16 Applications of the spectral theorem

16.1 Functions of operators

The theory of functions of operators is called *operational calculus*. There are several ways of defining a function of a linear operator. E.g., for any $A \in \mathcal{B}(\mathcal{H})$ (or for any element in a Banach algebra for that matter) we can define $\exp A$ by the following norm convergent series:

$$\exp A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

Here, we focus on functions of self-adjoint compact operators on a Hilbert space $\mathcal{H}$. If A is self-adjoint and compact, we can diagonalize it:

$$A = \sum_{n=0}^{\infty} \lambda_n P_n$$

where, $\{e_n\}$ is an o.n. basis of $\mathcal{H}$, $P_n = \langle e_n, \cdot \rangle e_n$, $\lambda_n \in [-\|A\|, \|A\|]$, $Ae_n = \lambda_n e_n$, and if $\dim \mathcal{H} = \infty$, $\lim_n \lambda_n = 0$.

Exercise 16.1 *Suppose* $\dim \mathcal{H} = \infty$. *If* $A \in \mathcal{B}(\mathcal{H})$ *is self-adjoint and compact and* $\{\lambda_n\}$ *is the set of eigenvalues of* A*, then the spectrum of* A*,* $\sigma(A)$*, is given by*

$$\sigma(A) = \overline{\{\lambda_n\}} = \{\lambda_n\} \cup \{0\}$$

It is obvious how to define $Q(A)$, for a polynomial of degree d given by $Q(x) = \sum_{n=0}^{d} Q_n x^n$:

$$Q(A) = \sum_{n=0}^{d} Q_n A^n \tag{16.1}$$

The spectral theorem, however, provides an alternative definition that gives $Q(A)$ in diagonal form:

$$Q(A) = \sum_{n=0}^{\infty} Q(\lambda_n) P_n \tag{16.2}$$

Exercise 16.2 a) *Suppose* $\dim H = \infty$. *Prove that the series (16.2) is convergent in norm iff* $Q(0) = 0$.

b) *In general, i.e., for an arbitrary value* $Q(0)$*, prove that the series (16.2) is convergent in the strong operator topology.*

Exercise 16.3 *Prove that (16.1) and (16.2) indeed define the same operator* $Q(A)$.

For a continuous function $f \in C([-\|A\|, \|A\|])$, we also have two options to define $f(A)$. We can make use of a uniformly convergent sequence of polynomials $Q_N \to f$, or we can directly extend (16.2) to continuous functions:

$$f(A) = \sum_{n=0}^{\infty} f(\lambda_n) P_n \tag{16.3}$$

Exercise 16.4 *Prove that the series in (16.3) is convergent in the strong operator topology for any $f \in C([-\|A\|, \|A\|])$, and that it converges uniformly if in addition $f(0) = 0$.*

To prove that the first option also works one needs the following property:

Exercise 16.5 *Let $f \in C([-\|A\|, \|A\|])$, and consider $f(A)$ defined by (16.3). Prove that*

$$\|f(A)\| = \sup\{|f(\lambda_n)| \mid n\}$$

Let (Q_N) be a sequence of polynomials of degree N, converging uniformly to f on $[-\|A\|, \|A\|]$. The existence of such a sequence is a consequence of the Weierstrass Approximation Theorem.

Exercise 16.6 *Prove that $(Q_N(A))$ converges in norm, and that its limit equals $f(A)$ as defined in (16.3).*

Exercise 16.7 *Let A be a self-adjoint compact operator on a Hilbert space $\mathcal{H}$, and let $f : [-\|A\|, \|A\|] \to \mathbb{C}$ be a continuous function. When exactly is $f(A)$ compact?*

Exercise 16.8 *Let A be a self-adjoint compact operator on a Hilbert space $\mathcal{H}$. Prove that*

$$e^{iA} = \cos A + i \sin A$$

Exercise 16.9 *Let A be a non-negative self-adjoint compact operator on a Hilbert space $\mathcal{H}$, i.e., for all $x \in \mathcal{H}$*

$$\langle x, Ax \rangle \geq 0$$

a) *Prove that there is a unique non-negative self-adjoint operator $B \in \mathcal{B}(\mathcal{H})$ such that $B^2 = A$, i.e., $B = A^{1/2}$.*
b) *Prove that $A^{1/2}$ is also compact.*

16.2 Matched filter

The following is a common variational problem on a Hilbert space $\mathcal{H}$: given a vector $0 \neq h \in \mathcal{H}$, and a positive operator W, i.e., $\langle x, Wx \rangle > 0$, for all $0 \neq x \in \mathcal{H}$, find all $0 \neq g^* \in \mathcal{H}$ such that

$$\frac{|\langle g^*, h \rangle|^2}{\langle g^*, Wg^* \rangle} = \sup_{0 \neq g \in \mathcal{H}} \frac{|\langle g, h \rangle|^2}{\langle g, Wg \rangle} \tag{16.4}$$

In the case of compact W we can solve this problen using the spectral theorem for s.a. compact operators. But let us first give an example of a concrete instance where this kind of variational problem arises.

Consider a *filter* L, acting on *signals* $S(t)$ as follows:

$$(LS)(t) = \int_0^t g(s)S(t-s)\,ds$$

Unfortunately, the input to the filter is expected to be contaminated by random *noise* $N(\omega, t)$, where ω takes values in a probability space Ω, with a probability measure $P(d\omega)$. Assume that both signal and noise are given by elements of $L^2([0,T])$. The standard definition of the *signal-to-noise ratio* Q, is

$$Q = \frac{\Sigma^2}{\sigma^2}$$

where

$$\Sigma^2 := (LS)(T)^2 = \left| \int_0^T g(s)S(t-s)\,ds \right|^2$$

is the square of the signal observed through the filter L at time T, and

$$\sigma^2 := \mathbb{E}[|(LN)(T)|^2]$$

is the same quantity for the noise averaged over the probability measure ($\mathbb{E}[\cdot]$ denotes the expectation with respect to the probability measure $P(d\omega)$), i.e.,

$$\begin{aligned} \sigma^2 &= \int_\Omega P(d\omega) \int_0^T ds_1 \int_0^T ds_2\, g(s_1)g(s_2)\overline{N(\omega, T-s_1)}N(\omega, T-s_2) \\ &= \int_0^T ds_1 \int_0^T ds_2\, g(s_1)g(s_2)w(s_1,s_2) \end{aligned}$$

with

$$w(s_1,s_2) := \int_\Omega P(d\omega)\overline{N(\omega, T-s_1)}N(\omega, T-s_2)$$

Q is a measure of the contrast between signal and noise (a good filter has a high Q, i.e., amplifies the signal and/or reduces the noise). Alternatively, ω could have the interpretation of a frequency.

Define an operator W on $L^2([0,T])$ by

$$Wg(t) = \int_0^T w(t,s)g(s)\,ds \tag{16.5}$$

Exercise 16.10 *Prove that W defined in (16.5) is a non-negative, self-adjoint Hilbert-Schmidt operator, if $N(\cdot,\cdot) \in L^2(\Omega \times [0,T])$.*

The signal-to-noise ratio can now be expressed as

$$Q = \frac{|\langle g, S(T-\cdot)\rangle|^2}{\langle g, Wg\rangle}$$

To find a filter (determined by $g \in L^2([0,T])$) that maximizes Q we need to solve a variational problem of the type (16.4).

We assume that W is invertible (does not have 0 as an eigenvalue). Then, $W^{1/2}$ is invertible too, and we can vary over $f = W^{1/2}g$ instead of g.

$$\frac{|\langle g,h\rangle|^2}{\langle g, Wg\rangle} = \frac{|\langle W^{1/2}g, W^{-1/2}h\rangle|^2}{\langle W^{1/2}g, W^{1/2}g\rangle} = \frac{|\langle f,k\rangle|^2}{\langle f,f\rangle}$$

where $k = W^{-1/2}h$ (which we assume to be well-defined). So we have to find f^* such that

$$\frac{|\langle f^*,k\rangle|^2}{\langle f^*,f^*\rangle} = \sup_{0\neq f\in\mathcal{H}} \frac{|\langle f,k\rangle|^2}{\langle f,f\rangle}$$

The Cauchy-Schwarz inequality gives

$$\frac{|\langle f,k\rangle|^2}{\langle f,f\rangle} \leq \langle k,k\rangle$$

for all $f \in L^2([0,T])$, with equality iff $f = ck$, $c \in \mathbb{C}$. Therefore, the solutions of (16.4) are given by

$$g^* = W^{-1/2}f^* = cW^{-1/2}k = cW^{-1}h$$

If one knows the eigensystem of W, g^* can be given more explicitly as:

$$g^* = c\sum_n (\lambda_n)^{-1}\langle e_n, h\rangle e_n$$

16.3 The mean ergodic theorem

The following theorem, due to von Neumann, is not an application of the spectral theorem, but it is another good example of Hilbert space methods.

Theorem 16.1 *Let U be a unitary operator on a Hilbert space $\mathcal{H}$, and let P be the orthogonal projection on the vectors $x \in \mathcal{H}$, invariant under U, i.e., such that $Ux = x$. Then, for all $x \in \mathcal{H}$ one has*

$$\lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} U^n x = Px \tag{16.6}$$

i.e., the average of the U^n converges strongly to P.

Proof: Consider the orthogonal decomposition of $\mathcal{H}$ determined by P:

$$\mathcal{H} = \ker P \oplus \mathrm{ran} P$$

(16.6) trivially holds for $x \in \mathrm{ran}P$.

We can complete the proof by showing that (16.6) also holds for $x \in \ker P$. Note that $\mathrm{ran}P = \ker(\mathbb{1} - U)$, and hence $\ker P = \ker(\mathbb{1} - U)^\perp = \ker(\mathbb{1} - U^*)^\perp = \overline{\mathrm{ran}(\mathbb{1} - U)}$, where we have used the fact that for a unitary U, $U^{-1} = U^*$ (which implies that $Ux = x \Leftrightarrow U^*x = x$), and also Lemma 13.5. So, any $x \in \ker P$, can be approximated by vectors of the form $(\mathbb{1} - U)y$, $y \in \mathcal{H}$. For such $x = (\mathbb{1} - U)y$, we have

$$\begin{aligned} \left\| \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} U^n x \right\| &= \left\| \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} U^n (\mathbb{1} - U) y \right\| \\ &= \lim_{N\to\infty} \frac{1}{N+1} \|y - U^{N+1} y\| = 0 \end{aligned}$$

If $x_k \to x \in \ker P$, with $x_k \in \mathrm{ran}(\mathbb{1} - U)$, for all k, we then have

$$\begin{aligned} \left\| \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} U^n x \right\| &\le \limsup_{N\to\infty} \left\| \frac{1}{N+1} \sum_{n=0}^{N} U^n (x - x_k) \right\| \\ &\quad + \limsup_{N\to\infty} \left\| \frac{1}{N+1} \sum_{n=0}^{N} U^n x_k \right\| \\ &\le \|x - x_k\| + 0 \end{aligned}$$

As k is arbitrary and $x_k \to x$,this proves that (16.6) holds for any $x \in \ker P$. It follows easily then, that (16.6) holds for all $x \in \mathcal{H}$, by using the orthogonal decomposition $x = x_0 + x_1$, with $x_0 \in \ker P$, and $x_1 \in \mathrm{ran}P$. ∎

A class of unitaries for which one applies this theorem is given in terms of bijective maps T of a probability space $(\Omega, P(d\omega))$ into itself, that satisfy

$$\int_A P(d\omega) = \int_{T^{-1}(A)} P(d\omega)$$

where $T^{-1}(A) := \{\omega \in \Omega \mid T(\omega) \in A\}$. Such T are called *measure preserving*. An operator $U : L^2(\Omega, P(d\omega)) \to L^2(\Omega, P(d\omega))$ is then defined by

$$Uf(\omega) = f(T(\omega)) \tag{16.7}$$

and one has

$$\int_\Omega Uf(\omega)P(d\omega) = \int_\Omega f(\omega)P(d\omega)$$

for all $f \in L^2(\Omega, P(d\omega))$.

Exercise 16.11 *Prove that U, defined in (16.7), is unitary.*

Definition 16.2 *The measure preserving bijection T is called ergodic if the only functions $f \in L^2(\Omega, P(d\omega))$, satisfying $f(\cdot) = f(T(\cdot))$, are the constant functions.*

For ergodic T time averages and phase-space averages coincide:

Theorem 16.3 *If T is ergodic then, for all $f \in L^2(\Omega, P(d\omega))$*

$$\lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} f(T^n(\cdot)) = \int_\Omega f(\omega)P(d\omega) \tag{16.8}$$

where the convergence is in the L^2 norm.

Proof: The proof is a direct consequence of Theorem 16.1 applied to the unitary U defined in (16.7). ∎

Compare this theorem with Theorem 9.1. There is another ergodic theorem (Birkhoff's ergodic theorem) stating that the averages on the LHS of (16.8) converge $P-$*almost everywhere* to the constant on the RHS for all $f \in L^1(\Omega, P(d\omega))$. This theorem is a bit harder to prove however.

Note that for (16.8) to hold for all f it is also necessary that T is ergodic.