Spectral Theory

François Genoud
TU Delft, Spring 2017*

*updated February 22, 2017

Contents

Foreword

These notes are based on the semester project of Yann Péquignot, *Théorie spectrale et évolution en mécanique quantique*, which was supervised by Prof. Boris Buffoni and myself at EPFL (Lausanne) in 2008. I am especially indebted to Yann for the exceptional quality of his work, and his permission to use it as teaching material.

The goal of the course is to provide a straightforward but comprehensive proof of the spectral theorem for unbounded selfadjoint operators in Hilbert spaces, and some applications to elementary quantum mechanics. The main focus will be on the decomposition of a selfadjoint operator onto its family of spectral projections. Some elements of functional calculus will also be given.

We will start in Chapter 1 by some recalls about bounded operators in Hilbert spaces and their spectra, as well as important properties of projections and positive operators, which will play a crucial role in the main proofs. Chapter 2 will be devoted to the proof of the spectral theorem for bounded selfadjoint operators, while Chapter 3 will present the extension to the unbounded case. In Chapter 4 we will apply the spectral theorem to discuss Stone's theorem, which characterizes strongly continuous one-parameter groups of unitary transformations of the Hilbert space. After a brief exposition of the main concepts of quantum mechanics formulated in the Hilbert space framework, we will use Stone's theorem to define the operators representing the energy and the momentum of a quantum particle.

The original work, as well as the present version, are based on classic books which are listed in the Bibliography. However, considerable effort is made here to present a unified and self-contained theory. The student with basic knowledge of functional analysis in Hilbert spaces should be sufficiently equipped to read these notes. Of course, as always in mathematics, we can only recommend very active reading (doing the exercises, reproducing the proofs mentally and on paper, etc.) in order to get familiar with the theory.

As I am translating this work to English and adding extra material and exercises, I of course take full responsibility should there be any mistakes or imprecisions in the text.

Delft, February 2015

Terminology We shall speak of *unbounded operators* when referring to general, not necessarily bounded, operators. The price to pay for this abuse of terminology is that bounded operators become a special case of unbounded operators! But we prefer to live with this rather than repeatedly using the awkward phrase 'general, not necessarily bounded, operators'.

The notion of selfadjointness for unbounded operators requires a careful definition (in particular the definition of the domain of the adjoint operator), while for a bounded operator A acting in a Hilbert space $\mathcal{H}$, it only amounts to requiring that A be symmetric, i.e. that $\langle Au, v\rangle = \langle u, Av\rangle$, for all $u, v \in \mathcal{H}$. If A is unbounded, it is also called symmetric provided the previous relation holds for all u, v in the domain of A. Throughout the course we shall reserve the term *selfadjoint* for unbounded operators, while bounded selfadjoint operators will merely be called *symmetric*.

Chapter 1

Preliminary notions

We start by recalling elements of the theory of linear operators acting in a Hilbert space $\mathcal{H}$. We present some basic results about bounded linear operators and some elementary properties of orthogonal projections.

1.1 Hilbert spaces

In these notes we will consider Hilbert spaces over the field $\mathbb{C}$ of complex scalars. The definition of Hilbert space is as follows.

Definition 1.1.1 A **pre-Hilbert space** $\mathcal{H}$ is a complex vector space endowed with an inner product $\langle \cdot, \cdot \rangle : \mathcal{H} \to \mathbb{C}$ satisfying

(i) for all $u \in \mathcal{H}$, $\langle u, u \rangle \geqslant 0$, and $\langle u, u \rangle = 0$ if and only if $u = 0$;

(ii) for all $u, v, w \in \mathcal{H}$, $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$;

(iii) for all $u, v \in \mathcal{H}$ and all $\lambda \in \mathbb{C}$, $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle$;

(iv) for all $u, v \in \mathcal{H}$, $\langle u, v \rangle = \overline{\langle v, u \rangle}$.

$\mathcal{H}$ is called a **Hilbert space** if it is a Banach space for the norm $\|x\| := \sqrt{\langle x, x \rangle}$, i.e. if the metric space $(\mathcal{H}, \mathrm{d})$, with $\mathrm{d}(x, y) := \|x - y\|$, is complete. ♦

We shall use the notations of Definition 1.1.1 throughout the text without further mention.

A typical example of Hilbert space is the following.

Example 1.1.2 (The Hilbert space $L^2[0,1]$) Consider the set of functions $u : [0,1] \to \mathbb{C}$ such that

$$\int_{[0,1]} |u(x)|^2 \, \mathrm{d}x < \infty,$$

where the integral is taken with respect to the Lebesgue measure on $[0,1]$. We define an equivalence relation by

$$u \sim v \quad \text{iff} \quad u = v \text{ almost everywhere.}$$

We denote by $L^2[0,1]$ the set of equivalence classes obtained in this way. This is a (complex) Hilbert space for the following operations:

$$\begin{aligned} (\lambda u)(x) &= \lambda u(x); \quad (\lambda \in \mathbb{C}) \\ (u+v)(x) &= u(x) + v(x); \\ \langle u, v \rangle &= \int_{[0,1]} u(x)\overline{v(x)} \, \mathrm{d}x, \end{aligned}$$

where u, v denote any elements of their respective equivalence classes. ⧫

In fact, Example 1.1.2 is generic, in the sense that all infinite-dimensional *separable* Hilbert spaces are isomorphic to $L^2[0,1]$.

The Cartesian product $\mathcal{H} \times \mathcal{H}$ has a natural Hilbert space structure given by the following operations:

$$\begin{aligned} \lambda(u_1, v_1) &= (\lambda u_1, \lambda v_1); \\ (u_1, v_1) + (u_2, v_2) &= (u_1 + u_2, v_1 + v_2); \\ \big\langle (u_1, v_1), (u_2, v_2) \big\rangle &= \langle u_1, u_2 \rangle + \langle v_1, v_2 \rangle, \end{aligned}$$

for all $(u_1, v_1), (u_2, v_2) \in \mathcal{H} \times \mathcal{H}$ and all $\lambda \in \mathbb{C}$. The metric topology induced by the inner product of $\mathcal{H} \times \mathcal{H}$ coincides with the product topology inherited from $\mathcal{H}$.

1.2 Linear operators

In these notes we are interested in linear operators between Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, that is mappings $A : \mathcal{H} \to \mathcal{H}'$ which preserve the vector space structures of $\mathcal{H}$ and $\mathcal{H}'$. In the first part we will consider linear operators which are bounded, i.e. continuous, and so also preserve the topology. Then we will deal with general, unbounded, operators.

Definition 1.2.1 A **linear operator** A from the Hilbert space $\mathcal{H}$ to the Hilbert space $\mathcal{H}'$ is a mapping $A : \mathcal{H} \to \mathcal{H}'$ such that:

(i) for all $u, v \in \mathcal{H}$, $A(u + v) = A(u) + A(v)$;

(ii) for all $\lambda \in \mathbb{C}$ and all $u \in \mathcal{H}$, $A(\lambda u) = \lambda A(u)$.

We merely write Au for the image $A(u)$ of the element $u \in \mathcal{H}$. The **range** of A is the subspace of $\mathcal{H}'$

$$\operatorname{rge} A = \{Au\,;\, u \in \mathcal{H}\},$$

and the **kernel** of A is the subspace of $\mathcal{H}$

$$\ker A = \{u \in \mathcal{H}\,;\, Au = 0\}.$$

If there is a constant $C \geqslant 0$ such that

$$\|Au\|_{\mathcal{H}'} \leqslant C\,\|u\|_{\mathcal{H}} \quad \text{for all } u \in \mathcal{H}, \tag{1.2.1}$$

we say that A is **bounded**. The set $\mathcal{B}(\mathcal{H}, \mathcal{H}')$ of bounded operators from $\mathcal{H}$ to $\mathcal{H}'$ is a Banach space for the operations:

$$\begin{aligned} (A + B)u &= Au + Bu \quad && \text{for all } A, B \in \mathcal{B}(\mathcal{H}, \mathcal{H}'),\ u \in \mathcal{H}; \\ (\lambda A)u &= \lambda Au && \text{for all } A \in \mathcal{B}(\mathcal{H}, \mathcal{H}'),\ \lambda \in \mathbb{C},\ u \in \mathcal{H}. \end{aligned}$$

The norm of $A \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$ is defined as the infimum of all $C \geqslant 0$ satisfying (1.2.1) and can be characterized as

$$\|A\|_{\mathcal{B}(\mathcal{H},\mathcal{H}')} = \sup_{\|u\|_{\mathcal{H}}=1} \|Au\|_{\mathcal{H}'}\,. \tag{1.2.2}$$

We will merely write $\|A\|$ when there is no risk of confusion. ♦

Remark 1.2.2 The above definitions can be made for $A : \mathcal{X} \to \mathcal{X}'$, where $\mathcal{X}$ and $\mathcal{X}'$ are merely normed vector spaces. The space of bounded operators $\mathcal{B}(\mathcal{X}, \mathcal{X}')$ is a Banach space for the norm $\|\cdot\|_{\mathcal{B}(\mathcal{H},\mathcal{H}')}$ defined in (1.2.2), provided the target space $\mathcal{X}'$ is Banach. Note that the bounded linear operators from $\mathcal{X}$ to $\mathcal{X}'$ are precisely those linear operators which are continuous with respect to the topologies of $\mathcal{X}$ and $\mathcal{X}'$, induced by the norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{X}'}$, respectively.

In these notes we shall be mainly concerned with the special case where $\mathcal{H}' = \mathcal{H}$. Then we simply write $\mathcal{B}(\mathcal{H})$ instead of $\mathcal{B}(\mathcal{H}, \mathcal{H})$.

The **graph** of a linear operator $A : \mathcal{H} \to \mathcal{H}$ is defined as the subspace

$$\boldsymbol{G}_A = \{(u, Au)\,;\, u \in \mathcal{H}\} \subseteq \mathcal{H} \times \mathcal{H}.$$

It is not difficult to see that $A : \mathcal{H} \to \mathcal{H}$ is bounded if and only if $\boldsymbol{G}_A$ is a closed subset of $\mathcal{H} \times \mathcal{H}$ (for the product topology inherited from $\mathcal{H}$).

We now recall a fundamental result due to Banach and Steinhaus. The proof can be found e.g. in [Fri82, Kre78, RSN90, Wei80].

Theorem 1.2.3 (Uniform Boundedness Principle) *Consider a sequence $(A_n)_{n\in\mathbb{N}}$ of bounded operators acting between Banach spaces $\mathcal{X}, \mathcal{X}'$. If, for all $u \in \mathcal{X}$, we have*

$$\sup_{n\in\mathbb{N}} \|B_n u\|_{\mathcal{X}} < \infty,$$

then there holds

$$\sup_{n\in\mathbb{N}} \|B_n\|_{\mathcal{B}(\mathcal{X},\mathcal{X}')} < \infty.$$

Remark 1.2.4 In Theorem 1.2.3, one can merely assume that $\mathcal{X}'$ is a normed vector space; completeness is not needed in the proof.

Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of bounded linear operators on the Hilbert space $\mathcal{H}$. If there is an operator $A \in \mathcal{B}(\mathcal{H})$ such that $\lim_{n\to\infty} \|A_n - A\|_{\mathcal{B}(\mathcal{H})} = 0$, we say that the sequence $(A_n)_{n\in\mathbb{N}}$ **converges (in norm)** to A. (This is the convergence in the topology of $\mathcal{B}(\mathcal{H})$.) If, for all $u \in \mathcal{H}$, the limit $\lim_{n\to\infty} A_n u$ exists, we say that the sequence $(A_n)_{n\in\mathbb{N}}$ is **stronlgy convergent**. The mapping $u \mapsto \lim_{n\to\infty} A_n u$ is then clearly linear. Moreover we have the following corollary of Theorem 1.2.3.

Corollary 1.2.5 *Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of bounded linear operators on the Hilbert space $\mathcal{H}$. If $(A_n)_{n\in\mathbb{N}}$ is strongly convergent then the operator $u \mapsto \lim_{n\to\infty} A_n u$ is bounded.*

Proof. Let us denote $Au = \lim_{n\to\infty} A_n u$ the limit operator. For all $u \in \mathcal{H}$, it follows by continuity of the norm (Problem 1.1) that

$$\lim_{n\to\infty} \|A_n u\| = \|Au\| \quad \text{exists.}$$

Therefore the sequence $(\|A_n u\|)_{n\in\mathbb{N}}$ is bounded for all $u \in \mathcal{H}$. It then follows from the Uniform Boundedness Principle that $(\|A_n\|)_{n\in\mathbb{N}}$ is bounded by a constant $C \geqslant 0$. Thus, for all $u \in \mathcal{H}$,

$$\|Au\| = \lim_{n\to\infty} \|A_n u\| \leqslant C \|u\|,$$

and so A is bounded. $\square$

A **(continuous) linear functional** on the Hilbert space $\mathcal{H}$ is an element of the **dual space** of $\mathcal{H}$, $\mathcal{H}^* = \mathcal{B}(\mathcal{H}, \mathbb{C})$. The linear functionals on a Hilbert space are characterized by the following theorem due to F. Riesz, the proof of which can be found e.g. in [Fri82, p. 206] or [Kre78, p. 188].

Theorem 1.2.6 (Riesz's Representation Theorem) *Let $\mathcal{H}$ be a Hilbert space. For any $u_0 \in \mathcal{H}$, the formula*

$$f(u) = \langle u, u_0 \rangle, \quad u \in \mathcal{H}, \tag{1.2.3}$$

defines a linear functional on $\mathcal{H}$, with $\|f\| = \|u_0\|$. Conversely, for any linear functional f on $\mathcal{H}$, there exists a unique $u_0 \in \mathcal{H}$ such that (1.2.3) *holds.*

We complete this section by recalling some spectral properties of linear operators.

Definition 1.2.7 Let $A \in \mathcal{B}(\mathcal{X})$ be a bounded linear operator acting in a complex normed vector space $\mathcal{X} \neq \{0\}$, and $I : \mathcal{X} \to \mathcal{X}$ be the identity. The **resolvent set** $\rho(A)$ of A is the set of all complex numbers λ such that $(A - \lambda I)$ is invertible and has a bounded inverse defined on a dense subspace of $\mathcal{X}$. The **spectrum** $\sigma(A)$ of A is defined as $\sigma(A) = \mathbb{C} \setminus \rho(A)$. It can be decomposed as

$$\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A), \tag{1.2.4}$$

where the sets $\sigma_p(A), \sigma_c(A), \sigma_r(A) \subseteq \mathbb{C}$ are defined as follows:

(i) $\lambda \in \sigma_p(A)$ iff $(A - \lambda I)$ is not invertible, i.e. λ is an eigenvalue of A;

(ii) $\lambda \in \sigma_c(A)$ iff $\mathrm{rge}(A - \lambda I)$ is dense in $\mathcal{X}$ and $(A - \lambda I)^{-1}$ exists but is not bounded;

(iii) $\lambda \in \sigma_r(A)$ iff $A - \lambda I$ is injective but the domain of $(A - \lambda I)^{-1}$ is not dense in $\mathcal{X}$.

The sets $\sigma_p(A)$, $\sigma_c(A)$ and $\sigma_r(A)$ are respectively called the **point spectrum**, the **continuous spectrum** and the **residual spectrum**. ♦

We now collect without proof some important properties of the spectrum; see e.g. [Kre78].

Theorem 1.2.8 (Spectrum) *Let $A \in \mathcal{B}(\mathcal{X})$ with $\mathcal{X} \neq \{0\}$ a complex Banach space. Then the spectrum $\sigma(A)$ of A is a compact non-empty subset of $\mathbb{C}$. Furthermore,*

$$\sigma(A) \subseteq \big\{ \lambda \in \mathbb{C}\,;\, |\lambda| \leqslant \|A\| \,\big\}.$$

Definition 1.2.9 The number $r_\sigma(A) = \sup_{\lambda \in \sigma(A)} |\lambda|$ is called the **spectral radius** of A. ♦

Theorem 1.2.10 (Spectral Radius I) *Let $A \in \mathcal{B}(\mathcal{X})$ with $\mathcal{X} \neq \{0\}$ a complex Banach space. Then the limit $\lim_{n\to\infty} \|A^n\|^{1/n}$ exists and is equal to $r_\sigma(A)$.*

1.3 Orthogonal projections

An orthogonal projection is a bounded linear transformation of the Hilbert space $\mathcal{H}$ that maps the whole space onto a closed linear subspace, parallelly to the directions orthogonal to this subspace. The essence of the Spectral Theorem is to decompose a selfadjoint operator as a linear combination of orthogonal projections, so these will play a prominent role throughout the course.

Two elements $u, v \in \mathcal{H}$ are **orthogonal** if $\langle u, v\rangle = 0$. Consider a subspace $M \subseteq \mathcal{H}$. An element $u \in \mathcal{H}$ is orthogonal to M if u is orthogonal to all elements of M. A subspace $N \subseteq \mathcal{H}$ is orthogonal to M if each element of N is orthogonal to M. The **orthogonal complement** $M^\perp$ of M is the *closed* subspace

$$M^\perp = \{u \in \mathcal{H}\,;\, \langle u, v\rangle = 0 \text{ for all } v \in M\} \subseteq \mathcal{H}.$$

The following theorem will allow us to define the notion of orthogonal projection. Its proof can be found e.g. in [Wei80, p. 31] or [Kre78, p. 146].

Theorem 1.3.1 *Let M be a closed subspace of the Hilbert space $\mathcal{H}$. For all $u \in \mathcal{H}$, there exists a unique $v \in M$ and a unique $w \in M^\perp$ such that $u = v + w$.*

If M is a closed subspace of the Hilbert space $\mathcal{H}$, the **orthogonal projection** or simply **projection** onto M is the bounded operator defined by $Pu = v$ for all $u \in \mathcal{H}$, where $u = v + w$ is the unique decomposition of u with $v \in M$ and $w \in M^\perp$. Note that any projection P satisfies $P^2 = P$ and, if $P \neq 0$ then $\|P\| = 1$. Furthermore, since $v = Pu$ and $w = (I - P)u$ are orthogonal, we have

$$\|u\|^2 = \|Pu\|^2 + \|(I-P)u\|^2 \implies \|Pu\| \leqslant \|u\|, \quad \text{for all } u \in \mathcal{H}.$$

We also recall the following results.

Theorem 1.3.2 *Let M and N be closed subspaces of the Hilbert space $\mathcal{H}$. Denote by P and Q the associated projections. Then one has:*

(a) for all $u, v \in \mathcal{H}$, $\langle Pu, v\rangle = \langle u, Pv\rangle$;

(b) $M = \operatorname{rge} P$ and $M^\perp = \ker P$;

(c) $I - P$ is the projection onto $M^\perp$;

(d) $M \subseteq N$ if and only if $PQ = QP = P$;

(e) for all $u \in \mathcal{H}$, $\langle Pu, u\rangle \leqslant \langle Qu, u\rangle$ if and only if $M \subseteq N$.

1.4 Symmetric operators

Symmetric operators on the Hilbert space are bounded operators having a peculiar behaviour with respect to the inner product. They are often referred to as 'selfadjoint' in the literature. However we shall reserve the term selfadjoint for the extension of the notion of symmetric operator to unbounded operators, where the definition of the adjoint operator requires extra care.

Definition 1.4.1 Consider a bounded operator A on the Hilbert space $\mathcal{H}$. The **adjoint** $A^* \in \mathcal{B}(\mathcal{H})$ of A is defined by

$$\langle Au, v\rangle = \langle u, A^* v\rangle\,, \quad \text{for all } u, v \in \mathcal{H}. \tag{1.4.1}$$

The operator A is called **symmetric** if $A = A^*$. We shall merely call **symmetric operator** a bounded symmetric operator. ♦

Remark 1.4.2 That (1.4.1) indeed defines a bounded operator A^* is ensured by Theorem 1.2.6.

The following lemma is a fundamental structural result about the adjoint operator.

Lemma 1.4.3 *Consider a bounded operator $A : \mathcal{H} \to \mathcal{H}$. Then*

$$\mathcal{H} = \ker A^* \oplus \overline{\operatorname{rge} A}.$$

Proof. See Problem 1.13. □

The following characterization of the residual spectrum is often useful to study the spectrum of an operator (see Problem 1.17).

Proposition 1.4.4 *The residual spectrum $\sigma_r(T)$ of a bounded operator $T : \mathcal{H} \to \mathcal{H}$ can be characterized as*

$$\sigma_r(T) = \{\lambda \in \mathbb{C}\,;\, \overline{\lambda} \in \sigma_p(T^*) \text{ but } \lambda \notin \sigma_p(T)\}.$$

Proof. We apply Lemma 1.4.3 with $A = T - \lambda I$. By definition of $\sigma_r(T)$ we have

$$\begin{aligned} \lambda \in \sigma_r(T) &\iff T - \lambda I \text{ is injective but } \overline{\operatorname{rge}(T - \lambda I)} \subsetneq \mathcal{H} \\ &\iff \ker(T^* - \overline{\lambda} I) \neq \{0\} = \ker(T - \lambda I), \end{aligned}$$

which proves the proposition. □

We now collect some basic properties of the adjoint operator. The proofs can be found e.g. in [Fri82] and [Kre78].

Proposition 1.4.5 *Consider two bounded linear operators $A, B : \mathcal{H} \to \mathcal{H}$ and any scalar $\lambda \in \mathbb{C}$. Then we have:*

(a) $(A+B)^* = A^* + B^*$;

(b) $(AB)^* = B^* A^*$;

(c) $(\lambda A)^* = \overline{\lambda} A^*$;

(d) $A^{**} = A$;

(e) $0^* = 0$ *and* $I^* = I$;

(f) $\|A^* A\| = \|AA^*\| = \|A\|^2$;

(g) if $T^{-1} \in \mathcal{B}(\mathcal{H})$ *then* T^* *is invertible,* $(T^*)^{-1} \in \mathcal{B}(\mathcal{H})$ *and* $(T^*)^{-1} = (T^{-1})^*$.

The **numerical range** of a bounded operator A is the set

$$\{\langle Au, u\rangle \ ; \ u \in \mathcal{H}, \ \|u\| = 1\} \subset \mathbb{C}.$$

Theorem 1.4.6 *An operator $A \in \mathcal{B}(\mathcal{H})$ is symmetric if and only if its numerical range is real.*

Proof. If A is symmetric then $\langle Au, u\rangle = \langle u, Au\rangle = \overline{\langle Au, u\rangle}$ for all $u \in \mathcal{H}$. Conversely, if $\langle Au, u\rangle$ is real for all $u \in \mathcal{H}$, we have $\langle Au, u\rangle = \langle A^* u, u\rangle$ and so

$$\langle (A - A^*)u, u\rangle = 0 \quad \text{for all } u \in \mathcal{H}.$$

It now follows by Problem 1.14 that $A - A^* = 0$. □

Remark 1.4.7 Note that it is essential for the 'if' part that $\mathcal{H}$ be complex. Indeed, in a real Hilbert space, the numerical range of any operator is real. (See also in Problem 1.14 why the last part of the above argument fails when $\mathcal{H}$ is real.)

Proposition 1.4.8 *The numerical range of any orthogonal projection $P \notin \{0, I\}$ is $[0, 1]$.*

We now recall without proof some important results about the spectrum of symmetric operators, which are discussed e.g. in [Fri82, pp. 218-221].

Theorem 1.4.9 *Consider a symmetric operator S on the Hilbert space $\mathcal{H}$. Then*

$$\|S\| = \sup_{\|u\|=1} |\langle Su, u\rangle| .$$

Definition 1.4.10 Let S be a symmetric operator on $\mathcal{H}$. The **lower bound** and **upper bound** of S are respectively defined as

$$m = \inf_{\|u\|=1} \langle Su, u \rangle \quad \text{and} \quad M = \sup_{\|u\|=1} \langle Su, u \rangle .$$

Remark that, by Theorem 1.4.9, $\|S\| = \max\{|m|, |M|\}$. ♦

Theorem 1.4.11 *Let S be a symmetric operator on $\mathcal{H}$ and let m, M be as in Definition 1.4.10. Then $\sigma(S)$ is real. In fact, we have*

$$\sigma(S) \subset [m, M] \quad \textit{and} \quad m, M \in \sigma(S).$$

Corollary 1.4.12 (Spectral Radius II) *Let S be a symmetric operator on $\mathcal{H}$. The spectral radius of S, introduced in Definition 1.2.9, satisfies $r_\sigma(S) = \|S\|$.*

Remark 1.4.13 It immediately follows from Theorem 1.4.11 and Proposition 1.4.4 that the residual spectrum of a symmetric operator is empty.

1.5 Positive operators

Our description of positive operators here essentially follows [Fri82].

Definition 1.5.1 A symmetric operator $S : \mathcal{H} \to \mathcal{H}$ is called **positive**, denoted $S \geqslant 0$, if

$$\langle Su, u \rangle \geqslant 0 \quad \text{for all } u \in \mathcal{H}.$$

We shall call **positive operator** a symmetric positive operator. For $S, T : \mathcal{H} \to \mathcal{H}$ symmetric, if $S - T \geqslant 0$ we say S is **larger** than T or T is **smaller** than S, and we write $S \geqslant T$ or $T \leqslant S$. ♦

Lemma 1.5.2 *Let $P : \mathcal{H} \to \mathcal{H}$ be a positive operator. There exists a sequence $(P_n)_{n=1}^\infty$ of operators which are polynomials in P with real coefficients, such that the partial sums $\left(\sum_{k=1}^n P_k^2\right)_{n=1}^\infty$ strongly converge to P:*

$$Pu = \sum_{n=1}^\infty P_n^2 u, \quad u \in \mathcal{H}.$$

Proof. If $P = 0$, the statement is trivial. Suppose $P \neq 0$ and define by induction the sequence of operators

$$B_1 = \tfrac{1}{\|P\|} P, \quad B_{n+1} = B_n - B_n^2, \qquad n = 2, 3, \ldots$$

Each B_n is a polynomial in P with real coefficients and so is symmetric by Problem 1.9. We now prove by induction that

$$0 \leqslant B_n \leqslant I \quad \text{for all } n \geqslant 1. \tag{1.5.1}$$

For $n = 1$, it follows from the positivity of P that, for all $u \in \mathcal{H}$,

$$\langle B_1 u, u \rangle = \frac{1}{\|P\|} \langle Pu, u \rangle \geqslant 0,$$

and so $B_1 \geqslant 0$. Moreover, for all $u \in \mathcal{H}$,

$$\langle (I - B_1)u, u \rangle = \langle u, u \rangle - \langle B_1 u, u \rangle = \langle u, u \rangle - \frac{1}{\|P\|} \langle Pu, u \rangle \geqslant 0,$$

since $\langle Pu, u \rangle \leqslant \|P\| \, \|u\|^2$ by Cauchy-Schwarz. Suppose then that (1.5.1) holds for some $m \geqslant 1$ and let us show that it holds for $m + 1$. For all $u \in \mathcal{H}$, since $B_m \geqslant 0$ we have, on the one hand,

$$\left\langle B_m (I - B_m)^2 u, u \right\rangle = \langle B_m (I - B_m) u, (I - B_m) u \rangle \geqslant 0,$$

and on the other, since $B_m \leqslant I$,

$$\left\langle B_m^2 (I - B_m) u, u \right\rangle = \langle (I - B_m) B_m u, B_m u \rangle \geqslant 0.$$

Hence $B_m (I - B_m)^2 \geqslant 0$ et $B_m^2 (I - B_m) \geqslant 0$. Therefore,

$$\begin{aligned} B_{m+1} &= B_m - B_m^2 \\ &= B_m (I - B_m)^2 + B_m^2 (I - B_m) \geqslant 0. \end{aligned}$$

Furthermore, since $B_m \leqslant I$, we have

$$I - B_{m+1} = (I - B_m) + B_m^2 \geqslant 0.$$

This conclude the proof of (1.5.1).

Let us now observe that

$$\sum_{k=1}^{n} B_k^2 = B_1 - B_{n+1}, \quad n \geqslant 1, \tag{1.5.2}$$

which implies

$$\sum_{k=1}^{n} B_k^2 \leqslant B_1, \quad n \geqslant 1.$$

Therefore, for all $u \in \mathcal{H}$,

$$\sum_{k=1}^{n} \langle B_k u, B_k u\rangle \leqslant \langle B_1 u, u\rangle\,, \quad n \geqslant 1.$$

It follows that $\sum_{n=1}^{\infty} \|B_n u\|^2 < \infty$ and so $\lim_{n\to\infty} \|B_n u\| = 0$. Hence, by (1.5.2),

$$\lim_{n\to\infty} \left\| B_1 u - \sum_{k=1}^{n} B_k^2 u \right\| = \lim_{n\to\infty} \|B_{n+1} u\| = 0.$$

On setting $P_n := \sqrt{\|P\|} B_n,\ n \geqslant 1$, the sequence $(P_n)_{n=1}^{\infty}$ does the job. □

Corollary 1.5.3 *Let P and Q be positive operators. If $PQ = QP$ then PQ is positive.*

Proof. Consider the sequence $(P_n)_{n=1}^{\infty}$ given by Lemma 1.5.2 for P. As each P_n is a polynomial in P, we have $P_n Q = Q P_n$ for all $n \geqslant 1$. Thus, for all $u \in \mathcal{H}$, it follows from the continuity of the inner product that

$$\langle PQu, u\rangle = \sum_{n=1}^{\infty} \left\langle P_n^2 Q u, u\right\rangle = \sum_{n=1}^{\infty} \langle P_n Q P_n u, u\rangle = \sum_{n=1}^{\infty} \langle Q P_n u, P_n u\rangle \geqslant 0,$$

by positivity of Q. □

Corollary 1.5.4 *Consider a sequence of symmetric operators $(S_n)_{n=1}^{\infty}$ and suppose there is a symmetric operator T such that:*

(i) if $1 \leqslant m \leqslant n$ then $S_m \leqslant S_n$ (increasing sequence);

(ii) for all $m, n \geqslant 1$, $S_m S_n = S_n S_m$;

(iii) for all $n \geqslant 1$, $T S_n = S_n T$;

(iv) for all $n \geqslant 1$, $S_n \leqslant T$ (upper bound).

Then $(S_n)_{n=1}^{\infty}$ converges strongly to a symmetric operator S.

Proof. Consider the operators $P_n = T - S_n$. Each P_n is positive by (iv) and the sequence $(P_n)_{n=1}^{\infty}$ is decreasing by (i). Indeed, if $1 \leqslant m \leqslant n$ then $P_m - P_n = S_n - S_m \geqslant 0$. Moreover, by (ii) and (iii), $P_n P_m = P_m P_n$ for all $m, n \geqslant 1$. Hence, Corollary 1.5.3 ensures that, if $1 \leqslant m \leqslant n$ then

$$(P_m - P_n) P_m \geqslant 0 \quad \text{and} \quad (P_m - P_n) P_n \geqslant 0.$$

We deduce that

$$\left\langle P_m^2 u, u\right\rangle \geqslant \langle P_m P_n u, u\rangle \geqslant \left\langle P_n^2 u, u\right\rangle \geqslant 0, \quad 1 \leqslant m \leqslant n,\ u \in \mathcal{H}. \tag{1.5.3}$$

Therefore, $(\langle P_n^2 u, u\rangle)_{n=1}^{\infty}$ is a decreasing sequence of positive numbers. Let $\alpha \in \mathbb{R}_+$ be its limit. By (1.5.3) we have, for all $1 \leqslant m < n$:

$$0 \leqslant \langle P_m P_n u, u\rangle - \alpha \leqslant \left\langle P_m^2 u, u\right\rangle - \alpha, \quad u \in \mathcal{H}.$$

It follows from the above identities that

$$\lim_{n,m\to\infty} \langle P_m P_n u, u\rangle = \lim_{m\to\infty} \left\langle P_m^2 u, u\right\rangle, \quad u \in \mathcal{H}.$$

Hence, for all $m, n \geqslant 1$ and all $u \in \mathcal{H}$:

$$\begin{aligned}
\|S_n u - S_m u\|^2 &= \|P_m u - P_n u\|^2 \\
&= \left\langle (P_m - P_n)^2 u, u\right\rangle \\
&= \left\langle P_m^2 u, u\right\rangle + \left\langle P_n^2 u, u\right\rangle - 2\langle P_m P_n u, u\rangle \xrightarrow{m,n\to\infty} 0.
\end{aligned}$$

It follows that, for all $u \in \mathcal{H}$, $(S_n u)_{n=1}^{\infty}$ is a Cauchy sequence in $\mathcal{H}$. It therefore has a limit, which we denote by $Su \in \mathcal{H}$. This defines a linear operator S, which is bounded by Corollary 1.2.5. Furthermore, S is symmetric. Indeed, by continuity of the inner product (Problem 1.1),

$$\langle Su, v\rangle = \left\langle \lim_{n\to\infty} S_n u, v\right\rangle = \lim_{n\to\infty} \langle S_n u, v\rangle = \lim_{n\to\infty} \langle u, S_n v\rangle = \left\langle u, \lim_{n\to\infty} S_n v\right\rangle = \langle u, Sv\rangle, \quad u, v \in \mathcal{H}.$$

The proof is complete. □

We now introduce the important notion of the square root of a positive operator.

Definition 1.5.5 Let P be a positive operator. A **square root** of P is a symmetric operator R such that $R^2 = P$. ♦

Theorem 1.5.6 *Let P be a positive operator. There exists a unique positive square root R of P. Furthermore, R commutes with any bounded operator which commutes with P.*

Proof. The proof is in two steps.

1. Existence. We need only show that all positive operator P such that $P \leqslant I$ has a positive square root. The general case can be deduced from this one by considering $\widetilde{P} = \varepsilon^2 P$, with $\varepsilon > 0$ such that $\varepsilon^2 \|P\| \leqslant 1$. So suppose that $P \leqslant I$ and define by induction the sequence of operators

$$\begin{aligned}
R_0 &= 0 \\
R_{n+1} &= R_n + \frac{1}{2}(P - R_n^2) \quad n = 1, 2, \ldots
\end{aligned} \tag{1.5.4}$$

Each R_n is a polynomial in P with real coefficients. Therefore, by Problem 1.9, R_n is symmetric and commutes with any bounded operator which commutes with P. It follows from the identity

$$I - R_{n+1} = \frac{1}{2}(I - R_n)^2 + \frac{1}{2}(I - P) \tag{1.5.5}$$

that $R_n \leqslant I$ for all $n \geqslant 0$. It then follows by subtraction that

$$\begin{aligned} R_{n+1} - R_n &= \frac{1}{2}(I - R_{n-1})^2 - \frac{1}{2}(I - R_n)^2 \\ &= \frac{1}{2}(R_{n-1}^2 - 2R_{n-1} + 2R_n - R_n^2) \\ &= \frac{1}{2}[(I - R_{n-1}) + (I - R_n)](R_n - R_{n-1}). \end{aligned}$$

Thanks to Corollary 1.5.3, this last identity allows one to show by induction that $R_{n+1} \geqslant R_n$ for all $n \geqslant 0$. In particular, since $R_0 = 0$, each R_n is positive. We can then apply Corollary 1.5.4 to the sequence $(R_n)_{n=0}^{\infty}$: there exists a symmetric operator R such that $\lim_{n\to\infty} R_n u = Ru$ for all $u \in \mathcal{H}$. On the other hand, the Uniform Boundedness Principle 1.2.3 yields a constant $C \geqslant 0$ such that $\|R_n\| \leqslant C$ for all $n \geqslant 1$, and so

$$\begin{aligned} \left\|R_n^2 u - R^2 u\right\| &= \left\|R_n^2 u - R_n R u + R_n R u - R^2 u\right\| \\ &\leqslant C \left\|R_n u - R u\right\| + \left\|R_n(Ru) - R(Ru)\right\| \xrightarrow{n\to\infty} 0. \end{aligned}$$

It follows that $\lim_{n\to\infty} R_n^2 u = R^2 u$ for all $u \in \mathcal{H}$. Therefore, letting $n \to \infty$ in (1.5.4), we have

$$Ru = Ru + \frac{1}{2}(P - R^2)u \quad \text{for all } u \in \mathcal{H}.$$

That is, $R^2 = P$. Since each R_n is positive, so is R by continuity of the inner product. Furthermore, since each R_n commutes with any bounded operator which commutes with P, this also holds for their limit R.

2. Uniqueness. Suppose that S is also a positive square root of P. Since $P = S^2$, S commutes with P and so with R. Pick any $u \in \mathcal{H}$ and let $v = (R - S)u$. Then

$$\langle Rv, v\rangle + \langle Sv, v\rangle = \langle (R + S)(R - S)u, v\rangle = \left\langle (R^2 - S^2)u, v\right\rangle = 0.$$

Since $\langle Rv, v\rangle \geqslant 0$ and $\langle Sv, v\rangle \geqslant 0$, it follows that $\langle Rv, v\rangle = \langle Sv, v\rangle = 0$. Now consider a positive square root T of R. By symmetry of T we have

$$\|Tv\|^2 = \left\langle T^2 v, v\right\rangle = \langle Rv, v\rangle = 0.$$

Hence $Tv = 0$ and we conclude that $Rv = T(Tv) = 0$. By the same argument we get $Sv = 0$. Finally,

$$\|Ru - Su\|^2 = \langle (R-S)^2 u, u \rangle = \langle (R-S)v, u \rangle = 0,$$

that is $Ru = Su$. Since u is arbitrary, we conclude that $R = S$. □

Lemma 1.5.7 *Let S and T be symmetric operators such that $ST = TS$ and $S^2 = T^2$. Denote by P the projection onto $L = \ker(S-T)$. We then have the following properties.*

(a) Any bounded operator which commutes with $S - T$ commutes with P.

(b) If $Su = 0$ then $Pu = u$.

(c) $P(S+T) = S+T$ and $P(S-T) = 0$.

Proof. Let B be a bounded operator which commutes with $S - T$. We observe that, if $v \in L$ then $Bv \in L$ since $(S-T)Bv = B(S-T)v = 0$. Hence $BPu \in L$ for all $u \in \mathcal{H}$. Therefore, $PBPu = BPu$ for all $u \in \mathcal{H}$, in other words $PBP = BP$. By Problem 1.11, the adjoint B^* also commutes with $S - T$. Hence $B^*P = PB^*P$, and it follows that

$$PB = (B^*P)^* = (PB^*P)^* = PBP = BP.$$

This proves (a).

Now suppose that $Su = 0$ for some $u \in \mathcal{H}$. Then

$$\|Tu\|^2 = \langle T^2 u, u \rangle = \langle S^2 u, u \rangle = \|Su\|^2 = 0,$$

and so $Tu = 0$. Therefore, $(S-T)u = 0$ as well and $u \in L$. It follows that $Pu = u$, proving (b).

To prove (c), observe that, since S and T commute we have, for all $u \in \mathcal{H}$,

$$(S-T)(S+T)u = (S^2 - T^2)u = 0.$$

Hence $(S+T)u \in L$, and so $P(S+T) = S+T$. The last statement follows from the orthogonality of $\ker(S-T)$ and $\operatorname{rge}(S-T)$ by Lemma 1.4.3. □

We are now in a position to prove the following lemma, which will play a crucial role in the spectral decomposition of bounded symmetric operators.

Definition 1.5.8 For a symmetric operator S, we denote by $|S|$ the unique positive square root of S^2. Note that one has $|S|\,S = S\,|S|$ and $|S| \geqslant 0$, for any symmetric operator S. ♦

Lemma 1.5.9 *Let S be a symmetric operator. The projection E_+ onto* $\ker(S-|S|)$ *has the following properties.*

(a) Every bounded operator which commutes with S commutes with E_+.

(b) $SE_+ \geqslant 0$ and $S(I - E_+) \leqslant 0$.

(c) If $Su = 0$ then $E_+u = u$.

Proof. Let C be bounded operator which commutes with S. Since $CS^2 = SCS = S^2C$, C also commutes with S^2. Then, by Theorem 1.5.6, C commutes with $|S|$ and so commutes with $S - |S|$ and finally, by Lemma 1.5.7, with E_+. This proves (a).

Since E_+ is the projection onto $\ker(S - |S|)$, we have that

$$SE_+ = |S|\, E_+. \tag{1.5.6}$$

Now $|S|$ commutes with E_+ by Lemma 1.5.7 and we conclude from Corollary 1.5.3 that $SE_+ \geqslant 0$. On the other hand, by Lemma 1.5.7, $S = (2E_+ - I)\,|S|$, which, together with (1.5.6) implies that

$$S(I - E_+) = -(I - E_+)\,|S|\,. \tag{1.5.7}$$

But $|S|$ also commutes with $I - E_+$ and so again by Corollary 1.5.3 we have $S(I - E_+) \leqslant 0$, finishing the proof of (b).

Part (c) follows directly from Lemma 1.5.7 (b). The proof is complete. □

A good mental picture of Lemma 1.5.9 is obtained from the following definition.

Definition 1.5.10 Let S be a symmetric operator and E_+ be the projection onto $\ker(S - |S|)$. The operator $S_+ = SE_+$ is called the **positive part** of S, while $S_- = S(I - E_+)$ is called the **negative part** of S. ♦

As S and $|S|$ commute with $S - |S|$, Lemma 1.5.7 (a) ensures that both commute with E_+ as well. It therefore follows from (1.5.6) that

$$E_+S = SE_+ = |S|\, E_+ = E_+\, |S|\,.$$

Then, using (1.5.7), one obtains

$$S_+ = \tfrac{1}{2}(S + |S|) \quad \text{and} \quad S_- = S - S_+ = \tfrac{1}{2}(S - |S|).$$

Problems

1. Show that the inner product of a Hilbert space is continuous in each variable, and that the norm is a continuous function.

2. Show that the norm $\|A\|_{\mathcal{B}(\mathcal{H},\mathcal{H}')}$ of an operator $A \in \mathcal{B}(\mathcal{H},\mathcal{H}')$ can be defined equivalently as the infimum of all $C \geqslant 0$ satisfying (1.2.1), or by (1.2.2).

3. Prove that a linear operator $A : \mathcal{H} \to \mathcal{H}$ is bounded if and only if its graph $\mathcal{G}(A)$ is a closed subset of $\mathcal{H} \times \mathcal{H}$.

4. Prove that the sets $\rho(A), \sigma_p(A), \sigma_c(A)$ and $\sigma_r(A)$ are pairwise disjoint, and that their union is the whole complex plane.

5. Let $A \in \mathcal{B}(\mathcal{X})$ and $\lambda \in \mathbb{C}$, as in Definition 1.2.7. Show that, if $\mathcal{X}$ is a Banach space and $\lambda \in \rho(A)$, then $(A - \lambda I)^{-1}$ is defined on the whole space $\mathcal{X}$.
Hint: Use the result proved in Problem 1.3.

6. Prove Theorem 1.3.2.

7. Show that a bounded operator $P : \mathcal{H} \to \mathcal{H}$ is a projection if and only if it is symmetric and idempotent (i.e. $P^2 = P$).

8. Prove Proposition 1.4.8.

9. Prove that, if S is a symmetric operator on $\mathcal{H}$, then any polynomial in S with real coefficients is also symmetric.

10. Let S be a symmetric operator on $\mathcal{H}$, and B a bounded operator on $\mathcal{H}$. Show that $T = B^*SB$ is symmetric.

11. Let S be a symmetric operator on $\mathcal{H}$, B a bounded operator on $\mathcal{H}$. First suppose that $SB = BS$ and show that $SB^* = B^*S$. Then, if both S and B are symmetric, show that SB is symmetric if and only if $SB = BS$.

12. Consider a sequence $(S_n)_{n\geqslant 1}$ of symmetric operators, such that $S_n \to S$ in $\mathcal{B}(\mathcal{H})$ as $n \to \infty$. Show that S is symmetric.
Hint: Apply the triangle inequality to $\|S - S^*\|$.

13. Prove Lemma 1.4.3.
Hint: (Prove and) use the fact that $\mathcal{X}^{\perp\perp} = \overline{\mathcal{X}}$ for any subspace $\mathcal{X} \subset \mathcal{H}$.

14. For $T \in \mathcal{B}(\mathcal{H})$, prove that
$$\langle Tu, u\rangle = 0 \ \forall u \in \mathcal{H} \implies T = 0.$$
Note that it is essential here that $\mathcal{H}$ be complex. Find a counterexample in the real case.
Hint: Write $u = v + \lambda w$ and use special values of $\lambda \in \mathbb{C}$.

15. On ℓ^2, let A be the multiplication operator defined by $(Au)_n = \theta_n u_n$, where $(\theta_n)_{n\geqslant 1} \subset \mathbb{C}$ is a given bounded sequence, with $M := \sup |\theta_n|$. Show that $\sigma_r(A) = \emptyset$ and $\sigma(A) = \overline{(\theta_n)}$. Determine under which condition A is symmetric.
Recall: ℓ^2 is the Hilbert space of all complex sequences $(u_1, u_2, \dots)$ such that $\sum_{n\geqslant 1} |u_n|^2 < \infty$, endowed with the inner product $\langle u, v\rangle = \sum_{n\geqslant 1} u_n \overline{v_n}$.

16. Find an operator $T : C[0,1] \to C[0,1]$ such that $\sigma(T) = [a, b]$, with $a < b$ given.

17. We define two operators $S, T : \ell^2 \to \ell^2$ by
$$(Su)_n = u_{n+1},\ n \geqslant 1 \quad \text{(left shift)}; \qquad (Tu)_1 = 0,\ (Tu)_n = u_{n-1},\ n \geqslant 2 \quad \text{(right shift)}.$$
   (a) Show that S and T are bounded. Compute $\|S\|$ and $\|T\|$.
   (b) Find S^* and T^*.
   (c) Find $\sigma_p(S), \sigma_r(S)$ and $\sigma_c(S)$.

18. Show that the multiplication operator $X : L^2[0,1] \to L^2[0,1]$ defined by
$$(Xu)(x) = xu(x), \quad x \in [0,1],$$
is a bounded symmetric operator without eigenvalues. Find the spectrum of X.

19. Let $S : \mathcal{H} \to \mathcal{H}$ be bounded and symmetric. Prove that all the eigenvalues (if any) of S are real. Show that eigenvectors of S corresponding to distinct eigenvalues are orthogonal.

20. Deduce Corollary 1.4.12 from Theorem 1.2.10.

21. Show that the relation $\geqslant$ introduced in Definition 1.5.1 defines a partial order among the symmetric operators on $\mathcal{H}$.
Recall: A partial order is a binary relation which is reflexive, transitive and antisymmetric.
Hint: Use Theorem 1.4.9.

22. For two projections P and Q of $\mathcal{H}$, show that $P \leqslant Q \iff \operatorname{rge} P \subseteq \operatorname{rge} Q$.
Hint: Use Theorem 1.3.2.

23. Let $A : \ell^2 \to \ell^2$ be defined by $(u_1, u_2, \dots) \mapsto (0, 0, u_1, u_2, \dots)$. Show that A is bounded and compute its norm. Is A symmetric? Find $B : \ell^2 \to \ell^2$ such that $A = B^2$.

24. Consider the operator $A : L^2[0,1] \to L^2[0,1]$ defined by

$$(Au)(x) = a(x)u(x), \quad x \in [0,1],$$

where $a \in L^\infty[0,1]$. Show that:

(a) $\sigma_p(A) = \big\{\lambda \in \mathbb{C}\,;\, \big|\{x \in [0,1]\,;\, a(x) = \lambda\}\big| > 0\big\}$;

(b) $\sigma_r(A) = \emptyset$;

(c) $\sigma(A) = \operatorname{essrge}(a)$.

Here, $|E|$ is the Lebesgue measure of a Borel set $E \subset [0,1]$, and $\operatorname{essrge}(a)$ is the *essential range* of a, defined as

$$\operatorname{essrge}(a) = \big\{\lambda \in \mathbb{C}\,;\, \text{for all } \epsilon > 0 : \big|\{x \in [0,1]\,;\, |a(x) - \lambda| < \epsilon\}\big| > 0\big\}.$$

Find a condition on a for A to be positive and, in this case, find the square root of A.

Chapter 2

The spectral decomposition of symmetric bounded operators

The goal of this chapter is to present the proof of the spectral theorem for (bounded) symmetric operators. The main idea is that a symmetric operator A can always be represented as a sum of projections, indexed by a real parameter running through $\sigma(A)$. In general, due to the presence of continuous spectrum, this representation involves a continuous sum, viz. a Riemann-Stieltjes integral. The projections are uniquely determined by A in the form of a spectral family. In the first part of the chapter we define the integral with respect to a general spectral family. Then we prove the spectral theorem, following [Fri82].

2.1 Integration with respect to a spectral family

We start with the definition of a spectral family.

Definition 2.1.1 A **spectral family** on $\mathcal{H}$ is a mapping $E : \mathbb{R} \to \mathcal{B}(\mathcal{H})$, denoted $(E_\lambda)_{\lambda\in\mathbb{R}}$ and satisfying the following properties.

(i) E_λ is a projection for all $\lambda \in \mathbb{R}$.

(ii) If $\lambda < \mu$, then $E_\lambda \leqslant E_\mu$.

(iii) The family $(E_\lambda)_{\lambda\in\mathbb{R}}$ is **strongly left-continuous**, i.e.

$$\lim_{\lambda\nearrow\mu} E_\lambda u = E_\mu u, \quad \text{for all } \mu \in \mathbb{R},\ u \in \mathcal{H}.$$

(iv) There exist $m, M \in \mathbb{R}$ such that $E_\lambda = 0$ for all $\lambda < m$ and $E_\lambda = I$ for all $\lambda > M$.

Two numbers $m, M \in \mathbb{R}$ satisfying (iv) are respectively called a **lower bound** and an **upper bound** of the spectral family. ♦

We consider a spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ with bounds $m, M \in \mathbb{R}$. Let f be a continuous complex-valued function defined on $[m, M]$. We extend f continuously to $[m, M+1]$ and we denote this extension by f. We now fix $0 < \varepsilon < 1$ and let Π be an arbitrary partition of $[m, M+\varepsilon]$, i.e. a finite sequence of numbers $(\lambda_k)_{k=0}^n$ such that

$$m = \lambda_0 < \lambda_1 < \cdots < \lambda_{n-1} < \lambda_n = M + \varepsilon.$$

We call **size** of the partition Π the number $|\Pi|$ defined as

$$|\Pi| = \max_{k=1,\ldots,n} \lambda_k - \lambda_{k-1}.$$

We then choose real numbers $\mu_1, \ldots, \mu_k$ such that

$$\mu_k \in [\lambda_{k-1}, \lambda_k] \quad \text{for each } k = 1, \cdots, n,$$

and form the sum

$$S_\Pi = \sum_{k=1}^{n} f(\mu_k)(E_{\lambda_k} - E_{\lambda_{k-1}}).$$

The following lemma ensures that S_Π converges in $\mathcal{B}(\mathcal{H})$ as $|\Pi| \to 0$, for any partition Π of $[m, M+\varepsilon]$.

Lemma 2.1.2 *Consider $(E_\lambda)_{\lambda\in\mathbb{R}}$ a spectral family with bounds $m, M \in \mathbb{R}$, and $f : [m, M] \to \mathbb{C}$ a continuous function. Let $0 < \varepsilon < 1$ and $f : [m, M+\varepsilon] \to \mathbb{C}$ a continuous extension of f. There is a unique bounded operator S such that, for all $\eta > 0$ there exists $\delta > 0$ such that for any partition Π of $[m, M+\varepsilon]$ satisfying $|\Pi| < \delta$, one has*

$$\|S_\Pi - S\| < \eta.$$

Furthermore, the operator S is independant of the extension chosen for f, of the choice of ε, and of the choice of the μ_ks.

Proof. Fix $\eta > 0$ arbitrary. Since f is uniformly continuous on the compact interval $[m, M+\varepsilon]$, there exists $\delta_\eta > 0$ such that

$$\text{for all } \lambda, \lambda' \in [m, M+\varepsilon],\ |\lambda - \lambda'| < \delta_\eta \implies |f(\lambda) - f(\lambda')| < \frac{\eta}{2}. \tag{2.1.1}$$

Part 1. We start by showing that, for any partitions Π and Π' of $[m, M+\varepsilon]$,

$$|\Pi|, |\Pi'| < \delta_\eta \implies \|S_\Pi - S_{\Pi'}\| < \eta, \tag{2.1.2}$$

independently of the points μ_k chosen to construct the sums S_Π and $S_{\Pi'}$.

Let us write $\Pi = (\lambda_k)_{k=0}^n$ and fix $\mu_1, \ldots, \mu_n$ with $\mu_i \in [\lambda_{i-1}, \lambda_i]$ for all $i = 1, \ldots, n$. We form a partition $\overline{\Pi} = (\bar{\lambda}_j)_{j=0}^{\bar{n}}$ as the union of Π and Π'. We denote by $0 = k_0 < k_1 < \cdots < k_j < k_n = \bar{n}$ the subsequence of indices satisfying $\bar{\lambda}_{k_i} = \lambda_i$ for $i = 0, \ldots, n$.

We then pick up numbers

$$\bar{\mu}_i \in [\bar{\lambda}_{i-1}, \bar{\lambda}_i] \quad i = 1, \cdots, k_n.$$

The sum associated with $\overline{\Pi}$ and the $\bar{\mu}_j$s is given by

$$S_{\overline{\Pi}} = \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} f(\bar{\mu}_j)(E_{\bar{\lambda}_j} - E_{\bar{\lambda}_{j-1}}).$$

Since, for all $i = 1, \ldots, n$,

$$\sum_{j=k_{i-1}+1}^{k_i} E_{\bar{\lambda}_j} - E_{\bar{\lambda}_{j-1}} = E_{\bar{\lambda}_{k_i}} - E_{\bar{\lambda}_{k_{i-1}}} = E_{\lambda_i} - E_{\lambda_{i-1}},$$

we can write S_Π as

$$S_\Pi = \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} f(\mu_i)(E_{\bar{\lambda}_j} - E_{\bar{\lambda}_{j-1}}).$$

Since $|\Pi| < \delta_\eta$, it now follows from (2.1.1) that, for all $i = 1, \ldots, n$ and all $j = k_{i-1} + 1, \ldots, k_i$,

$$|\mu_i - \bar{\mu}_j| < \lambda_{k_i} - \lambda_{k_{i-1}} < \delta_\eta \implies |f(\mu_i) - f(\bar{\mu}_j)| < \frac{\eta}{2}.$$

Moreover, since $E_m = 0$ and $E_{M+\varepsilon} = I$ we have, for all $u \in \mathcal{H}$ such that $\|u\| = 1$:

$$\begin{aligned} |\langle (S_\Pi - S_{\overline{\Pi}})u, u\rangle| &= \left| \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} [f(\mu_i) - f(\bar{\mu}_j)] \left\langle (E_{\bar{\lambda}_j} - E_{\bar{\lambda}_{j-1}})u, u \right\rangle \right| \\ &\leqslant \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} |f(\mu_i) - f(\bar{\mu}_j)| \left\langle (E_{\bar{\lambda}_j} - E_{\bar{\lambda}_{j-1}})u, u \right\rangle \\ &< \frac{\eta}{2} \left\langle \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} (E_{\bar{\lambda}_j} - E_{\bar{\lambda}_{j-1}})u, u \right\rangle \\ &= \frac{\eta}{2} \langle (E_{M+\varepsilon} - E_m)u, u\rangle = \frac{\eta}{2} \|u\|^2 = \frac{\eta}{2}. \end{aligned}$$

By Theorem 1.4.9 it follows that $\|S_\Pi - S_{\overline{\Pi}}\| < \frac{\eta}{2}$. Similarly we have $\|S_{\Pi'} - S_{\overline{\Pi}}\| < \frac{\eta}{2}$ and so

$$\|S_\Pi - S_{\Pi'}\| \leqslant \|S_\Pi - S_{\overline{\Pi}}\| + \|S_{\overline{\Pi}} - S_{\Pi'}\| < \eta,$$

as advertised.

Part 2. Let us now consider a sequence $(\Pi_n)_{n=1}^\infty$ of partitions of $[m, M+\varepsilon]$ such that $\lim_{n\to\infty} |\Pi_n| = 0$. The operators $(S_{\Pi_n})_{n=1}^\infty$ then form a Cauchy sequence in $\mathcal{B}(\mathcal{H})$. Indeed, there is an $N \geqslant 1$ such that $|\Pi_n| < \delta_\eta$ for all $n \geqslant N$. Hence, by (2.1.2), for all $n, n' \geqslant N$ there holds $\left\|S_{\Pi_n} - S_{\Pi_{n'}}\right\| < \eta$. Since $\mathcal{B}(\mathcal{H})$ is complete, the sequence $(S_{\Pi_n})_{n=1}^\infty$ has a limit $S \in \mathcal{B}(\mathcal{H})$. In particular, there exists $N_\eta \geqslant 1$ such that $|\Pi_{N_\eta}| < \delta_{\frac{\eta}{2}}$ and $\left\|S_{\Pi_{N_\eta}} - S\right\| < \frac{\eta}{2}$. Therefore, in view of (2.1.2), any partition Π smaller than $\delta_{\frac{\eta}{2}}$ satisfies

$$\|S_\Pi - S\| \leqslant \left\|S_\Pi - S_{\Pi_{N_\eta}}\right\| + \left\|S_{\Pi_{N_\eta}} - S\right\| < \eta.$$

The limit S does not depend on the extension of f, neither on ε, for $E_\lambda - E_\mu = 0$ for all $M < \mu < \lambda$. □

Thanks to Lemma 2.1.2, we can now make the following definition.

Definition 2.1.3 Consider a spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ with bounds $m, M \in \mathbb{R}$ and $f : [m, M] \to \mathbb{C}$ continuous. The limit operator $S \in \mathcal{B}(\mathcal{H})$ obtained from Lemma 2.1.2 is called the **integral of f with respect to the spectral family** $(E_\lambda)_{\lambda\in\mathbb{R}}$. It is written

$$S = \int_m^{M+\varepsilon} f(\lambda)\,\mathrm{d}E_\lambda. \ \blacklozenge$$

Observe that, for any spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$, one has

$$\int_m^{M+\varepsilon} \mathrm{d}E_\lambda = E_{M+\varepsilon} - E_m = I.$$

2.2 The spectral theorem for symmetric operators

We are now ready to prove the spectral theorem for bounded symmetric operators.

Theorem 2.2.1 (Spectral Theorem I) *Let S be a bounded symmetric operator. There exists a unique spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ with the following properties.*

(a) The lower and upper bounds m, M of S are respectively lower and upper bounds for $(E_\lambda)_{\lambda\in\mathbb{R}}$.

(b) Every bounded operator which commutes with S commutes with E_λ, for all $\lambda \in \mathbb{R}$.

(c) For all $u \in \mathcal{H}$ there exists

$$E_{\mu+0}u = \lim_{\lambda \searrow \mu} E_\lambda u.$$

(d) The operator S has the representation

$$S = \int_m^{M+\varepsilon} \lambda \, \mathrm{d}E_\lambda.$$

*The family $(E_\lambda)_{\lambda\in\mathbb{R}}$ is called **the spectral family of** S.*

Proof. For all $\lambda \in \mathbb{R}$, let $E_+(\lambda)$ be the projection onto $\ker[(S - \lambda I) - |S - \lambda I|]$, which was studied in Lemma 1.5.9. Observe that $E_+(\lambda)$ is uniquely determined in this way. We will show that the projections $E_\lambda = I - E_+(\lambda)$ form a spectral family satisfying (a)–(d).

Since $E_+(\lambda)$ commutes with all bounded operators which commute with S, this is also true for E_λ. Hence, (b) is readily satisfied. In particular, $E_\mu E_\lambda = E_\lambda E_\mu$ for all $\mu, \lambda \in \mathbb{R}$.

Le us now show that, if $\lambda < \mu$ then $E_\lambda \leqslant E_\mu$, i.e. part (ii) of Definition 2.1.1 holds. Suppose $\lambda < \mu$ and let $P = E_\lambda(I - E_\mu)$. We will show that $P = 0$. Firstly, we have

$$E_\lambda P = P, \qquad (I - E_\mu)P = P. \tag{2.2.1}$$

Moreover, by definition of E_λ and E_μ, Lemma 1.5.9 implies that

$$(S - \lambda I)E_\lambda \leqslant 0, \qquad (S - \mu I)(I - E_\mu) = (S - \mu I)E_+(\mu) \geqslant 0. \tag{2.2.2}$$

Choosing $u \in \mathcal{H}$ and letting $v = Pu$, it follows from (2.2.1) that

$$E_\lambda v = E_\lambda P u = P u = v.$$

Similarly, $(I - E_\mu)v = v$. Therefore, by (2.2.2),

$$\begin{aligned} \langle (S - \lambda I)v, v\rangle &= \langle (S - \lambda I)E_\lambda v, v\rangle \leqslant 0, \\ \langle (S - \mu I)v, v\rangle &= \langle (S - \mu I)(I - E_\mu)v, v\rangle \geqslant 0. \end{aligned}$$

We deduce that

$$(\mu - \lambda)\,\langle v, v\rangle = \langle (S - \lambda I)v, v\rangle - \langle (S - \mu I)v, v\rangle \leqslant 0.$$

However, since $\mu > \lambda$, we must have $Pu = v = 0$. Since u is arbitrary, we have shown that $P = 0$. By definition of P, this means that $E_\lambda = E_\lambda E_\mu$ which, by Theorem 1.3.2, is equivalent to $E_\lambda \leqslant E_\mu$. The family $(E_\lambda)_{\lambda\in\mathbb{R}}$ therefore satisfies condition (ii) of Definition 2.1.1.

We now show that $(E_\lambda)_{\lambda\in\mathbb{R}}$ satisfies (c). Let $u \in \mathcal{H}$. Since $E_\lambda \leqslant E_\mu$ for $\lambda < \mu$, $\langle E_\lambda u, u\rangle$ is a positive increasing function of λ. Therefore, for all $\mu \in \mathbb{R}$, it has a limit from the left:

$$\lim_{\lambda\nearrow\mu} \langle E_\lambda u, u\rangle = \sup_{\lambda<\mu} \langle E_\lambda u, u\rangle = l_\mu.$$

Hence, for all $\eta > 0$ there exists $\delta > 0$ such that $0 < \mu - \lambda < \delta$ implies $l_\mu - \langle E_\lambda u, u\rangle < \frac{1}{2}\eta$. It follows that, for $\mu - \delta < \lambda < \nu < \mu$,

$$\|E_\nu u - E_\lambda u\|^2 = \left\langle (E_\nu - E_\lambda)^2 u, u\right\rangle = \langle (E_\nu - E_\lambda)u, u\rangle \leqslant |\langle E_\nu u, u\rangle - l_\mu| + |l_\mu - \langle E_\lambda u, u\rangle| < \eta.$$

Therefore, since $\mathcal{H}$ is complete, the limit

$$\lim_{\lambda\nearrow\mu} E_\lambda u = E_{\mu-0} u \quad \text{exists for all } u \in \mathcal{H}.$$

Similarly, for all $u \in \mathcal{H}$, the limit

$$\lim_{\lambda\searrow\mu} E_\lambda u = E_{\mu+0} u$$

exists as well. This proves (c) and half of part (iii) of Definition 2.1.1.

We now complete the proof of part (iii) of Definition 2.1.1, i.e. we show that $(E_\lambda)_{\lambda\in\mathbb{R}}$ is strongly left-continuous. If $\lambda < \mu$, we write $E_\Delta = E_\mu - E_\lambda$ and we have

$$E_\mu E_\Delta = E_\Delta \quad \text{ et } \quad (I - E_\lambda)E_\Delta = E_\mu - E_\lambda - E_\lambda E_\mu + E_\lambda^2 = E_\Delta. \tag{2.2.3}$$

Using (2.2.2), $E_\Delta \geqslant 0$, and the fact that a composition of commuting positive operators is positive (Corollary 1.5.3),

$$\begin{aligned} (S - \mu I)E_\Delta &= (S - \mu I)E_\mu E_\Delta \leqslant 0,\\ (S - \lambda I)E_\Delta &= (S - \lambda I)(I - E_\lambda)E_\Delta \geqslant 0. \end{aligned}$$

In consequence, we have the following inequalities

$$\lambda E_\Delta \leqslant SE_\Delta \leqslant \mu E_\Delta \qquad \text{if } \lambda < \mu. \tag{2.2.4}$$

Consider now the operator $E_{\Delta_0} = E_\mu - E_{\mu-0}$. We want to show that $E_{\Delta_0} = 0$. Observe that, for all $u \in \mathcal{H}$,

$$\lim_{\lambda\nearrow\mu} E_\Delta u = E_{\Delta_0} u.$$

Thus, letting $\lambda \nearrow \mu$ in (2.2.4), we get

$$\mu E_{\Delta_0} \leqslant SE_{\Delta_0} \leqslant \mu E_{\Delta_0}.$$

In consequence, $\langle (S-\mu I)E_{\Delta_0}u, u\rangle = 0$ for all $u \in \mathcal{H}$ and Theorem 1.4.9 then ensures that

$$(S-\mu I)E_{\Delta_0} = 0.$$

Let us now fix $u \in \mathcal{H}$ and set $v = E_{\Delta_0}u$, so that $(S-\mu I)v = 0$ and, by part (c) of Lemma 1.5.9, we have

$$E_\mu E_{\Delta_0}u = E_\mu v = (I - E_+(\mu))v = 0.$$

Finally, it follows from (2.2.3) that

$$E_{\Delta_0}u = \lim_{\lambda\nearrow\mu} E_\Delta u = \lim_{\lambda\nearrow\mu} E_\mu E_\Delta u = E_\mu E_{\Delta_0}u = 0.$$

Since $u \in \mathcal{H}$ is arbitrary, we have indeed shown that $E_{\Delta_0} = 0$.

Let us now prove part (a), thereby showing that $(E_\lambda)_{\lambda\in\mathbb{R}}$ satisfies condition (iv) of Definition 2.1.1. Suppose by contradiction that $\lambda < m$ and $E_\lambda \neq 0$. Then there exists $u \in \mathcal{H}$ such that $E_\lambda u \neq 0$ and we let $v = E_\lambda u$. We can suppose that $\|v\| = 1$, and we obtain from (2.2.2) that

$$\langle Sv, v\rangle - \lambda = \langle (S-\lambda I)v, v\rangle = \langle (S-\lambda I)E_\lambda u, u\rangle \leqslant 0.$$

Therefore, by definition of the lower bound of S,

$$m \leqslant \langle Sv, v\rangle \leqslant \lambda,$$

a contradiction. Hence $E_\lambda = 0$ for all $\lambda < m$. Suppose now by contradiction that $\lambda > M$ and $E_\lambda \neq I$. Then there is $u \in \mathcal{H}$ such that $w = (I - E_\lambda)u \neq 0$. We can again suppose that $\|w\| = 1$. Hence, by (2.2.2),

$$\langle Sw, w\rangle - \lambda = \langle (S-\lambda I)w, w\rangle = \langle (S-\lambda I)(I-E_\lambda)u, u\rangle \geqslant 0.$$

Therefore, by definition of the upper bound of S,

$$\lambda \leqslant \langle Sw, w\rangle \leqslant M,$$

a contradiction. Hence $E_\lambda = I$ for all $\lambda > M$.

We now prove part (d). Consider a sequence of partitions $(\Pi_l)_{l=1}^{\infty}$, denoted as

$$\Pi_l : \quad m = \lambda_0^l < \lambda_1^l < \cdots < \lambda_{n_l-1}^l < \lambda_{n_l}^l = M + \varepsilon,$$

and satisfying $\lim_{l\to\infty} |\Pi_l| = 0$. For all $l \geqslant 1$ and all $k = 1, \ldots, n_l$, writing $E_{\Delta_k^l} = E_{\lambda_k^l} - E_{\lambda_{k-1}^l}$, it follows from (2.2.4) that

$$\lambda_{k-1}^l E_{\Delta_k^l} \leqslant S E_{\Delta_k^l} \leqslant \lambda_k^l E_{\Delta_k^l}.$$

Hence, for fixed $l \geqslant 1$, since $\sum_{k=1}^{n_l} E_{\Delta_k^l} = I$, summation over $k = 1, \ldots, n_l$ yields

$$S_{\Pi_l} \leqslant S \leqslant \widetilde{S}_{\Pi_l},$$

where the sum S_{Π_l} is taken over the partition Π_l, with the function $f(\lambda) = \lambda$ and the choice of $\mu_k = \lambda_{k-1}$, while $\widetilde{S}_{\Pi_l}$ is computed with $\widetilde{\mu}_k = \lambda_{k-1}$. Hence, letting $l \to \infty$, Lemma 2.1.2 gives

$$\int_m^{M+\varepsilon} \lambda \, \mathrm{d}E_\lambda \leqslant S \leqslant \int_m^{M+\varepsilon} \lambda \, \mathrm{d}E_\lambda.$$

It follows that, for all $u \in \mathcal{H}$,

$$\left\langle \left(S - \int_m^{M+\varepsilon} \lambda \, \mathrm{d}E_\lambda \right) u, u \right\rangle = 0.$$

Therefore, by Theorem 1.4.9 and Problem 1.12,

$$S = \int_m^{M+\varepsilon} \lambda \, \mathrm{d}E_\lambda.$$

It remains to prove the uniqueness of the spectral family of S. This will be done in the problems using the following lemma. □

Lemma 2.2.2 *Consider a symmetric operator S and let $(E_\lambda)_{\lambda \in \mathbb{R}}$ be a corresponding spectral family satisfying parts (a) to (d) of the theorem. For any real polynomial[1] p, there holds*

$$p(S) = \int_m^{M+\varepsilon} p(\lambda) \, \mathrm{d}E_\lambda.$$

Proof. We need only show that the conclusion holds for any monomial $p(\lambda) = \lambda^l$ with $l \geqslant 0$. We already know that this is true for $l = 0$, and the theorem gives the result for $l = 1$. We prove the result by induction, assuming it holds for $p(\lambda) = \lambda^l$ and inferring it holds for $p(\lambda) = \lambda^{l+1}$. Fix $0 < \eta < 1$ arbitrary. Using the theorem and the induction hypothesis, there exists $\delta > 0$ such that for any partition $\Pi = (\lambda_k)_{k=0}^n$ satisfying $|\Pi| < \delta$, there holds

$$\|S - \textstyle\sum_{k=1}^n \lambda_k E_{\Delta_k}\| < \eta \quad \text{and} \quad \left\|S^l - \textstyle\sum_{k=1}^n \lambda_k^l E_{\Delta_k}\right\| < \eta,$$

[1] A real polynomial is a polynomial with real coefficients.

where $E_{\Delta_k} = E_{\lambda_k} - E_{\lambda_{k-1}}$. Let us write $T = \sum_{k=1}^n \lambda_k E_{\Delta_k}$ and $T^{(l)} = \sum_{k=1}^n \lambda_k^l E_{\Delta_k}$. Observe that

$$\left\| \Big(S^l - \sum_{k=1}^n \lambda_k^l E_{\Delta_k} \Big) \Big(S - \sum_{k=1}^n \lambda_k E_{\Delta_k} \Big) \right\| \le \left\| S^l - T^{(l)} \right\| \|S - T\| < \eta^2.$$

Hence,

$$\left\| S^{l+1} + \sum_{k=1}^n \lambda_k^l E_{\Delta_k} \sum_{k=1}^n \lambda_k E_{\Delta_k} - S^l T - S T^{(l)} \right\| < \eta^2.$$

Now for all k,

$$E_{\Delta_k}^2 = E_{\lambda_k}^2 - 2E_{\lambda_k} E_{\lambda_{k-1}} + E_{\lambda_{k-1}}^2 = E_{\Delta_k},$$

while for all $i \neq j$,

$$E_{\Delta_i} E_{\Delta_j} = E_{\lambda_i} E_{\lambda_j} - E_{\lambda_i} E_{\lambda_{j-1}} - E_{\lambda_{i-1}} E_{\lambda_j} + E_{\lambda_{i-1}} E_{\lambda_{j-1}} = 0.$$

Therefore, $T^{(l)} T = \sum_{k=1}^n \lambda_k^{l+1} E_{\Delta_k}$. Using

$$\left\| S^l T - S^{l+1} \right\| \le \left\| S^l \right\| \|T - S\| \le \|S\|^l \eta$$

and

$$\left\| S T^{(l)} - S^{l+1} \right\| \le \|S\| \left\| T^{(l)} - S^l \right\| \le \|S\| \eta,$$

we have

$$\begin{aligned} \left\| \sum_{k=1}^n \lambda_k^{l+1} E_{\Delta_k} - S^{l+1} \right\| &= \left\| (S^{l+1} + T^{(l)} T - S^l T - S T^{(l)}) + (S^l T - S^{l+1}) + (S T^{(l)} - S^{l+1}) \right\| \\ &\le \eta^2 + \|S\|^l \eta + \|S\| \eta \\ &< \eta(\eta + \|S\|^l + \|S\|). \end{aligned}$$

It then follows Lemma 2.1.2 that

$$S^{l+1} = \int_m^{M+\varepsilon} \lambda^{l+1} \, \mathrm{d}E_\lambda,$$

which concludes the proof. □

Corollary 2.2.3 *Let S be a symmetric operator and $(E_\lambda)_{\lambda \in \mathbb{R}}$ a corresponding spectral family satisfying parts (a) to (d) of the theorem. For all $u, v \in \mathcal{H}$ and all real polynomial p, we have*

$$\langle p(S) u, v \rangle = \int_m^{M+\varepsilon} p(\lambda) \, \mathrm{d}\langle E_\lambda u, v \rangle \,. \tag{2.2.5}$$

Proof. See Problem 2.1. □

The right-hand side of (2.2.5) is a Riemann-Stieltjes integral (cf. Appendix A). For all $u \in \mathcal{H}$ the function $\lambda \mapsto \langle E_\lambda u, u\rangle$ is increasing and so, in particular, of bounded variations in $[m, M+\varepsilon]$. Its total variation in $[m, M+\varepsilon]$ is equal to $\|u\|^2$. Furthermore, the value of the right-hand side does not depend on the choice of continuous extension of p nor on the value of ε. This follows again from the fact that, for $\lambda > \mu > M$, $E_\lambda - E_\mu = 0$.

2.3 Further properties of the spectral family

We conclude this chapter by discussing the relation between the spectral family of a symmetric operator S and the two components of its spectrum, $\sigma_p(S)$ and $\sigma_c(S)$. Notice that, by Remark 1.4.13, $\sigma_r(S) = \emptyset$ for any symmetric operator S.

Theorem 2.3.1 *Let S be a symmetric operator and $(E_\lambda)_{\lambda\in\mathbb{R}}$ the corresponding spectral family. The real number λ_0 is an eigenvalue of S if and only if the mapping $\lambda \mapsto E_\lambda$ is discontinuous at $\lambda = \lambda_0$ (i.e. $E_{\lambda_0} \neq E_{\lambda_0+0}$). In this case,*

$$\ker(S - \lambda_0 I) = \mathrm{rge}(E_{\lambda_0+0} - E_{\lambda_0}). \tag{2.3.1}$$

Proof. See Problem 2.6. □

The next result completes our discussion by a remarkable characterization of the resolvent set.

Theorem 2.3.2 *Let S be a symmetric operator and $(E_\lambda)_{\lambda\in\mathbb{R}}$ the corresponding spectral family. Then a real number λ_0 belongs to the resolvent set $\rho(S)$ of S if and only if there exists $\epsilon > 0$ such that the mapping $\lambda \mapsto E_\lambda$ is constant on the interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$.*
We call such a λ_0 a point of constancy *of $(E_\lambda)_{\lambda\in\mathbb{R}}$.*

Our final result follows immediately from the two previous theorems, and Remark 1.4.13

Corollary 2.3.3 *Let S be a symmetric operator and $(E_\lambda)_{\lambda\in\mathbb{R}}$ the corresponding spectral family. A real number λ_0 belongs to the continuous spectrum $\sigma_c(S)$ of S if and only if λ_0 is neither a point of constancy, nor a point of discontinuity of $(E_\lambda)_{\lambda\in\mathbb{R}}$.*

Problems

1. Prove Corollary 2.2.3.
   Hint: Use Lemma 2.2.2 and Theorem A.2.1.

2. Prove the uniqueness of the spectral family in Theorem 2.2.1.
   Hint: Use the Weierstrass approximation theorem, Corollary 2.2.3 and Theorems A.3.3 and A.3.4.

3. We consider again the operator $X : L^2[0,1] \to L^2[0,1]$, $(Xu)(x) = xu(x)$. In Problem 1.18 we showed that $\sigma(X) = \sigma_c(X) = [0,1]$. Prove that the spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ of X is given by

$$E_\lambda u = \begin{cases} 0 & \text{if} \quad \lambda \leqslant 0, \\ \chi_{[0,\lambda]} u & \text{if} \quad \lambda \in (0,1], \\ u & \text{if} \quad \lambda > 1, \end{cases}$$

   where $\chi_{[0,\lambda]}$ is the characteristic function of the interval $[0,\lambda]$.
   Hint: Start by showing that $|X - \lambda I|$ is the operator $T_\lambda : L^2[0,1] \to L^2[0,1]$ defined by $(T_\lambda u)(x) = |x - \lambda| u(x)$, $x \in [0,1]$. Then find the corresponding projection $E_+(\lambda)$ appearing in the proof of Theorem 2.2.1.

4. From Theorems 2.2.1, 2.3.1 and 2.3.2, deduce the structure of the spectral family and the spectral decomposition (a) of a (finite-dimensional) Hermitian matrix, (b) of a compact operator having infinitely many eigenvalues.

5. Verify that the spectral family obtained in Problem 2.3 satisfies the conclusions of Theorems 2.3.1 and 2.3.2.

6. The goal of this problem is to prove Theorem 2.3.1. First, the whole proof can be reduced to checking (2.3.1). Explain why. Then, to prove (2.3.1), one can proceed as follows.
   (a) To prove that $\ker(S - \lambda_0 I) \supset \operatorname{rge}(E_{\lambda_0+0} - E_{\lambda_0})$, use inequality (2.2.4).
   (b) The other inclusion is more involved. We need to show that, if $u \in \ker(S - \lambda_0 I)$ then $F_0 u = u$, where we have put $F_0 = E_{\lambda_0+0} - E_{\lambda_0}$ — explain why this is enough. To do this, use Corollary 2.2.3 with $p(\lambda) = (\lambda - \lambda_0)^2$ to prove that $\langle E_{\lambda_0-\varepsilon} u, u\rangle = \langle u - E_{\lambda_0+\varepsilon} u, u\rangle = 0$, for any $\varepsilon > 0$.

7. In this problem we shall prove Theorem 2.3.2. We will use the fact that $\lambda_0 \in \rho(S)$ if and only if there exists $\gamma > 0$ such that

$$\|(S - \lambda_0 I)u\| \geqslant \gamma \|u\|, \quad u \in \mathcal{H}. \tag{2.3.2}$$

(a) To prove that the constancy condition implies $\lambda_0 \in \rho(S)$, use Corollary 2.2.3 with $p(\lambda) = (\lambda - \lambda_0)^2$, and (2.3.2).

(b) To show that λ_0 is a point of constancy if it is in the resolvent set, proceed by contradiction using again (2.3.2) and Corollary 2.2.3, with $p(\lambda) = (\lambda - \lambda_0)^2$ and a suitably chosen u. *Hint:* The identities $E_\lambda E_\mu = E_\mu E_\lambda = E_\lambda,\ \lambda \leqslant \mu$, can be useful here.

Chapter 3

The spectral decomposition of selfadjoint operators

In this chapter we will extend the spectral theorem to general (unbounded) selfadjoint operators. We start by introducing the properties of unbounded operators that will be relevant to our analysis.

3.1 Unbounded linear operators

Some natural operators in a Hilbert space $\mathcal{H}$ fail to be bounded. Those are typically only defined on a subspace of $\mathcal{H}$, called their *domain*.

Definition 3.1.1 Let $\mathcal{H}$ be a Hilbert space. A **linear operator** (or simply **operator**) in $\mathcal{H}$ is a mapping $T : \mathfrak{D}_T \subseteq \mathcal{H} \to \mathcal{H}$ satisfying:

(i) $\mathfrak{D}_T$ is a subspace of $\mathcal{H}$ called the **domain** of T;

(ii) for all $u, v \in \mathfrak{D}_T$,
$$T(u + v) = Tu + Tv;$$

(iii) for all $\lambda \in \mathbb{C}$ and all $u \in \mathfrak{D}_T$,
$$T(\lambda u) = \lambda T u. \blacklozenge$$

We also adapt in the obvious way the notions of range and kernel. Let T be a linear operator acting in $\mathcal{H}$. Its **range** is defined as

$$\operatorname{rge}(T) = \{Tu\, ;\, u \in \mathfrak{D}_T\},$$

and its **kernel** as

$$\ker(T) = \{u \in \mathfrak{D}_T \,;\, Tu = 0\}.$$

Consider two operators T and T' in $\mathcal{H}$, with respective domains $\mathfrak{D}_T$ and $\mathfrak{D}_{T'}$. If

$$\mathfrak{D}_T \subseteq \mathfrak{D}_{T'} \quad \text{and} \quad Tu = T'u \quad \text{for all } u \in \mathfrak{D}_T,$$

we say that T' is an **extension** of T, and we write

$$T \subseteq T' \text{ or } T' \supseteq T.$$

This is easily seen to define a partial order on the operators of $\mathcal{H}$. In particular, two operators T and T' are equal if and only if $\mathfrak{D}_T = \mathfrak{D}_{T'}$ and $Tu = T'u$ for all $u \in \mathfrak{D}_T$.

As will be seen in Problem 3.1, an operator which is bounded on its domain can always be extended to the whole of $\mathcal{H}$, so Definition 3.1.1 doesn't bring anything new in this case. Its role is really to allow for *unbounded operators*.

Similarly to the bounded case, we define the **graph** of an operator $T : \mathfrak{D}_T \subseteq \mathcal{H} \to \mathcal{H}$ as

$$\boldsymbol{G}_T = \{(x, Tx) \,;\, x \in \mathfrak{D}_T\} \subset \mathcal{H} \times \mathcal{H},$$

and we equip the product Hilbert space $\mathcal{H} \times \mathcal{H}$ with the natural inner product inherited from $\mathcal{H}$:

$$\Big\langle (u,v), (u',v') \Big\rangle := \langle u, u' \rangle + \langle v, v' \rangle \,, \quad (u,v), (u',v') \in \mathcal{H}.$$

We say that T is **closed** if its graph is a closed subset of $\mathcal{H} \times \mathcal{H}$. If T_1 and T_2 are linear operators, then obviously $T_1 \subseteq T_2$ is equivalent to $\boldsymbol{G}_{T_1} \subseteq \boldsymbol{G}_{T_2}$.

Now consider an operator $T : \mathfrak{D}_T \subseteq \mathcal{H} \to \mathcal{H}$ and let I denote the identity on $\mathcal{H}$. We define the **resolvent set** $\rho(T)$ of T as the set of all complex numbers λ such that $T - \lambda I$ is a bijection of $\mathfrak{D}_T$ onto $\mathcal{H}$, with a bounded inverse. The (point/continuous/residual) **spectrum** is then defined exactly as in the bounded case.

Note that, for a complex number λ to belong to $\rho(T)$, several conditions must be satisfied, which in general are not independent. For instance if T is closed, it follows from the Closed Graph Theorem that, if $T - \lambda I$ is a bijection from $\mathfrak{D}_T$ onto $\mathcal{H}$, then its inverse is automatically bounded. For other relations, see Problem 3.2.

Consider two linear operators $T_1 : \mathfrak{D}_{T_1} \subseteq \mathcal{H} \to \mathcal{H}$ and $T_2 : \mathfrak{D}_{T_2} \subseteq \mathcal{H} \to \mathcal{H}$, and a scalar $\lambda \in \mathbb{C}$. The sum of T_1 and T_2 is defined as:

$$\begin{aligned} T_1 + T_2 : \mathfrak{D}_{T_1} \cap \mathfrak{D}_{T_2} &\longrightarrow \mathcal{H} \\ u &\longmapsto T_1 u + T_2 u. \end{aligned}$$

The multiplication of T_1 by λ is defined by:

$$\begin{aligned} \lambda T_1 : \mathfrak{D}_{T_1} &\longrightarrow \mathcal{H} \\ u &\longmapsto \lambda T_1 u. \end{aligned}$$

Finally, the composition (or product) of T_1 and T_2 is the operator defined as:

$$\begin{aligned} T_1 T_2 : \{u \in \mathfrak{D}_{T_2}\,;\, T_2 u \in \mathfrak{D}_{T_1}\} &\longrightarrow \mathcal{H} \\ u &\longmapsto T_1(T_2 u). \end{aligned}$$

We also introduce the notion of inverse for unbounded operators. If an operator $T : \mathfrak{D}_T \to \mathcal{H}$ is one-to-one, we call **inverse** of T the mapping T^{-1} defined by:

$$\begin{aligned} T^{-1} : \mathrm{rge}(T) &\longrightarrow \mathcal{H} \\ Tu &\longmapsto u. \end{aligned}$$

Observe that $\mathrm{rge}(T^{-1}) = \mathfrak{D}_T$. Furthermore, if T is one-to-one, there holds

$$TT^{-1} \subseteq I \quad \text{and} \quad T^{-1}T \subseteq I.$$

3.2 The adjoint operator

We shall now extend the notion of an adjoint operator to unbounded operators having a dense domain. In the case of a bounded T, recall that the adjoint T^* of T was merely defined by

$$\langle Tu, v\rangle = \langle u, T^* v\rangle \text{ for all } u, v \in \mathcal{H}. \tag{3.2.1}$$

This definition relies on the Riesz Representation Theorem 1.2.6, using the fact that, for all $v \in \mathcal{H}$, the mapping

$$u \mapsto \langle Tu, v\rangle \tag{3.2.2}$$

is a continuous linear functional. In case T is not bounded, two problems require extra care. Firstly, the mapping (3.2.2) can a priori be defined only on the domain of T. Then, in general, this mapping will not be bounded, certainly not for all $v \in \mathcal{H}$. Observe that even if it is bounded for some v, then if the domain $\mathfrak{D}_T$ is not dense there will be several possible continuous extensions, and so several points $v^* \in \mathcal{H}$ satisfying:

$$\langle Tu, v\rangle = \langle u, v^*\rangle \text{ for all } u \in \mathfrak{D}_T.$$

To overcome these difficulties, we make the following definition.

Definition 3.2.1 Consider an operator $T : \mathfrak{D}_T \subseteq \mathcal{H} \to \mathcal{H}$ such that $\mathfrak{D}_T$ is dense in $\mathcal{H}$. The domain of the **adjoint** T^* of T is defined as

$$\mathfrak{D}_{T^*} = \{v \in \mathcal{H}\,;\text{ there exists } C \geqslant 0 \text{ s.t. } |\langle Tu, v\rangle| \leqslant C\,\|u\| \text{ for all } u \in \mathfrak{D}_T\}\,.$$

So for all $v \in \mathfrak{D}_{T^*}$, since $\mathfrak{D}_T$ is dense in $\mathcal{H}$, there exists a unique continuous extension to $\mathcal{H}$ of the functional $u \mapsto \langle Tu, v\rangle$. The Riesz Representation Theorem 1.2.6 allows us to define, for all $v \in \mathfrak{D}_{T^*}$, T^*v as the unique element of $\mathcal{H}$ satsifying

$$\langle Tu, v\rangle = \langle u, T^*v\rangle \text{ for all } u \in \mathfrak{D}_T. \blacklozenge$$

Observe that this defines a linear operator. Let $u_1, u_2 \in \mathfrak{D}_{T^*}$ and $\lambda_1, \lambda_2 \in \mathbb{C}$. Then indeed, for all $u \in \mathfrak{D}_T$, we have

$$\begin{aligned}\langle Tu, \lambda_1 u_1 + \lambda_2 u_2\rangle &= \overline{\lambda}_1 \langle Tu, u_1\rangle + \overline{\lambda}_2 \langle Tu, u_2\rangle\\ &= \overline{\lambda}_1 \langle u, T^*u_1\rangle + \overline{\lambda}_2 \langle u, T^*u_2\rangle\\ &= \langle u, \lambda_1 T^*u_1 + \lambda_2 T^*u_2\rangle\,.\end{aligned}$$

It follows that $\lambda_1 u_1 + \lambda_2 u_2 \in \mathfrak{D}_{T^*}$ and we have:

$$T^*(\lambda_1 u_1 + \lambda_2 u_2) = \lambda_1 T^*u_1 + \lambda_2 T^*u_2.$$

An operator T with dense domain immediately satisfies

$$\langle Tu, v\rangle = \langle u, T^*v\rangle \text{ for all } u \in \mathfrak{D}_T \text{ and all } v \in \mathfrak{D}_{T^*}.$$

We also remark that

$$\operatorname{rge}(T)^\perp = \ker(T^*).$$

Indeed, since $\mathfrak{D}_T$ is dense,

$$v \in \ker(T^*) \iff v \in \mathfrak{D}_{T^*} \;\&\; T^*v = 0 \iff \langle Tu, v\rangle = 0\ \forall u \in \mathfrak{D}_T \iff v \in \operatorname{rge}(T)^\perp.$$

The following proposition gives a fundamental property of the adjoint.

Proposition 3.2.2 *Consider an operator T with dense domain $\mathfrak{D}_T \subseteq \mathcal{H}$. Then the adjoint T^* is closed.*

Proof. Consider a sequence $(u_n, T^*u_n)_{n=1}^\infty$ in the graph of T^* which converges to a point $(u, v) \in \mathcal{H}\times\mathcal{H}$. Let us show that $u \in \mathfrak{D}_{T^*}$ and $v = T^*u$. It follows from

$$\begin{aligned}\|(u_n, T^*u_n) - (u, v)\|^2_{\mathcal{H}\times\mathcal{H}} &= \langle u_n - u, u_n - u\rangle + \langle T^*u_n - v, T^*u_n - v\rangle\\ &= \|u_n - u\|^2 + \|T^*u_n - v\|^2\end{aligned}$$

that $u_n \to u$ and $T^*u_n \to v$ as $n \to \infty$. Then, by the continuity of the inner product, for all $w \in \mathfrak{D}_T$:

$$\langle Tw, u\rangle = \lim_{n\to\infty} \langle Tw, u_n\rangle = \lim_{n\to\infty} \langle w, T^*u_n\rangle = \langle w, v\rangle .$$

Hence, $w \mapsto \langle Tw, u\rangle$ is bounded on $\mathfrak{D}_T$ and so $u \in \mathfrak{D}_{T^*}$, $v = T^*u$. □

The adjoint operator also enjoys the following elementary properties.

Properties 3.2.3 Consider two operators T_1 and T_2 with dense domains, and $\lambda \in \mathbb{C}$. Then:

(a) $(\lambda T)^* = \overline{\lambda} T^*$;

(b) if $\mathfrak{D}_{T_1+T_2}$ is dense, then $T_1^* + T_2^* \subseteq (T_1 + T_2)^*$;

(c) if $\mathfrak{D}_{T_2T_1}$ is dense, then $T_1^*T_2^* \subseteq (T_2T_1)^*$;

(d) if $T_1 \subseteq T_2$, then $T_1^* \supseteq T_2^*$.

3.3 Commutativity and reduction

We define here what it means for a bounded operator to commute with an unbounded operator.

Definition 3.3.1 Consider a bounded operator B and another operator T. We say that B commutes with T, and we write $B \smile T$, if

$$BT \subseteq TB. \ \blacklozenge$$

More generally, the commutativity between two unbounded *selfadjoint* operators will be defined later through the commutativity of their respective spectral families. This is the notion of commutativity which is required by quantum mechanics.

Lemma 3.3.2 *Let P be a projection, $Q = I - P$ the projection onto the orthogonal complement of* $\operatorname{rge} P$*, and T be any operator. If $P \smile T$, then the closed subspaces* $\operatorname{rge} P$ *and* $\operatorname{rge} Q$ *reduce the operator T in the sense that, on the one hand,*

$$PTP = TP \quad \text{and} \quad QTQ = TQ,$$

and, on the other hand,

$$T = TP + TQ.$$

Proof. Since $P \smile T$, i.e. $PT \subseteq TP$, we have

$$PTP = (PT)P \subseteq (TP)P = TP,$$

and since the operators PTP and TP have same domain,

$$PTP = TP.$$

Furthermore, since

$$QT = (I - P)T = T - PT \subseteq T - TP = T(I - P),$$

i.e. $Q \smile T$, we obtain similarly that $QTQ = TQ$. On the other hand, we have

$$T = (P + Q)T = PT + QT \subseteq TP + TQ \subseteq T(P + Q) = T,$$

and so

$$T = TP + TQ.$$

□

3.4 More on operator graphs

We define two linear operators $\mathbf{U}, \mathbf{V} : \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H}$ by

$$\mathbf{U}(u, v) = (v, u) \quad \text{and} \quad \mathbf{V}(u, v) = (v, -u), \quad (u, v) \in \mathcal{H} \times \mathcal{H}.$$

They are Hilbert space automorphisms of $\mathcal{H} \times \mathcal{H}$ (i.e. they are bijective and preserve the inner product). Furthermore, they satisfy the following identities, where $\mathbf{I}$ denotes the identity on $\mathcal{H} \times \mathcal{H}$:

$$\mathbf{UV} = -\mathbf{VU} \quad \text{and} \quad -\mathbf{V}^2 = \mathbf{U}^2 = \mathbf{I}.$$

The following observation will be useful.

Lemma 3.4.1 *Let T be a densely defined operator in $\mathcal{H}$. The graphs of T and T^* satisfy*

$$\boldsymbol{G}_{T^*} = \left(\mathbf{V}\overline{\boldsymbol{G}_T}\right)^{\perp} \text{ or, equivalently } \quad \mathbf{V}\overline{\boldsymbol{G}_T} = (\boldsymbol{G}_{T^*})^{\perp}.$$

Proof. For all $u \in \mathfrak{D}_T$ and all $v \in \mathfrak{D}_{T^*}$ we have

$$\langle Tu, v\rangle = \langle u, T^*v\rangle\,,$$

which can be written as

$$\Big\langle \mathbf{V}(u, Tu), (v, T^*v)\Big\rangle = 0. \tag{3.4.1}$$

Hence, we already remark that any element of $\boldsymbol{G}_{T^*}$ is orthogonal to all elements of $\mathbf{V}\,\boldsymbol{G}_T$. Now the proof is in two steps.

$\subseteq$. Consider $(w, T^*w) \in \boldsymbol{G}_{T^*}$ and a sequence $(u_n, v_n)_{n=1}^{\infty} \subset \mathbf{V}\,\boldsymbol{G}_T$ which converges to $(u, v) \in \overline{\mathbf{V}\,\boldsymbol{G}_T} = \mathbf{V}\overline{\boldsymbol{G}_T}$. It then follows from (3.4.1) that

$$\Big\langle (u, v), (w, T^*w)\Big\rangle = \lim_{n\to\infty} \Big\langle (u_n, v_n), (v, T^*v)\Big\rangle = 0.$$

Hence, $\boldsymbol{G}_{T^*} \subseteq \Big(\mathbf{V}\ \overline{\boldsymbol{G}_T}\Big)^{\perp}$.

$\supseteq$. Conversely, let (u, v) be an element of the orthogonal complement of $\mathbf{V}\overline{\boldsymbol{G}_T}$. Then in particular, for all $w \in \mathfrak{D}_T$:

$$0 = \Big\langle \mathbf{V}(w, Tw), (u, v)\Big\rangle = \langle Tw, u\rangle - \langle w, v\rangle\,. \tag{3.4.2}$$

Therefore, the mapping $w \mapsto \langle Tw, u\rangle$ coincides on $\mathfrak{D}_T$ with the linear functional $w \mapsto \langle w, v\rangle$, which is bounded on $\mathcal{H}$. It follows that the mapping $w \mapsto \langle Tw, u\rangle$ is bounded on $\mathfrak{D}_T$. Hence, $u \in \mathfrak{D}_{T^*}$ and $T^*u = v$, i.e. $(u, v) \in \boldsymbol{G}_{T^*}$. Therefore, $\boldsymbol{G}_{T^*} \supseteq \Big(\mathbf{V}\overline{\boldsymbol{G}_T}\Big)^{\perp}$. ☐

For closed operators, the previous lemma has the following important consequence.

Theorem 3.4.2 *Let T be a closed operator with dense domain. Then the domain of T^* is also dense in $\mathcal{H}$, and so $T^{**} = (T^*)^*$ does exist. Furthermore, $T^{**} = T$.*

Proof. Suppose by contradiction that $\mathfrak{D}_{T^*}$ is not dense in $\mathcal{H}$. Then there exists $h \in \mathcal{H} \setminus \{0\}$ which is orthogonal to $\mathfrak{D}_{T^*}$. It follows that, for all $v \in \mathfrak{D}_{T^*}$,

$$\Big\langle (0, h), \mathbf{V}(v, T^*v)\Big\rangle = \langle 0, T^*v\rangle - \langle h, v\rangle = 0,$$

and so $(0, h) \in \big(\mathbf{V}\,\boldsymbol{G}_{T^*}\big)^{\perp}$. However, we know from Lemma 3.4.1 that $\big(\boldsymbol{G}_{T^*}\big)^{\perp} = \mathbf{V}\overline{\boldsymbol{G}_T}$. Hence, since $\mathbf{V}$ is unitary,

$$\big(\mathbf{V}\,\boldsymbol{G}_{T^*}\big)^{\perp} = \mathbf{V}\big(\boldsymbol{G}_{T^*}\big)^{\perp} = \mathbf{V}^2\overline{\boldsymbol{G}_T} = -\mathbf{I}\overline{\boldsymbol{G}_T} = \overline{\boldsymbol{G}_T} = \boldsymbol{G}_T, \tag{3.4.3}$$

by closedness of T. Therefore, $\big(\mathbf{V}\,\boldsymbol{G}_{T^*}\big)^{\perp} = \boldsymbol{G}_T$ and so $(0, h) \in \boldsymbol{G}_T$, $h = T0 = 0$, which contradicts $h \neq 0$. Hence, $\mathfrak{D}_{T^*}$ is dense in $\mathcal{H}$, and T^{**} exists. But $\boldsymbol{G}_{T^{**}}$ is the orthogonal complement of $\mathbf{V}\,\boldsymbol{G}_{T^*}$, so by (3.4.3)

$$\boldsymbol{G}_{T^{**}} = \boldsymbol{G}_T.$$

That is, $T^{**} = T$. □

The next theorem states a rather surprising fact, which will play an important role in the proof of the spectral theorem for selfadjoint operators.

Theorem 3.4.3 *Let T be a closed operator with dense domain. The operators*

$$\mathsf{B} = (I + T^*T)^{-1} \quad \text{and} \quad \mathsf{C} = T(I + T^*T)^{-1}$$

are defined and bounded on $\mathcal{H}$, with

$$\|\mathsf{B}\| \leqslant 1 \quad \text{and} \quad \|\mathsf{C}\| \leqslant 1.$$

Furthermore, B is symmetric and positive.

Proof. Since T is closed, Theorem 3.4.2 implies that $\mathfrak{D}_{T^*}$ is dense in $\mathcal{H}$ and that $T = T^{**}$. It then follows from Lemma 3.4.1 that $\boldsymbol{G}_T$ and $\mathbf{V}\boldsymbol{G}_{T^*}$ are orthogonal complements in $\mathcal{H} \times \mathcal{H}$. Hence, for all $h \in \mathcal{H}$, there is a unique $u_h \in \mathfrak{D}_T$ and a unique $v_h \in \mathfrak{D}_{T^*}$ such that

$$(h, 0) = (u_h, Tu_h) + (T^*v_h, -v_h), \tag{3.4.4}$$

or, by components,

$$\begin{cases} h = u_h + T^*v_h, \\ 0 = Tu_h - v_h. \end{cases}$$

This first observation allows us to define two linear maps B and C by

$$\mathsf{B}h = u_h \quad \text{and} \quad \mathsf{C}h = v_h.$$

These operators are defined on $\mathcal{H}$, and satisfy

$$\begin{cases} I = \mathsf{B} + T^*\mathsf{C}, \\ 0 = T\mathsf{B} - \mathsf{C}, \end{cases}$$

from which we deduce that

$$\mathsf{C} = T\mathsf{B} \quad \text{and} \quad I = \mathsf{B} + T^*T\mathsf{B} = (I + T^*T)\mathsf{B}. \tag{3.4.5}$$

Furthermore, the two terms in the right-hand side of (3.4.4) are orthogonal, so

$$\begin{aligned} \|h\|^2 &= \|(h,0)\|^2_{\mathcal{H}\times\mathcal{H}} = \|(u_h, Tu_h)\|^2_{\mathcal{H}\times\mathcal{H}} + \|(T^*v_h, -v_h)\|^2_{\mathcal{H}\times\mathcal{H}} \\ &= \|u_h\|^2 + \|Tu_h\|^2 + \|T^*v_h\|^2 + \|v_h\|^2 . \end{aligned}$$

Therefore,

$$\|\mathsf{B}h\|^2 + \|\mathsf{C}h\|^2 = \|u_h\|^2 + \|v_h\|^2 \leqslant \|h\|^2 ,$$

and it follows that

$$\|\mathsf{B}\| \leqslant 1 \quad \text{and} \quad \|\mathsf{C}\| \leqslant 1.$$

We now observe that, for any element u in the domain of T^*T, we have

$$\langle (I + T^*T)u, u\rangle = \langle u, u\rangle + \langle Tu, Tu\rangle \geqslant \langle u, u\rangle = \|u\|^2 \geqslant 0.$$

Hence, if $(I + T^*T)u = 0$ then $u = 0$. This shows that $I + T^*T$ is one-to-one and so has an inverse $(I+T^*T)^{-1}$. But now the right identity in (3.4.5) implies that $\operatorname{rge}(I+T^*T) = \mathcal{H}$, and it follows that

$$\mathsf{B} = (I + T^*T)^{-1}.$$

Finally, B is symmetric and positive. Indeed, for all $u, v \in \mathcal{H}$,

$$\begin{aligned}\langle \mathsf{B}u, v\rangle = \langle \mathsf{B}u, (I + T^*T)\mathsf{B}v\rangle &= \langle \mathsf{B}u, \mathsf{B}v\rangle + \langle \mathsf{B}u, T^*T\mathsf{B}v\rangle \\ &= \langle \mathsf{B}u, \mathsf{B}v\rangle + \langle T^*T\mathsf{B}u, \mathsf{B}v\rangle = \langle (I + T^*T)\mathsf{B}u, \mathsf{B}v\rangle = \langle u, \mathsf{B}v\rangle .\end{aligned}$$

Furthermore, for all $u \in \mathcal{H}$,

$$\langle \mathsf{B}u, u\rangle = \langle \mathsf{B}u, (I + T^*T)\mathsf{B}u\rangle = \langle \mathsf{B}u, \mathsf{B}u\rangle + \langle T\mathsf{B}u, T\mathsf{B}u\rangle = \|\mathsf{B}u\|^2 + \|T\mathsf{B}u\|^2 \geqslant 0.$$

This concludes the proof. □

A useful corollary of Theorem 3.4.3 will be stated in the next section.

3.5 Symmetric and selfadjoint operators

We will now introduce the terms *symmetric* and *selfadjoint* in the framework of unbounded operators. The two notions coincide for bounded operators.

Definition 3.5.1 We call **symmetric** an operator $T : \mathfrak{D}_T \subseteq \mathcal{H} \to \mathcal{H}$ such that $\mathfrak{D}_T$ is dense in $\mathcal{H}$ and

$$T \subseteq T^*. \ \blacklozenge$$

Proposition 3.5.2 *An operator $T : \mathfrak{D}_T \subseteq \mathcal{H} \to \mathcal{H}$ is symmetric if and only if $\mathfrak{D}_T$ is dense and*

$$\langle Tu, v\rangle = \langle u, Tv\rangle \quad \textit{for all } u, v \in \mathfrak{D}_T. \tag{3.5.1}$$

Proof. The 'only if' part is trivial. For the 'if' part, we suppose that $\mathfrak{D}_T$ is dense, and we pick up $v \in \mathfrak{D}_T$. Then (3.5.1) shows that the linear functional $u \mapsto \langle Tu, v\rangle$ is bounded on $\mathfrak{D}_T$, so that $v \in \mathfrak{D}_{T^*}$. Thus, $\mathfrak{D}_T \subseteq \mathfrak{D}_{T^*}$. Then, for all $v \in \mathfrak{D}_T$, the definition of T^* and (3.5.1) imply that $\langle u, T^*v - Tv\rangle = 0$ for all $u \in \mathfrak{D}_T$. Since D_T is dense, it follows that $T^*v - Tv = 0$. This shows that T^* indeed coincides with T on $\mathfrak{D}_T$. □

If T is a symmetric operator, then the domain of T^* is also dense in $\mathcal{H}$. Indeed,

$$\mathcal{H} = \overline{\mathfrak{D}_T} \subseteq \overline{\mathfrak{D}_{T^*}} \subseteq \mathcal{H}.$$

In this case, T^* also has an adjoint, T^{**}. The operator T^{**} is a closed extension of T since, by Lemma 3.4.1,

$$\boldsymbol{G}_{T^{**}} = \left(\mathbf{V}\,\boldsymbol{G}_{T^*}\right)^{\perp} = \mathbf{V}^2\overline{\boldsymbol{G}_T} = \overline{\boldsymbol{G}_T} \supseteq \boldsymbol{G}_T. \tag{3.5.2}$$

Furthermore, T^{**} is symmetric. Indeed, $T \subseteq T^{**}$ implies that $\mathfrak{D}_{T^{**}}$ is dense. Moreover, since $T \subseteq T^*$ we have $T^* \supseteq T^{**}$ and so, by Theorem 3.4.2,

$$T^{**} \subseteq T^* = (T^*)^{**} = (T^{**})^*.$$

We have thus shown that any symmetric operator has a closed symmetric extension.

Definition 3.5.3 An operator T is said to be **closable** provided the closure of its graph is the graph of an operator. This operator is then called the **closure** of T, denoted $\overline{T}$:

$$\overline{\boldsymbol{G}_T} = \boldsymbol{G}_{\overline{T}}. \ \blacklozenge$$

Remark 3.5.4 It follows from the previous discussion that any symmetric operator T is closable, with $\overline{T} = T^{**}$. In fact, as will be seen in Problem 3.4, this relation between closure and adjoint always holds true (as long as T^* is densely defined).

If a bounded operator B is symmetric in the sense of Definition 3.5.1, then it is symmetric in the sense of bounded operators, i.e. $B^* = B$. However, for a densely defined operator T, being symmetric does not imply $T^* = T$.

Definition 3.5.5 An operator $T : \mathfrak{D}_T \subseteq \mathcal{H} \to \mathcal{H}$ is called **selfadjoint** if $\mathfrak{D}_T$ is dense in $\mathcal{H}$ and

$$T = T^*. \ \blacklozenge$$

It follows immediately from Proposition 3.2.2 that a selfadjoint operator is closed. It is noteworthy that a non-selfadjoint symmetric operator (even closed) does not necessarily possess a selfadjoint extension. A symmetric operator T is called **maximal symmetric** if it has no proper symmetric extension, i.e. if there is no symmetric operator S such that $T \subseteq S$ and $T \neq S$. Observe that any selfadjoint A is maximal symmetric. Indeed,

$$A \subseteq T \quad \text{and} \quad T \subseteq T^*$$

implies

$$A = A^* \supseteq T^* \supseteq T \supseteq A,$$

and so $T = A$.

We say that a symmetric operator A is **essentially selfadjoint** if $\overline{A}$ is selfadjoint. The proof of the next theorem can be found in [Wei80, p. 108].

Theorem 3.5.6 *A symmetric operator A is essentially selfadjoint if and only if the subspaces* $\mathrm{rge}(A \pm iI)$ *are dense in $\mathcal{H}$.*

We conclude this section with the following corollary of Theorem 3.4.3.

Corollary 3.5.7 *If A is selfadjoint, then the operators $\mathsf{B} = (I + A^2)^{-1}$ and $\mathsf{C} = A\mathsf{B}$ given by Theorem 3.4.3 have the following additional properties:*

(a) $\mathsf{B}(\mathfrak{D}_A) = \mathfrak{D}_{A^3}$;

(b) $\mathsf{B}\smile A$, *i.e.* $\mathsf{B}A \subseteq A\mathsf{B}$;

(c) $\mathsf{C}\mathsf{B} = \mathsf{B}\mathsf{C}$;

(d) any operator $T \in \mathcal{B}(\mathcal{H})$ such that $T\smile A$ satisfies $T\mathsf{B} = \mathsf{B}T$.

Proof. Let us first prove that $\mathsf{B}(\mathfrak{D}_A) = \mathfrak{D}_{A^3}$. Since $(I + A^2)\mathsf{B} = I$, we have for all $u \in \mathfrak{D}_A$,

$$\begin{aligned}(A - \mathsf{C})u &= Au - A\mathsf{B}u \\ &= A(I - \mathsf{B})u \\ &= A\big((I + A^2)\mathsf{B} - \mathsf{B}\big)u \\ &= A^3\mathsf{B}u.\end{aligned}$$

Since $\mathfrak{D}_{A-\mathsf{C}} = \mathfrak{D}_A$, it follows that $\mathfrak{D}_A \subseteq \mathsf{B}(\mathfrak{D}_A)$. Conversely, if $v \in \mathfrak{D}_{A^3} \subseteq \mathfrak{D}_{A^2} = \mathfrak{D}_{\mathsf{B}^{-1}}$, then

$$u = \mathsf{B}^{-1}v = (I + A^2)v \in \mathfrak{D}_A.$$

We now show that $\mathsf{B}A \subseteq A\mathsf{B}$. For all $u \in \mathfrak{D}_A$, since $\mathsf{B}u \in \mathfrak{D}_{A^3}$, we have

$$(I + A^2)A\mathsf{B}u = A(I + A^2)\mathsf{B}u = Au.$$

We then apply B and, since $h = \mathsf{B}(I + A^2)h$ for all $h \in \mathfrak{D}_{A^2}$, we get

$$A\mathsf{B}u = \mathsf{B}(I + A^2)A\mathsf{B}u = \mathsf{B}Au, \quad \text{for all } u \in \mathfrak{D}_A.$$

It follows that

$$\mathsf{BC} = (\mathsf{B}A)\mathsf{B} \subseteq (A\mathsf{B})\mathsf{B} = \mathsf{CB},$$

and, since BC is defined everywhere, $\mathsf{BC} = \mathsf{CB}$.

Finally, for $T \in \mathcal{H}$ such that $T\smile A$, we have in particular that

$$TA^2 \subseteq ATA \subseteq A^2T, \tag{3.5.3}$$

and so

$$T\mathsf{B}^{-1} = T(I + A^2) \subseteq (I + A^2)T = \mathsf{B}^{-1}T. \tag{3.5.4}$$

It also follows from (3.5.3) that, if $u \in \mathfrak{D}_{A^2}$ then $TA^2u = A^2Tu$, and so $Tu \in \mathfrak{D}_{A^2}$. Consider $u \in \mathfrak{D}_A$ arbitrary. By part (a), $\mathsf{B}u \in \mathfrak{D}_{A^3} \subseteq \mathfrak{D}_{A^2}$ and so $T\mathsf{B}u \in \mathfrak{D}_{A^2} = \mathfrak{D}_{\mathsf{B}^{-1}}$. It then follows by (3.5.4) that

$$T\mathsf{B}u = \mathsf{B}\mathsf{B}^{-1}T\mathsf{B}u = \mathsf{B}T\mathsf{B}^{-1}\mathsf{B}u = \mathsf{B}Tu.$$

Since $\mathfrak{D}_A$ is dense in $\mathcal{H}$ and since $T\mathsf{B}$ and $\mathsf{B}T$ are continuous, it follows that $T\mathsf{B} = \mathsf{B}T$. □

3.6 Integration with respect to a spectral family

In this section we extend the notion of spectral family in a natural way to deal with unbounded operators and, using Lebesgue integration, we define the integral with respect to a spectral family for a large class of functions.

Definition 3.6.1 A **spectral family** on $\mathcal{H}$ is a mapping $E : \mathbb{R} \to \mathcal{B}(\mathcal{H})$, denoted $(E_\lambda)_{\lambda\in\mathbb{R}}$ which satisfies (i)–(iii) of Definition 2.1.1 and, instead of property (iv) of Definition 2.1.1,

(iv)' for all $u \in \mathcal{H}$:

$$\lim_{\lambda\to-\infty} E_\lambda u = 0 \quad \text{and} \quad \lim_{\lambda\to\infty} E_\lambda u = u. \ ♦$$

The following preliminary result was essentially proved during the proof of the spectral theorem for bounded symmetric operators.

Lemma 3.6.2 *Let $E : \mathbb{R} \to \mathcal{B}(\mathcal{H})$ satisfy parts (i) and (ii) of Definition 2.1.1. Then for all $\mu \in \mathbb{R}$, there exist projections $E_{\mu-0}$ and $E_{\mu+0}$ such that, for all $u \in \mathcal{H}$,*

$$\lim_{\lambda \nearrow \mu} E_\lambda u = E_{\mu-0} u \quad \text{and} \quad \lim_{\lambda \searrow \mu} E_\lambda u = E_{\mu+0} u.$$

Now consider a spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$. For all $u \in \mathbb{R}$, the function

$$\begin{aligned} F_u : \mathbb{R} &\longrightarrow \mathbb{R} \\ \lambda &\longmapsto F_u(\lambda) := \langle E_\lambda u, u\rangle = \|E_\lambda u\|^2, \end{aligned}$$

is increasing, left-continuous, and satisfies

$$\lim_{\lambda\to-\infty} F_u(\lambda) = 0 \quad \text{and} \quad \lim_{\lambda\to\infty} F_u(\lambda) = \|u\|^2.$$

We can in particular associate to each function F_u the corresponding Lebesgue-Stieltjes measure μ_{F_u} which we denote by $\mu_{\|E_\lambda u\|^2}$ (cf. Section B.1). We then make the following definition.

Definition 3.6.3 Consider a spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$. We shall say that a function $f : \mathbb{R} \to \mathbb{C}$ is **E-measurable** if f is $\mu_{\|E_\lambda u\|^2}$-measurable for all $u \in \mathcal{H}$. ♦

The scope of this definition is very large. Indeed, all (Lebesgue-)measurable functions are E-measurable for any spectral family E.

To define the integral with respect to a spectral family, we start with a step function

$$t = \sum_{k=0}^{n} c_k \chi_{I_k},$$

where the c_ks are complex numbers and the χ_{I_k}s are characteristic functions of real non-empty intervals of one of the following forms:

$$(a_k, b_k), \quad [a_k, b_k), \quad (a_k, b_k], \quad [a_k, b_k].$$

The integral of t with respect to the spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ is now defined as

$$\int_{\mathbb{R}} t(\lambda)\, \mathrm{d}E_\lambda = \sum_{k=0}^{n} c_k E(I_k),$$

where

$$E\big((a,b)\big) = E_b - E_{a+0}, \quad E\big([a,b)\big) = E_b - E_a, \quad E\big((a,b]\big) = E_b - E_{a+0}, \quad E\big([a,b]\big) = E_b - E_a.$$

For any step function t we have

$$\begin{aligned}\left\| \int t(\lambda)\,\mathrm{d}E_\lambda u \right\|^2 &= \left\langle \sum_{k=0}^n c_k E(I_k)u, \sum_{j=0}^n c_j E(I_j)u \right\rangle = \sum_{k=0}^n \sum_{j=0}^n c_k \bar{c}_j \left\langle E(I_k)u, E(I_j)u \right\rangle \\ &= \sum_{k=0}^n |c_k|^2 \left\langle E(I_k)u, u \right\rangle = \int |t(\lambda)|^2 \,\mathrm{d}\mu_{\|E_\lambda u\|^2},\end{aligned}$$

where the integral in the right-hand side is a Lebesgue-Stieltjes integral (cf. Appendix B). Here and henceforth, unless otherwise specified, the integrals are taken over the whole of $\mathbb{R}$.

Consider now a spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ and an E-measurable function $f : \mathbb{R} \to \mathbb{C}$. By Theorem B.3.1, for all $u \in \mathcal{H}$ such that $f \in L^2(\mathbb{R}, \mu_{\|E_\lambda u\|^2})$, there is a sequence of step functions $(t_n)_{n=1}^\infty$ converging to f in $L^2(\mathbb{R}, \mu_{\|E_\lambda u\|^2})$. Hence, for all $\varepsilon > 0$, there exists $N \geqslant 1$ such that, for all $m, n \geqslant N$,

$$\left\| \int t_n(\lambda)\,\mathrm{d}E_\lambda u - \int t_m(\lambda)\,\mathrm{d}E_\lambda u \right\|^2 = \int |t_n(\lambda) - t_m(\lambda)|^2 \,\mathrm{d}\mu_{\|E_\lambda u\|^2} < \varepsilon^2.$$

Therefore, the sequence $\left(\int t_n(\lambda)\,\mathrm{d}E_\lambda u\right)_{n=0}^\infty$ is Cauchy in $\mathcal{H}$, hence convergent, and we write

$$\int f(\lambda)\,\mathrm{d}E_\lambda u = \lim_{n\to\infty} \int t_n(\lambda)\,\mathrm{d}E_\lambda u,$$

which is independent of the choice of the sequence $(t_n)_{n=1}^\infty$. Letting

$$\mathfrak{D}_{E(f)} = \left\{ u \in \mathcal{H}\,;\, f \in L^2(\mathbb{R}, \mu_{\|E_\lambda u\|^2}) \right\},$$

we have thus defined a mapping

$$\begin{aligned} E(f) : \mathfrak{D}_{E(f)} &\longrightarrow \mathcal{H} \\ u &\longmapsto \int f(\lambda)\,\mathrm{d}E_\lambda u. \end{aligned} \tag{3.6.1}$$

Definition 3.6.4 We denote the linear operator (3.6.1) by $\int f(\lambda)\,\mathrm{d}E_\lambda$ and we call it the **integral of f with respect to** $(E_\lambda)_{\lambda\in\mathbb{R}}$. ♦

The following theorem gives the main properties of the integral.

Theorem 3.6.5 *Consider a spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ and $f : \mathbb{R} \to \mathbb{C}$ an E-measurable function. Then the operator $E(f)$ is **normal**, i.e. $E(f)E(f)^* = E(f)^*E(f)$. Furthermore:*

(a) $u \in \mathfrak{D}_{E(f)} \iff \|E(f)u\|^2 = \int |f|^2 \,\mathrm{d}\mu_{\|E_\lambda u\|^2} < \infty$.

(b) *If f is bounded, then* $\mathfrak{D}_{E(f)} = \mathcal{H}$, $E(f) \in \mathcal{B}(\mathcal{H})$, *and* $\|E(f)\| \leqslant \operatorname{ess\,sup}_{\lambda\in\mathbb{R}} |f(\lambda)|$.

(c) *If* $f(\lambda) = 1$ *for all* $\lambda \in \mathbb{R}$, *then* $E(f) = I$.

(d) *For all* $u \in \mathfrak{D}_{E(f)}$,

$$\langle E(f)u, u\rangle = \int f(\lambda)\,\mathrm{d}\mu_{\|E_\lambda u\|^2}.$$

(e) *For* $a, b \in \mathbb{C}$ *and any E-measurable function* $g : \mathbb{R} \to \mathbb{C}$,

$$aE(f) + bE(g) \subseteq E(af + bg) \quad \text{and} \quad \mathfrak{D}_{E(f)+E(g)} = \mathfrak{D}_{E(|f|+|g|)}.$$

(f) *For any E-measurable function* $g : \mathbb{R} \to \mathbb{C}$,

$$E(f)E(g) \subseteq E(fg) \quad \text{and} \quad \mathfrak{D}_{E(f)E(g)} = \mathfrak{D}_{E(g)} \cap \mathfrak{D}_{E(fg)};$$

(g) *We have*

$$E(\bar{f}) = E(f)^* \quad \text{and} \quad \mathfrak{D}_{E(f)^*} = \mathfrak{D}_{E(f)}.$$

(h) *For all* $\mu \in \mathbb{R}$,

$$E_\mu E(f) \subseteq E(f)E_\mu,$$

with equality if f is bounded.

Proof. See Problem 3.10. □

We now observe that, given a bounded symmetric operator A on $\mathcal{H}$, Theorem 2.2.1 yields a unique spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ in the sense of Definition 2.1.1 — which is, *a fortiori*, a spectral family in the sense of Definition 3.6.1 — such that $A = \int \lambda \,\mathrm{d}E_\lambda$. (The equivalence between the present theory of integration and that of Chapter 2 in the case where the spectral family has finite bounds is studied in Problem 3.11.) Hence, for a bounded symmetric operator A and its spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$, we can now make the following definition. Let $f : [m, M] \to \mathbb{C}$ be a function whose extension to $\mathbb{R}$ by $f(\lambda) = 0$ for all $\lambda \in \mathbb{R}\backslash[m, M]$ is E-measurable. The operator $f(A)$ is defined as

$$f(A) = E(f) = \int f(\lambda)\,\mathrm{d}E_\lambda.$$

The functions of A have the following properties.

Theorem 3.6.6 *Consider a bounded symmetric operator A and its spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ given by Theorem 2.2.1. Let $f, g : \mathbb{R} \to \mathbb{C}$ be E-measurable functions such that $f(\lambda) = g(\lambda) = 0$ for all $\lambda \in \mathbb{R}\backslash[m, M]$. Then the following holds.*

(a) If f is bounded, then $f(A) \in \mathcal{B}(\mathcal{H})$ and

$$\|f(A)\| \leqslant \sup_{\lambda\in[m,M]} |f(\lambda)| \,.$$

(b) If f and g are bounded, then

$$f(A)g(A) = \int_m^{M+\varepsilon} f(\lambda)g(\lambda)\,\mathrm{d}E_\lambda.$$

(c) $f(A)^* = \bar{f}(A)$.

(d) Any bounded operator which commutes with A commutes with $f(A)$.

Proof. Parts (a) to (c) are direct consequences of Theorem 3.6.5, and part (d) will be proved in Problem 3.12. □

3.7 The spectral theorem for selfadjoint operators

Our proof of the spectral theorem follows that of Riesz and Lorch, as presented in [RSN90]. Another proof due to von Neumann, based on the Cayley transform, can also be found in [RSN90]. The main idea in Riesz's proof is to reduce the problem to the case of bounded symmetric operators by a limit procedure. With start with the following result.

Lemma 3.7.1 *Let*

$$\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_i, \ldots$$

be a sequence of closed, pairwise orthogonal, subspaces of the Hilbert space $\mathcal{H}$, such that $\oplus_i\mathcal{H}_i = \mathcal{H}$. We denote by u_i the projection of any $u \in \mathcal{H}$ onto $\mathcal{H}_i$. Consider a sequence

$$A_1, A_2, \ldots, A_i, \ldots$$

of operators in $\mathcal{H}$ such that the restriction $A_i|_{\mathcal{H}_i}$ is a bounded symmetric operator mapping $\mathcal{H}_i$ into itself, for all $i \geqslant 1$.

Then there exists a unique selfadjoint operator $A : \mathfrak{D}_A \subseteq \mathcal{H} \to \mathcal{H}$ *which coincides with* A_i *on* $\mathcal{H}_i$*, for all* $i \geqslant 1$*. The domain of* A *is defined by*

$$\mathfrak{D}_A = \left\{ u \in \mathcal{H} \, ; \, \sum_{i=1}^{\infty} \|A_i u_i\|^2 < \infty \right\} \tag{3.7.1}$$

and, for all $u \in \mathfrak{D}_A$*,*

$$Au = \sum_{i=1}^{\infty} A_i u_i. \tag{3.7.2}$$

Proof. Let us first observe that the operator defined by (3.7.1)–(3.7.2) is linear. Furthermore $\mathfrak{D}_A$ is dense in $\mathcal{H}$ since, for all $u \in \mathcal{H}$ and all $\varepsilon > 0$, there exists $N \geqslant 1$ such that

$$\left\| u - \sum_{i=1}^{N} u_i \right\| < \varepsilon,$$

and clearly $\sum_{i=1}^{N} u_i \in \mathfrak{D}_A$. Moreover, A is symmetric since, by linearity and continuity of the inner product, and by the pairwise orthogonality of the $\mathcal{H}_i$s we have, for all $u, v \in \mathfrak{D}_A$,

$$\langle Au, v \rangle = \sum_{i=1}^{\infty} \langle A_i u_i, v_i \rangle = \sum_{i=1}^{\infty} \langle u_i, A_i v_i \rangle = \langle u, Av \rangle \, .$$

To prove that A is, in fact, selfadjoint, we need only show that $\mathfrak{D}_A = \mathfrak{D}_{A^*}$. To this end, consider an arbitrary $v \in \mathfrak{D}_{A^*}$. Then, for any $u \in \mathfrak{D}_A$,

$$\langle Au, v \rangle = \langle u, A^* v \rangle \, ,$$

and so

$$\sum_{i=1}^{\infty} \langle A_i u_i, v_i \rangle = \sum_{i=1}^{\infty} \langle u_i, (A^* v)_i \rangle \, .$$

In particular, for any $j \geqslant 1$ and $u \in \mathcal{H}_j$ we have $u_i = 0$ for $i \neq j$ and $u_j = u$, so that

$$\langle A_j u, v_j \rangle = \langle u, (A^* v)_j \rangle \, .$$

But A_j is bounded and symmetric on $\mathcal{H}_j$, so it follows that

$$(A^* v)_j = A_j v_j.$$

Then the Pythagoras theorem implies

$$\sum_{i=1}^{\infty} \|A_i v_i\|^2 = \sum_{i=1}^{\infty} \|(A^* v)_i\|^2 = \|A^* v\|^2 \,,$$

showing that $v \in \mathfrak{D}_A$. Therefore, $\mathfrak{D}_A = \mathfrak{D}_{A^*}$.

To see that A is unique, consider A' selfadjoint which coincides with A_i on $\mathcal{H}_i$, for all $i \geqslant 1$. Since A' is selfadjoint, it is closed, and so well defined at every $u \in \mathcal{H}$ for which the series

$$\sum_{i=1}^{\infty} A' u_i \tag{3.7.3}$$

is convergent. Furthermore, for all such u, the series converges to $A'u$. But $A'u_i = A_i u_i = Au_i$ and, since all terms are pairwise orthogonal, the series (3.7.3) converges if and only if $\sum_{i=1}^{\infty} \|Au_i\|^2 < \infty$. Therefore, $\mathfrak{D}_A \subseteq \mathfrak{D}_{A'}$ and, for all $u \in \mathfrak{D}_A$, $A'u = Au$. That is, $A \subseteq A'$. But A is selfadjoint and hence maximal symmetric, so we must have $A' = A$. The proof is complete. □

The operator A constructed in Lemma 3.7.1 is unbounded unless all the operators A_i possess a common bound. We have thus found a way to construct unbounded selfadjoint operators from bounded symmetric ones. A remarkable result is that, in fact, any selfadjoint operator can be constructed in this way.

Lemma 3.7.2 *Consider a selfadjoint operator A acting in $\mathcal{H}$. There exists a sequence*

$$\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_i, \ldots$$

of closed, pairwise orthogonal, subspaces of $\mathcal{H}$, such that $\oplus_i \mathcal{H}_i = \mathcal{H}$ and the restriction $A|_{\mathcal{H}_i}$ is a bounded symmetric operator mapping $\mathcal{H}_i$ into itself, for all $i \geqslant 1$. Moreover, the restriction to $\mathcal{H}_i$ of any bounded operator T such that $T \smile A$ is a bounded operator mapping $\mathcal{H}_i$ into itself, for all $i \geqslant 1$.

Proof. Consider the operators of Theorem 3.4.3:

$$\mathsf{B} = (I + A^2)^{-1} \quad \text{and} \quad \mathsf{C} = A\mathsf{B} = A(I + A^2)^{-1}.$$

B is bounded symmetric and satisfies $0 \leqslant \mathsf{B} \leqslant I$. By the spectral theorem for bounded symmetric operators, B has a unique spectral family $(F_\lambda)_{\lambda \in \mathbb{R}}$ with bounds 0 and 1, such that

$$\mathsf{B} = \int_0^{1+\varepsilon} \lambda \, \mathrm{d}F_\lambda.$$

We now show that F is continuous at $\lambda = 0$, i.e. for all $u \in \mathcal{H}$, $F_{0+0}u = \lim_{\lambda\searrow 0} F_\lambda u = F_0 u = 0$. We start by recalling that, by Theorem 2.2.1, the strong limit F_{0+0} exists. But for all $\lambda > 0$, there is a sequence $(\mu_n)_{n=1}^\infty \subset \mathbb{R}$ which converges to 0 and such that $0 < \mu_n < \lambda$ for all $n \geqslant 1$. Since $F_\lambda F_{\mu_n} = F_{\mu_n}$ for all $n \geqslant 1$, the continuity of F_λ then implies, for all $u \in \mathcal{H}$,

$$F_\lambda F_{0+0}u = \lim_{n\to\infty} F_\lambda F_{\mu_n} u = \lim_{n\to\infty} F_{\mu_n} u = F_{0+0}u.$$

It follows that, for any partition $\Pi = (\lambda_k)_{k=0}^m$ of $[0, 1+\varepsilon]$,

$$\begin{aligned} S_\Pi F_{0+0} &= \sum_{k=1}^m \lambda_k (F_{\lambda_k} - F_{\lambda_{k-1}}) F_{0+0} \\ &= \lambda_1 (F_{0+0} - F_0) + \sum_{k=2}^m \lambda_k (F_{0+0} - F_{0+0}) \\ &= \lambda_1 F_{0+0}. \end{aligned}$$

Thus, letting $|\Pi| \to 0$, we get $\mathsf{B}F_{0+0} = 0$. Since $\mathsf{B}\mathsf{B}^{-1} = I$, it follows that

$$F_{0+0} = \mathsf{B}^{-1}\mathsf{B}F_{0+0} = 0. \tag{3.7.4}$$

We then define the projections

$$P_1 = F_{1+\varepsilon} - F_{\frac{1}{2}}, \quad P_i = F_{\frac{1}{i}} - F_{\frac{1}{i+1}} \text{ for } i \geqslant 2,$$

and we let $\mathcal{H}_i = \operatorname{rge} P_i$. We observe that the closed subspaces $\mathcal{H}_i$ are pairwise orthogonal and satisfy $\oplus_i \mathcal{H}_i = \mathcal{H}$ since $P_i P_j = 0$ for $i \neq j$ and, by (3.7.4),

$$\sum_{i=1}^\infty P_i = F_{1+\varepsilon} - F_{0+0} = I.$$

We now need to show that the restriction of A to $\mathcal{H}_i$ is bounded and symmetric for all $i \geqslant 1$. By Lemma 3.3.2, if $P_i \smile A$, then $P_i A P_i = A P_i$. In this case, if we show that AP_i is defined everywhere and bounded, the selfadjointness of A will imply that the restriction of A to $\mathcal{H}_i$ is a symmetric operator on $\mathcal{H}_i$. Thus, we need only show $P_i \smile A$ and $AP_i \in \mathcal{B}(\mathcal{H})$. To this end, we invoke Corollary 3.5.7: $\mathsf{B} \smile A$ and $\mathsf{CB} = \mathsf{BC}$. Since C commutes with B, it follows by part (b) of Theorem 2.2.1 that C commutes with each F_λ and hence with each P_i. By part (d) of Theorem 3.6.6, C also commutes with all F-measurable functions of B. Consider then the bounded functions $s_i : [0,1] \to \mathbb{R}$ defined for all $i \geqslant 1$ by

$$s_i(\lambda) = \begin{cases} \frac{1}{\lambda} & \text{if } \lambda \in \left[\frac{1}{i+1}, \frac{1}{i}\right), \\ 0 & \text{otherwise.} \end{cases}$$

By Theorem 3.6.6,

$$s_i(\mathsf{B}) = \int_0^{1+\varepsilon} s_i(\lambda)\,\mathrm{d}F_\lambda \in \mathcal{B}(\mathcal{H})$$

and

$$s_i(\mathsf{B})\mathsf{B} = \mathsf{B}s_i(\mathsf{B}) = \int_0^{1+\varepsilon} \lambda s_i(\lambda)\,\mathrm{d}F_\lambda = \int \chi_{[\frac{1}{i+1},\frac{1}{i})}\,\mathrm{d}F_\lambda = P_i.$$

It follows that

$$AP_i = A\mathsf{B}s_i(\mathsf{B}) = \mathsf{C}s_i(\mathsf{B}),$$

showing that $AP_i \in \mathcal{B}(\mathcal{H})$. On the other hand, $\mathsf{B}\smile A$ and we have

$$P_iA = s_i(\mathsf{B})\mathsf{B}A \subseteq s_i(\mathsf{B})A\mathsf{B} = s_i(\mathsf{B})\mathsf{C}.$$

Since C commutes with $s_i(\mathsf{B})$, we have shown that

$$P_iA \subseteq s_i(\mathsf{B})\mathsf{C} = \mathsf{C}s_i(\mathsf{B}) = AP_i,$$

that is, $P_i \smile A$, as expected.

Finally, consider $T \in \mathcal{B}(\mathcal{H})$ such that $T\smile A$. It follows from part (d) of Corollary 3.5.7 that $T\mathsf{B} = \mathsf{B}T$. By Theorem 3.6.6, T commutes with any function of B, and so with each $P_i = \mathsf{B}s_i(\mathsf{B})$. Therefore, we indeed have $P_iTP_i = TP_i$. This concludes the proof of the lemma. □

Theorem 3.7.3 (Spectral Theorem II) *Let A be a selfadjoint operator. There exists a unique spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ such that*

$$A = \int \lambda\,\mathrm{d}E_\lambda.$$

Furthermore, any bounded operator T such that $T\smile A$ commutes with each E_λ.

We call $(E_\lambda)_{\lambda\in\mathbb{R}}$, the **spectral family of** A.

Proof. By Lemma 3.7.2, there exist pairwise orthogonal closed subspaces

$$\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_i, \ldots$$

such that $\oplus_i \mathcal{H}_i = \mathcal{H}$, and for which the restriction A_i of A to $\mathcal{H}_i$ is a bounded symmetric operator on $\mathcal{H}_i$, $i = 1, 2, \ldots$. By the spectral theorem for bounded symmetric operators, there exists for each $i \geqslant 1$ a unique spectral family $(E_{\lambda,i})_{\lambda\in\mathbb{R}}$ on $\mathcal{H}_i$ such that

$$A_i = \int_{m_i}^{M_i+\varepsilon} \lambda\,\mathrm{d}E_{\lambda,i}.$$

Then, for all $\lambda \in \mathbb{R}$, each $E_{\lambda,i}$ is a bounded symmetric operator on $\mathcal{H}_i$. Therefore, by Lemma 3.7.1, there exists a selfadjoint operator E_λ which reduces to $E_{\lambda,i}$ in each subspace $\mathcal{H}_i$. Moreover, for all $\lambda \in \mathbb{R}$, the domain of E_λ is $\mathcal{H}$. Indeed, for all $u \in \mathcal{H}$, denoting by u_i the projection of u onto $\mathcal{H}_i$, it follows from Pythagoras' theorem that

$$\sum_{i=1}^{m} \|E_{\lambda,i} u_i\|^2 \leqslant \sum_{i=1}^{m} \|u_i\|^2 = \Big\| \sum_{i=1}^{m} u_i \Big\|^2 .$$

Hence, letting $m \to \infty$, we get

$$\sum_{i=1}^{\infty} \|E_{\lambda,i} u_i\|^2 \leqslant \|u\|^2 ,$$

and so $u \in \mathfrak{D}_{E_\lambda}$. Since E_λ is selfadjoint its graph is closed and, since it is defined on $\mathcal{H}$, it is bounded. E_λ is therefore a bounded symmetric operator, for all $\lambda \in \mathbb{R}$.

Next we prove that $E_\lambda^2 = E_\lambda$, so that E_λ is a projection. For all $i \geqslant 1$ we have

$$E_\lambda^2 u_i = E_{\lambda,i}^2 u_i = E_{\lambda,i} u_i = E_\lambda u_i .$$

Hence, the selfadjoint operators E_λ^2 and E_λ coincide on $\mathcal{H}_i$ for all $i \geqslant 1$ and so $E_\lambda^2 = E_\lambda$.

We now show that, for $\lambda < \mu$, $E_\lambda \leqslant E_\mu$. For any $u \in \mathcal{H}$,

$$\Big\langle (E_\lambda - E_\mu) \sum_{i=1}^{m} u_i, \sum_{i=1}^{m} u_i \Big\rangle = \sum_{i=1}^{m} \langle (E_{\lambda,i} - E_{\mu,i}) u_i, u_i \rangle \geqslant 0,$$

so that, letting $m \to \infty$,

$$\langle (E_\lambda - E_\mu) u, u \rangle \geqslant 0.$$

We now prove that the family $(E_\lambda)_{\lambda\in\mathbb{R}}$ is strongly left-continuous. By Lemma 3.6.2, for all $\mu \in \mathbb{R}$ the left pointwise limit $E_{\mu-0}$ is a projection. We need only show that $E_{\mu-0}$ and E_μ coincide on each $\mathcal{H}_i$. But this is clear since, for all $u_i \in \mathcal{H}_i$,

$$E_{\mu-0} u_i = \lim_{\lambda \nearrow \mu} E_\lambda u_i = \lim_{\lambda \nearrow \mu} E_{\lambda,i} u_i = E_{\mu,i} u_i = E_\mu u_i .$$

To finish proving that $(E_\lambda)_{\lambda\in\mathbb{R}}$ is a spectral family, it only remains to show that, for all $u \in \mathcal{H}$,

$$\lim_{\lambda \to -\infty} E_\lambda u = 0 \quad \text{and} \quad \lim_{\lambda \to \infty} E_\lambda u = u. \tag{3.7.5}$$

Let us first recall that, for all $i \geqslant 1$, the spectral family $(E_{\lambda,i})_{\lambda\in\mathbb{R}}$ is bounded by the lower and upper bounds m_i and M_i of A_i. Now fix $u \in \mathcal{H}$ and consider $\varepsilon > 0$. There exists $N_\varepsilon \geqslant 1$ such that

$$\Big\| \sum_{i=N_\varepsilon+1}^{\infty} u_i \Big\| < \varepsilon$$

and, since $\|E_\lambda\| = 1$,

$$\|E_\lambda u\| \leqslant \Big\| \sum_{i=1}^{N_\varepsilon} E_{\lambda,i} u_i \Big\| + \Big\| E_\lambda \sum_{i=N_\varepsilon+1}^{\infty} u_i \Big\| < \Big\| \sum_{i=1}^{N_\varepsilon} E_{\lambda,i} u_i \Big\| + \varepsilon. \tag{3.7.6}$$

Putting $m_\varepsilon = \min_{1\leqslant i \leqslant N_\varepsilon} m_i$, we have $\sum_{i=1}^{N_\varepsilon} E_{\lambda,i} u_i = 0$ for all $\lambda \leqslant m_\varepsilon$. It then follows from (3.7.6) that

$$\|E_\lambda u\| < \varepsilon \quad \text{for all } \lambda \leqslant m_\varepsilon,$$

showing the first statement of (3.7.5).

On the other hand, for all $\varepsilon > 0$, there exists $N_\varepsilon \geqslant 1$ such that $\big\| \sum_{i=N_\varepsilon+1}^{\infty} u_i \big\| < \frac{1}{2}\varepsilon$ and so

$$\begin{aligned} \|E_\lambda u - u\| &\leqslant \Big\| \sum_{i=1}^{N_\varepsilon} E_{\lambda,n} u_i - u_i \Big\| + \Big\| (E_\lambda - I) \sum_{i=N_\varepsilon+1}^{\infty} u_i \Big\| \\ &\leqslant \Big\| \sum_{i=1}^{N_\varepsilon} E_{\lambda,i} u_i - u_i \Big\| + 2 \Big\| \sum_{i=N_\varepsilon+1}^{\infty} u_i \Big\| \\ &< \Big\| \sum_{i=1}^{N_\varepsilon} E_{\lambda,i} u_i - u_i \Big\| + \varepsilon. \end{aligned} \tag{3.7.7}$$

Hence, letting $M_\varepsilon = \max_{1\leqslant i \leqslant N_\varepsilon} M_i$, we have

$$\sum_{i=1}^{N_\varepsilon} E_{\lambda,i} u_i = \sum_{i=1}^{N_\varepsilon} u_i \quad \text{for all } \lambda \geqslant M_\varepsilon,$$

and it then follows from (3.7.7) that

$$\|E_\lambda u - u\| < \varepsilon \quad \text{for all } \lambda \geqslant M_\varepsilon.$$

This proves the second statement in (3.7.5).

We now show that the spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ satisfies

$$A = \int \lambda \, \mathrm{d}E_\lambda.$$

Theorem 3.6.5 ensures that $\int \lambda \, \mathrm{d}E_\lambda$ is a selfadjoint operator since $f(\lambda) = \lambda$ is real-valued. By Lemma 3.7.1 it thus suffices to show that $\int \lambda \, \mathrm{d}E_\lambda$ and A coincide on each $\mathcal{H}_i$. To see this, consider a sequence of step functions $(t_l)_{l=1}^{\infty}$ of the form

$$t_l = \sum_{k=1}^{p_l} r_k^l \chi_{I_k^l}, \quad l \geqslant 1,$$

which converges to $f(\lambda) = \lambda$ in $L^2(\mathbb{R}, \mu_{\|E_\lambda u_i\|^2})$ as $l \to \infty$. Then

$$\begin{aligned}\int \lambda \, \mathrm{d}E_\lambda u_i &= \lim_{l\to\infty} \int t_l(\lambda)\, \mathrm{d}E_\lambda u_i = \lim_{l\to\infty} \sum_{k=1}^{p_l} r_k^l E(I_k^l) u_i \\ &= \lim_{l\to\infty} \sum_{k=1}^{p_l} r_k^l E_i(I_k^l) u_i = \int \lambda \, \mathrm{d}E_{\lambda,i} u_i \\ &= \int_{m_i}^{M_i+\varepsilon} \lambda \, \mathrm{d}E_{\lambda,i} u_i = A_i u_i = A u_i,\end{aligned}$$

where we have used the fact that if, for instance, $I = [a, b]$, then

$$E(I)u_i = (E_{b+0} - E_a)u_i = (E_{b+0,i} - E_{a,i})u_i = E_i(I)u_i.$$

Finally, consider an operator $T \in \mathcal{B}(\mathcal{H})$ such that $T \smile A$. By Lemma 3.7.2, for each $i \geqslant 1$, the restriction T_i of T to $\mathcal{H}_i$ satisfies $T_i \in \mathcal{B}(\mathcal{H}_i)$. For all $i \geqslant 1$, since T_i commutes with A_i, it follows by the spectral theorem for bounded symmetric operators that T_i commutes with $E_{\lambda,i}$ for all $\lambda \in \mathbb{R}$. Therefore, for all $\lambda \in \mathbb{R}$ and all $u \in \mathcal{H}$,

$$TE_\lambda u = T \lim_{i\to\infty} \sum_{k=1}^{i} E_{\lambda,i} u_i = \lim_{i\to\infty} \sum_{k=1}^{i} E_{\lambda,i} T_i u_i = E_\lambda T u.$$

To complete the proof, it only remains to show that the spectral family associated with A is unique, which will be done in Problem 3.13. □

Definition 3.7.4 Similarly to the case of bounded symmetric operators, thanks to Theorem 3.7.3 we can now define a function of any selfadjoint operator A. Let $(E_\lambda)_{\lambda\in\mathbb{R}}$ be the spectral family of A. For any E-measurable function f, we write

$$f(A) = E(f) = \int f(\lambda) \, \mathrm{d}E_\lambda. \ \blacklozenge$$

Problems

1. Consider an operator $T : \mathfrak{D}_T \subset \mathcal{H} \to \mathcal{H}$, and suppose that T is bounded on $\mathfrak{D}_T$, in the sense that there is a constant $C \geqslant 0$ such that $\|Tu\| \leqslant C \|u\|$, $u \in \mathfrak{D}_T$. Show that T can be extended to a bounded linear operator on $\mathcal{H}$.

2. Let $S : \mathfrak{D}_S \subset \mathcal{H} \to \mathcal{H}$ be a one-to-one operator. Consider the following additional properties:
   (i) S is closed.
   (ii) $\operatorname{rge} S$ is dense.
   (iii) $\operatorname{rge} S$ is closed.
   (iv) There is a constant C such that $\|Su\| \geqslant C\,\|u\|$ for all $u \in \mathfrak{D}_S$.

   (a) Prove that (i)–(iii) imply (iv). *Hint:* Apply the Closed Graph Theorem to S^{-1}.
   (b) Prove that (ii)–(iv) imply (i).
   (c) Prove that (i) and (iv) imply (iii).

3. Prove that, if T^{-1}, T^* and $(T^{-1})^*$ exist, then $(T^*)^{-1}$ also exists, and $(T^{-1})^* = (T^*)^{-1}$.

4. (a) Show that a densely defined operator T is closable if and only if T^* is densely defined, in which case $\overline{T} = T^{**}$.
   (b) Prove that, if a densely defined operator T is closable, then $\left(\overline{T}\right)^* = T^*$.

5. We define an operator T on $L^2(\mathbb{R})$ by $(Tu)(x) = \phi(x)u(x)$, where $\phi : \mathbb{R} \to \mathbb{C}$ is a bounded function. Show that T is bounded and compute its norm. Find T^*. Under what condition is $T = T^*$? If S is defined by $(Su)(x) = \psi(x)u(x)$ with $\psi : \mathbb{R} \to \mathbb{C}$ bounded, find TS and $(TS)^*$.

6. For T defined as in the previous problem, suppose now $\lim_{x\to\infty} |\phi(x)| = \infty$. What is the domain of definition of T? Show that T is unbounded and find T^*.

7. Let $\mathcal{H} = L^2[0,1]$ and consider the differential operator $D_1 : \mathfrak{D}_{D_1} \subset \mathcal{H} \to \mathcal{H}$ defined by

$$D_1 := i\frac{\mathrm{d}}{\mathrm{d}x}, \quad \mathfrak{D}_{D_1} := \left\{u \in AC[0,1]\,;\, u' \in L^2[0,1] \text{ and } u(0) = u(1) = 0\right\}.$$

   Recall: $AC[0,1]$ is the space of *absolutely continuous* functions in $[0,1]$; see [KF80, Sec. 33]. The Lebesgue version of the fundamental theorem of calculus states the following:
   ⋆ If $f \in AC[0,1]$ then it is differentiable a.e. in $[0,1]$, $f' \in L^1[0,1]$ and $f(x) = f(0) + \int_0^x f'(y)\,\mathrm{d}y$.
   ⋆ If $g \in L^1[0,1]$ and $f(x) = f(0) + \int_0^x g(y)\,\mathrm{d}y$, then $f \in AC[0,1]$ and $f' = g$ a.e.

   (a) Show that D_1 is unbounded and symmetric.
   (b) Prove that the adjoint ${D_1}^*$ of D_1 is given by

$${D_1}^* = i\frac{\mathrm{d}}{\mathrm{d}x}, \quad \mathfrak{D}_{{D_1}^*} = \mathfrak{D} := \left\{u \in AC[0,1]\,;\, u' \in L^2[0,1]\right\}.$$

   Hint: To prove that $\mathfrak{D}_{{D_1}^*} \subset \mathfrak{D}$, consider $v \in \mathfrak{D}_{{D_1}^*}$ and make clever use of the relation $\langle D_1 u, v\rangle = \langle u, D_1^* v\rangle$, for a well chosen $u \in \mathfrak{D}_{D_1}$ (constructed using v and $D_1^* v$).

(c) Prove that ${D_1}^{**} = D_1$.

(d) Now consider $\mathcal{H} = L^2(\mathbb{R})$ and the operator $D_2 : \mathfrak{D}_{D_2} \subset \mathcal{H} \to \mathcal{H}$ defined by $D_2 := i\frac{\mathrm{d}}{\mathrm{d}x}$ and

$$\mathfrak{D}_{D_2} := \left\{u \in L^2(\mathbb{R})\,;\, u' \in L^2(\mathbb{R}) \text{ and } u \in AC[a,b] \text{ for any } -\infty < a < b < \infty\right\}.$$

Prove that D_2 is selfadjoint.
Hint: To prove that $\mathfrak{D}_{{D_2}^*} \subset \mathfrak{D}_{D_2}$, extend the arguments in (b) to the case of an arbitrary finite interval $[a,b]$.

8. Let again $\mathcal{H} = L^2[0,1]$ and consider now $D_3 : \mathfrak{D}_{D_3} \subset \mathcal{H} \to \mathcal{H}$ defined by

$$D_3 := i\frac{\mathrm{d}}{\mathrm{d}x}, \quad \mathfrak{D}_{D_3} := C_0^1(0,1).$$

(a) Show that D_3 is unbounded.

(b) Prove that the adjoint ${D_3}^*$ of D_3 exists and is given by

$$ {D_3}^* = i\frac{\mathrm{d}}{\mathrm{d}x}, \quad \mathfrak{D}_{{D_3}^*} = \left\{u \in AC[0,1]\,;\, u' \in L^2[0,1]\right\}.$$

Hint: Consider any subinterval $[a,b] \subseteq [0,1]$, and construct a sequence $(u_n) \subset \mathfrak{D}_{D_3}$ which converges to $\chi_{[a,b]}$. Use this sequence to compute $\int_a^b ({D_3}^* u)(x)\,\mathrm{d}x$ for any $u \in \mathfrak{D}_{{D_3}^*}$.

(c) Observe the inclusion relations between $\mathfrak{D}_{D_3}, \mathfrak{D}_{{D_3}^*}$ and $\mathfrak{D}_{D_1}, \mathfrak{D}_{{D_1}^*}$.

(d) Prove that the closure $\overline{D_3}$ of D_3 is given by

$$\overline{D_3} = i\frac{\mathrm{d}}{\mathrm{d}x}, \quad \mathfrak{D}_{\overline{D_3}} = \left\{u \in AC[0,1]\,;\, u' \in L^2[0,1] \text{ and } u(0) = u(1) = 0\right\}.$$

9. Consider the multiplication operator in $\mathcal{H} = L^2(\mathbb{R})$ defined by

$$(Xu)(x) =: xu(x), \quad \mathfrak{D}_X := \left\{u \in L^2(\mathbb{R})\,;\, xu(x) \in L^2(\mathbb{R})\right\}.$$

Show that X is an unbounded selfadjoint operator. Find its spectrum and its spectral family.

10. Prove Theorem 3.6.5.

11. Show that, given a spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ with finite lower and upper bounds $m < M$, the integral of any continuous function $f : [m,M] \to \mathbb{C}$ in the sense of Definition 3.6.4 coincides with that of Definition 2.1.3.

12. Prove part (d) of Theorem 3.6.6.

13. Prove the uniqueness of the spectral family in Theorem 3.7.3.

Chapter 4

Applications to quantum mechanics

In this chapter we will define and study elementary properties of the basic observables of the quantum mechanical particle on the real line: energy, position, momentum. In quantum mechanics, physical observables are represented by selfadjoint operators acting in a Hilbert space $\mathcal{H}$, the elements of which represent the possible states of the system. The fundamental postulates of quantum mechanics will be formulated in Section 4.2.

As usual in physics, but perhaps even more in quantum mechanics, an important role is played by the symmetries of the system. Indeed, by Noether's theorem, each symmetry group gives rise to a conserved physical quantity. Some of the most important symmetries of quantum mechanical systems can be expressed by the action on the Hilbert space $\mathcal{H}$ of one-parameter unitary groups.

4.1 Representation of one-parameter unitary groups

An operator $U : \mathcal{H} \to \mathcal{H}$ is called **unitary** if it is surjective and, for all $u, v \in \mathcal{H}$,

$$\langle Uu, Uv \rangle = \langle u, v \rangle .$$

It then follows easily that $\|U\| = 1$, and that U is invertible with $U^{-1} = U^*$ (Problem 4.1).

Definition 4.1.1 We call **one-parameter unitary group** a mapping $U : \mathbb{R} \to \mathcal{B}(\mathcal{H})$ such that $U(t) \equiv U_t$ is unitary for all $t \in \mathbb{R}$, with

$$U_0 = I \quad \text{and} \quad U_t U_s = U_{t+s} \quad \text{for all } t, s \in \mathbb{R}.$$

A one-parameter unitary group is said to be **strongly continuous** if, for all $u \in \mathcal{H}$, the mapping $t \mapsto U_t u$ is continuous on $\mathbb{R}$. ♦

It follows directly from the definition that a one-parameter unitary group $(U_t)_{t\in\mathbb{R}}$ satisfies

$$U_{-t} = (U_t)^* = (U_t)^{-1}.$$

We also observe that if $(U_t)_{t\in\mathbb{R}}$ is **weakly continuous**, in the sense that, for all $u, v \in \mathcal{H}$, $t \mapsto \langle U_t u, v\rangle$ is continuous on $\mathbb{R}$, then $(U_t)_{t\in\mathbb{R}}$ is in fact strongly continuous (Problem 4.2). We now define the central object of the theory of one-parameter unitary groups.

Definition 4.1.2 Let $(U_t)_{t\in\mathbb{R}}$ be a strongly continuous one-parameter unitary group. The **infinitesimal generator** of $(U_t)_{t\in\mathbb{R}}$ is the operator G defined on the domain

$$\mathfrak{D}_G = \left\{u \in \mathcal{H}\,;\, \lim_{t\to 0} \frac{1}{t}(U_t - I)u \text{ exists}\right\}$$

by

$$Gu = \lim_{t\to 0} \frac{1}{t}(U_t - I)u. \blacklozenge$$

The following result is a first step towards the characterization of strongly continuous (one-parameter) unitary groups.

Theorem 4.1.3 *Let A be a selfadjoint operator acting in $\mathcal{H}$ and $(E_\lambda)_{\lambda\in\mathbb{R}}$ be its spectral family. Then*

$$U_t = \mathrm{e}^{itA} = \int \mathrm{e}^{it\lambda}\,\mathrm{d}E_\lambda, \quad t \in \mathbb{R}, \tag{4.1.1}$$

defines a strongly continuous unitary group with infinitesimal generator iA. Furthermore, if $u \in \mathfrak{D}_A$ then $U_t u \in \mathfrak{D}_A$ for all $t \in \mathbb{R}$.

Proof. Since the function $f(\lambda) = \mathrm{e}^{it\lambda}$ is bounded on $\mathbb{R}$ for all $t \in \mathbb{R}$, it follows by Theorem 3.6.5 that, for all $t \in \mathbb{R}$, $U_t \in \mathcal{B}(\mathcal{H})$ and

$$U_t^* = \int \overline{\mathrm{e}^{it\lambda}}\,\mathrm{d}E_\lambda = \int \mathrm{e}^{-it\lambda}\,\mathrm{d}E_\lambda = U_{-t}.$$

Hence, by Theorem 3.6.5,

$$U_t^* U_t = U_{-t}U_t = \int \mathrm{e}^{-it\lambda}\mathrm{e}^{it\lambda}\,\mathrm{d}E_\lambda = \int 1\,\mathrm{d}E_\lambda = I.$$

Thus, by Problem 4.1, every U_t is unitary. Moreover, using again Theorem 3.6.5,

$$U_t U_s = \int \mathrm{e}^{it\lambda}\mathrm{e}^{is\lambda}\,\mathrm{d}E_\lambda = \int \mathrm{e}^{i(t+s)\lambda}\,\mathrm{d}E_\lambda = U_{t+s} \quad \text{for all } t, s \in \mathbb{R}.$$

Therefore, $(U_t)_{t\in\mathbb{R}}$ is a one-parameter unitary group. We now show that it is strongly continuous. Since for all $x, y \in \mathbb{R}$ we have

$$\left|\mathrm{e}^{ix} - \mathrm{e}^{iy}\right| = 2\left|\sin\frac{x-y}{2}\right|,$$

it follows by Theorem 3.6.5 that, for any given $u \in \mathcal{H}$ and all $s, t \in \mathbb{R}$,

$$\|U_t u - U_s u\|^2 = \left\|\int (\mathrm{e}^{it\lambda} - \mathrm{e}^{is\lambda})\,\mathrm{d}E_\lambda u\right\|^2 = \int \left|\mathrm{e}^{it\lambda} - \mathrm{e}^{is\lambda}\right|^2 \mathrm{d}\mu_{\|E_\lambda u\|^2} = 4\int \left|\sin\frac{(t-s)\lambda}{2}\right|^2 \mathrm{d}\mu_{\|E_\lambda u\|^2}.$$

But since

$$\left|\sin\frac{(t-s)\lambda}{2}\right|^2 \leqslant 1 \quad \text{and} \quad \lim_{s\to t}\left|\sin\frac{(t-s)\lambda}{2}\right|^2 = 0,$$

it follows by dominated convergence (Theorem B.2.5) that

$$\lim_{s\to t}\|U_t u - U_s u\| = 0,$$

showing that $(U_t)_{t\in\mathbb{R}}$ is indeed strongly continuous.

Let us now show that the infinitesimal generator G of $(U_t)_{t\in\mathbb{R}}$ is equal to iA. We first observe that, for all $u \in \mathfrak{D}_A$ and all $t \neq 0$,

$$\left\|\left[\frac{1}{t}(U_t - I) - iA\right]u\right\|^2 = \int \left|\frac{1}{t}(\mathrm{e}^{it\lambda} - 1) - i\lambda\right|^2 \mathrm{d}\mu_{\|E_\lambda u\|^2}.$$

But since

$$\lim_{t\to 0}\frac{1}{t}(\mathrm{e}^{it\lambda} - 1) = i\lambda,$$

it follows that

$$\lim_{t\to 0}\left|\frac{1}{t}(\mathrm{e}^{it\lambda} - 1) - i\lambda\right|^2 = 0. \tag{4.1.2}$$

On the other hand, by the mean-value theorem, $|\frac{1}{t}(\mathrm{e}^{it\lambda} - 1)| \leqslant |\lambda|$, and so

$$\left|\frac{1}{t}(\mathrm{e}^{it\lambda} - 1) - i\lambda\right|^2 \leqslant (|\lambda| + |\lambda|)^2 = 4\lambda^2. \tag{4.1.3}$$

Moreover, the function $4\lambda^2$ is $\mu_{\|E_\lambda u\|^2}$-integrable for any $u \in \mathfrak{D}_A$ since

$$\int \lambda^2\,\mathrm{d}\mu_{\|E_\lambda u\|^2} = \|Au\|^2 < \infty.$$

It therefore follows from (4.1.2), (4.1.3) and the dominated convergence theorem that, for all $u \in \mathfrak{D}_A$,

$$\lim_{t\to 0}\left\|\left[\frac{1}{t}(U_t - I) - iA\right]u\right\|^2 = \lim_{t\to 0}\int\left|\frac{1}{t}(\mathrm{e}^{it\lambda} - 1) - i\lambda\right|^2 \mathrm{d}\mu_{\|E_\lambda u\|^2} = 0,$$

and so

$$Gu = \lim_{t\to 0}\frac{1}{t}(U_t - I)u = iAu \quad \text{for all } u \in \mathfrak{D}_A,$$

showing that $G \supseteq iA$. To see that $G = iA$, it remains to show that $\mathfrak{D}_G \subseteq \mathfrak{D}_A$. To this end, consider $u \in \mathfrak{D}_G$, i.e. $u \in \mathcal{H}$ such that the limit

$$\lim_{t\to 0}\frac{1}{t}(U_t - I)u \quad \text{exists.}$$

In this case we have

$$\|Gu\|^2 = \lim_{t\to 0}\left\|\frac{1}{t}(U_t - I)u\right\|^2 = \lim_{t\to 0}\left\|\int\frac{1}{t}(\mathrm{e}^{it\lambda} - 1)\,\mathrm{d}E_\lambda u\right\|^2 = \lim_{t\to 0}\int\left|\frac{1}{t}(\mathrm{e}^{it\lambda} - 1)\right|^2 \mathrm{d}\mu_{\|E_\lambda u\|^2}.$$

Now, since

$$\lim_{t\to 0}\left|\frac{1}{t}(\mathrm{e}^{it\lambda} - 1)\right|^2 = \lambda^2,$$

it follows by Fatou's lemma (Theorem B.2.4) that

$$\int \lambda^2\,\mathrm{d}\mu_{\|E_\lambda u\|^2} \leqslant \liminf_{t\to 0}\int\left|\frac{1}{t}(\mathrm{e}^{it\lambda} - 1)\right|^2 \mathrm{d}\mu_{\|E_\lambda u\|^2} = \|Gu\|^2.$$

The function $f(\lambda) = \lambda$ therefore belongs to $L^2(\mathbb{R}, \mu_{\|E_\lambda u\|^2})$ and we conclude that $u \in \mathfrak{D}_A$.

We finally show that, if $u \in \mathfrak{D}_A$ then $U_t u \in \mathfrak{D}_A$ for all $t \in \mathbb{R}$. First, by part (h) of Theorem 3.6.5, $E_\lambda U_t = U_t E_\lambda$ for all $t \in \mathbb{R}$ and all $\lambda \in \mathbb{R}$. Hence, since every U_t is unitary, we have for all $\lambda \in \mathbb{R}$ and all $t \in \mathbb{R}$ that

$$\|E_\lambda U_t u\|^2 = \langle E_\lambda U_t u, E_\lambda U_t u\rangle = \langle U_t E_\lambda u, U_t E_\lambda u\rangle = \|E_\lambda u\|^2.$$

It follows that

$$\int \lambda^2 \mathrm{d}\mu_{\|E_\lambda (U_t u)\|^2} = \int \lambda^2 \mathrm{d}\mu_{\|E_\lambda u\|^2} < \infty,$$

and so $U_t u \in \mathfrak{D}_A$, which concludes the proof of the theorem. □

The next theorem, due to Marshall H. Stone, shows that, in fact, any strongly continuous one-parameter unitary group is of the form (4.1.1), for a uniquely defined selfadjoint A.

Theorem 4.1.4 (Stone's Theorem) *Let $(U_t)_{t\in\mathbb{R}}$ be a strongly continuous one-parameter unitary group. There exists a unique selfadjoint operator A such that*

$$U_t = \mathrm{e}^{itA} \quad \text{for all } t \in \mathbb{R}.$$

Furthermore, $U_t \smile A$ for all $t \in \mathbb{R}$.

Proof. See Problem 4.3. □

4.2 A glimpse of quantum mechanics

A self-contained exposition of the theory of quantum mechanics would largely exceed the scope of these notes. We refer the interested reader to the classic works [Dir58, Mes99, VN55]. Let us just mention that quantum mechanics typically describes microscopic systems, in the order of atomic length scales and below. At this level, the description of physical systems cannot rely any more on the macroscopic concepts we experience in our everyday lives. In fact, atoms and subatomic particles turn out to have very peculiar properties when we try to express them in terms of our macroscopic perception of the world. For instance, a most striking feature is that it is not possible in general to assign a precise position in space to a quantum mechanical particle, nor a precise speed (or rather momentum = mass · speed). The Heisenberg uncertainty principle, which is one of the cornerstones of quantum mechanics, indeed states that the respective standard deviations Δq and Δp of the position q and the momentum p of a particle satisfy

$$\Delta q \Delta p \geqslant \hbar/2 \tag{4.2.1}$$

where $\hbar = h/2\pi$ is the reduced Planck constant, with the Planck constant $h \simeq 6.62610^{-34}\,\mathrm{J}\cdot\mathrm{s}$. And this is not a matter of 'not being able to' measure things more precisely, but rather a fundamental obstruction of Nature. In fact, as we shall see below, this restriction is due to the status of physical observables in quantum theory. The constant h is named after Max Planck, who was the first to propose a physical model in which electromagnetic radiation could only be emitted as integer multiples — called *quanta* — of the fundamental unit h (1900). In fact, Planck introduced this assumption as a trick to resolve an apparent paradox in the description of the black-body radiation, and did not himself believe in this 'quantized' emission process in real physical terms. The corpuscular nature of light was later fully assumed by Einstein in his work on the photoelectric effect (1905) which owed him the Nobel prize in physics in 1921.[1]

[1] Ironically, Einstein later became one of the fiercest opponents of quantum physics.

We shall soon see that the conceptual framework of quantum mechanics is in sharp contrast with that of classical mechanics. In classical mechanics, the observables of a point-like particle are (smooth) functions $f(\mathbf{q}, \mathbf{p})$, where $(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^3 \times \mathbb{R}^3$ are the position and the momentum of the particle. For example, the Hamiltonian (energy)

$$H(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{p} \cdot \mathbf{p}}{2m} + V(\mathbf{q})$$

describes a particle in a potential field $V : \mathbb{R}^3 \to \mathbb{R}$. It emerged from various attempts to obtain a satisfactory theory of atoms at the beginning of the last century (most notably from the works of Max Planck, Albert Einstein, Niels Bohr, Werner Heisenberg, Erwin Schrödinger, Max Born, Pascual Jordan, Wolfgang Pauli and Paul Dirac) that this classical description of the world dramatically fails to account for microscopic phenomena, such as the very stability of atoms[2], and their emitting of light only at certain universal discrete frequencies (Balmer series). To cut a long story short, on the one hand Heisenberg realized that atomic observables were best represented by (infinite) matrices, a theory now referred to as *matrix mechanics* (1925). A stunning consequence of this new paradigm is that the product of quantum observables is non-commutative. We shall see that this results in particular in the famous *uncertainty principle* mentioned above. On the other hand, Schrödinger derived a theory describing the evolution of a quantum particle through a 'wave function',[3] therefore known as *wave mechanics* (1926). It should be noted that Schrödinger's theory is a dynamical one, where the central equation — known as the *Schrödinger equation* — governs the time evolution of the wave function. On the other hand, Heisenberg's matrix mechanics only describes the stationary states of the system. Apart from this difference, the complete equivalence of the two theories was proved mathematically by Dirac shortly thereafter[4] and is presented in his famous book *The Principles of Quantum Mechanics* (1930) [Dir58].

Thus, following three decades of intense creative work from a bunch of brilliant theoretical physicists, a new coherent theory was finally born. And it very soon turned out that a natural mathematical framework to describe the theory was that of (unbounded) operators acting in a Hilbert space. A complete set of axioms for the new physics was formulated by von Neumann in *Mathematical Foundations of Quantum Mechanics* (1932) [VN55], while parts of the mathematical theory was developed independently by Marshall H. Stone in his great book *Linear Transformations in Hilbert Space* (1932) [Sto32]. We shall now state the basic postulates of quantum mechanics in this Hilbert space formalism.

[2]According to Newtonian mechanics and Maxwell's theory of electromagnetics, the electron radiating electromagnetic energy while orbiting about the nucleus would very quickly collapse to it.

[3]The idea of representing a particle as an oscillatory phenomenon was motivated by earlier considerations from Einstein and de Broglie, pointing that both light and matter can behave either like waves or like particles depending on the experimental setting, the now famous 'wave/particle duality'.

[4]This is at least the official version, see the Handout for an interesting discussion on the history and equivalence of the two theories.

The basic postulates of quantum mechanics

The following postulates pertain to *quantum systems*, that is, systems which are best described by the laws of quantum mechanics. We shall see that underlying lies the essential notion of *measurement apparatus*. Indeed, since we don't have any macroscopic intuition of quantum systems, the only thing the theory describes/predicts is the results of measurements performed on the system under given experimental conditions. The so-called 'Copenhagen interpretation' of quantum mechanics — largely personified by Bohr — pushed this new paradigm so far as to claim that it is irrelevant to attempt any interpretation of the theory in terms of elements of *physical reality* (a notion that was the object of intense debates, notably between Einstein and Bohr), outside the scope of measurement devices. For instance, there exists no such thing as the position of an electron, so long as the electron is not detected experimentally by an external *observer* (another notion prone to vast philosophical debate)!

Postulate I At any given time, the state of the system is represented by a vector $\psi \neq 0$ of a complex separable Hilbert space $\mathcal{H}$. Furthermore, for all $c \in \mathbb{C}\backslash\{0\}$, the vector $c\psi$ represents the same state as ψ. Thus, the states of the system are in one-to-one correspondence with the rays

$$\{c\psi\,;\, c \in \mathbb{C}\} \subseteq \mathcal{H}, \quad \psi \in \mathcal{H}\backslash\{0\},$$

or, equivalently, with the orthogonal projections P_ψ onto these one-dimensional subspaces.

Postulate II Every **observable** $\mathcal{A}$ is represented by a selfadjoint operator A acting in $\mathcal{H}$.

Postulate III The result of a **measurement** of the observable $\mathcal{A}$ can only be a real number λ, eigenvalue of the operator A.

Postulate IV If the system is in state ψ at time t, then the **probability** of observing the value λ when measuring the observable $\mathcal{A}$ at time t is given by

$$\text{Prob}_\psi\{\text{meas. of } \mathcal{A} \text{ yields } \lambda\} = \frac{\langle\psi, P_\lambda\psi\rangle}{\langle\psi,\psi\rangle}, \tag{4.2.2}$$

where P_λ is the projection on the eigenspace of A corresponding to λ.

Postulat V The **mean value** of A, calculated over a large number of systems all prepared in state ψ, is given by

$$\langle A\rangle_\psi = \frac{\langle\psi, A\psi\rangle}{\langle\psi,\psi\rangle}. \tag{4.2.3}$$

Postulate VI If the system is in state ψ, then immediately after a measurement of $\mathcal{A}$ yielding the value λ, the system is in state $\phi = P_\lambda \psi$, and so ϕ is an eigenvector of A corresponding to the eigenvalue λ.

Postulate VIII There exists a selfadjoint operator H, called the **Hamiltonian**, representing the energy of the system and such that the time evolution of the system is given by the **Schrödinger equation**

$$i\hbar \partial_t \psi_t = H\psi_t,$$

where ψ_t is the state of the system at time t and $\hbar = h/2\pi$ is the reduced Planck constant.

Remark 4.2.1 If the state ψ is **normalized** so that $\|\psi\| = 1$, the formulas (4.2.2) and (4.2.3) simply become

$$\mathrm{Prob}_\psi\{\text{meas. of } \mathcal{A} \text{ yields } \lambda\} = \langle \psi, P_\lambda \psi \rangle \,, \tag{4.2.4}$$

and

$$\langle A \rangle_\psi = \langle \psi, A\psi \rangle \,. \tag{4.2.5}$$

Distribution function of an observable

Consider a quantum system and let $\mathcal{H}$ be the state space. Suppose the state $\psi \in \mathcal{H}$ of the system at a given time is normalized. Then consider an observable $\mathcal{A}$ represented by a selfadjoint operator A, with spectral family $(E_\lambda)_{\lambda \in \mathbb{R}}$, so that

$$A = \int \lambda \, \mathrm{d}E_\lambda.$$

The normalization of ψ yields

$$1 = \|\psi\|^2 = \int \mathrm{d}\mu_{\|E_\lambda \psi\|^2},$$

showing that $\mu_{\|E_\lambda \psi\|^2}$ is a probability measure on $\mathbb{R}$. From part (d) of Theorem 3.6.5, the mean value of $\mathcal{A}$, defined in Postulate V, can be written as

$$\langle A \rangle_\psi = \langle \psi, A\psi \rangle = \int \lambda \, \mathrm{d}\mu_{\|E_\lambda \psi\|^2},$$

which is precisely the expectation value of the probability measure $\mu_{\|E_\lambda \psi\|^2}$. The function

$$\lambda \mapsto F_\psi(\lambda) := \|E_\lambda \psi\|^2 = \langle E_\lambda \psi, \psi \rangle$$

thus represents the **distribution function** of the observable $\mathcal{A}$ for our system in state ψ. It is indeed increasing, left-continuous[5], satisfies

$$F_\psi(\lambda) = \int_{-\infty}^{\lambda} \mathrm{d}\mu_{\|E_\lambda \psi\|^2}$$

and, in particular,

$$\lim_{\lambda\to-\infty} F_\psi(\lambda) = 0, \qquad \lim_{\lambda\to+\infty} F_\psi(\lambda) = 1.$$

Furthermore, if F_ψ is absolutely continuous on any finite interval, then the measure $\mu_{\|E_\lambda\psi\|^2}$ has a probability density φ_ψ satisfying

$$\mathrm{d}\mu_{\|E_\lambda\psi\|^2} = \varphi_\psi(\lambda)\,\mathrm{d}\lambda, \qquad \varphi_\psi(\lambda) = F'_\psi(\lambda) \quad \text{a.e. } \lambda \in \mathbb{R}.$$

With this probabilistic interpretation of the function F_ψ, it follows that

$$\begin{aligned}
\mathrm{Prob}_\psi\{\text{meas. of } \mathcal{A} < \lambda\} &= \langle E_\lambda \psi, \psi\rangle\,, \\
\mathrm{Prob}_\psi\{\text{meas. of } \mathcal{A} \in [a,b]\} &= \langle (E_{b+0} - E_a)\psi, \psi\rangle \\
\text{and} \quad \mathrm{Prob}_\psi\{\text{meas. of } \mathcal{A} = \lambda\} &= \langle (E_{\lambda+0} - E_\lambda)\psi, \psi\rangle \\
&= \begin{cases} 0 & \text{if } \lambda \text{ is not an eigenvalue of } A, \\ \langle P_\lambda \psi, \psi\rangle & \text{if } \lambda \text{ is an eigenvalue of } A, \end{cases}
\end{aligned} \tag{4.2.6}$$

where P_λ is the projection on the eigenspace corresponding to the eigenvalue λ. In particular, we see that Postulates III and IV are logical consequences of Postulate V and the spectral theorem.

Time evolution: Stone's theorem and the existence of dynamics

We say that a quantum system with state space $\mathcal{H}$ is **invariant under time translations** if its evolution in time is governed by a strongly continuous one-parameter unitary group $(U_t)_{t\in\mathbb{R}}$ on $\mathcal{H}$. That is, if the system at time $t = 0$ is in the **initial state** ψ_0, then its state at time t is

$$\psi_t = U_t \psi_0.$$

Then indeed the system "does not see time translations" since we can change the origin in time arbitrarily as

$$t \to t' = t + \tau, \quad \text{i.e.} \quad t_0 = 0 \to t'_0 = \tau,$$

[5]The convention in probability theory is usually to assume right continuity, but the theory can also be developed with left continuity instead.

so that the new initial state is given by $\psi_{t'_0} = \psi_\tau = U_\tau\psi_0$ and, after a time interval $t = t' - t'_0$ has elapsed, we have

$$\psi_{t'} = \psi_{t+\tau} = U_{t+\tau}\psi_0 = U_t U_\tau \psi_0 = U_t \psi_{t'_0},$$

showing that the law of evolution remains identical in the new time frame.

Now for a system invariant under time translations, Stone's theorem ensures that there exists a selfadjoint operator G which is the infinitesimal generator of the group $(U_t)_{t\in\mathbb{R}}$. We recall that the domain of G is given by

$$\mathfrak{D}_G = \Big\{\psi \in H\,;\, \lim_{t\to 0} \frac{1}{\hbar}(U_t - I)\psi \text{ exists}\Big\},$$

and that

$$G\psi = \lim_{t\to 0} \frac{1}{t}(U_t - I)\psi \quad \text{for all } \psi \in \mathfrak{D}_G.$$

We define the selfadjoint **Hamiltonian** H by

$$H = i\hbar G.$$

By Theorem 4.1.3, if $\psi \in \mathfrak{D}_G = \mathfrak{D}_H$ then $U_t\psi \in \mathfrak{D}_H$ for all $t \in \mathbb{R}$. Therefore, given any initial state $\psi_0 \in \mathfrak{D}_H$, we have for all $t \in \mathbb{R}$

$$\partial_t \psi_t = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(\psi_{t+\varepsilon} - \psi_t) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}(U_\varepsilon - I)\psi_t = G\psi_t = -\frac{i}{\hbar}H\psi_t.$$

This is the Schrödinger equation corresponding to the Hamiltonian H, as given in Postulate VIII. We thus conclude that (provided the Hamiltonian operator does not depend on time) postulate VIII is equivalent to requiring that the evolution is invariant under time translations.

We now verify that, **if the system is invariant under time translations then the energy is a constant of the motion**, i.e. that the mean value of H does not change under the time evolution of the system.[6] Indeed, by Stone's theorem, $U_t H\psi = HU_t\psi$ for all $\psi \in \mathfrak{D}_H$. Hence, if $\psi_0 \in \mathfrak{D}_H$, we have for all $t \in \mathbb{R}$

$$\langle H\rangle_{\psi_t} = \langle \psi_t, H\psi_t\rangle = \langle U_t\psi_0, HU_t\psi_0\rangle = \langle U_t\psi_0, U_t H\psi_0\rangle = \langle \psi_0, H\psi_0\rangle = \langle H\rangle_{\psi_0}.$$

In view of Postulate III, the only values of the energy of a system that can be measured in a laboratory are the eigenvalues of the Hamiltonian H. The corresponding states (eigenvectors) given by Postulate VI are called **bound states**. This is the surprising discovery physicists from the early 20th century made by studying spectral rays of light emission by atoms, and observing only discrete sets of frequencies (characteristic of each atom). It is thus of utmost importance in the framework of quantum mechanics to understand the spectral structure of Hamiltonians describing atomic systems. A thorough analysis of this issue can be found in [RS]. (In Volume I a proof of Stone's theorem can be found, as well as interesting historical remarks about the mathematical formalization of quantum mechanics.) Basic examples of Hamiltonians will be given at the end of the next section.

[6] This conservation principle is an instance, in the quantum setting, of Noether's theorem.

4.3 The quantum particle on $\mathbb{R}$

The quantum particle on the real line $\mathbb{R}$ is the system characterized by the following properties.

(A) To any (Borel) subset $\Delta \subseteq \mathbb{R}$ one can associate a measurement device, i.e. an observable $\mathcal{P}_\Delta$ represented by a selfadjoint operator P_Δ, called **particle detector**, taking the value 0 or 1 depending on whether the particle is in Δ or not, respectively.

(B) The set of all operators P_Δ, for $\Delta \subseteq \mathbb{R}$, forms a family of pairwise commuting selfadjoint operators.

(C) To every $a \in \mathbb{R}$ one can associate a translation of the detectors

$$\tau_a P_\Delta = P_{\Delta - a}, \quad \text{where } \Delta - a = \{q \in \mathbb{R}\,;\, q + a \in \Delta\}.$$

Equivalently (see below), the translation can be interpreted as acting on the system rather than on the detectors:

$$U_a \psi = \psi_a, \quad \psi \in \mathcal{H}.$$

(D) The only observables commuting with all the P_Δ are functions of them.

A natural choice of separable Hilbert space satisfying these properties is $L^2(\mathbb{R})$. Then the state of the system at any time is given by a so-called **wave function** $\psi \in L^2(\mathbb{R})$, and the detector $\mathcal{P}_\Delta$ is represented by the projection

$$\big(P_\Delta \psi\big)(q) = \chi_\Delta(q)\psi(q), \quad \psi \in \mathcal{H}, \quad \text{a.e. } q \in \mathbb{R}.$$

Interpretation of the wave function ψ, observable position

For any $\Delta \subseteq \mathbb{R}$, the probability of finding the particle in a normalized state ψ in Δ is the mean value of the observable $\mathcal{P}_\Delta$. By Postulate V it is given by

$$\langle P_\Delta \rangle = \langle \psi, P_\Delta \psi \rangle = \int_\Delta |\psi(q)|^2 \, \mathrm{d}q.$$

Therefore, the function $|\psi(q)|^2$ is interpreted as the density of probability of observing the particle in state ψ.

The observable position is then naturally represented by the multiplication operator

$$(Q\psi)(q) = q\psi(q)$$

on the domain

$$\mathfrak{D}_Q = \{\psi \in L^2(\mathbb{R})\,;\, q\psi(q) \in L^2(\mathbb{R})\}.$$

Indeed, in view of the above probabilistic interpretation of $|\psi(q)|^2$, the mean value of the position of the particle in state ψ is then the expectation value of the probability measure $|\psi(q)|^2\,\mathrm{d}q$,

$$\int_{\mathbb{R}} q|\psi(q)|^2\,\mathrm{d}q.$$

Physically, it represents the 'averaged position' of the particle, when the measurement of position is performed on a large number of copies of the system in the same state ψ. The selfadjoint operator Q is therefore called **position operator**.

Observable momentum

In a similar way to how we associated the Hamiltonian operator to the group of time translations, we shall now apply Stone's theorem to obtain the momentum operator as the infinitesimal generator of space translations on the line.

To each translation $a \in \mathbb{R}$, we associate a transformation acting on the states of the system as mentioned in assumption (C) above. This transformation is explicitly given by

$$\big(U_a\psi\big)(q) = \psi_a(q) = \psi(q-a), \quad q \in \mathbb{R}. \tag{4.3.1}$$

It is not difficult to check (see Problem 4.5) that $(U_a)_{a\in\mathbb{R}}$ defines a strongly continuous one-parameter unitary group on $L^2(\mathbb{R})$. By Stone's theorem, there exists a selfadjoint operator A such that

$$U_a = \mathrm{e}^{iaA}, \quad \text{for all } a \in \mathbb{R}.$$

Moreover,

$$iA\psi = \frac{\mathrm{d}}{\mathrm{d}a}\Big|_{a=0} U_a\psi = \lim_{a\to 0}\frac{1}{a}(U_a - I)\psi, \quad \psi \in \mathfrak{D}_A.$$

Hence, for $\psi \in \mathfrak{D}_A$,

$$\begin{aligned}(A\psi)(q) &= \frac{1}{i}\lim_{a\to 0}\frac{\psi(q-a)-\psi(q)}{a}\\ &= -\frac{1}{i}\lim_{a\to 0}\frac{\psi(q-a)-\psi(q)}{-a}\\ &= -\frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}q}\psi(q).\end{aligned}$$

For reasons of physical dimensions, the **momentum operator** P is defined by $\mathfrak{D}_P = \mathfrak{D}_A$ and

$$(P\psi)(q) = \frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}q}\psi(q), \quad q \in \mathbb{R},$$

so that

$$U_a = \mathrm{e}^{-\frac{i}{\hbar}aP}, \quad \text{for all } a \in \mathbb{R}.$$

Commutation relation and uncertainty principle

For the particle in a normalized state ψ, given an observable represented by a selfadjoint operator A, we define the **variance** of A by

$$\mathrm{var}_\psi(A) = \left\langle [A - \langle A\rangle_\psi I]^2 \right\rangle_\psi = \int_{\mathbb{R}} \left[A - \langle A\rangle_\psi I\right]^2 \psi(q)\overline{\psi(q)}\,\mathrm{d}q$$

(where $\langle\cdot\rangle_\psi$ is the mean value defined in Postulate V) and its **standard deviation** by

$$\Delta_\psi(A) = \sqrt{\mathrm{var}_\psi(A)}.$$

The following theorem is of fundamental importance in the algebraic structure of quantum mechanics. It will be proved in Problem 4.8.

Theorem 4.3.1 *Let A and B be selfadjoint operators acting in the Hilbert space $L^2(\mathbb{R})$. Then the* ***commutator*** *$C = AB - BA$ satisfies*

$$|\langle C\rangle_\psi| \leqslant 2\Delta_\psi(A)\Delta_\psi(B),$$

for all $\psi \in \mathfrak{D}_C$.

We will see in Problem 4.7 that the position and momentum operators satisfy the **Heisenberg commutation relation**

$$QP - PQ = i\hbar I, \tag{4.3.2}$$

where I is the identity operator on the domain $\mathfrak{D}(QP - PQ) = \mathfrak{D}(QP) \cap \mathfrak{D}(PQ)$. We then have the following immediate corollary of Theorem 4.3.1.

Corollary 4.3.2 (Heisenberg's Uncertainty Principle) *The position and momentum operators of the quantum particle on the real line satisfy*

$$\Delta_\psi(Q)\Delta_\psi(P) \geqslant \frac{\hbar}{2}.$$

Particle in a potential

The quantum Hamiltonian for the particle in a potential field $V : \mathbb{R} \to \mathbb{R}$ is constructed by replacing the classical momentum p by $\frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}q}$ in the total mechanical energy $E = \frac{p^2}{2m} + V(q)$, where m is the mass of the particle. This yields

$$(H\psi)(q) = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}q^2}\psi(q) + V(q)\psi(q), \quad \psi \in \mathfrak{D}_H. \tag{4.3.3}$$

The Schrödinger equation then becomes

$$i\hbar\partial_t\psi_t(q) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2}\psi_t(q) + V(q)\psi_t(q).$$

To solve it, one can seek a solution in the form of a **standing wave** $\psi_t(q) = \mathrm{e}^{-\frac{i}{\hbar}Et}\varphi(q)$, which yields the **stationary Schrödinger equation** for φ:

$$E\varphi = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}q^2}\varphi(q) + V(q)\varphi(q),$$

where $E \in \mathbb{R}$ is therefore an eigenvalue of the Hamiltonian H, and φ a corresponding eigenvector.[7]

The spectral theory of **Schrödinger operators** of the form (4.3.3) is very well known. In the typical case of a continuous attractive potential $V < 0$, with $\lim_{|x|\to\infty} V(x) = 0$, the spectrum of H consists of a finite number of negative isolated eigenvalues of multiplicity one, and the continuous spectrum $\sigma_c(H) = [0, \infty)$. But many other scenarios can occur, as discussed e.g. in [RS]. For instance, in three dimensions, the Hamiltonian of the electron in the hydrogen atom is

$$H\psi = -\frac{\hbar^2}{2m}\Delta\psi + V(|\mathbf{q}|)\psi, \quad \text{where } V(|\mathbf{q}|) = -\frac{e^2}{4\pi\epsilon_0|\mathbf{q}|}, \quad \mathbf{q} \in \mathbb{R}^3.$$

Here e is the electric charge of the electron, ϵ_0 is the vacuum electric permittivity, and $|\mathbf{q}|$ is the distance of the electron to the (fixed) nucleus. In this case there are infinitely many negative eigenvalues ($\propto -1/n^2,\ n = 1, 2, \dots$) corresponding to the bound states of the electron.

In the present one-dimensional case, even though it does not belong to $L^2(\mathbb{R})$, it is instructive to study a solution in the form of a **plane wave** $\varphi(q) = \mathrm{e}^{ikq}$ for $k \in \mathbb{R}$. Then our standing wave becomes

$$\psi_t(q) = \mathrm{e}^{i(kq-\omega t)}, \quad \text{where } \omega = E/\hbar.$$

Some simple properties of plane wave solutions will be studied in Problem 4.9, in the case where V is a 'potential barrier', notably displaying the famous *tunnel effect*.

[7]Note that the same Ansatz yields the stationary Schrödinger equation $E\varphi = H\varphi$ for any time-independent Hamiltonian H.

Problems

1. Show that a unitary operator $U : \mathcal{H} \to \mathcal{H}$ satisfies: (a) $\|U\| = 1$; (b) U is invertible with $U^{-1} = U^*$ unitary. Then show that, in fact, $U \in \mathcal{B}(\mathcal{H})$ is unitary if and only if $U^*U = I$.

2. Show that a weakly continuous one-parameter unitary group is strongly continuous.

3. Prove Theorem 4.1.4.

   Strategy: Let $A := -iG$ where G is the infinitesimal generator of $(U_t)_{t\in\mathbb{R}}$, and consider the set $\mathfrak{D}$ of all finite linear combinations of elements of the form

$$u_\phi = \int_{\mathbb{R}} \phi(t) U_t u \,\mathrm{d}t,$$

   for $u \in \mathcal{H}$ and $\phi \in C_0^\infty(\mathbb{R})$ (see Appendix C for the meaning of this integral). The theorem is then proved in three steps.

   (i) Using the results of Section C.2, show that $\mathfrak{D}$ is dense in $\mathcal{H}$, and contained in $\mathfrak{D}_A$.

   (ii) Show that A is symmetric and, in fact, essentially selfadjoint (use Theorem 3.5.6 for this).

   (iii) By differentiating $\|(U_t - \mathrm{e}^{it\bar{A}})u\|^2$ with respect to t, show that $U_t = \mathrm{e}^{it\bar{A}}$. Conclude.

4. Using (4.2.6) and the spectral family of the position operator Q obtained in Problem 3.9, recover the statement that the probability of finding the particle in a normalized state ψ in the interval $\Delta = [a, b]$ is given by $\int_a^b |\psi(q)|^2 \,\mathrm{d}q$.

5. Show that the group of translations acting on $L^2(\mathbb{R})$ as defined in (4.3.1) is a strongly continuous one-parameter unitary group.

6. Prove that the domain $\mathfrak{D}_P$ of the momentum operator P given by Stone's theorem coincides with the domain $\mathfrak{D}_{D_2}$ in Problem 3.7 (d), i.e. that $P = -\hbar D_2$.

7. Prove (4.3.2).

8. Prove Theorem 4.3.1.
   Hint: First show that $C = ST - TS$ where $S = A - \langle A \rangle_\psi$ and $T = B - \langle B \rangle_\psi$.

9. Consider the stationary Schrödinger equation with a **potential barrier** defined in three regions of $\mathbb{R}$ by: (A) $V(q) = 0$ for $q < 0$, (B) $V(q) = V_0 > 0$ for $0 \leqslant q \leqslant a$, (C) $V(q) = 0$ for $q > a$.

Seek solutions in the three regions with the same given energy $E > 0,\ E \neq V_0$, in the following forms:

$$\begin{aligned}\varphi_A(q) &= A_r \mathrm{e}^{ik_A x} + A_l \mathrm{e}^{-ik_A x},\\ \varphi_B(q) &= B_r \mathrm{e}^{ik_B x} + B_l \mathrm{e}^{-ik_B x},\\ \varphi_C(q) &= C_r \mathrm{e}^{ik_C x} + C_l \mathrm{e}^{-ik_C x},\end{aligned}$$

with k_A, k_B, k_C to be determined as functions of E. The indices 'r' and 'l' stand for the direction of the velocity vector of each wave component, respectively 'right' and 'left'.

Now the global solution is obtained by 'gluing' together the solutions $\varphi_A, \varphi_B, \varphi_C$. Find the relations between the coefficients so that the global solution is continuous everywhere, as well as its derivative.

Tunnel effect: We define the **transmission coefficient** t and the **reflection coefficient** r as follows. Consider the solution for which $A_r = 1$ (particle coming from the left), $A_l = r$ (reflection), $C_l = 0$ (no particle coming from the right), and $C_r = t$ (transmission). The number $|t|^2$ is the probability for the particle to be transmitted through the barrier and $|s|^2$ the probability of it being reflected by the barrier. Find the explicit expressions of t and r, in terms of k_A, k_B, k_C and a. Verify that $|t|^2 + |r|^2 = 1$ (conservation of the particle).

Now discuss the two cases $E < V_0$ and $E > V_0$. Decide if the particle is always transmitted or always reflected in each case. How do the results contrast with the situation of a classical particle with the same energy?

Appendix A

The Riemann-Stieltjes integral

In this appendix we will extend the Riemann integral to measures generated by functions of ‘bounded variation’.

A.1 Functions of bounded variation

Let $[a, b] \subset \mathbb{R}$ be a non-empty compact interval. A **partition** Π of $[a, b]$ is a finite sequence of numbers $(\lambda_k)_{k=0}^n$ such that

$$a = \lambda_0 < \lambda_1 < \cdots < \lambda_n = b.$$

The size of Π is the positive number

$$|\Pi| := \max_{k=1,\ldots,n} \lambda_k - \lambda_{k-1}.$$

We denote $\mathfrak{P}[a, b]$ the set of all partitions of $[a, b]$.

A function $\phi : [a, b] \to \mathbb{R}$ is said to have **bounded variation** if there exists a constant $C \geqslant 0$ such that

$$\sum_{k=1}^{n} |\phi(\lambda_k) - \phi(\lambda_{k-1})| \leqslant C,$$

for any partition $\Pi = (\lambda_k)_{k=0}^n \in \mathfrak{P}[a, b]$. If ϕ has bounded variation in $[a, b]$ we define for all $x, y \in [a, b]$ such that $x < y$ the **total variation** of ϕ in $[x, y]$ as the positive number

$$V_x^y(\phi) := \sup_{\Pi \in \mathfrak{P}[x,y]} \sum_{k=1}^{n} |\phi(\lambda_k) - \phi(\lambda_{k-1})| \,.$$

The functions of bounded variation in $[a, b]$ form a vector space, denoted $VB\,[a, b]$.

The following theorems are proved in [KF80], pp. 329 and 330.

Theorem A.1.1 *Let $\phi \in VB[a,b]$. If $a < c < d \leqslant b$ then*

$$V_a^d(\phi) = V_a^c(\phi) + V_c^d(\phi).$$

Theorem A.1.2 *Let $\phi \in VB[a,b]$. If ϕ is left-continuous at a point $x_0 \in [a,b]$, then the function $x \mapsto V_a^x(\phi)$ is also left-continuous at x_0.*

A.2 Definition of the integral

Let us now consider a continuous function $f : [a,b] \to \mathbb{R}$ and a function of bounded variation $\phi \in VB[a,b]$. For any partition $\Pi = (\lambda_k)_{k=0}^n$ and any real numbers

$$\mu_k \in [\lambda_k, \lambda_{k-1}], \quad k = 1, \dots, n,$$

we let

$$S_\Pi = \sum_{k=1}^n f(\mu_k)\big(\phi(\lambda_k) - \phi(\lambda_{k-1})\big).$$

Theorem A.2.1 *Let $f : [a,b] \to \mathbb{R}$ continuous and $\phi \in VB[a,b]$ of bounded variation. There exists a unique number $I \in \mathbb{R}$ satisfying: for all $\varepsilon > 0$ there exists $\delta > 0$ such that if Π is a partition of $[a,b]$ satisfying $|\Pi| < \delta$, then*

$$|S_\Pi - I| < \varepsilon.$$

Furthermore, the number I is independent of the choice of points μ_k.

Proof. Let $\varepsilon > 0$. Since f continuous in a compact set, it is uniformly continuous. So there exists δ_ε such that for all $x, y \in [a,b]$ $|x - y| < \delta_\varepsilon \implies$

$$|f(x) - f(y)| \leqslant \frac{\varepsilon}{2V_a^b(\phi)},$$

where $V_a^b(\phi)$ is the total variation of ϕ in $[a,b]$.

Let us now show that for any two partitions $\Pi, \Pi' \in \mathfrak{P}[a,b]$ we have

$$|\Pi|, |\Pi'| < \delta_\varepsilon \implies |S_\Pi - S_{\Pi'}| \leqslant \varepsilon, \tag{A.2.1}$$

independently of the choice of points μ_k used to compute the partial sums S_Π et $S_{\Pi'}$. We denote by $(\lambda_k)_{k=0}^n$ the partition Π and we consider the partition $\overline{\Pi} = \Pi \cup \Pi'$, which we index as follows:

$$\bar\lambda_0 = \lambda_0 < \bar\lambda_1 < \cdots < \bar\lambda_{k_1} = \lambda_1 < \bar\lambda_{k_1+1} < \cdots < \bar\lambda_{k_2} = \lambda_2 < \mathrm{e} \cdots \mathrm{e} < \bar\lambda_{k_n} = \lambda_n.$$

Fixing arbitrarily $\mu_i \in [\lambda_i, \lambda_{i-1}],\ k = 1, \ldots, n$, and $\bar{\mu}_j \in [\bar{\lambda}_j, \bar{\lambda}_{j-1}],\ j = 1, \ldots, k_n$, we have

$$S_{\overline{\Pi}} = \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} f(\bar{\mu}_j) \left[\phi(\bar{\lambda}_j) - \phi(\bar{\lambda}_{j-1})\right].$$

Furthermore, since for all $i = 1, \ldots, n$

$$\sum_{j=k_{i-1}+1}^{k_i} \phi(\bar{\lambda}_j) - \phi(\bar{\lambda}_{j-1}) = \phi(\bar{\lambda}_{k_i}) - \phi(\bar{\lambda}_{k_{i-1}}) = \phi(\lambda_i) - \phi(\lambda_{i-1}),$$

we can write S_Π as

$$S_\Pi = \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} f(\mu_i) \left[\phi(\bar{\lambda}_j) - \phi(\bar{\lambda}_{j-1})\right].$$

Since $|\Pi| < \delta_\varepsilon$, we now have that, for all $i = 1, \ldots, n$ and all $j = k_{i-1} + 1, \ldots, k_i$,

$$|\bar{\mu}_j - \mu_i| < \lambda_i - \lambda_{i-1} < \delta_\varepsilon \implies |f(\bar{\mu}_j) - f(\mu_i)| < \frac{\varepsilon}{2V_a^b(\phi)}.$$

Therefore,

$$\begin{aligned} |S_\Pi - S_{\overline{\Pi}}| &\leqslant \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} |f(\bar{\mu}_j) - f(\mu_i)| \left|\phi(\bar{\lambda}_j) - \phi(\bar{\lambda}_{j-1})\right| \\ &\leqslant \frac{\varepsilon}{2V_a^b(\phi)} \sum_{j=1}^{k_n} \left|\phi(\bar{\lambda}_j) - \phi(\bar{\lambda}_{j-1})\right| \leqslant \frac{\varepsilon}{2}. \end{aligned}$$

Similarly, we also have $|S_{\Pi'} - S_{\overline{\Pi}}| \leqslant \varepsilon/2$. Thus, as expected,

$$|S_\Pi - S_{\Pi'}| \leqslant |S_\Pi - S_{\overline{\Pi}}| + |S_{\overline{\Pi}} - S_{\Pi'}| \leqslant \varepsilon.$$

Let us now consider a sequence of partitions $(\Pi_n)_{n=1}^\infty$ satisfying $\lim_{n\to\infty} |\Pi_n| = 0$. There exists $M \geqslant 1$ such that $|\Pi_n| < \delta_\varepsilon$ for all $n \geqslant M$. It then follows from (A.2.1) that $|S_{\Pi_n} - S_{\Pi_m}| < \varepsilon$, for all $n, m \geqslant M$. Hence $(S_{\Pi_n})_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{R}$. So there exists $I \in \mathbb{R}$ and $N \geqslant 1$ such that $|S_{\Pi_N} - I| < \varepsilon/2$. Finally, by (A.2.1), any partition Π smaller than $\delta_{\varepsilon/2}$ satisfies

$$|S_\Pi - I| \leqslant |S_\Pi - S_{\Pi_N}| + |S_{\Pi_N} - I| \leqslant \varepsilon.$$

This concludes the proof of the theorem. □

The limit I in Theorem A.2.1 is called the **Riemann-Stieltjes integral** of f with respect to ϕ. We denote it as

$$\int_a^b f(x)\,\mathrm{d}\phi(x) \quad \text{ore merely as} \quad \int_a^b f\,\mathrm{d}\phi.$$

A.3 Some properties of the Riemann-Stieltjes integral

The following elementary properties are direct consequences of the definition and of Theorem A.2.1.

Properties A.3.1 Consider two continuous functions $f, g : [a,b] \to \mathbb{R}$, two functions of bounded variation $\phi, \psi \in VB[a,b]$, $c \in (a,b)$ and $\alpha \in \mathbb{R}$. Then there holds:

(a) $\int_a^b f \, \mathrm{d}\phi = \int_a^c f \, \mathrm{d}\phi + \int_c^b f \, \mathrm{d}\phi$;

(b) $\int_a^b \alpha f \, \mathrm{d}\phi = \alpha \int_a^b f \, \mathrm{d}\phi$;

(c) $\int_a^b (f+g) \, \mathrm{d}\phi = \int_a^b f \, \mathrm{d}\phi + \int_a^b g \, \mathrm{d}\phi$;

(d) $\int_a^b f \, \mathrm{d}\phi + \int_a^b f \, \mathrm{d}\psi = \int_a^b f \, \mathrm{d}(\phi + \psi)$. ♦

We next state the integral form of the mean-value theorem for Riemann-Stieltjes integrals.

Theorem A.3.2 *Consider $\phi \in VB[a,b]$ and $f : [a,b] \to \mathbb{R}$ continuous. Then*

$$\left| \int_a^b f \, \mathrm{d}\phi \right| \leqslant \sup_{x \in [a,b]} |f(x)| \, V_a^b(\phi).$$

Proof. Let $\Pi = (\lambda_k)_{k=0}^n$ be a partition of $[a,b]$. Then

$$\left| \sum_{k=1}^n f(\lambda_k) \big[\phi(\lambda_k) - \phi(\lambda_{k-1}) \big] \right| \leqslant \sup_{x \in [a,b]} |f(x)| \sum_{k=1}^n |\phi(\lambda_k) - \phi(\lambda_{k-1})| \leqslant \sup_{x \in [a,b]} |f(x)| \, V_a^b(\phi).$$

The result follows by considering a sequence of partitions, the size of which converges to zero. □

Here is now the analogue of the uniform convergence theorem for Riemann integrals.

Theorem A.3.3 *Let $\phi \in VB[a,b]$ and consider a sequence $(f_n)_{n=1}^\infty$ of real continuous functions defined on $[a,b]$, converging uniformly to a function f. Then there holds*

$$\lim_{n \to \infty} \int_a^b f_n \, \mathrm{d}\phi = \int_a^b f \, \mathrm{d}\phi.$$

Proof. Let us first observe that, since f is continuous, its Riemann-Stieltjes integral w.r.t. ϕ exists. Now consider $\varepsilon > 0$ arbitrary. Since $(f_n)_{n=1}^\infty$ converges uniformly to f, there exists $N \geqslant 1$ such that, for all $n \geqslant N$,

$$\sup_{x \in [a,b]} |f_n(x) - f(x)| \leqslant \frac{\varepsilon}{V_a^b(\phi)}.$$

Then take a seqeunce of partitions $(\Pi_m)_{m=1}^{\infty}$ of $[a,b]$ such that $\lim_{m\to\infty} |\Pi_m| = 0$. Denoting by $(\lambda_k^m)_{k=1}^{l_m}$ partition Π_m we have, for all $n \geqslant N$ and all $m \geqslant 1$,

$$\begin{aligned}\Big|\sum_{k=1}^{l_m} f_n(\lambda_k^m)\big[\phi(\lambda_k^m) - \phi(\lambda_{k-1}^m)\big] - \sum_{k=1}^{l_m} f(\lambda_k^m)\big[\phi(\lambda_k^m) - \phi(\lambda_{k-1}^m)\big]\Big| \\ \leqslant \sum_{k=1}^{l_m} |f_n(\lambda_k^m) - f(\lambda_k^m)|\,\big|\phi(\lambda_k^m) - \phi(\lambda_{k-1}^m)\big| \\ \leqslant \sup_{x\in[a,b]} |f_n(x) - f(x)|\, V_a^b(\phi) \leqslant \varepsilon.\end{aligned}$$

For any fixed $n \geqslant N$, we let $m \to \infty$ and we get, thanks to Theorem A.2.1,

$$\left|\int_a^b f_n \,\mathrm{d}\phi - \int_a^b f \,\mathrm{d}\phi\right| \leqslant \varepsilon.$$

□

The following theorem is used in Chapter 2.

Theorem A.3.4 *Let $\phi \in VB[a,b]$ and suppose that ϕ is left-continuous in $[a,b]$. If, for any continuous function $f : [a,b] \to \mathbb{R}$, there holds*

$$\int_a^b f \,\mathrm{d}\phi = 0,$$

then $\phi(x) = \phi(a)$ for all $x \in [a,b]$.

Proof. Fix $x_0 \in (a,b)$ and choose $N \in \mathbb{N}$ so that $x_0 - 1/N > a$. For any $n \geqslant N$, we define a continuous function $f_n : [a,b] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} 1 & \text{if} \quad x \in [a, x_0 - \frac{1}{n}], \\ -nx + nt & \text{if} \quad x \in [x_0 - \frac{1}{n}, x_0], \\ 0 & \text{if} \quad x \in [x_0, b]. \end{cases}$$

Then, for all $n \geqslant N$,

$$\begin{aligned} 0 = \int_a^b f_n \,\mathrm{d}\phi &= \int_a^{x_0 - \frac{1}{n}} 1 \,\mathrm{d}\phi + \int_{x_0 - \frac{1}{n}}^{x_0} (-nx + nx_0) \,\mathrm{d}\phi(x) + \int_{x_0}^b 0 \,\mathrm{d}\phi \\ &= \phi(x_0 - \tfrac{1}{n}) - \phi(a) + \int_{x_0 - \frac{1}{n}}^{x_0} (-nx + nx_0) \,\mathrm{d}\phi(x). \end{aligned} \tag{A.3.1}$$

However by Theorem A.3.2 we have, for all $n \geqslant N$,

$$0 \leqslant \left| \int_{x_0 - \frac{1}{n}}^{x_0} (-nx + nx_0) \, \mathrm{d}\phi(x) \right| \leqslant V_{x_0 - \frac{1}{n}}^{x_0}(\phi).$$

Furthermore, by Theorem A.1.1 we can express the total variation of ϕ in $[x_0 - \frac{1}{n}, x_0]$ as

$$V_{x_0 - \frac{1}{n}}^{x_0}(\phi) = V_a^{x_0}(\phi) - V_a^{x_0 - \frac{1}{n}}(\phi).$$

Since ϕ is left-continuous, Theorem A.1.2 yields

$$\lim_{n \to \infty} V_{x_0 - \frac{1}{n}}^{x_0}(\phi) = V_a^{x_0}(\phi) - \lim_{n \to \infty} V_a^{x_0 - \frac{1}{n}}(\phi) = 0.$$

Thus,

$$\lim_{n \to \infty} \int_{x_0 - \frac{1}{n}}^{x_0} (-nx + nx_0) \, \mathrm{d}\phi(x) = 0.$$

Hence, letting $n \to \infty$ in (A.3.1), the left continuity of ϕ implies that

$$0 = \lim_{n \to \infty} \phi(x_0 - \tfrac{1}{n}) - \phi(a) = \phi(x_0) - \phi(a).$$

Since $x_0 \in (a, b)$ is arbitrary, we have proved that $\phi(x) = \phi(a)$ for all $x \in [a, b)$, and it follows that

$$\phi(b) = \lim_{x \nearrow b} \phi(x) = \phi(a),$$

which finishes the proof. □

Appendix B

The Lebesgue-Stieltjes integral

B.1 Lebesgue-Stieltjes measures

A measure can be associated with any increasing left-continuous function $F : \mathbb{R} \to \mathbb{R}$ in the following way. Write $F(\lambda + 0) = \lim_{\mu \searrow \lambda} F(\mu)$ for all $\lambda \in \mathbb{R}$. The following theorem guarantees the existence of a measure μ_F defined on the Borel sets $\mathcal{B}$ of $\mathbb{R}$ such that

$$\begin{aligned}\mu_F[a, b) &= F(b) - F(a)\\ \mu_F(a, b] &= F(b + 0) - F(a + 0)\\ \mu_F(a, b) &= F(b) - F(a + 0)\end{aligned}$$

for any $a, b \in \mathbb{R}$ with $a < b$, and

$$\mu_F[a, b] = F(b + 0) - F(a),$$

for $a, b \in \mathbb{R}$ with $a \leqslant b$.

Theorem B.1.1 *Consider an increasing left-continuous function $F : \mathbb{R} \to \mathbb{R}$. There exists a unique measure μ_F defined on $\mathcal{B}$ such that $\mu_F[a, b) = F(b) - F(a)$ for any $a, b \in \mathbb{R}$ with $a < b$.*

Proof. See [SS05, p. 282]. □

We call **Lebesgue-Stieltjes measure** associated to F the measure μ_F. This measure can be extended to a complete measure on the σ-algebra $\Sigma_F \supset \mathcal{B}$ of all μ_F-measurable sets.

Remark B.1.2 Any measure μ on $\mathcal{B}$ which is finite on bounded intervals is, in fact, a Lebesgue-Stieltjes measure. Indeed $\mu = \mu_F$, where

$$F(x) := \begin{cases} -\mu[-x,0) & \text{if} \quad x<0, \\ 0 & \text{if} \quad x=0, \\ \mu[0,x) & \text{if} \quad x>0, \end{cases}$$

is increasing and left-continuous. See [SS05, p 282] for details.

In the remainder of this section, F will always denote an increasing left-continuous function.

B.2 The Lebesgue-Stieltjes integral

The Lebesgue-Stieltjes integral with respect to F is merely defined as the Lebesgue integral associated with the measure μ_F. We shall briefly recall here one possible construction in the context of interest to these notes. We essentially follow [KF80], where more details can be found.

The first step is to define the integral of **simple functions**, i.e. functions of the form

$$s = \sum_{k=1}^{\infty} c_k \chi_{A_k}, \tag{B.2.1}$$

where χ_A denotes the characteristic function of A, the sets A_k are measurable and pairwise disjoint, and the numbers $c_k \in \mathbb{R}$, $k \geqslant 1$, are distincts. Then s is called Lebesgue-Stieltjes integrable with respect to F if the series

$$\sum_{k=1}^{\infty} c_k \mu_F(A_k)$$

is absolutely convergent. In this case the Lebesgue-Stieltjes integral of s with respect to F is defined as

$$\int s \,\mathrm{d}\mu_F = \sum_{k=1}^{\infty} c_k \mu_F(A_k).$$

Hence, the definition of the integral for simple functions is quite straightforward. To extend it to general measurable functions, one then benefits from the following theorem. Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is called **measurable** if

$$f^{-1}(A) \text{ is measurable for any measurable set } A.$$

Theorem B.2.1 *A function f is μ_F-measurable if and only if there exists a sequence $(s_n)_{n=1}^{\infty}$ of simple functions converging uniformly to f.*

Definition B.2.2 A μ_F-measurable function f is called **Lebesgue-Stieltjes integrable with respect to** F, or simply μ_F**-integrable**, if there is a sequence $(s_n)_{n=1}^{\infty}$ of μ_F-integrable simple functions which converges uniformly to f. In this case we call **Lebesgue-Stieltjes integral with respect to** F of the function f the limit

$$\int f \, \mathrm{d}\mu_F = \lim_{n\to\infty} \int_A s_n d\mu_F.$$

One shows in particular that the above limit exists, is finite, and does not depend on the choice of the sequence of simple functions $(s_n)_{n=1}^{\infty}$, see [KF80, p. 296] for more details.

For a complex function $f : \mathbb{R} \to \mathbb{C}$, writing $f = g + ih$ with $g = \Re f$ and $h = \Im f$, we say that f is Lebesgue-Stieltjes integrable with respect to F if both g and h are. In this case, the integral of f is simply defined as

$$\int f \, \mathrm{d}\mu_F = \int g \, \mathrm{d}\mu_F + i \int h \, \mathrm{d}\mu_F.$$

The following theorem ensures that the Riemann–Stieltjes integral and the Lebesgue–Stieltjes integral coincide for continuous functions.

Theorem B.2.3 *Let $f : [a, b] \to \mathbb{R}$ be a continuous function and $F : \mathbb{R} \to \mathbb{R}$ be increasing and left-continuous. Then*

$$\int_a^b f \, \mathrm{d}F = \int_{[a,b)} f \, \mathrm{d}\mu_F.$$

Proof. Consider a sequence $(\Pi_n)_{n=1}^{\infty}$ of partitions of $[a, b]$ given as

$$\Pi_n : \quad a = \lambda_0^n < \lambda_1^n < \cdots < \lambda_{m_n-1}^n < \lambda_{m_n}^n = b,$$

and such that $\lim_{n\to\infty} |\Pi_n| = 0$. For all $n \geqslant 1$ we define a simple function $\psi_n : [a, b) \to \mathbb{R}$ by

$$\psi_n(x) = f(\lambda_k^n) \quad \text{if } x \in [\lambda_k^n, \lambda_{k-1}^n), \quad \text{for } k = 0, \ldots, m_n - 1.$$

Each ψ_n is μ_F-integrable over $[a, b)$ and it follows from the uniform continuity of f that the sequence $(\psi_n)_{n=1}^{\infty}$ converges uniformly to f. Furthermore,

$$S_{\Pi_n} = \int_{[a,b)} \psi_n \, \mathrm{d}\mu_F$$

and so, letting $n \to \infty$, Theorem A.2.1 implies that

$$\int_a^b f \, \mathrm{d}F = \int_{[a,b)} f \, \mathrm{d}\mu_F.$$

□

We now state, without proofs, some important convergence results. We say that a sequence of functions $(f_n)_{n=1}^{\infty}$ **converges μ_F-almost everywhere** to a function f, and we simply write $f_n \to f$ a.e. if the context is clear, provided there exists a μ_F-measurable set N such that $\mu_F(N) = 0$ and

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{for all } x \in \mathbb{R}\backslash N.$$

Similarly, any pointwise property holding everywhere except on a set N such that $\mu_F(N) = 0$ will be said to hold **μ_F-almost everywhere**, or merely almost everywhere if the context is clear.

Theorem B.2.4 (Fatou's Lemma) *Let $(f_n)_{n=1}^{\infty}$ be a sequence of μ_F-measurable functions such that $f_n \geqslant 0$ a.e. for all $n \geqslant 1$. If $f_n \to f$ a.e. then*

$$\int f \, \mathrm{d}\mu_F \leqslant \liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu_F.$$

The proof of this result can be found on p. 61 of [SS05].

Theorem B.2.5 (Dominated Convergence) *Let $(f_n)_{n=1}^{\infty}$ be a sequence of μ_F-measurable functions such that $f_n \to f$ a.e. If there exists a μ_F-integrable function g such that $|f_n| \leqslant g$ for all $n \geqslant 1$, then*

$$\lim_{n\to\infty} \int |f_n - f| \, \mathrm{d}\mu_F = 0$$

and so

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}\mu_F = \int f \, \mathrm{d}\mu_F.$$

The proof of this result can be found on p. 67 of [SS05]. An equivalent formulation is given on p. 303 of [KF80].

B.3 The Hilbert space $L^2(\mathbb{R}, \mu_F)$

The space $L^2(\mathbb{R}, \mu_F)$ is defined similarly to the case where μ_F is the Lebesgue measure on $\mathbb{R}$. It consists of equivalence classes of functions μ_F-measurable $f : \mathbb{R} \to \mathbb{C}$ such that $|f|^2$ is μ_F-integrable over $\mathbb{R}$, the equivalence relation being given by

$$f \sim_{\mu_F} g \iff f = g \ \mu_F\text{-a.e.}$$

Endowed with the inner product

$$\langle f, g\rangle_{L^2(\mathbb{R},\mu_F)} = \int f\bar{g} \, \mathrm{d}\mu_F, \quad f, g \in L^2(\mathbb{R}, \mu_F),$$

it is a Hilbert space.

The definition of the integral with respect to a spectral family in Chapter 3 is based on the following result. A proof can be found in [Wei80, p. 25]. It can also be deduced from the proof of the separability of $L^2(\mathbb{R})$ in [SS05, p. 160].

We call **step function** a function that can be written as

$$t = \sum_{k=0}^{n} c_k \chi_{I_k}$$

where the sets I_k are intervals of the form

$$(a_k, b_k), \quad [a_k, b_k), \quad (a_k, b_k], \quad [a_k, b_k],$$

with $a_k, b_k \in \mathbb{R}$, $a_k \leqslant b_k$, and $c_k \in \mathbb{C}$.

Theorem B.3.1 *Step functions are dense in* $L^2(\mathbb{R}, \mu_F)$.

Appendix C

The Banach space-valued Riemann integral

Throughout this appendix we consider a Banach space $\mathcal{X}$ and a continuous function $F : [a, b] \to \mathcal{X}$ where $a, b \in \mathbb{R}$ with $a < b$.

C.1 Definition and existence

The function F being continuous, it is uniformly continuous over the compact interval $[a, b]$. That is, for all $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $t, s \in [a, b]$,

$$|t - s| < \delta \implies \|F(t) - F(s)\| < \varepsilon.$$

On the other hand, $F : [a, b] \to \mathcal{X}$ is bounded, that is,

$$\sup_{t\in[a,b]} \|F(t)\| < \infty.$$

In fact, this supremum is achieved at some point in $[a, b]$.

In perfect analogy with the real-valued Riemann integral, for any partition Π of $[a, b]$,

$$\Pi : \quad a = \lambda_0 < \lambda_1 < \cdots < \lambda_{m-1} < \lambda_m = b,$$

and any sequence of points

$$\mu_k \in [\lambda_{k-1}, \lambda_k], \quad k = 1, \ldots, m,$$

we consider the sum

$$S_\Pi = \sum_{k=1}^{m} (\lambda_k - \lambda_{k-1}) F(\mu_k).$$

Theorem C.1.1 *Let $F : [a, b] \to \mathcal{X}$ be continuous. There exists a unique $Y \in \mathcal{X}$ such that, for all $\varepsilon > 0$, there exists $\delta > 0$ such that, for any partition Π of $[a, b]$ satisfying $|\Pi| < \delta$, we have*

$$\|S_\Pi - Y\| < \varepsilon.$$

The element $Y \in \mathcal{X}$ given by Theorem C.1.1 is called the **Riemann integral** of F over $[a, b]$, and we write

$$Y = \int_a^b F(t)\,\mathrm{d}t.$$

The proof of this result is virtually identical to those of Theorem A.2.1 and Lemma 2.1.2. We therefore leave as an easy revision exercise the obvious changes of notation required.

The integral defined by Theorem C.1.1 is of course linear. That is, if $F, G : [a, b] \to \mathcal{X}$ are continuous and $\alpha, \beta \in \mathbb{R}$, then

$$\int_a^b \big(\alpha F(t) + \beta G(t)\big)\,\mathrm{d}t = \alpha \int_a^b F(t)\,\mathrm{d}t + \int_a^b \beta G(t)\,\mathrm{d}t.$$

Furthermore, it is easily seen that

$$\left\|\int_a^b F(t)\,\mathrm{d}t\right\| \leqslant \int \|F(t)\|\,\mathrm{d}t.$$

C.2 Some useful results

Lemma C.2.1 *Let $T \in \mathcal{B}(\mathcal{X})$ and $F : [a, b] \to \mathcal{X}$ continuous. The composition $T \circ F$ is continuous and we have*

$$T \int_a^b F(t)\,\mathrm{d}t = \int_a^b TF(t)\,\mathrm{d}t.$$

Proof. The continuity of the composition follows from

$$\|TF(t) - TF(s)\| = \|T(F(t) - F(s)\| \leqslant \|T\|\,\|F(t) - F(s)\|\,.$$

Consider now a sequence $(\Pi_n)_{n=1}^\infty$ of partitions of $[a, b]$, given as

$$\Pi_n : \quad a = \lambda_0^n < \lambda_1^n < \cdots < \lambda_{m_n-1}^n < \lambda_{m_n}^n = b,$$

and such that $\lim_{n\to\infty} |\Pi_n| = 0$. We then have

$$TS_{\Pi_n} = T\sum_{k=1}^{m_n}(\lambda_k^n - \lambda_{k-1}^n)F(\lambda_k^n) = \sum_{k=1}^{m_n}(\lambda_k^n - \lambda_{k-1}^n)TF(\lambda_k^n),$$

and so by continuity of T and Theorem C.1.1, the result follows by letting $n \to \infty$. □

Theorem C.2.2 *Let $\phi : [a,b] \to \mathbb{R}$ and $F : [a,b] \to \mathcal{X}$ be continuous functions. The product ϕF defined by $\phi F(t) = \phi(t)F(t)$ is continuous and we have*

$$\left\| \int \phi(t)F(t)\,\mathrm{d}t \right\| \leq \int |\phi(t)| \,\mathrm{d}t \sup_{s\in[a,b]} \|F(s)\| .$$

Proof. The continuity of ϕF follows from

$$\begin{aligned} \|\phi(t)F(t) - \phi(s)F(s)\| &\leqslant \|\phi(t)F(t) - \phi(s)F(t)\| + \|\phi(s)F(t) - \phi(s)F(s)\| \\ &\leqslant |\phi(t) - \phi(s)| \, \|F(t)\| + |\phi(s)| \, \|F(t) - F(s)\| . \end{aligned}$$

Consider now a sequence $(\Pi_n)_{n=1}^{\infty}$ of partitions of $[a,b]$, given as

$$\Pi_n : \quad a = \lambda_0^n < \lambda_1^n < \cdots < \lambda_{m_n-1}^n < \lambda_{m_n}^n = b,$$

and such that $\lim_{n\to\infty} |\Pi_n| = 0$. We then have

$$\begin{aligned} \left\| S_{\Pi_n} \right\| = \left\| \sum_{k=1}^{m_n} (\lambda_k^n - \lambda_{k-1}^n)\phi(\lambda_k^n)F(\lambda_k^n) \right\| &\leqslant \sum_{k=1}^{m_n} (\lambda_k^n - \lambda_{k-1}^n) \, |\phi(\lambda_k^n)| \, \|F(\lambda_k^n)\| \\ &\leqslant \sum_{k=1}^{m_n} (\lambda_k^n - \lambda_{k-1}^n) \, |\phi(\lambda_k^n)| \sup_{s\in[a,b]} \|F(s)\| , \end{aligned}$$

and so letting $n \to \infty$ the result follows by continuity of the norm, by Theorem C.1.1, and the fact that

$$\lim_{n\to\infty} \sum_{k=1}^{m_n} (\lambda_k^n - \lambda_{k-1}^n) \, |\phi(\lambda_k^n)| = \int_a^b |\phi(t)| \,\mathrm{d}t.$$

□

Finally, we have the following uniform convergence theorem.

Theorem C.2.3 *Consider a sequence $(F_n)_{n=1}^{\infty}$ of continuous functions $F_n : [a,b] \to \mathcal{X}$ and a continuous function $F : [a,b] \to \mathcal{X}$. Suppose $(F_n)_{n=1}^{\infty}$ converges uniformly to F. That is, for all $\varepsilon > 0$ there exists $N \geqslant 1$ such that*

$$n \geqslant N \implies \sup_{s\in[a,b]} \|F_n(s) - F(s)\| < \varepsilon.$$

Then we have

$$\lim_{n\to\infty} \int_a^b F_n(t)\,\mathrm{d}t = \int_a^b F(t)\,\mathrm{d}t.$$

Proof. Applying Theorem C.2.2 to $\phi(s) \equiv 1$, we have, for all $n \geqslant 1$,

$$\left\| \int_a^b F_n(t)\,\mathrm{d}t - \int_a^b F(t)\,\mathrm{d}t \right\| \leqslant (b-a) \sup_{t \in [a,b]} \|F_n(s) - F(s)\| \,.$$

The result then follows from the uniform convergence of $(F_n)_{n=1}^{\infty}$ to F. □

Bibliography

[Dir58] Dirac, P. A. M. *The Principles of Quantum Mechanics.* Clarendon Press Oxford, 1958.

[Fri82] Friedman, A. *Foundations of Modern Analysis.* Dover Publications, 1982.

[KF80] Kolmogorov, A. and Fomin, S. *Introductory Real Analysis.* Dover Publications, 1980.

[Kre78] Kreyszig, E. *Introductory Functional Analysis.* John Wiley & Sons. Inc., 1978.

[Mes99] Messiah, A. *Quantum Mechanics.* Dover Publications, 1999.

[RS] Reed, M. and Simon, B. *Methods of Modern Mathematical Physics.* Volumes I-IV.

[RSN90] Riesz, F. and Sz.-Nagy, B. *Functional Analysis.* Dover Publications, 1990.

[SS05] Stein, E. M. and Shakarchi, R. *Real Analysis, Volume III. Measure Theory, Integration, and Hilbert Spaces.* Princeton University Press, 2005.

[Sto32] Stone, M. H. *Linear Transformations in Hilbert Space.* American Mathematical Society, 1932.

[VN55] Von Neumann, J. *Mathematical Foundations of Quantum Mechanics.* Princeton University Press, 1955.

[Wei80] Weidmann, J. *Linear Operators in Hilbert Spaces.* Springer Verlag, 1980.