state some properties of adjoint.

Theorem 4.14. *Let $A, A_1, A_2 \in B(H)$ and A^*, A_1^*, A_2^* be adjoints of A, A_1, A_2 respectively. We have*

1. *$(\alpha A_1 + \beta A_2)^* = \bar{\alpha} A_1^* + \bar{\beta} A_2^*$ for all $\alpha, \beta \in \mathbb{F}$.*
2. *$(AB)^* = B^* A^*$.*
3. *$(A^*)^* = A$.*
4. *$\|A^*\| = \|A\|$.*
5. *$\|A^* A\| = \|A\|^2$.*
6. *If A^{-1} exists, then $(A^*)^{-1}$ also exists, and $(A^*)^{-1} = (A^{-1})^*$.*

Theorem 4.15. *Consider $A \in B(H)$. Then, $Ker(A^*) = Ran(A)^{\perp}$ and $Ker(A) = Ran(A^*)^{\perp}$.*

Corollary 4.16. *Consider $A \in B(H)$. Then, $Ker(A^*)^{\perp} = \overline{Ran(A)}$ and $Ker(A)^{\perp} = \overline{Ran(A^*)}$.*

Definition 4.17. *Consider $A \in B(H)$.*

1. *A is **normal** if $A^* A = A A^*$.*
2. *A is **symmetric** or **self-adjoint** if $A^* = A$, i.e. $(Ax, y) = (x, Ay)$ for all $x, y \in H$.*
3. *A is **unitary** if $AA^* = A^* A = I$, i.e. $A^* = A^{-1}$.*

Remark. It is the same for a bounded operator to be symmetric and to be self-adjoint, yet it might not be the case for unbounded operators.

From the definition, we have the following simple result.

Proposition 4.18. *A self-adjoint operator is normal, and an unitary operator is also normal.*

4.2 Unbounded Operators and Their Adjoints*

Here, we will be looking at unbounded operators and their adjoints. They will be discussed informally, and they are not examinable. The only reason for bringing them up is that a lot of interesting examples of operators and their adjoints arise only for unbounded operators, such as the differential operators. It would be a shame and slightly boring to not mention them when we study spectral theory.

Consider an unbounded operator $A : D_A \to H$ where H is a Hilbert space and D_A is a domain for A. We will use $f_{A,y}(x)$ to denote the inner product (Ax, y). Suppose that for a given y, we can represent $f_{A,y}(x) = (x, h)$ for all $x \in D_A$ for some h, then we say that $y \in D_{A^*}$ and $A^*y = h$. This provides us with a definition for the adjoint of an unbounded operator.

The first question we should ask is, is h unique, or is this well-defined? It is. To see this, suppose there exists h_1, h_2 such that

$$(Ax, y) = f_{A,y}(x) = (x, h_1) = (x, h_2)$$

for all $x \in D_A$. This means we have

$$(x, h_1) - (x, h_2) = (x, h_1 - h_2) = 0$$

for all $x \in D_A$, which means $h_1 - h_2 \in D_A^\perp$. Domain D_A is dense in H, so its orthogonal complement $D_A^\perp = \{0\}$, and therefore $h_1 - h_2 = 0 \implies h_1 = h_2$.

Next question we may ask is what is the relationship between D_A and D_{A^*}. This is very case-by-case, and we will study these two domains for many examples in the following.

We have mentioned three types of bounded operators in Hilbert spaces - normal, symmetric / self-adjoint, and unitary. For unbounded operators, we will only be looking at symmetric and self-adjoint. These two terms mean the same thing in the bounded setting, but different things in the unbounded setting.

Definition 4.19. *Consider an unbounded operator $A : D_A \to H$.*

1. *A is **symmetric** if $(Ax, y) = (x, Ay)$ for all $x, y \in D_A$.*
2. *A is **self-adjoint** if $A^* = A$, and in particular $D_{A^*} = D_A$.*

Remark. It can be observed that A is self-adjoint implies that A is symmetric. However, if A is symmetric, it implies that $D_{A^*} \supseteq D_A$ and $A^*|_{D_A} = A|_{D_A}$.

Definition 4.20. *For an operator A on a Hilbert space, the **quadratic form** of A is $Q_A : H \to \mathbb{F}$ where $Q_A(x) = (Ax, x)$.*

Next, we will study (a lot of) examples of unbounded operators.

1) Consider $H = L^2(\mathbb{R})$ and $Af := f'$ with domain $D_A = \{f \in L^2, f' \in L^2\} = H^1(\mathbb{R}) = W^{1,2}(\mathbb{R})$. Here, H^1 and $W^{1,2}$ are two notations for Sobolev spaces. The superscript of H is to denote that it contains those with first derivatives, whereas the superscripts for W denote it contains the first derivatives for the first number, and it is a L^2 space for the second number.

We claim that $f \in H^1(\mathbb{R}) \implies \lim_{x\to\pm\infty} f(x) = 0$.

Using integration by parts, we have

$$(Af, g) = \int_{-\infty}^{\infty} f'\overline{g} = -\int_{-\infty}^{\infty} f\overline{g'} + [fg]_{-\infty}^{\infty} = -\int_{-\infty}^{\infty} f\overline{g'} = -(f, Ag).$$

So, $D_{A^*} \supseteq D_A$, and $A^*|_{D_A} = A|_{D_A}$. In fact, $D_{A^*} = D_A$. This, however, is slightly awkward as $A^* = -A$, so it is not exactly self-adjoint.

2) Consider $H = L^2(\mathbb{R})$ and $Af := if'$ with domain $D_A = H^1(\mathbb{R})$. We have

$$(Af, g) = \int_{-\infty}^{\infty} if'\overline{g} = -i\int_{-\infty}^{\infty} f\overline{g'} = \int_{-\infty}^{\infty} f\overline{ig'} = (f, Ag).$$

So, $A^* = A$.

3) Consider $H = L^2([0,1])$ and $Af := if'$ with domain $D_A = H^1([0,1])$. We have

$$(Af, g) = \int_0^1 if'\overline{g} = \int_0^1 f\overline{ig'} + [if\overline{g}]_0^1 = (f, Ag) + if(1)\overline{g}(1) - if(0)\overline{g}(0).$$

We would want the boundary terms to cancel out, so we have the following domain for adjoint, $D_{A^*} = \{g \in H^1([0,1]), g(0) = g(1) = 0\}$.

4) Consider $H = L^2([0,1])$ and $Af := if'$ with domain $D_A = \{f \in H^1([0,1]), f(0) = f(1) = 0\}$. Then, $D_{A^*} = H^1([0,1])$.

5) Consider $H = L^2([0,1])$ and $Af := if'$ with domain $D_A = \{f \in H^1([0,1]), f(0) = 0\}$. Then, $D_{A^*} = \{g \in H^1([0,1]), g(1) = 0\} \neq D_A$.

6) Consider $H = L^2([0,1])$ and $Af := if'$ with domain $D_A = \{f \in H^1([0,1]), f(0) = f(1)\}$. Then, $D_{A^*} = \{g \in H^1([0,1]), g(0) = g(1)\} = D_A$. So, this A with domain D_A is self-adjoint.

7) Consider $H = L^2([0,1])$ and $Af := if'$ with domain $D_A = \{f \in H^1([0,1]), f(0) = e^{i\theta}f(1)\}$. Then, $D_{A^*} = \{g \in H^1([0,1]), g(0) = e^{i\theta}g(1)\} = D_A$. Here, $\theta \in \mathbb{R}$ is a constant. So, again, this A with domain D_A is self-adjoint.

8) Consider $H = L^2([0,\infty))$ and $Af := if'$ with domain $D_A = H^1([0,\infty))$. We have

$$(Af, g) = \int_0^{\infty} if'\overline{g} = \int_0^{\infty} f\overline{ig'} + [if\overline{g}]^{\infty} = (f, Ag) - if(0)\overline{g}(0).$$

So, $D_{A^*} = \{g \in H^1, g(0) = 0\}$. And there is no way we can adjust the domain to make sure that the operator is self-adjoint.

9) Consider $H = L^2(\mathbb{R})$ and $Af := -f''$ with domain $D_A = \{f \in L^2, f'' \in L^2, f' \in L^2\} = H^2(\mathbb{R})$. We have

$$(Af, g) = \int_{\mathbb{R}} -f''\bar{g} = \int_{\mathbb{R}} f'\overline{g'} = \int_{\mathbb{R}} f\overline{-g''} = (f, Ag)$$

So A with this domain is self-adjoint. Notice that it will still be self-adjoint if we have $Af := f''$ with the same domain instead. The extra minus sign is to make sure that its quadratic form is non-negative.

10) Consider $H = L^2([0,1])$ and $Af := -f''$ with domain $D_A = \{f \in L^2, f'' \in L^2, f' \in L^2\} = H^2(\mathbb{R})$. We have

$$\begin{aligned}(Af, g) &= \int_0^1 -f''\bar{g} = \int_0^1 f'\bar{g}' + [-f'\bar{g}]_0^1 \\ &= \int_0^1 f\overline{-g''} + [-f'\bar{g}]_0^1 + [f\overline{g'}]_0^1 \\ &= (f, Ag) - f'(1)\bar{g}(1) + f'(0)\bar{g}(0) + f(1)\overline{g'}(1) - f(0)\overline{g'}(0)\end{aligned}$$

So, $D_{A^*} = \{g \in H^2, g(0) = g(1) = g'(0) = g'(1) = 0\}$.

Clearly, there are many ways we can adjust the domains to make the operator self-adjoint.

- (Periodic) $f(0) = f(1), f'(0) = f'(1)$.
- (Quasi-Periodic) $f(0) = e^{0\theta} f(1), f'(0) = e^{0\theta} f'(1)$ for some fixed $\theta \in \mathbb{R}$.
- (Robin) $f(0) = \lambda f'(0), f(1) = \mu f'(1)$ for some constant $\lambda, \mu \in \mathbb{R}$.
- (Dirichlet) $f(0) = f(1) = 0$.
- (Neumann) $f'(0) = f'(1) = 0$.

11) Consider $H = L^2(\mathbb{R}^d)$ and $Af := -\Delta f$, where Δ is the Laplacian, with domain $D_A = \{f \in L^2, \partial_\alpha f \in L^2, \partial_\alpha\partial_\beta f \in L^2 \forall \alpha, \beta\} = H^2(\mathbb{R}^d)$. We have

$$(Af, g) = \int_{\mathbb{R}^d} -\Delta f\bar{g} = \int_{\mathbb{R}^d} \nabla f \cdot \overline{\nabla g} = -\int_{\mathbb{R}^d} f\overline{\Delta g},$$

so $D_A = D_{A^*}$, $A^* = A$, and A is self-adjoint.

12) Consider some region $\Omega \subset \mathbb{R}^d$ with smooth boundary $\Gamma = \partial\Omega$. $H = L^2(\Omega)$ and $Af := -\Delta f$. There, $D_A = H^2(\Omega)$. We have, using Green's theorem,

$$(Af, g) = \int_\Omega -\Delta f\bar{g} = \int_\Omega \nabla f \cdot \overline{\nabla g} - \int_\Gamma \partial_n f \cdot \bar{g} = -\int_\Omega f\overline{\Delta g} - \int_\Gamma \partial_n f \cdot \bar{g} + \int_\Gamma f \cdot \partial_n \bar{g}.$$

where ∂n is the derivative in the direction of a normal n of Ω. So, $D_{A^*} = \{g \in H^2(\Omega), g|_\Gamma = 0, \partial_n g|_\Gamma = 0\}$. Same as before, we can make adjustments to the domain to make the operator self-adjoint.

- (Dirichlet) $g|_\Gamma = 0$.
- (Neumann) $\partial_n g|_\Gamma = 0$.
- (Robin) $\partial_n g|_\Gamma = \lambda g|_\Gamma$ for some continuous function $\lambda : \Gamma \to \mathbb{R}$.

We have studied three types of operators in Hilbert spaces - normal operators, unitary operators, and self-adjoint operators. We will extend our study of these types of operators here. But first, let us recap the definitions of these three kinds of operators.

Definition 4.21. *Consider $A \in B(H)$.*

1. *A is **normal** if $A^*A = AA^*$.*
2. *A is **symmetric** or **self-adjoint** if $A^* = A$, i.e. $(Ax, y) = (x, Ay)$ for all $x, y \in H$.*
3. *A is **unitary** if $AA^* = A^*A = I$, i.e. $A^* = A^{-1}$.*

4.3 Normal Operators

Theorem 4.22. *Consider $A \in B(H)$ where H is a Hilbert space. A is normal if and only if $\|Ax\| = \|A^*x\|$ for all $x \in H$.*

Proof. The forward direction is simple. Since A is normal, we have $\|Ax\|^2 = (Ax, Ax) = (A^*Ax, x) = (AA^*x, x) = (A^*x, A^*x) = \|A^*x\|^2$, as desired.

The backward direction is slightly more involved, but simple nonetheless. Suppose $\|Ax\| = \|A^*x\|$ for all $x \in H$, i.e. $(Ax, Ax) = (A^*x, A^*x)$ for all x. Then, we will use the polarisation identity of Hilbert space. If $\mathbb{F} = \mathbb{R}$, we have $4(Ax, Ay) = (Ax + Ay, Ax + Ay) - (Ax - Ay, Ax - Ay) = (A^*x + A^*y, A^*x + A^*y) - (A^*x - A^*y, A^*x - A^*y) = 4(A^*x, A^*y)$, therefore $(Ax, Ay) = (A^*x, A^*y)$ for all $x, y \in H$. If $\mathbb{F} = \mathbb{C}$, it would be a similar story and we would get $(Ax, Ay) = (A^*x, A^*y)$ for all $x, y \in H$. So, $(AA^*x, y) = (A^*Ax, y)$ and we have $((AA^* - A^*A)x, y) = 0$ for all $x, y \in H$. If we take $y = (AA^* - A^*A)x$ and $x \neq 0$, we would get $AA^* - A^*A = 0$ so $A^*A = AA^*$, as desired. $\square$

Theorem 4.23. *Let A be a normal operator. Then, we have*

1. *$(RanA^*)^{\perp} = KerA = KerA^* = (RanA)^{\perp}$*
2. *$Ax = \lambda x \implies A^*x = \bar{\lambda}x$*
3. *Given $Ax = \lambda_1 x, Ay = \lambda_2 y$. If $\lambda_1 \neq \lambda_2$, then $(x, y) = 0$*

Proof. (i) According to the previous theorem, we have $\|Ax\| = 0$ for some $x \in \text{Ker}A$, and $\|Ax\| = \|A^*x\| = 0 \implies x \in \text{Ker}A^*$. The other equalities follow from the definition.

(ii) $Ax = \lambda x \iff x \in \text{Ker}(A - \lambda I) = \text{Ker}[(A - \lambda I)^*] = \text{Ker}(A^* - \bar{\lambda} I) \iff A^*x = \bar{\lambda}x$.

(iii) $\lambda_1(x, y) = (\lambda_1 x, y) = (Ax, y) = (x, A^*y) = (x, \overline{\lambda_2} y) = \lambda_2(x, y)$. Since $\lambda_1 \neq \lambda_2$, we have $(x, y) = 0$, as desired. $\square$

4.4 Unitary Operators

Theorem 4.24. *Consider $U \in B(H)$, then the following are equivalent:*

1. *U is unitary*
2. *$(Ux, Uy) = (x, y)$ for all x, y and $RanU = H$*
3. *$\|Ux\| = \|x\|$ for all x and $RanU = H$*

Proof. (1) $\implies$ (2) U is unitary so $UU^* = I$ and this means $\text{Ran}U = H$. Also, $(Ux, Uy) = (U^*Ux, y) = (x, y)$.

(2) $\implies$ (3) It is obvious if we let $y = x$.

(3) $\implies$ (2) This follows from the polarisation identity.

(2) $\implies$ (1) $(Ux, Uy) = (x, y) = (U^*Ux, y)$. So, we have $((U^*U - I)x, y) = 0$ which means $U^*U - I = 0$ by letting $y = (U^*U - I)x$ and $x \neq 0$. So, $U^*U = I$ and U^* is the left inverse of U. In addition, since (2) implies (3), we have $\text{Ker}U = \{0\}$ as $\|Ux\| = \|x\|$. So, since $\text{Ran}U = H$, U is invertible, and $U^* = U^{-1}$. □

Theorem 4.25. *If $0 \notin \sigma(A)$, then $\sigma(A^{-1}) = 1/\sigma(A) = \{1/\lambda | \lambda \in \sigma(A)\}$.*

Proof. $0 \neq \lambda \notin \sigma(A) \iff (A - \lambda I)^{-1}$ exists $\iff A^{-1} - \lambda^{-1}I = A^{-1}\lambda^{-1}(\lambda I - A)$ is invertible $\iff \lambda^{-1} \notin \sigma(A^{-1})$. □

Theorem 4.26. *Let U be unitary. Then, $\sigma(U) \subseteq \{\lambda \in \mathbb{C} |\ |\lambda| = 1\}$.*

Proof. $\|U\| = 1$, so $\sigma(U) \subseteq B_C(0,1)$. Additionally, $\|U^{-1}\| = 1$, so $\sigma(U^{-1}) \subseteq B_C(0,1)$. Using the previous theorem, we have $\sigma(U) = 1/\sigma(U^{-1}) = \{\lambda |\ |\lambda| \geq 1\}$. So, taking the intersection of these two, we have $\sigma(U) \subseteq \{\lambda \in \mathbb{C} |\ |\lambda| = 1\}$. □

Definition 4.27. *Let $A, B \in B(H)$. They are **unitary equivalent** if there exists unitary U such that $A = UBU^{-1}$. They are **similar** if there exists invertible S with $S, S^{-1} \in B(H)$ such that $A = SBS^{-1}$.*

4.5 Self-Adjoint Operators

Theorem 4.28. *We have*

1. *A is self-adjoint, then $\lambda \in \mathbb{R} \implies \lambda A$ is self-adjoint.*
2. *A, B are self-adjoint implies $A + B$ is self-adjoint.*
3. *A, B are self-adjoint implies AB is self-adjoint if and only if $AB = BA$.*
4. *$\{A_n\}$ are self-adjoint. $\|A_n - A\| \to 0$ implies A is self-adjoint.*

Theorem 4.29. *Let $\mathbb{F} = \mathbb{C}$ and $A \in B(H)$. Then A is self-adjoint is equivalent to $Q_A(x) = (Ax, x) \in \mathbb{R}$ for all $x \in H$.*

Proof. The forward direction is simple. $(Ax, x) = (x, A^*x) = \overline{(A^*x, x)} = \overline{(Ax, x)}$, so $Ax, x) \in \mathbb{R}$.

For the backward direction, we have $(Ax, x) = \overline{(Ax, x)} = (x, Ax) = (A^*x, x)$ for all x as $Q_A(x) \in \mathbb{R}$. This would help us to derive $(Ax, y) = (A^*x, y)$ using the polarisation identity. The derivation is slightly technical, so we will omit it here. Thus, $A = A^*$. □

Corollary 4.30. *If A is self-adjoint, $\sigma_p(A) \subseteq \mathbb{R}$.*

Proof. $Ax = \lambda x \implies (Ax, x) = \lambda \|x\|^2 \in \mathbb{R}$ by the previous result. Next, as $\|x\|^2$ is real, $\lambda \in \mathbb{R}$. □

Theorem 4.31. *Let $P \in B(H)$ and $P^2 = P$. The following are equivalent:*

1. *P is self-adjoint.*
2. *P is normal.*

3. $RanP = (KerP)^{\perp}$.
4. $(Px, x) = \|Px\|^2$ for all $x \in H$.

Then, we call P an ***orthogonal projection****.*

Proof. (1) $\implies$ (2) This is obvious.

(2) $\implies$ (3) $\text{Ker}P = \text{Ker}P^* = (\text{Ran}P)^{\perp}$ as P is normal. Thus, $(\text{Ker}P)^{\perp} = (\text{Ran}P)^{\perp\perp} = \overline{\text{Ran}P} = \text{Ran}P$ as P is a projection so its range is closed.

(3) $\implies$ (1) $\text{Ran}P = (\text{Ker}P)^{\perp}$, so $(Px, (I-P)y) = 0$ for all $x, y \in H$. Therefore, we have

$$(Px, y) = (Px, y) - (Px, (I-P)y) = (Px, Py) = (Px, Py) + ((I-P)x, Py) = (x, Py).$$

So, $P = P^*$.

(3) $\implies$ (4) From above, we have obtained $(Px, y) = (Px, Py)$. Let $y = x$, we would have $(Px, x) = (Px, Px) = \|Px\|^2$, as desired.

(4) $\implies$ (3) Consider $\mathbb{F} = \mathbb{C}$. By the previous result, we know that $Q_P(x) \in \mathbb{R} \iff P$ is self-adjoint. So, as $(Px, x) = \|Px\|^2 \in \mathbb{R}$, P is self-adjoint, and (1) implies (3). Consider $\mathbb{F} = \mathbb{R}$, let $x \in \text{Ker}P$ and $y \in \text{Ran}P$, set $z = x + y$. Now, $Pz = Py = y$, so

$$\|Pz\|^2 = (Pz, z) \iff \|y\|^2 = (y, x) + \|y\|^2 \iff (y, x) = 0 \iff \text{Ran}P = (\text{Ker}P)^{\perp}.$$

$\square$

4.6 Numerical Range

Definition 4.32. *Let $A \in B(H)$, then the* ***numerical range*** *of A, denoted by $Num(A)$, is*

$$Num(A) = \{Q_A(x) = (Ax, x) \mid \|x\| = 1\} = \left\{ \frac{(Ax, x)}{\|x\|^2} \;\middle|\; x \neq 0 \right\}.$$

Since $|(Ax, x)| \le \|Ax\| \cdot \|x\| = \|Ax\| \le \|A\|$, $Num(A) \subseteq B_C(0, \|A\|)$.

Theorem 4.33. *$Num(A)$ is convex.*

Theorem 4.34. *$\sigma(A) \subseteq \overline{Num(A)}$.*

We will establish an auxiliary lemma first.

Lemma 4.35. *Let $A \in B(H)$. Suppose there exists $c > 0$ such that $|(Ax, x)| \ge c\|x\|^2$ for all x, then A^{-1} exists.*

Proof. Using Cauchy-Schwartz and the condition, we have

$$\|Ax\|\|x\| \ge |(Ax, x)| \ge c\|x\|^2 \implies \|Ax\| \ge c\|x\| \implies \text{Ker}A = \{0\} \& \text{ Ran}A \text{ is closed.}$$

Suppose $x \in (\text{Ran}A)^{\perp}$, then $(Ax, x) = 0$ so $x = 0$. Thus, $(\text{Ran}A)^{\perp} = \{0\} \implies \text{Ran}A$ is dense in $H \implies \overline{\text{Ran}A} = H$. But $\text{Ran}A$ is closed, so $\text{Ran}A = H$. Thus, A is invertible. $\square$

Now we prove the theorem.

Proof. Suppose $\lambda \notin \overline{Num(A)}$. Take $z \in H$ and $\|z\| = 1$, then $|((A - \lambda I)z, z)| = |(Az, z) - \lambda(z, z)| = |(Az, z) - \lambda| \geq c = c\|z\|^2$ for some constant $c > 0$, since (Az, z) is in $Num(A)$ and λ is not in the closure of $Num(A)$. Given this form, we can lift the $\|z\|^2 = 1$ condition, and get $|((A - \lambda I)z, z)| \geq c\|z\|^2$ for all z, so by the previous lemma, $(A - \lambda I)^{-1}$ exists, and $\lambda \notin \sigma(A)$. Thus, $\lambda \notin \overline{Num(A)}. \implies \lambda \notin \sigma(A)$, as desired. □

Corollary 4.36. $A = A^* \implies \sigma(A) \subseteq \mathbb{R}$.

Proof. If $\mathbb{F} = \mathbb{C}$, then a self-adjoint operator A will have real quadratic forms, and $\overline{Num(A)} \subseteq \mathbb{R}$. If $\mathbb{F} = \mathbb{R}$, we would have the same property about its numerical range. So, $\sigma(A) \subseteq \overline{Num(A)} \subseteq \mathbb{R}$. □

Notice that the map $u(z) = (z - i)/(z + i)$ maps the upper half plane to the unit disc. For $A = A^*$, if we define $U := (A - i)(A + i)^{-1}$, we would have U being unitary. This is called the **Cayley transform**. Similarly, if U is unitary and $1 \notin \sigma(U)$, then $A = i(I + U)(I - U)^{-1}$ is self-adjoint.

Definition 4.37. *The* ***numerical radius*** *of A, denoted by $q(A)$, is $q(A) := \sup_{\|x\|=1} |(Ax, x)|$.*

The numerical radius has certain properties. Notice that $|(Ax, x)| \leq q(A)\|x\|^2$ for all x. We would then have $\|A\| \geq q(A) \geq r(A)$ for all operators A. This chain of inequalities can be straightened to equalities if we have some more conditions of A.

Theorem 4.38. *If A is self-adjoint, $\|A\| = q(A)$.*

Proof. We already have $\|A\| \geq q(A)$, to obtain the desired result we just need to establish $\|A\| \leq q(A)$, i.e. for all $x \in H$ and $\|x\| = 1$, we would have $\|Ax\| \leq q(A)$.

We have

$$(A(x + y), x + y) - (A(x - y), x - y) = 2(Ax, y) + 2(Ay, x) = 2(Ax, y) + 2\overline{(Ax, y)} = 4Re(Ax, y).$$

This means, we have

$$4Re(Ax, y) \leq |4Re(Ax, y)| \leq q(A)(\|x + y\|^2 + \|x - y\|^2) = 2q(A)(\|x\|^2 + \|y\|^2).$$

If $\|x\| = \|y\| = 1$, we would therefore get $4Re(Ax, y) \leq 4q(A)$. If we take $y = Ax/\|Ax\|$ when $\|Ax\| \neq 0$, then $Re(Ax, y) = \|Ax\|^2/\|Ax\| = \|Ax\| \leq q(A)$. If $\|Ax\| = 0$, the desired inequality would be trivially true. So, we have the desired inequality, thus the desired result. □

Theorem 4.39. *Let $A = A^* \in B(H)$. Denote $m := \inf_{\|x\|=1}(Ax, x)$ and $M := \sup_{\|x\|=1}(Ax, x)$. Then, we have (i) $\sigma(A) \subseteq [m, M]$, (ii) $m, M \in \sigma(A)$. Moreover, if there exists x such that $\|x\| = 1$ and $(Ax, x) = m$, i.e. the infimum is attained, then $Ax = mx$. Similarly, if the supremum is attained by some x, then we have $Ax = Mx$.*

Proof. The numerical range of A will be an interval $[m, M]$, and we would shift it so that it is symmetrical about the origin by adding some constant to the operator A. We let $\alpha := (M+m)/2$ and $\beta := (M - m)/2$, and $B := A - \alpha I$. This means the numerical range of B is $[-\beta, \beta]$, and $\|B\| = q(B) = \beta$ using Theorem 4.37 above.

Suppose $(Bx, x) = \beta$ for some x with $\|x\| = 1$, we have

$$\begin{aligned}\|(B - \beta I)x\|^2 &= ((B - \beta I)x, (B - \beta I)x) \\ &= (Bx, Bx) - \beta(x, Bx) - \beta(Bx, x) + \beta^2(x, x) \\ &= \|Bx\|^2 - 2\beta^2 + \beta^2 \\ &\leq \beta^2 - \beta^2 = 0.\end{aligned}$$

So, as a square is non-negative, $\|(B - \beta I)x\| = 0$, which implies $(B - \beta I)x = 0$.

Since $\beta = \sup_{\|x\|=1}(Bx, x)$, there exists a sequence $\{x_n\}$ such that $\|x_n\} = 1$ and $(Bx_n, x_n) \to \beta$. Then,

$$\begin{aligned}\|(B - \beta I)x_n\|^2 &= ((B - \beta I)x_n, (B - \beta I)x_n) \\ &= (Bx_n, Bx_n) - \beta(x_n, Bx_n) - \beta(Bx_n, x_n) + \beta^2(x_n, x_n) \\ &= \|Bx_n\|^2 - 2\beta(Bx_n, x_n) + \beta^2 \\ &\leq 2\beta^2 - 2\beta(Bx_n, x_n) \to 0\end{aligned}$$

as $n \to \infty$. So, $(B - \beta I)x_n \to 0$ as $n \to \infty$.

Suppose now that $(B - \beta I)^{-1}$ exists so it is bounded, then $\|x_n\| = \|(B - \beta I)^{-1}(B - \beta I)x_n\| \leq \|(B-\beta I)^{-1}\| \cdot \|(B-\beta I)x_n\| \to 0$. However, $\|x_n\| = 1$ by construction, so we have a contradiction. Thus, $\beta \in \sigma(B)$.

Therefore, if we reverse the shifting, we get $M \in \sigma(A)$. The case is similar for $-\beta$ and thus m. □

Corollary 4.40. *If $A = A^* \in B(H)$ and $\sigma(A) = \{0\}$, then $A = 0$.*

Proof. Since A is self-adjoint so normal, we have $\|A\| = r(A)$ (see theorem below). Furthermore, we have

$$r(A) = \sup\{|\lambda|\} = \max\{|\lambda|\} = 0 = \|A\|$$

so $A = 0$. □

Corollary 4.41. *If $A = A^* \in B(H)$, then $\sigma(A) \subseteq \{\lambda \in \mathbb{R}, \lambda \geq 0\} \iff (Ax, x) \geq 0$ for all x. We call such an operator A as* ***positive****.*

Remark. This result still holds even when A is unbounded.

Theorem 4.42. *Let $A \in B(H)$ be normal, i.e. $AA^* = A^*A$, then $r(A) = \|A\|$. This means $r(A) = q(A) = \|A\|$.*

Proof. A is normal so $\|Ax\| = \|A^*x\|$ for all x. So, $\|AAx\| = \|A^2x\| = \|A^*Ax\|$ for al x, and $\|A^2\| = \sup_{\|x\|=1} \|A^2x\| = \sup_{\|x\|=1} \|A^*Ax\| = \|A^*A\| = \|A\|^2$. Then, we have $\|A^{2k}\| = \|A\|^{2k}$ for any positive integer k. Thus,

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{k\to\infty} \|A^{2k}\|^{1/2k} = \lim_{k\to\infty} \|A\| = \|A\|,$$

as desired. □

4.7 Hilbert-Schmidt

Suppose now $A = A^* \in Com(H)$. Then all non-zero points in $\sigma(A)$ are real eigenvalues. Let us list them including multiplicities: $\lambda_1, \lambda_2, \dots$ such that $\lambda_j \in \mathbb{R}$ and $|\lambda_1| \geq |\lambda_2| \geq \cdots$. If there are infinite many eigenvalues, then $\lim_{n\to\infty} \lambda_n = 0$. Let $\{e_n\}$ be an orthonormal system of corresponding eigenvectors with $Ae_n = \lambda_n e_n$.

We denote $L := \overline{\text{span}\{e_n\}_{n=1}^N}$, $N \in \mathbb{N} \cup \{\infty\}$. Then, for all $x \in H$, we can write

$$x = \sum_{n=1}^{N} c_n e_n + y$$

with $y \in L^\perp$. Then, $Ax = \sum_{n=1}^N c_n \lambda_n e_n$, i.e. $y \in \text{Ker}A$. This last line is known as the Hilbert-Schmidt Theorem.

Theorem 4.43 (Hilbert-Schmidt Theorem)**.** $L^\perp = KerA$.

Proof. We know that $L^\perp \supseteq \text{Ker}A$, and this is a standard result. We want to have $L^\perp \subseteq \text{Ker}A$ to get the desired result. Suppose $y \in L^\perp$, then $(Ay, e_n) = (y, Ae_n) = \lambda_n(y, e_n) = 0$. So $Ay \in L^\perp$.

Thus, $AL^\perp \subseteq L^\perp$. Consider $A|_{L^\perp}$. This is compact. If $k \in \sigma(A|_{L^\perp})$ with $k \neq 0$, then k is an eigenvalue of $A|_{L^\perp}$, yet all eigenvalues of it are 0. So, $\sigma(A|_{L^\perp}) = \{0\}$. Also, $A|_{L^\perp}$ is self-adjoint, so its spectrum is zero implying the operator is zero, so $L^\perp \subseteq \text{Ker}A$. Done. $\square$

Consider $x = \sum_{n=1}^N c_n e_n + y$ and $Ax = \sum_{n=1}^N c_n \lambda_n e_n$. Given a function $f : \mathbb{R} \to \mathbb{C}$, we can define $f(A)$ by

$$f(A)x = \sum_{n=1}^{N} c_n f(\lambda_n) e_n + f(0)y.$$

In particular, if A is positive, so $\lambda_n > 0$ for all n, then we can define $\sqrt{A}$ as an compact operator such that

$$\sqrt{A}x = \sum_{n=1}^{N} c_n \sqrt{\lambda_n} e_n$$

whenever $x = \sum_{n=1}^N c_n e_n + y$.

Let $\{e_n\}$ and $\{f_k\}$ be two complete orthonormal systems and $T \in B(H)$. Then,

$$\sum_{n=1}^{\infty} \|Te_n\|^2 = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} |(Te_n, f_k)|^2 = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} |(e_n, T^* f_k)|^2 = \sum_{k=1}^{\infty} \|T^* f_k\|^2.$$

Thus, $\sum_{n=1}^\infty \|Te_n\|^2 = \sum_{n=1}^\infty \|Te_n'\|^2$ for any orthonormal system $\{e_n'\}$.

Definition 4.44. *A **Hilbert-Schmidt norm** of T is* $\|T\|_{HS} = \sqrt{\sum_{n=1}^\infty \|Te_n\|^2}$. *If* $\|T\|_{HS} < \infty$, *then T is called a **Hilbert-Schmidt operator**.*

We have $\|T\|^2 = \sup_{\|x\|=1} \|Tx\|^2 = \sup_{\|e_1\|=1} \|Te_1\|^2 \leq \sup_{e_1,e_2,\dots} \sum \|Te_n\|^2 = \|T\|_{HS}$.

Theorem 4.45. *Hilbert-Schmidt Operators are compact.*

Proof. For $x = \sum_{n=1}^{N} a_n e_n$, we define $T_N x$ by $T_N x = x = \sum_{n=1}^{N} a_n T e_n$ and T_N is finite rank so compact. Next,

$$\begin{aligned}
\|(T - T_N)x\| &= \left\| \sum_{n=N+1}^{\infty} a_n T e_n \right\| \\
&\le \sum_{n=N+1}^{\infty} |a_n| \cdot \|Te_n\| \\
&\le \sqrt{\sum_{n=1}^{\infty} |a_n|^2} \sqrt{\sum_{n=N+1}^{\infty} \|Te_n\|^2} \\
&= \|x\| \sqrt{\sum_{n=N+1}^{\infty} \|Te_n\|^2}
\end{aligned}$$

for all x. So,

$$\|T - T_n\| \le \sqrt{\sum_{n=N+1}^{\infty} \|Te_n\|^2} \to 0$$

as $n \to \infty$. Thus, T is compact. $\square$

Let us do an example. Consider $H = L_2[0,1]$ and $k : [0,1]^2 \to \mathbb{F}$ satisfying

$$\int_0^1 \int_0^1 |k(t,\tau)|^2 \, d\tau dt < \infty,$$

i.e. $k \in L_2[0,1]^2$. Consider

$$(kf)(t) = \int_0^1 k(t,\tau) f(\tau) \, d\tau.$$

Proposition 4.46. *$k : H \to H$ is Hilbert-Schmidt and $\|k\|_{HS} = \sqrt{\int_0^1 \int_0^1 |k(t,\tau)|^2 \, d\tau dt}$.*

Proof. Denote $k_t(\tau) = k(t,\tau)$. Let $\{e_n\}$ be a complete orthonormal system and we have

$$(ke_n)(t) = \int_0^1 k(t,\tau) e_n(t) d\tau = (k_t, \overline{e_n}).$$

Therefore,

$$\|ke_n\|^2 = \int_0^1 |(ke_n)(t)|^2 \, dt = \int_0^1 |(k_t, \overline{e_n})|^2 \, dt$$

and

$$\begin{aligned}
\|k\|_{HS}^2 = \sum_{n=1}^{\infty} \|ke_n\|^2 &= \int_0^1 \sum_{n=1}^{\infty} |(k_t, \overline{e_n})|^2 \, dt \\
&= \int_0^1 \|k_t\|^2 \, dt = \int_0^1 \int_0^1 |k_t(\tau)|^2 \, d\tau dt \\
&= \int_0^1 \int_0^1 |k(t,\tau)|^2 \, d\tau dt.
\end{aligned}$$

$\square$

4.8 Schatten-von Neumann Class

Let $T \in Com(H)$. Consider T^*T. We denote its eigenvalues by

$$s_1^2(T) \geq s_2^2(T) \geq \cdots$$

with $s_j \geq 0$. The sequence $\{s_j(T)\}$ is called the **s-numbers**, or **approximate numbers**, of T.

Definition 4.47. *We define the **Schatten-von Neumann class** $\mathfrak{S}_p = \{T | \sum_{j=1}^{\infty} s_j(T)^p < \infty\}$. We denote $\|T\|_p = (\sum_{j=1}^{\infty} s_j(T)^p)^{1/p}$. Here, $p \geq 1$. Also, $\|T\|_2 = \|T\|_{HS}$.*

*Operators $T \in \mathfrak{S}_1$ are called the **trace-class operators**. For them,*

$$\sum_{j=1}^{\infty} (Te_j, e_j) = trT = \sum \lambda_j(T)$$

makes sense.

Chapter 5

Additional Topics

5.1 Sturm-Liouville Operators*

Consider $H = L^2[0,1]$, and operator A defined by $Af = -f'' + qf$ where $q \in C[0,1]$ is real-valued and is commonly known as the potential.The domain of this operator, with Dirichlet boundary condition, is $D_A = \{f \in H^2[0,1], f(0) = f(1) = 0\}$, which means the operator is self-adjoint. The quadratic form of A is

$$Q_A(f) = \int [|f'|^2 + qf^2] \geq \int f^2$$

if $q \geq 1$. This means $Num(A) \subseteq [1, \infty)$, so $\sigma(A) \subseteq [1, \infty)$ as A is self-adjoint. So, $0 \notin \sigma(A)$.

We would now like to find the resolvent of A. We will solve the equation $Au = 0$, which is a second-order ODE with two unique solutions after we have some initial conditions. We denote the two solutions of the ODE as u_1 and u_2, and we have

$$\begin{aligned} u_1(0) &= 0, \qquad u_1'(0) = 1, \qquad Au_1 = 0. \\ u_2(1) &= 0, \qquad u_2'(1) = 1, \qquad Au_2 = 0. \end{aligned}$$

So, for $j = 1, 2$, we have $u_j'' = qu_j$.

The Wronskian of this equation is thus

$$W(t) = \det \begin{bmatrix} u_1(t) & u_2(t) \\ u_1'(t) & u_2'(t) \end{bmatrix} = u_1(t)u_2'(t) - u_2(t)u_1'(t).$$

Firstly, $W'(t) = 0$. To see this, we have

$$W' = u_1'u_2' + u_1u_2'' - u_2'u_1' - u_2u_1'' = u_1u_2'' - u_2u_1'' = qu_1u_2 - qu_1u_2 = 0.$$

So, $W(t)$ is a constant, and this is a non-zero constant. To see this, suppose $W(t) = 0$, consider $t = 0$ and $W(0) = 0$. We have

$$W(0) = u_1(0)u_2'(0) - u_2(0)u_1'(0) = -u_2(0) = 0.$$

However, u_2 is not a zero function since $u_2'(1) = 1$. So, $u_2 \in D_A$ and $u_2 \neq 0$ with $Au_2 = 0$. This means, u_2 is an eigenfunction of A with eigenvalue 0, implying $0 \in \sigma(A)$ and this is a contradiction.

Definition 5.1. *The **Green's function** is defined as*

$$k(t,\tau) := \begin{cases} -c^{-1}u_1(t)u_2(\tau) & t \le \tau \\ -c^{-1}u_2(t)u_1(\tau) & t > \tau. \end{cases}$$

Notice that $k(t,\tau) = 0$ whenever either t or τ is 0 or 1.

We will define the operator K as

$$(Kf)(t) = \int_0^1 k(t,\tau)f(\tau)d\tau.$$

We claim that $K = A^{-1}$. The verification of this claim is slightly technical, thus omitted here.

We notice that the spectrum of K is

$$\sigma(K) = \{k_1 \ge k_2 \ge k_3 \ge \cdots .k_j \to 0\}$$

and the spectrum of A is

$$\sigma(K) = \{\lambda_1 \ge \lambda_2 \ge \lambda_3 \ge \cdots .\lambda_j \to \infty\}.$$

Definition 5.2. *Let $A = A^*$ and $\lambda \in \sigma(A)$. We say that λ belongs to the **discrete spectrum** of A, denoted by $\sigma_d(A)$, if λ is an eigenvalue of finite multiplicity isolated from the rest of the spectrum of A.*

*We will define the **essential spectrum** σ_{ess} as the complement of the discrete spectrum in the whole spectrum, i.e. $\sigma_{ess}(A) = \sigma(A)\backslash\sigma_d(A)$.*

Note that the essential spectrum is special since it is preserved under compact perturbation, i.e. $\sigma_{ess}(A) = \sigma_{ess}(A+K)$ where K is a compact operator.

5.2 Variational Definition of Eigenvalues*

Suppose $A = A^*$ and $A > 0$ with $\sigma_{ess}(A) = \emptyset$. The spectrum is A is all discrete, i.e. $\sigma(A) = \sigma_d(A) = \{\lambda_1 \le \lambda_2 \le \cdots\}$ with $\lambda_j \to \infty$.

Given that the quadratic form $Q_A(f) = (Af, f)$, we can have a variational definition of eigenvalues

$$\lambda_1 := \inf_{f\neq 0} \frac{Q_A(f)}{\|f\|^2}$$

and we also have

$$\lambda_n := \inf_{\substack{\dim L = n \\ L \subset D_A}} \sup_{\substack{f\in L \\ f\neq 0}} \frac{Q_A(f)}{\|f\|^2}.$$