Linear Functional Analysis (Chapters 8 and 9)

Joan Cerdà

Universitat de Barcelona

E-mail address: `jcerda@ub.edu`

2000 *Mathematics Subject Classification.* 46-01

Chapter 8

Banach algebras

Some important Banach spaces are equipped in a natural way with a continuous product that determines a Banach algebra structure.[1] Two basic examples are $\mathcal{C}(K)$ with the pointwise multiplication and $\mathcal{L}(E)$ with the product of operators if E is a Banach space. It can be useful for the reader to retain $\mathcal{C}(K)$ as a simple reference model.

The first work devoted to concrete Banach algebras is contained in some papers by J. von Neumann and beginning in 1930. The advantage of considering algebras of operators was clear in his contributions, but it was the abstract setting of Banach algebras which proved to be convenient and which allowed the application of similar ideas in many directions.

The main operator on these algebras is the Gelfand transform[2] $\mathcal{G} : a \mapsto \widehat{a}$, which maps a unitary commutative Banach algebra A on $\mathbf{C}$ to the space $\mathcal{C}(\Delta)$ of all complex continuous functions on the spectrum Δ of A, which is the set of all nonzero elements $\chi \in A'$ that are multiplicative. Here Δ is endowed with the restriction of the w^*-topology and it is compact. As seen in Example 8.14, Δ is the set of all the evaluations δ_t $(t \in K)$ if $A = \mathcal{C}(K)$, and $\widehat{f}(\delta_t) = \delta_t(f) = f(t)$, so that in this case one can consider $\widehat{f} = f$.

But we will be concerned with the spectral theory of operators in a complex Hilbert space H. If T is a bounded normal operator in H, so that

[1]Banach algebras were first introduced in 1936 with the name of "linear metric rings" by the Japanese mathematician Mitio Nagumo. He extended Cauchy's function theory to the functions with values in such an algebra to study the resolvent of an operator. They were renamed "Banach algebras" by Charles E. Rickart in 1946.

[2]Named after the Ukrainian mathematician Israel Moiseevich Gelfand, who is considered the creator, in 1941, of the theory of commutative Banach algebras. Gelfand and his colleagues created this theory which included the spectral theory of operators and proved to be an appropriate setting for harmonic analysis.

T and the adjoint T^* commute, then the closed Banach subalgebra $A = \langle T \rangle$ of $\mathcal{L}(H)$ generated by I, T, and T^* is commutative.

It turns out that the Gelfand theory of commutative Banach algebras is especially well suited in this setting. Through the change of variables $z = \widehat{T}(\chi)$ one can consider $\sigma(T) \equiv \Delta$, and the Gelfand transform is a bijective mapping that allows us to define a functional calculus $g(T)$ by $\widehat{g(T)} = g(\widehat{T})$ if g is a continuous function on the spectrum of T.

For this continuous functional calculus there is a unique operator-valued measure E on $\sigma(T)$ such that

$$g(T) = \int_{\sigma(T)} g(\lambda)\, dE(\lambda),$$

and the functional calculus is extended by

$$f(T) = \int_{\sigma(T)} f(\lambda)\, dE(\lambda)$$

to bounded measurable functions f.

The Gelfand transform, as a kind of abstract Fourier operator, is also a useful tool in harmonic analysis and in function theory. The proof of Wiener's 1932 lemma contained in Exercise 8.15 is a nice unexpected application discovered by Gelfand in 1941, and generalizations of many theorems of Tauberian type and applications to the theory of locally compact groups have also been obtained with Gelfand's methods. We refer the reader to the book by I. M. Gelfand, D. A. Raikov and G. E. Chilov [**?**] for more information.

8.1. Definition and examples

We say that A is a complex Banach algebra or, simply, a **Banach algebra** if it is a complex Banach space with a bilinear multiplication and the norm satisfies

$$\|xy\| \leq \|x\|\|y\|,$$

so that the multiplication is continuous since, if $(x_n, y_n) \to (x, y)$, then

$$\|xy - x_n y_n\| \leq \|x\|\|y - y_n\| + \|x - x_n\|\|y_n\| \to 0.$$

Real Banach algebras are defined similarly.

The Banach algebra A is said to be **unitary** if it has a unit, which is an element e such that $xe = ex = x$ for all $x \in A$ and $\|e\| = 1$. This unit is unique since, if also $e'x = xe' = x$, then $e = ee' = e'$.

We will only consider unitary Banach algebras. As a matter of fact, every Banach algebra can be embedded in a unitary Banach algebra, as shown in Exercise 8.1.

Example 8.1. (a) If X is a nonempty set, $\mathbf{B}(X)$ will denote the unitary Banach algebra of all complex bounded functions on X, with the pointwise multiplication and the uniform norm $\|f\|_X := \sup_{x\in X} |f(x)|$. The unit is the constant function 1.

(b) If K is a compact topological space, then $\mathcal{C}(K)$ is the closed subalgebra of $\mathbf{B}(K)$ that contains all the continuous complex functions on K. It is a unitary Banach subalgebra of $\mathbf{B}(K)$, since $1 \in \mathcal{C}(K)$.

(c) The **disc algebra** is the unitary Banach subalgebra $A(D)$ of $\mathcal{C}(\bar{D})$. Since the uniform limits of analytic functions are also analytic, $A(D)$ is closed in $\mathcal{C}(\bar{D})$.

Example 8.2. If Ω is a σ-finite measure space, $L^\infty(\Omega)$ denotes the unitary Banach algebra of all measurable complex functions on Ω with the usual norm $\|\cdot\|_\infty$ of the essential supremum. As usual, two functions are considered equivalent when they are equal a.e.

Example 8.3. Let E be any nonzero complex Banach space. The Banach space $\mathcal{L}(E) = \mathcal{L}(E; E)$ of all bounded linear operators on E, endowed with the usual product of operators, is a unitary Banach algebra. The unit is the identity map I.

8.2. Spectrum

Throughout this section, A **denotes a unitary Banach algebra**, possibly not commutative. An example is $\mathcal{L}(E)$, if E is a complex Banach space.

A **homomorphism** between A and a second unitary Banach algebra B is a homomorphism of algebras $\Psi : A \to B$ such that $\Psi(e) = e$ if e denotes the unit both in A and in B.

The notion of the spectrum of an operator is extended to any element of A:

The **spectrum** of $a \in A$ is the subset of $\mathbf{C}$

$$\sigma_A(a) = \sigma(a) := \{\lambda \in \mathbf{C};\ \lambda e - a \notin G(A)\},$$

where $G(A)$ denotes the multiplicative group of all invertible elements of A.

Note that, if B is a unitary Banach subalgebra of A and $b \in B$, an inverse of $\lambda e - b$ in B is also an inverse in A, so that $\sigma_A(b) \subset \sigma_B(b)$.

Example 8.4. If E is a complex Banach space and $T \in \mathcal{L}(E)$, we denote $\sigma(T) = \sigma_{\mathcal{L}(E)}(T)$. Thus, $\lambda \in \sigma(T)$ if and only if $T - \lambda I$ is not bijective, by the Banach-Schauder theorem. Recall that the eigenvalues of T, and also the approximate eigenvalues, are in $\sigma(T)$. Cf. Subsection **??**.

Example 8.5. If E is an infinite-dimensional Banach space and $T \in \mathcal{L}(E)$ is compact, the Riesz-Fredholm theory shows that $\sigma(T) \setminus \{0\}$ can be arranged in a sequence of nonzero eigenvalues (possibly finite), all of them with finite multiplicity, and $0 \in \sigma(T)$, by the Banach-Schauder theorem.

Example 8.6. The spectrum of an element f of the Banach algebra $\mathcal{C}(K)$ is its image $f(K)$.

Indeed, the continuous function $f - \lambda$ has an inverse if it has no zeros, that is, if $f(t) \neq \lambda$ for all $t \in K$. Hence, $\lambda \in \sigma(f)$ if and only if $\lambda \in f(K)$.

Let us consider again a general unitary Banach algebra A.

Theorem 8.7. *If $p(\lambda) = \sum_{n=0}^{N} c_n \lambda^n$ is a polynomial and $a \in A$, then $\sigma(p(a)) = p(\sigma(a))$.*

Proof. We assume that $p(a) = c_0 e + c_1 a + \cdots + c_N a^N$, and we exclude the trivial case of a constant polynomial $p(\lambda) \equiv c_0$.

For a given $\mu \in \mathbf{C}$, by division we obtain $p(\mu) - p(\lambda) = (\mu - \lambda) q(\lambda)$ and $p(\mu)e - p(a) = (\mu e - a) q(a)$. If $\mu e - a \notin G(A)$, then also $p(\mu)e - p(a) \notin G(A)$. Hence, $p(\sigma(a)) \subset \sigma(p(a))$.

Conversely, if $\mu \in \sigma(p(a))$, by factorization we can write

$$\mu - p(\lambda) = \alpha(\lambda_1 - \lambda) \cdot \ldots \cdot (\lambda_N - \lambda)$$

with $\alpha \neq 0$. Then $\mu e - p(a) = \alpha(\lambda_1 e - a) \cdot \ldots \cdot (\lambda_N e - a)$, where $\mu e - p(a) \notin G(A)$, so that $\lambda_i e - a \notin G(A)$ for some $1 \leq i \leq N$. Thus, $\lambda_i \in \sigma(a)$ and we have $p(\lambda_i) = \mu$, which means that $\mu \in p(\sigma(a))$. $\square$

The **resolvent** of an element $a \in A$ is the function $R_a : \sigma(a)^c \to A$ such that $R_a(\lambda) = (\lambda e - a)^{-1}$. It plays an important role in spectral theory.

Note that, if $\lambda \neq 0$,

$$R_a(\lambda) = -(a - \lambda e)^{-1} = \lambda^{-1}(e - \lambda^{-1} a)^{-1}.$$

To study the basic properties of R_a, we will use some facts from function theory.

As in the numerical case and with the same proofs, a vector-valued function $F : \Omega \to A$ on an open subset Ω of $\mathbf{C}$ is said to be analytic or holomorphic if every point $z_0 \in \Omega$ has a neighborhood where F is the sum of a convergent power series:

$$F(z) = \sum_{n=0}^{\infty} (z - z_0)^n a_n \qquad (a_n \in A).$$

The series is absolutely convergent at every point of the convergence disc, which is the open disc in $\mathbf{C}$ with center z_0 and radius

$$R = \frac{1}{\limsup_n \|a_n\|^{1/n}} > 0.$$

The Cauchy theory remains true without any change in this setting, and F is analytic if and only if, for every $z \in \Omega$, the complex derivative

$$F'(z) = \lim_{h \to 0} \frac{F(z+h) - F(z)}{h}$$

exists.

We will show that $\sigma(a)$ is closed and bounded and, to prove that R_a is analytic on $\sigma(a)^c$, we will see that $\mathbf{R}'_a(\lambda)$ exists whenever $\lambda \notin \sigma(a)$.

Let us first show that R_a is analytic on $|\lambda| > \|a\|$.

Theorem 8.8. (a) *If* $\|a\| < 1$, $e - a \in G(A)$ *and*

$$(e-a)^{-1} = \sum_{n=0}^{\infty} a^n \qquad (a^0 := e).$$

(b) *If* $|\lambda| > \|a\|$, *then* $\lambda \notin \sigma(a)$ *and*

$$R_a(\lambda) = \sum_{n=0}^{\infty} \lambda^{-n-1} a^n.$$

(c) *Moreover,*

$$\|R_a(\lambda)\| \leq \frac{1}{|\lambda| - \|a\|},$$

and $\lim_{|\lambda| \to \infty} R_a(\lambda) = 0$.

Proof. (a) As in (**??**), the Neumann series $\sum_{n=0}^{\infty} a^n$ is absolutely convergent ($\|a^m\| \leq \|a\|^m$ and $\|a\| < 1$), so that $z = \sum_{n=0}^{\infty} a^n \in A$ exists, and it is easy to check that z is the right and left inverse of $e - a$. For instance,

$$\lim_{N \to \infty} (e-a) \sum_{n=0}^{N} a^n = (e-a)z,$$

since the multiplication by $e - a$ is linear and continuous, so that

$$(e-a) \sum_{n=0}^{N} a^n = \sum_{n=0}^{N} a^n - \sum_{n=1}^{N+1} a^n = e - a^{N+1} \to e \quad \text{if } N \to \infty.$$

(b) Note that

$$R_a(\lambda) = \lambda^{-1}(e - \lambda^{-1}a)^{-1}$$

and, if $\|\lambda^{-1}a\| < 1$, we obtain the announced expansion from (a).

(c) Finally,

$$\|R_a(\lambda)\| = |\lambda|^{-1}\|\sum_{n=0}^{\infty}\lambda^{-n}a^n\| \le \frac{1}{|\lambda| - \|a\|}.$$

□

The **spectral radius** of $a \in A$ is the number

$$r(a) := \sup\{|\lambda|;\ \lambda \in \sigma(a)\}.$$

From Theorem 8.8 we have that $r(a) \le \|a\|$, an inequality that can be strict.

The following estimates are useful.

Lemma 8.9. (a) *If* $\|a\| < 1$,

$$\|(e-a)^{-1} - e - a\| \le \frac{\|a\|^2}{1 - \|a\|}.$$

(b) *If* $x \in G(A)$ *and* $\|h\| < 1/(2\|x^{-1}\|)$, *then* $x + h \in G(A)$ *and*

$$\|(x+h)^{-1} - x^{-1} + x^{-1}hx^{-1}\| \le 2\|x^{-1}\|^3\|h\|^2.$$

Proof. To check (a), we only need to sum the right-hand side series in

$$\|(e-a)^{-1} - e - a\| = \|\sum_{n=2}^{\infty} a^n\| \le \sum_{n=2}^{\infty}\|a\|^n.$$

To prove (b) note that $x + h = x(e + x^{-1}h)$, and we have

$$\|x^{-1}h\| \le \|x^{-1}\|\|h\| < 1/2.$$

If we apply (a) to $a = -x^{-1}h$, since $\|a\| < 1/2$, we obtain that $x+h \in G(A)$, and

$$\|(x+h)^{-1} - x^{-1} + x^{-1}hx^{-1}\| \le \|(e-a)^{-1} - e - a\|\|x^{-1}\|$$

with $\|(e-a)^{-1} - e - a\| \le \|x^{-1}h\|^2/(1 - \|a\|) \le 2\|x^{-1}h\|^2$. □

Theorem 8.10. (a) $G(A)$ *is an open subset of* A *and* $x \in G(A) \mapsto x^{-1} \in G(A)$ *is continuous.*

(b) R_a *is analytic on* $\sigma(a)^c$ *and zero at infinity.*

(c) $\sigma(a)$ *is a nonempty compact subset of* $\mathbf{C}$ *and*[3]

$$r(a) = \lim_{n\to\infty}\|a^n\|^{1/n} = \inf_n \|a^n\|^{1/n}.$$

[3]This spectral radius formula and the analysis of the resolvent have a precedent in the study by Angus E. Taylor (1938) of operators which depend analytically on a parameter. This formula was included in the 1941 paper by I. Gelfand on general Banach algebras.

Proof. (a) According to Lemma 8.9(b), for every $x \in G(A)$,

$$B\big(x, \frac{1}{2\|x^{-1}\|}\big) \subset G(A)$$

and $G(A)$ is an open subset of A.

Moreover

$$\|(x+h)^{-1} - x^{-1}\| \le \|(x+h)^{-1} - x^{-1} + x^{-1}hx^{-1}\| + \|x^{-1}hx^{-1}\| \to 0$$

if $\|h\| \to 0$, and $x \in G(A) \mapsto x^{-1} \in G(A)$ is continuous.

(b) On $\sigma(a)^c$,

$$R'_a(\lambda) = \lim_{\mu\to 0} \mu^{-1}[((\lambda+\mu)e - a)^{-1} - (\lambda e - a)^{-1}] = -R_a(\lambda)^2$$

follows from an application of Lemma 8.9(b) to $x = \lambda e - a$ and $h = \mu e$. In this case $x^{-1}hx^{-1} = \mu x^{-1}x^{-1}$ and, writing $x^{-2} = x^{-1}x^{-1}$, we obtain

$$\mu^{-1}[(x+\mu e)^{-1} - x^{-1}] = \mu^{-1}[(x+\mu e)^{-1} - x^{-1} + x^{-1}hx^{-1}] - x^{-2} \to x^{-2}$$

as $\mu \to 0$, since

$$\|\mu^{-1}[(x+\mu e)^{-1} - x^{-1} + x^{-1}hx^{-1}]\| \le |\mu|^{-1}2\|x^{-1}\|^3|\mu|^2 \to 0.$$

By Theorem 8.8(b), $\|R_a(\lambda)\| \le 1/(|\lambda| - \|a\|) \to 0$ if $|\lambda| \to \infty$.

(c) Recall that $\sigma(a) \subset \{\lambda;\, |\lambda| \le r(a)\}$ and $r(a) \le \|a\|$. This set is closed in $\mathbf{C}$, since $\sigma(a)^c = F^{-1}(G(A))$ with $F(\lambda) := \lambda e - x$, which is a continuous function from $\mathbf{C}$ to A, and $G(A)$ is an open subset of A. Hence $\sigma(a)$ is a compact subset of $\mathbf{C}$.

If we suppose that $\sigma(a) = \emptyset$, we will arrive at a contradiction. The function R_a would be entire and bounded, with $\lim_{|\lambda|\to\infty} R_a(\lambda) = 0$, and the Liouville theorem is also true in the vector-valued case: for every $u \in A'$, $u \circ R_a$ would be an entire complex function and $\lim_{|\lambda|\to\infty} u(R_a(\lambda)) = 0$, so that $u(R_a(\lambda)) = 0$ and by the Hahn-Banach theorem $R_a(\lambda) = 0$, a contradiction to $R_a(\lambda) \in G(A)$.

Let us calculate the spectral radius. Since

$$R_a(\lambda) = \lambda^{-1} \sum_{n=0}^{\infty} \lambda^{-n} a^n$$

if $|\lambda| > r(a)$, the power series $\sum_{n=0}^{\infty} z^n a^n$ is absolutely convergent when $|z| = |\lambda|^{-1} < 1/r(a)$, and the convergence radius of $\sum_{n=0}^{\infty} \|a^n\||z|^n$ is

$$R = (\limsup_{n\to\infty} \|a^n\|^{1/n})^{-1} \ge 1/r(a).$$

Then, $r(a) \ge \limsup_{n\to\infty} \|a^n\|^{1/n}$.

Conversely, if $\lambda \in \sigma(a)$, then $\lambda^n \in \sigma(a^n)$ by Theorem 8.7, so that $|\lambda^n| \leq \|a^n\|$ and

$$|\lambda| \leq \inf_n \|a^n\|^{1/n} \leq \liminf \|a^n\|^{1/n}.$$

Then it follows that $r(a) = \lim_{n\to\infty} \|a^n\|^{1/n} = \inf_n \|a^n\|^{1/n}$. □

As an important application of these results, let us show that **C** is the unique Banach algebra which is a field, in the sense that if A is a field, then $\lambda \mapsto \lambda e$ is an isometric isomorphism from **C** onto A. The inverse isometry is the **canonical isomorphism**:

Theorem 8.11 (Gelfand-Mazur[4]). *If every nonzero element of the unitary Banach algebra A is invertible (i.e., $G(A) = A \setminus \{0\}$), then $A = \mathbf{C}e$, and $\lambda \mapsto \lambda e$ is the unique homomorphism of unitary algebras between* **C** *and A.*

Proof. Let $a \in A$ and $\lambda \in \sigma(a)$ ($\sigma(a) \neq \emptyset$). Then $\lambda e - a \notin G(A)$ and it follows from the hypothesis that $a = \lambda e$. A homomorphism $\mathbf{C} \to A = \mathbf{C}e$ maps $1 \to e$ and necessarily $\lambda \to \lambda e$. □

8.3. Commutative Banach algebras

In this section A **represents a commutative unitary Banach algebra**. Some examples are $\mathbf{C}$, $\mathbf{B}(X)$, $\mathcal{C}(K)$, and $L^\infty(\Omega)$. Recall that $\mathcal{L}(E)$ (if $\dim E > 1$) is not commutative, and the convolution algebra $L^1(\mathbf{R})$ is not unitary.

8.3.1. Maximal ideals, characters, and the Gelfand transform. A **character** of A is a homomorphism $\chi : A \to \mathbf{C}$ of unitary Banach algebras (hence $\chi(e) = 1$). We use $\Delta(A)$, or simply Δ, to denote the set of all characters of A. It is called **the spectrum of** A.

An **ideal**, J, of A is a linear subspace such that $AJ \subset J$ and $J \neq A$. It cannot contain invertible elements, since $x \in J$ invertible would imply $e = xx^{-1} \in J$ and then $A = Ae \subset J$, a contradiction to $J \neq A$.

Note that, if J is an ideal, then $\bar{J}$ is also an ideal, since it follows from $J \cap G(A) = \emptyset$ that $e \notin \bar{J}$ and $\bar{J} \neq A$. The continuity of the operations implies that $\bar{J} + \bar{J} \subset \bar{J}$ and $A\bar{J} \subset \bar{J}$.

This shows that every maximal ideal is closed.

[4] According to a result announced in 1938 by Stanislaw Mazur, a close collaborator of Banach who made important contributions to geometrical methods in linear and nonlinear functional analysis, and proved by Gelfand in 1941.

Theorem 8.12. (a) *The kernel of every character is a maximal ideal and the map* $\chi \mapsto \operatorname{Ker}\chi$ *between characters and maximal ideals of* A *is bijective.*

(b) *Every character* $\chi \in \Delta(A)$ *is continuous and*

$$\|\chi\| = \sup_{\|a\|_A \leq 1} |\chi(a)| = 1.$$

(c) *An element* $a \in A$ *is invertible if and only if* $\chi(a) \neq 0$ *for every* $\chi \in \Delta$.

(d) $\sigma(a) = \{\chi(a);\, \chi \in \Delta(A)\}$, *and* $r(a) = \sup_{\chi\in\Delta} |\chi(a)|$.

Proof. (a) The kernel M of any $\chi \in \Delta(A)$ is an ideal and, as the kernel of a nonzero linear functional, it is a hyperplane; that is, the complementary subspaces of M in A are one-dimensional, since χ is bijective on them, and M is maximal.

If M is a maximal ideal, the quotient space A/M has a natural structure of unitary Banach algebra, and it is a field. Indeed, if $\pi : A \to A/M$ is the canonical mapping and $\pi(x) = \widetilde{x}$ is not invertible in A/M, then $J = \pi(xA) \neq A/M$ is an ideal of A/M, and $\pi^{-1}(J) \neq A$ is an ideal of A which is contained in a maximal ideal that contains M. Thus, $\pi^{-1}(J) = M$, so that $\pi(xA) \subset \pi(M) = \{0\}$ and $\widetilde{x} = 0$.

Let $\widetilde{\chi} : A/M = \mathbf{C}\tilde{e} \to \mathbf{C}$ be the canonical isometry, so that M is the kernel of the character $\chi_M := \widetilde{\chi} \circ \pi_M$. Any other character χ_1 with the same kernel M factorizes as a product of π_M with a bijective homomorphism between A/M and $\mathbf{C}$ which has to be the canonical mapping $\mathbf{C}\tilde{e} \to \mathbf{C}$, and then $\chi_1 = \chi_M$.

(b) If $\chi = \chi_M \in \Delta(A)$, then $\|\chi\| \leq \|\pi_M\|\|\widetilde{\chi}\| = \|\pi_M\| \leq 1$ and $\|\chi\| \geq \chi(e) = 1$.

(c) If $x \in G(A)$, we have seen that it does not belong to any ideal. If $x \notin G(A)$, then xA does not contain e and is an ideal, and by Zorn's lemma every ideal is contained in a maximal ideal. So $x \in G(A)^c$ if and only if x belongs to a maximal ideal or, equivalently, $\chi(x) \neq 0$ for every character χ.

(d) Finally, $\lambda e - a \notin G(A)$ if and only if $\chi(\lambda e - a) = 0$, that is , $\lambda = \chi(a)$ for some $\chi \in \Delta(A)$. □

We associate to every element a of the unitary commutative algebra A the function $\widehat{a}$ which is the restriction of $\langle a, \cdot\rangle$ to the characters,[5] so that

$$\widehat{a} : \Delta(A) \to \mathbf{C}$$

is such that $\widehat{a}(\chi) = \chi(a)$. On $\Delta(A) \subset \bar{B}_{A'}$ we consider the **Gelfand topology**, which is the restriction of the weak-star topology $w^* = \sigma(A', A)$ of A'.

[5] Recall that $\langle a, u\rangle = u(a)$ was defined for every u in the dual A' of A as a Banach space.

In this way, $\widehat{a} \in \mathcal{C}(\Delta(A))$, and

$$\mathcal{G} : a \in A \mapsto \widehat{a} \in \mathcal{C}(\Delta(A))$$

is called the **Gelfand transform**.

Theorem 8.13. *Endowed with the Gelfand topology, $\Delta(A)$ is compact and the Gelfand transform $\mathcal{G} : A \mapsto \mathcal{C}(\Delta(A))$ is a continuous homomorphism of commutative unitary Banach algebras.*

Moreover $\|\widehat{a}\| = r(a) \leq \|a\|$ and $\mathcal{G}e = 1$, so that $\|\mathcal{G}\| = 1$.

Proof. For the first part we only need to show that $\Delta \subset \bar{B}_{A'}$ is weakly closed, since $\bar{B}_{A'}$ is weakly compact, by the Alaoglu theorem. But

$$\Delta = \Big\{\xi \in \bar{B}_{A'};\, \xi(e) = 1,\, \xi(xy) = \xi(x)\xi(y)\, \forall x, y \in A\Big\}$$

is the intersection of the weakly closed sets of $\bar{B}_{A'}$ defined by the conditions $\langle xy, \cdot\rangle - \langle x, \cdot\rangle\langle y, \cdot\rangle = 0$ $(x, y \in A)$ and $\langle e, \cdot\rangle = 1$.

It is clear that it is a homomorphism of commutative unitary Banach algebras. For instance, $\widehat{e}(\chi) = \chi(e) = 1$ and $\widehat{xy}(\chi) = \chi(x)\chi(y) = \widehat{x}(\chi)\widehat{y}(\chi)$.

Also, $\|\widehat{a}\| = \sup_{\chi\in\Delta} |\chi(a)| \leq \|a\|$, according to Theorem 8.12(d). □

Example 8.14. If K is a compact topological space, then $\mathcal{C}(K)$ is a unitary commutative Banach algebra whose characters are the evaluation maps δ_t at the different points $t \in K$, and $t \in K \mapsto \delta_t \in \Delta$ is a homeomorphism.

Obviously $\delta_t \in \Delta$. Conversely, if $\chi = \chi_M \in \Delta$, we will show that there is a common zero for all $f \in M$. If not, for every $t \in K$ there would exist some $f_t \in M$ such that $f_t(t) \neq 0$, and $|f_t| \geq \varepsilon_t > 0$ on a neighborhood $U(t)$ of this point t. Then, $K = U(t_1) \cup \cdots \cup U(t_N)$, and the function

$$f = |f_{t_1}|^2 + \cdots + |f_{t_N}|^2 = f_{t_1}\overline{f_{t_1}} + \cdots + f_{t_N}\overline{f_{t_N}},$$

which belongs to M, would be invertible, since it has no zeros.

Hence, there exists some $t \in K$ such that $f(t) = 0$ for every $f \in M$. But M is maximal and contains all the functions $f \in \mathcal{C}(K)$ such that $f(t) = 0$.

Both K and Δ are compact spaces and $t \in K \mapsto \delta_t \in \Delta$, being continuous, is a homeomorphism.

8.3.2. Algebras of bounded analytic functions. Suppose that Ω is a bounded domain of $\mathbf{C}$ and denote by $H^\infty(\Omega)$ the algebra of bounded analytic functions in Ω, which is a commutative Banach algebra under the uniform norm

$$\|f\|_\infty = \sup_{z\in\Omega} |f(z)|.$$

It is a unitary Banach subalgebra of $\mathbf{B}(\Omega)$.

The Gelfand transform $\mathcal{G} : H^\infty(\Omega) \to \mathcal{C}(\Delta)$ is an isometric isomorphism, since $\|f^2\| = \|f\|^2$ and $r(f) = \|f\|$ for every $f \in A$, and we can see $H^\infty(\Omega)$ is a unitary Banach subalgebra of $\mathcal{C}(\Delta)$ (cf. Exercise 8.18).

For every $\zeta \in \Omega$, the evaluation map δ_ζ is the character of $H^\infty(\Omega)$ uniquely determined by $\delta_\zeta(z) = \zeta$, where z denotes the coordinate function.

Indeed, if $\chi \in \Delta$ satisfies the condition $\chi(z) = \zeta$ and if $f \in H^\infty(\Omega)$, then

$$f(z) = f(\zeta) + \frac{f(z) - f(\zeta)}{z - \zeta}(z - \zeta)$$

and

$$\chi(f) = f(\zeta) + \chi\Big(\frac{f - f(\zeta)}{z - \zeta}\Big)\chi(z - \zeta) = f(\zeta).$$

It can be shown that the embedding $\Omega \hookrightarrow \Delta$ such that $\zeta \mapsto \delta_\zeta$ is a homeomorphism from Ω onto an open subset of Δ (see Exercise 8.14 where we consider the case $\Omega = U$, the unit disc) and, for every $f \in H^\infty(\Omega)$, it is convenient to write $\widehat{f}(\delta_\zeta) = f(\zeta)$ if $\zeta \in \Omega$.

Suppose now that $\xi \in \partial\Omega$, a boundary point. Note that $z - \xi$ is not invertible in $H^\infty(\Omega)$, so that

$$\Delta_\xi := \Big\{\chi \in \Delta;\ \chi(z - \xi) = 0\Big\} = \Big\{\chi \in \Delta;\ \chi(z) = \xi\Big\} = (\widehat{z})^{-1}(\xi)$$

is not empty.

For every $\chi \in \Delta$, $\chi(z - \chi(z)) = 0$ and $z - \chi(z)$ is not invertible, so that $\chi(z) \in \bar{\Omega}$ and $\chi \in \Omega$ or $\chi \in \Omega_\xi$ for some $\xi \in \partial\Omega$. That is,

$$\Delta = \Omega \cup \Big(\bigcup_{\xi \in \partial\Omega} \Delta_\xi\Big) \tag{8.1}$$

and we can imagine Δ as the domain Ω with a compact fiber $\Delta_\xi = (\widehat{z})^{-1}(\xi)$ lying above every $\xi \in \partial\Omega$.

The corona problem asks whether Ω is dense in Δ for the Gelfand topology, and it admits a more elementary equivalent formulation in terms of function theory:

Theorem 8.15. *For the Banach algebra $H^\infty(\Omega)$, the domain Ω is dense in Δ if and only if the following condition holds:*

If $f_1, \dots, f_n \in H^\infty(\Omega)$ and if

$$|f_1(\zeta)| + \cdots + |f_n(\zeta)| \geq \delta > 0 \tag{8.2}$$

for every $\zeta \in \Omega$, then there exist $g_1, \dots, g_n \in H^\infty(\Omega)$ such that

$$f_1 g_1 + \cdots + f_n g_n = 1. \tag{8.3}$$

Proof. Suppose that Ω is dense in Δ. By continuity, if $|f_1| + \cdots + |f_n| \geq \delta$ on Ω, then also $|\widehat{f_1}| + \cdots + |\widehat{f_n}| \geq \delta$ on Δ, so that $\{f_1, \ldots, f_n\}$ is contained in no maximal ideal and

$$1 \in H^\infty(\Omega) = f_1 H^\infty(\Omega) + \cdots + f_n H^\infty(\Omega).$$

Conversely, suppose Ω is not dense in Δ and choose $\chi_0 \in \Delta$ with a neighborhood V disjoint from Ω. The Gelfand topology is the w^*-topology and this neighborhood has the form

$$V = \Big\{\chi;\ \max_{j=1,\ldots,n} |\chi(h_j) - \chi_0(h_j)| < \delta,\ h_1, \ldots, h_n \in H^\infty(\Omega)\Big\}.$$

The functions $f_j = h_j - \chi_0(h_j)$ are in V and they satisfy (8.2) because $\delta_\zeta \notin V$ and then $|f_j(\zeta)| \geq \delta$. But (8.3) is not possible because $f_1, \ldots, f_n \in \operatorname{Ker} \chi_0$ and $\chi_0(1) = 1$. □

Starting from the above equivalence, in 1962 Carleson[6] solved the corona problem for the unit disc, that is, D is dense in $\Delta(H^\infty(D))$.

The version of the corona theorem for the disc algebra is much easier. See Exercise 8.3.

8.4. C^*-algebras

We are going to consider a class of algebras whose Gelfand transform is a bijective and isometric isomorphism. Gelfand introduced his theory to study these algebras.

8.4.1. Involutions. A **C^*-algebra** is a unitary Banach algebra with an **involution**, which is a mapping $x \in A \mapsto x^* \in A$ that satisfies the following properties:

(a) $(x + y)^* = x^* + y^*$,
(b) $(\lambda x)^* = \bar{\lambda} x^*$,
(c) $(xy)^* = y^* x^*$,
(d) $x^{**} = x$, and
(e) $e^* = e$

for any $x, y \in A$ and $\lambda \in \mathbf{C}$, and such that $\|x^* x\| = \|x\|^2$ for every $x \in A$.

[6] The Swedish mathematician Lennart Carleson, awarded the Abel Prize in 2006, has solved some outstanding problems such as the corona problem (1962) and the almost everywhere convergence of Fourier series of any function in $L^2(\mathbf{T})$ (1966) and in complex dynamics. To quote Carleson "The corona construction is widely regarded as one of the most difficult arguments in modern function theory. Those who take the time to learn it are rewarded with one of the most malleable tools available. Many of the deepest arguments concerning hyperbolic manifolds are easily accessible to those who understand well the corona construction."

An involution is always bijective and it is its own inverse. It is isometric, since $\|x\|^2 = \|x^*x\| \le \|x^*\|\|x\|$, so that $\|x\| \le \|x^*\|$ and $\|x^*\| \le \|x^{**}\| = \|x\|$.

Throughout this section, A **will be a C^*-algebra**.

If H is a complex Hilbert, $\mathcal{L}(H)$ is a C^*-algebra with the involution $T \mapsto T^*$, where T^* denotes the adjoint of T. It has been proved in Theorem **??** that $\|T^*T\| = \|TT^*\| = \|T\|^2$.

Let A and B be two C^*-algebras. A **homomorphism of C^*-algebras** is a homomorphism $\Psi : A \to B$ of unitary Banach algebras such that $\Psi(x^*) = \Psi(x)^*$ (and, of course, $\Psi(e) = e$).

We say that $a \in A$ is **hermitian** or **self-adjoint** if $a = a^*$. The orthogonal projections of H are hermitian elements of $\mathcal{L}(H)$. We say that $a \in A$ is **normal** if $aa^* = a^*a$.

Example 8.16. If $a \in A$ is normal and $\langle a\rangle$ denotes the closed subalgebra of A generated by a, a^*, and e, then $\langle a\rangle$ contains all elements of A that can be obtained as the limits of sequences of polynomials in a, a^* and e. With the restriction of the involution of A, $\langle a\rangle$ is a commutative C^*-algebra.

Lemma 8.17. *Assume that A is commutative.*

(a) *If $a = a^* \in A$, then $\sigma_A(a) \subset \mathbf{R}$.*

(b) *For every $a \in A$ and $\chi \in \Delta(A)$, $\chi(a^*) = \overline{\chi(a)}$.*

Proof. If $t \in \mathbf{R}$, since $\|\chi\| = 1$,

$$\begin{aligned} |\chi(a+ite)|^2 &\le \|a+ite\|^2 = \|(a+ite)^*(a+ite)\| \\ &= \|(a-ite)(a+ite)\| = \|a^2+t^2e\| \le \|a\|^2 + t^2. \end{aligned}$$

Let $\chi(a) = \alpha + i\beta$ ($\alpha, \beta \in \mathbf{R}$). Then

$$\|a\|^2 + t^2 \ge |\alpha + i\beta + it|^2 = \alpha^2 + \beta^2 + 2\beta t + t^2,$$

i.e., $\|a\|^2 \ge \alpha^2 + \beta^2 + 2\beta t$, and it follows that $\beta = 0$ and $\chi_A(a) = \alpha \in \mathbf{R}$.

For any $a \in A$, if $x = (a + a^*)/2$ and $y = (a - a^*)/2i$, we obtain $a = x + iy$ with x, y hermitian, $\chi(x), \chi(y) \in \mathbf{R}$, and $a^* = x - iy$. Hence, $\chi(a) = \chi(x) + i\chi(y)$ and $\chi(a^*) = \chi(x) - i\chi(y) = \overline{\chi(a)}$. $\square$

Theorem 8.18. *If B is a closed unitary subalgebra of A such that $b^* \in B$ for every $b \in B$, then $\sigma_B(b) = \sigma_A(b)$ for every $b \in B$.*

Proof. First let $b^* = b$. From Lemma 8.17 we know that $\sigma_{\langle b\rangle} \subset \mathbf{R}$ and, obviously,

$$\sigma_A(b) \subset \sigma_B(b) \subset \sigma_{\langle b\rangle}(b) = \partial\sigma_{\langle b\rangle}(b).$$

To prove the inverse inclusions, it is sufficient to show that $\partial\sigma_{\langle b\rangle}(b) \subset \sigma_A(b)$. Let $\lambda \in \partial\sigma_{\langle b\rangle}(b)$ and suppose that $\lambda \notin \sigma_A(b)$. There exists $x \in A$ so that $x(b-\lambda e) = (b-\lambda e)x = e$ and the existence of $\lambda_n \notin \sigma_{\langle b\rangle}(b)$ such that $\lambda_n \to \lambda$ follows from $\lambda \in \partial\sigma_{\langle b\rangle}(b)$. Thus we have

$$(b-\lambda_n e)^{-1} \in \langle b\rangle \subset A, \quad b-\lambda_n e \to b-\lambda e, \text{ and } (b-\lambda_n e)^{-1} \to (b-\lambda e)^{-1} = x.$$

Hence $x \in \langle b\rangle$, in contradiction to $\lambda \in \sigma_{\langle b\rangle}(b)$.

In the general case we only need to prove that if $x \in B$ has an inverse y in A, then $y \in B$ also. But it follows from $xy = e = yx$ that $(x^*x)(yy^*) = e = (yy^*)(x^*x)$, and x^*x is hermitian. In this case we have seen above that x^*x has its unique inverse in B, so that $yy^* = (x^*x)^{-1} \in B$ and $y = y(y^*x^*) = (yy^*)x^* \in B$. □

8.4.2. The Gelfand-Naimark theorem and functional calculus. We have proved in Theorem 8.13 that the Gelfand transform satisfies $\|\widehat{a}\|_\Delta = r(a) \le \|a\|$, but in the general case it may not be injective. This is not the case for C^*-algebras.

Theorem 8.19 (Gelfand-Naimark)**.** *If A is a commutative C^*-algebra, then the Gelfand transform $\mathcal{G} : A \to \mathcal{C}(\Delta(A))$ is a bijective isometric isomorphism of C^*-algebras.*

Proof. We have $\widehat{a^*}(\chi) = \overline{\chi(a)} = \overline{\widehat{a}(\chi)}$ and $\mathcal{G}(a^*) = \overline{\mathcal{G}(a)}$.

If $x^* = x$, then $r(x) = \lim_n \|x^{2^n}\|^{1/2^n} = \|x\|$, since $\|x^2\| = \|xx^*\| = \|x\|^2$ and, by induction, $\|x^{2^{(n+1)}}\| = \|(x^{2^n})^2\| = (\|x\|^{2^n})^2 = \|x\|^{2^{(n+1)}}$.

If we take $x = a^*a$, then $\|\widehat{a^*a}\|_\Delta = \|a^*a\|$, so

$$\|a\|^2 = \|a^*a\| = \|\widehat{a^*a}\|_\Delta = \|\overline{\widehat{a}}\widehat{a}\|_\Delta = \|\widehat{a}\|_\Delta^2$$

and $\|a\| = \|\widehat{a}\|_\Delta$.

Since $\mathcal{G}$ is an isometric isomorphism, $\mathcal{G}(A)$ is a closed subalgebra of $\mathcal{C}(\Delta(A))$. This subalgebra contains the constant functions ($\widehat{e} = 1$) and it is self-conjugate and separates points (if $\chi_1 \neq \chi_2$, there exists $a \in A$ such that $\chi_1(a) \neq \chi_2(a)$, i.e., $\widehat{a}(\chi_1) \neq \widehat{a}(\chi_2)$). By the complex form of the Stone-Weierstrass theorem, the image is also dense, so $\mathcal{G}(A) = \mathcal{C}(\Delta(A))$ and $\mathcal{G}$ is bijective. □

Theorem 8.20. *Let a be a normal element of the C^*-algebra A, let $\Delta = \Delta\langle a\rangle$ be the spectrum of the subalgebra $\langle a\rangle$, and let $\mathcal{G} : \langle a\rangle \to \mathcal{C}(\Delta)$ be the Gelfand transform. The function $\widehat{a} : \Delta \to \sigma_A(a) = \sigma_{\langle a\rangle}(a)$ is a homeomorphism.*

Proof. We know $\sigma(a) = \widehat{a}(\Delta)$. If $\chi_1, \chi_2 \in \Delta$, from $\widehat{a}(\chi_1) = \widehat{a}(\chi_2)$ we obtain $\chi_1(a) = \chi_2(a)$, $\chi_1(a^*) = \overline{\chi_1(a)} = \overline{\chi_2(a)} = \chi_2(a^*)$, and $\chi_1(e) = 1 = \chi_2(e)$, so that $\chi_1(x) = \chi_2(x)$ for all $x \in \langle a\rangle$; hence, $\chi_1 = \chi_2$ and $\widehat{a} : \Delta \to \sigma(a)$ is bijective and continuous between two compact spaces, and then the inverse is also continuous. □

The homeomorphism $\widehat{a} : \Delta \to \sigma(a)$ $(\lambda = \widehat{a}(\chi))$ allows us to define the isometric isomorphism of C^*-algebras $\tau = \circ\widehat{a} : \mathcal{C}(\sigma(a)) \to \mathcal{C}(\Delta)$ such that $[g(\lambda)] \mapsto [G(\chi)] = [g(\widehat{a}(\chi))]$.

By Theorem 8.19, the composition

$$\Phi_a = \mathcal{G}^{-1} \circ \tau : \mathcal{C}(\sigma(a)) \to \mathcal{C}(\Delta) \to \langle a \rangle \subset A,$$

such that $g \in \mathcal{C}(\sigma(a)) \mapsto \mathcal{G}^{-1}(g(\widehat{a})) \in \langle a \rangle$, is also an isometric isomorphism of C^*-algebras. If $g \in \mathcal{C}(\sigma(a))$, then the identity $\widehat{\Phi_a(g)} = g \circ \widehat{a} = g(\widehat{a})$ suggests that we may write $g(a) := \Phi_a(g)$.

So, we have the isometric isomorphism of C^*-algebras

$$\Phi_a : g \in \mathcal{C}(\sigma(a)) \mapsto g(a) \in \langle a \rangle \subset A$$

such that, if $g_0(\lambda) = \lambda$ is the identity on $\sigma(a)$, then $\widehat{\Phi_a(g_0)} = \widehat{a}$ and $g_0(a) = a$, since $\tau(g_0) = \widehat{a} = \mathcal{G}(a)$. Also $\bar{g}_0(a) = a^*$ and

$$p(a) = \sum_{0 \le j,k \le N} c_{j,k} a^j (a^*)^k \text{ if } p(z) = \sum_{0 \le j,k \le N} c_{j,k} z^j \bar{z}^k. \tag{8.4}$$

We call Φ_a the **functional calculus with continuous functions**. It is the unique homomorphism $\Phi : \mathcal{C}(\sigma(a)) \to A$ of C^*-algebras such that

$$\Phi(p) = \sum_{0 \le j,k \le N} c_{j,k} a^j (a^*)^k (a)$$

if $p(z) = \sum_{0 \le j,k \le N} c_{j,k} z^j \bar{z}^k$.

Indeed, it follows from the Stone-Weierstrass theorem that the subalgebra $\mathcal{P}$ of all polynomials $p(z)$ considered in (8.4) is dense in $\mathcal{C}(\sigma(a))$ and, if $g = \lim_n p_n$ in $\mathcal{C}(\sigma(a))$ with $p_n \in \mathcal{P}$, then

$$\Phi(g) = \lim_n p_n(a) = \Phi_a(g).$$

These facts are easily checked and justify the notation $g(a)$ for $\Phi_a(g)$.

8.5. Spectral theory of bounded normal operators

In this section we are going to consider normal operators $T \in \mathcal{L}(H)$. By Theorem 8.18,

$$\sigma(T) = \sigma_{\mathcal{L}(H)}(T) = \sigma_{\langle T \rangle}(T)$$

and it is a nonempty compact subset of $\mathbf{C}$.

From now on, by $\mathbf{B}(\sigma(T))$ we will denote the C^*-algebra of all bounded Borel measurable functions $f : \sigma(T) \to \mathbf{C}$, endowed with the involution $f \mapsto \bar{f}$ and with the uniform norm. Obviously $\mathcal{C}(\sigma(T))$ is a closed unitary subalgebra of $\mathbf{B}(\sigma(T))$.

An application of the Gelfand-Naimark theorem to the commutative C^*-algebra $\langle T \rangle$ gives an isometric homomorphism from $\langle T \rangle$ onto $\mathcal{C}(\Delta\langle T \rangle)$.

The composition of this homomorphism with the change of variables $\lambda = \widehat{T}(\chi)$ ($\lambda \in \sigma(T)$ and $\chi \in \Delta\langle T\rangle$) defines the functional calculus with continuous functions on $\sigma(T)$, $g \in \mathcal{C}(\sigma(T)) \mapsto g(T) \in \langle T\rangle \subset \mathcal{L}(H)$, which is an isometric homomorphism of C^*-algebras.

If $x, y \in H$ are given, then

$$u_{x,y}(g) := (g(T)x, y)_H$$

defines a continuous linear form on $\mathcal{C}(\sigma(T))$ and, by the Riesz-Markov representation theorem,

$$(g(T)x, y)_H = u_{x,y}(g) = \int_{\sigma(T)} g\, d\mu_{x,y}$$

for a unique complex Borel measure $\mu_{x,y}$ on $\sigma(T)$.

We will say that $\{\mu_{x,y}\}$ is **the family of complex spectral measures of** T. For any bounded Borel measurable function f on $\sigma(T)$, we can define

$$u_{x,y}(f) := \int_{\sigma(T)} f\, d\mu_{x,y}$$

and in this way we extend $u_{x,y}$ to a linear form on these functions. Note that $|u_{x,y}(g)| = |(g(T)x, y)_H| \leq \|x\|_H \|y\|_H \|g\|_{\sigma(T)}$.

8.5.1. Functional calculus of normal operators. Now our goal is to show that it is possible to define $f(T) \in \mathcal{L}(H)$ for every f in the C^*-algebra $\mathbf{B}(\sigma(T))$ of all bounded Borel measurable functions on $\sigma(T) \subset \mathbf{C}$, equipped with the uniform norm and with the involution $f \mapsto \bar{f}$, so that

$$(f(T)x, y)_H = u_{x,y}(f) = \int_{\sigma(T)} f\, d\mu_{x,y}$$

in the hope of obtaining a functional calculus $f \mapsto f(T)$ for bounded but not necessarily continuous functions.

Theorem 8.21. *Let $T \in \mathcal{L}(H)$ be a normal operator ($TT^* = T^*T$) and let $\{\mu_{x,y}\}$ be its family of complex spectral measures. Then there exists a unique homomorphism of C^*-algebras*

$$\Phi_T : \mathbf{B}(\sigma(T)) \to \mathcal{L}(H)$$

such that

$$(\Phi_T(f)x, y)_H = \int_{\sigma(T)} f\, d\mu_{x,y} \qquad (x, y \in H).$$

It is an extension of the continuous functional calculus $g \mapsto g(T)$, and $\|\Phi_T(f)\| \leq \|f\|_{\sigma(T)}$.

Proof. Note that, if μ_1 and μ_2 are two complex Borel measures on $\sigma(T)$ and if $\int g\,d\mu_1 = \int g\,d\mu_2$ for all real $g \in \mathcal{C}(\sigma(T))$, then $\mu_1 = \mu_2$, by the uniqueness in the Riesz-Markov representation theorem.

If $g \in \mathcal{C}(\sigma(T))$ is a real function, then $g(T)$ is self-adjoint, since $g(T)^* = \bar{g}(T)$. Hence, $(g(T)x, y)_H = \overline{(g(T)y, x)_H}$ and then

$$\int_{\sigma(T)} g\,d\mu_{x,y} = \overline{\int_{\sigma(T)} g\,d\mu_{y,x}} = \int_{\sigma(T)} g\,d\bar{\mu}_{y,x},$$

so that

$$\mu_{x,y} = \bar{\mu}_{y,x}.$$

Obviously, $(x, y) \mapsto \int_{\sigma(T)} g\,d\mu_{x,y} = (g(T)x, y)_H$ is a continuous sesquilinear form and, from the uniqueness in the Riesz-Markov representation theorem, the map $(x, y) \mapsto \mu_{x,y}(B)$ is also sesquilinear, for any Borel set $B \subset \sigma(T)$. For instance, $\mu_{x,\lambda y} = \bar{\lambda}\mu_{x,y}$, since for continuous functions we have

$$\int_{\sigma(T)} g\,d\mu_{x,\lambda y} = \bar{\lambda}(g(T)x, y)_H = \int_{\sigma(T)} g\,d\lambda\mu_{x,y}.$$

With the extension $u_{x,y}(f) := \int_{\sigma(T)} f\,d\mu_{x,y}$ of $u_{x,y}$ to functions f in $\mathbf{B}(\sigma(T))$, it is still true that

$$|u_{x,y}(f)| \le \|x\|_H \|y\|_H \|f\|_{\sigma(T)}.$$

For every $f \in \mathbf{B}(\sigma(T))$,

$$(x, y) \mapsto B_f(x, y) := \int_{\sigma(T)} f\,d\mu_{x,y}$$

is a continuous sesquilinear form on $H \times H$ and $B_f(y, x) = \overline{B_f(x, y)}$, since

$$\overline{\int_{\sigma(T)} f\,d\mu_{x,y}} = \int_{\sigma(T)} f\,d\mu_{y,x}$$

($\overline{\mu_{x,y}(B)} = \mu_{y,x}(B)$ extends to simple functions). Let us check that an application of the Riesz representation theorem produces a unique operator $\Phi_T(f) \in \mathcal{L}(H)$ such that $B_f(x, y) = (\Phi_T(f)x, y)_H$.

Note that $B_f(\cdot, x) \in H'$ and there is a unique $\Phi_T(f)x \in H$ so that $B_f(y, x) = (y, \Phi_T(f)x)_H$ for all $y \in H$. Then

$$(\Phi_T(f)x, y)_H = \overline{B_f(y, x)} = B_f(x, y) = \int_{\sigma(T)} f\,d\mu_{x,y} \qquad (x, y \in H).$$

It is clear that $B_f(x, y)$ is linear in f and that we have defined a bounded linear mapping $\Phi_T : \mathbf{B}(\sigma(T)) \to \mathcal{L}(H)$ such that

$$|(\Phi_T(f)x, y)_H| \le \|f\|_{\sigma(T)} \|x\|_H \|y\|_H$$

and $\|\Phi_T(f)\| \leq \|f\|_{\sigma(T)}$. Moreover, with this definition, Φ_T extends the functional calculus with continuous functions, $g \mapsto g(T)$.

To prove that Φ_T is a continuous homomorphism of C^*-algebras, all that remains is to check its behavior with the involution and with the product.

If f is real, then from $\mu_{x,y} = \bar{\mu}_{y,x}$ we obtain $(\Phi_T(f)x, y)_H = \overline{(\Phi_T(f)y, x)}_H$ and $\Phi_T(f)^* = \Phi_T(f)$. In the case of a complex function, f, $\Phi_T(f)^* = \Phi_T(\bar{f})$ follows by linearity.

Finally, to prove that $\Phi_T(f_1f_2) = \Phi_T(f_1)\Phi_T(f)(f_2)$, we note that, on continuous functions,

$$\int_{\sigma(T)} hg\, d\mu_{x,y} = (h(T)g(T)x, y)_H = \int_{\sigma(T)} h\, d\mu_{g(T)x,y}$$

and $g\, d\mu_{x,y} = d\mu_{g(T)x,y}$ $(x, y \in H)$. Hence, also

$$\int_{\sigma(T)} f_1 g\, d\mu_{x,y} = \int_{\sigma(T)} f_1\, d\mu_{g(T)x,y}$$

if f_1 is bounded, and then

$$\begin{aligned} \int_{\sigma(T)} f_1 g\, d\mu_{x,y} &= (\Phi_T(f_1)g(T)x, y)_H = (g(T)x, \Phi_T(f_1)^*y)_H \\ &= \int_{\sigma(T)} g\, d\mu_{x,\Phi(f_1)^*y}. \end{aligned}$$

Again $f_1 d\mu_{x,y} = d\mu_{x,\Phi_T(f_1)^*y}$, and also $\int_{\sigma(T)} f_1 f_2\, d\mu_{x,y} = \int_{\sigma(T)} f_2\, d\mu_{x,f_1(T)^*y}$ if f_1 and f_2 are bounded. Thus,

$$\begin{aligned} (\Phi_T(f_1f_2)x, y)_H &= \int_{\sigma(T)} f_1 f_2\, d\mu_{x,y} \\ &= \int_{\sigma(T)} f_2\, d\mu_{x,\Phi_T(f_1)^*y} = (\Phi_T(f_1)\Phi_T(f_2)x, y)_H \end{aligned}$$

and $\Phi_T(f)$ is multiplicative. □

As in the case of the functional calculus for continuous functions, if $f \in \mathbf{B}(\sigma(T))$, we will denote the operator $\Phi_T(f)$ by $f(T)$; that is,

$$(f(T)x, y)_H = \int_{\sigma(T)} f\, d\mu_{x,y} \qquad (x, y \in H).$$

8.5.2. Spectral measures. For a given Hilbert space, H, a **spectral measure**, or a **resolution of the identity**, on a locally compact subset K of $\mathbf{C}$ (or of $\mathbf{R}^n$), is an operator-valued mapping defined on the Borel σ-algebra $\mathcal{B}_K$ of K,

$$E : \mathcal{B}_K \longrightarrow \mathcal{L}(H),$$

that satisfies the following conditions:

(1) Each $E(B)$ is an orthogonal projection.

(2) $E(\emptyset) = 0$ and $E(K) = I$, the identity operator.

(3) If $B_n \in \mathcal{B}_K$ $(n \in \mathbf{N})$ are disjoint, then

$$E(\biguplus_{n=1}^{\infty} B_n)x = \sum_{n=1}^{\infty} E(B_n)x$$

for every $x \in H$, and it is said that

$$E(\biguplus_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} E(B_n)$$

for the strong convergence, or that E is strongly σ-additive.

Note that E also has the following properties:

(4) If $B_1 \cap B_2 = \emptyset$, then $E(B_1)E(B_2) = 0$ (orthogonality).

(5) $E(B_1 \cap B_2) = E(B_1)E(B_2) = E(B_2)E(B_1)$ (multiplicativity).

(6) If $B_1 \subset B_2$, then $\operatorname{Im} E(B_1) \subset \operatorname{Im} E(B_2)$ (usually represented by $E(B_1) \leq E(B_2)$).

(7) If $B_n \uparrow B$ or $B_n \downarrow B$, then $\lim_n E(B_n)x = E(B)x$ for every $x \in H$ (it is said that $E(B_n) \to E(B)$ strongly).

Indeed, to prove (4), if $y = E(B_2)x$, the equality

$$(E(B_1) + E(B_2))^2 = E(B_1 \uplus B_2)^2 = E(B_1) + E(B_2)$$

and the condition $B_1 \cap B_2 = \emptyset$ yield

$$E(B_1)E(B_2)x + E(B_2)E(B_1)x = 0,$$

that is, $E(B_1)y + y = 0$ and, applying $E(B_1)$ to both sides, $E(B_1)y = 0$.

Now (5) follows from multiplying the equations

$$E(B_1) = E(B_1 \cap B_2) + E(B_1 \setminus B_1 \cap B_2), \quad E(B_2) = E(B_1 \cap B_2) + E(B_2 \setminus B_1 \cap B_2)$$

and taking into account (4).

If $B_n \uparrow B$, then $\lim_n E(B_n)x = E(B)x$ follows from (3), since

$$B = B_1 \uplus (B_2 \setminus B_1) \uplus (B_3 \setminus B_2) \uplus \cdots .$$

The decreasing case $B_n \downarrow B$ reduces to $K \setminus B_n \uparrow K \setminus B$.

It is also worth noticing that the spectral measure E generates the family of complex measures $E_{x,y}$ $(x, y \in H)$ defined as

$$E_{x,y}(B) := (E(B)x, y)_H.$$

If $x \in H$, then $E_x(B) := E(B)x$ defines a vector measure $E_x : \mathcal{B}_K \to H$, i.e., $E_x(\emptyset) = 0$ and $E_x(\biguplus_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} E_x(B_n)$ in H.

Note that for every $x \in H$, $E_{x,x}$ is a (positive) measure such that

$$E_{x,x}(B) = (E(B)x, y)_H = \|E(B)x\|_H^2, \quad E_{x,x}(K) = \|x\|_H^2,$$

a probability measure if $\|x\|_H = 1$, and that operations with the complex measures $E_{x,y}$, by polarization, reduce to operations with positive measures:

$$E_{x,y}(B) = \frac{1}{4}\Big(E_{x+y,x+y}(B) - E_{x-y,x-y}(B) + iE_{x+iy,x+iy}(B) - iE_{x-iy,x-iy}(B)\Big).$$

The notions E**-almost everywhere** (E-a.e.) and E**-essential supremum** have the usual meaning. In particular, if f is a real measurable function,

$$E\text{-}\sup f = \inf\{M \in \mathbf{R};\ f \le M \ E\text{-a.e.}\}.$$

Note that $E(B) = 0$ if and only if $E_{x,x}(B) = 0$ for every $x \in H$. Thus, if $B_1 \subset B_2$ and $E(B_2) = 0$, then $E(B_1) = 0$, and the class of E-null sets is closed under countable unions.

The **support of a spectral measure** E on K is defined as the least closed set $\operatorname{supp} E$ such that $E(K \setminus \operatorname{supp} E) = 0$. The support consists precisely of those points in K for which every neighborhood has nonzero E-measure and $E(B) = E(B \cap \operatorname{supp} E)$ for every Borel set $B \subset K$.

The existence of the support is proved by considering the union V of the open sets V_α of K such that $E(V_\alpha) = 0$. Since there is a sequence V_n of open sets in K such that $V_\alpha = \bigcup_{\{n;\, V_n \subset V_\alpha\}} V_n$, then also $V = \bigcup_{\{n;\, V_n \subset V\}} V_n$ and $E(V) = 0$. Then $\operatorname{supp} E = K \setminus V$.

We write

$$R = \int_K f\, dE$$

to mean that

$$(Rx, y)_H = \int_K f\, dE_{x,y} \qquad (x, y \in H).$$

It is natural to ask whether the family $\{\mu_{x,y}\}$ of complex measures associated to a normal operator T is generated by a single spectral measure E associated to T. The next theorem shows that the answer is affirmative, allowing us to rewrite the functional calculus of Theorem 8.21 as $f(T) = \int_{\sigma(T)} f\, dE$.

Theorem 8.22 (Spectral resolution[7]). *If $T \in \mathcal{L}(H)$ is a normal operator, then there exists a unique spectral measure $E : \mathcal{B}_{\sigma(T)} \to \mathcal{L}(H)$ which satisfies*

$$T = \int_{\sigma(T)} \lambda \, dE(\lambda). \tag{8.5}$$

Furthermore,

$$f(T) = \int_{\sigma(T)} f(\lambda) \, dE(\lambda) \qquad (f \in \mathbf{B}(\sigma(T))), \tag{8.6}$$

and

$$E(B) = \chi_B(T) \qquad (B \in \mathcal{B}_{\sigma(T)}). \tag{8.7}$$

Proof. If $\Phi_T : \mathbf{B}(\sigma(T)) \to \mathcal{L}(H)$ is the homomorphism that defines the functional calculus, then to obtain (8.6) we must define E by condition (8.7),

$$E(B) := \Phi_T(\chi_B) \qquad (B \in \mathcal{B}_{\sigma(T)}),$$

and then check that E is a spectral measure with the convenient properties.

Obviously, $E(B) = E(B)^2$ and $E(B)^* = E(B)$ (χ_B is real), so that $E(B)$ is an orthogonal projection (Theorem **??**).

Moreover, it follows from the properties of the functional calculus for continuous functions that $E(\sigma(T)) = \Phi(1) = 1(T) = I$ and $E(\emptyset) = \Phi(0) = 0$ and, since Φ is linear, E is additive. Also, from

$$(E(B)x, y)_H = (\Phi(\chi_B)x, y)_H = \mu_{x,y}(B),$$

we obtain that

$$\int_{\sigma(T)} g \, dE = g(T), \quad \int_{\sigma(T)} f \, dE = \Phi(f) \qquad \Big(g \in \mathcal{C}(\sigma(T)),\ f \in \mathbf{B}(\sigma(T))\Big).$$

Finally, E is strongly σ-additive since, if B_n ($n \in \mathbf{N}$) are disjoint Borel sets, $E(B_n)E(B_m) = 0$ if $n \neq m$, so that the images of the projections $E(B_n)$ are mutually orthogonal (if $y = E(B_m)x$, we have $y \in \operatorname{Ker} E(B_n)$ and $y \in E(B_n)(H)^{\perp}$) and then, for every $x \in H$, $\sum_n E(B_n)x$ is convergent to some $Px \in H$ since

$$\sum_n \|E(B_n)x\|_H^2 \leq \|x\|_H^2,$$

this being true for partial sums, $\|E(\biguplus_{n=1}^N B_n)x\|_H^2 \leq \|x\|_H^2$.

[7] In their work on integral equations, D. Hilbert for a self-adjoint operator on ℓ^2 and F. Riesz on L^2 used the Stieltjes integral

$$\int_{\sigma(T)} t \, dE(t) = \lim \sum_k t_k (E(t_k) - E(t_{k-1})) \qquad (\text{here } E(t_k) - E(t_{k-1}) = E(t_{k-1}, t_k])$$

to obtain this spectral theorem.

But then

$$(Px,y)_H = \sum_{n=1}^{\infty}(E(B_n)x,y)_H = \mu_{x,y}\big(\biguplus_{n=1}^{\infty} B_n\big) = (E\big(\biguplus_{n=1}^{\infty} B_n\big)x,y)_H$$

and $\sum_n E(B_n)x = Px = E\big(\biguplus_n B_n\big)x$.[8]

The uniqueness of E follows from the uniqueness for the functional calculus for continuous functions Φ_T and from the uniqueness of the measures $E_{x,y}$ in the Riesz-Markov representation theorem. □

Remark 8.23. A more general spectral theorem due to John von Neumann in 1930 can also be obtained from the Gelfand-Naimark theorem: any commutative family of normal operators admits a single spectral measure which simultaneously represents all operators of the family as integrals $\int_K g\,dE$ for various functions g.

8.5.3. Applications. There are two special instances of normal operators that we are interested in: self-adjoint operators and unitary operators.

Recall that an operator $U \in \mathcal{L}(H)$ is said to be **unitary** if it is a bijective isometry of H. This means that

$$UU^* = U^*U = I$$

since $U^*U = I$ if and only if $(Ux,Uy)_H = (x,y)$ and U is an isometry. If it is bijective, then $(U^{-1}x,U^{-1}y)_H = (x,y)_H$ and $((U^{-1})^*U^{-1}x,y)_H = (x,y)_H$, where $(U^{-1})^*U^{-1}x = (U^*)^{-1}U^{-1}x = (UU^*)^{-1}x$ and then $((UU^*)^{-1}x,y)_H = (x,y)_H$, so that $UU^* = I$. Conversely, if $UU^* = I$, then U is exhaustive.

The Fourier transform is an important example of a unitary operator of $L^2(\mathbf{R}^n)$.

Knowing the spectrum allows us to determine when a normal operator is self-adjoint or unitary:

Theorem 8.24. *Let $T \in \mathcal{L}(H)$ be a normal operator.*

(a) *T is self-adjoint if and only if $\sigma(T) \subset \mathbf{R}$.*

(b) *T is unitary if and only if $\sigma(T) \subset \mathbf{S} = \{\lambda;\ |\lambda| = 1\}$.*

Proof. We will apply the continuous functional calculus Φ_T for T to the identity function $g(\lambda) = \lambda$ on $\sigma(T)$, so that $g(T) = T$ and $\bar{g}(T) = T^*$.

From the injectivity of Φ_T, $T = T^*$ if and only if $g = \bar{g}$, meaning that $\lambda = \bar{\lambda} \in \mathbf{R}$ for every $\lambda \in \sigma(T)$.

Similarly, T is unitary if and only if $TT^* = T^*T = I$, i.e., when $g\bar{g} = 1$, which means that $|\lambda| = 1$ for all $\lambda \in \sigma(T)$. □

[8] See also Exercise 8.21.

Positivity can also be described through the spectrum:

Theorem 8.25. *Suppose $T \in \mathcal{L}(H)$. Then*

$$(Tx, x)_H \geq 0 \qquad (x \in H) \tag{8.8}$$

if and only if

$$T = T^* \text{ and } \sigma(T) \subset [0, \infty). \tag{8.9}$$

Such an operator is said to be positive.

Proof. It follows from (8.8) that $(Tx, x)_H \in \mathbf{R}$ and then

$$(Tx, x)_H = (x, Tx)_H = (T^*x, x)_H.$$

Let us show that then $S := T - T^* = 0$.

Indeed, $(Sx, y)_H + (Sy, x) = 0$ and, replacing y by iy, $-i(Sx, y)_H + i(Sy, x) = 0$. Now we multiply by i and add to obtain $(Sx, y)_H = 0$ for all $x, y \in H$, so that $S = 0$.

Thus $\sigma(T) \subset \mathbf{R}$. To prove that $\lambda < 0$ cannot belong to $\sigma(T)$, we note that the condition (8.8) allows us to set

$$\|(T - \lambda I)x\|_H^2 = \|Tx\|_H^2 - 2\lambda(Tx, x)_H + \lambda^2\|x\|_H^2.$$

This shows that $T_\lambda := T - \lambda I : H \to F = \Im(T - \lambda I)$ has a continuous inverse with domain F, which is closed. This operator is easily extended to a left inverse R of T_λ by defining $R = 0$ on $F^\perp$. But T_λ is self-adjoint and $RT_\lambda = I$ also gives $T_\lambda R^* = I$, T_λ is also right invertible, and $\lambda \notin \sigma(T)$.

Suppose now that $T = T^*$ and $\sigma(T) \subset [0, \infty)$. In the spectral resolution

$$(Tx, x)_H = \int_{\sigma(T)} \lambda \, dE_{x,x}(\lambda) \geq 0,$$

since $E_{x,x}$ is a positive measure and $\lambda \geq 0$ on $\sigma(T) \subset [0, \infty)$. $\square$

Let us now give an application of the functional calculus with bounded functions:

Theorem 8.26. *If $T = \int_{\sigma(T)} \lambda \, dE(\lambda)$ is the spectral resolution of a normal operator $T \in \mathcal{L}(H)$ and if $\lambda_0 \in \sigma(T)$, then*

$$\operatorname{Ker}(T - \lambda_0 I) = \operatorname{Im} E\{\lambda_0\},$$

so that λ_0 is an eigenvalue of T if and only if $E(\{\lambda_0\}) \neq 0$.

Proof. The functions $g(\lambda) = \lambda - \lambda_0$ and $f = \chi_{\{\lambda_0\}}$ satisfy $fg = 0$ and $g(T)f(T) = 0$. Since $f(T) = E(\{\lambda_0\})$,

$$\operatorname{Im} E(\{\lambda_0\}) \subset \operatorname{Ker}(g(T)) = \operatorname{Ker}(T - \lambda_0 I).$$

Conversely, let us take

$$G = \sigma(T) \setminus \{\lambda_0\} = \biguplus_n B_n$$

with $d(\lambda_0, B_n) > 0$ and define the bounded functions

$$f_n(\lambda) = \frac{\chi_{B_n}(\lambda)}{\lambda - \lambda_0}.$$

Then $f_n(T)(T - \lambda_0 I) = E(B_n)$, and $(T - \lambda_0 I)x = 0$ implies $E(B_n)x = 0$ and $E(G)x = \sum_n E(B_n)x = 0$. Hence, $x = E(G)x + E(\{\lambda_0\})x = E(\{\lambda_0\})x$, i.e., $x \in \operatorname{Im} E(\{\lambda_0\})$. □

As shown in Section **??**, if T is compact, then every nonzero eigenvalue has finite multiplicity and $\sigma(T) \setminus \{0\}$ is a finite or countable set of eigenvalues with finite multiplicity with 0 as the only possible accumulation point. If T is normal, the converse is also true:

Theorem 8.27. *If $T \in \mathcal{L}(T)$ is a normal operator such that $\sigma(T)$ has no accumulation point except possibly* 0 *and* $\dim \operatorname{Ker}(T - \lambda I) < \infty$ *for every $\lambda \neq 0$, then T is compact.*

Proof. Let $\sigma(T) \setminus \{0\} = \{\lambda_1, \lambda_2, \ldots\}$ and $|\lambda_1| \geq |\lambda_2| \geq \cdots$. We apply the functional calculus to the functions g_n defined as

$$g_n(\lambda) = \lambda \text{ if } \lambda = \lambda_k \text{ and } k \leq n$$

and $g_n(\lambda) = 0$ at the other points of $\sigma(T)$ to obtain the compact operator with finite-dimensional range

$$g_n(T) = \sum_{k=1}^{n} \lambda_k E(\{\lambda_k\}).$$

Then

$$\|T - g_n(T)\| \leq \sup_{\lambda \in \sigma(T)} |\lambda - g_n(\lambda)| \leq |\lambda_n|$$

and $|\lambda_n| \to 0$ as $n \to \infty$ if $\sigma(T) \setminus \{0\}$ is an infinite set. This shows that T is compact as a limit of compact operators. □

8.6. Exercises

Exercise 8.1. Show that every Banach algebra A without a unit element can be considered as a Banach subalgebra of a unitary Banach algebra A_1 constructed in the following fashion. On $A_1 = A \times \mathbf{C}$, which is a vector space, define the multiplication $(a, \lambda)(b, \mu) := (ab + \lambda b + \mu a, \lambda\mu)$ and the norm $\|(a, \lambda)\| := \|a\| + |\lambda|$. The unit is $\delta = (0, 1)$.

The map $a \mapsto (a, 0)$ is an isometric homomorphism (i.e., linear and multiplicative), so that we can consider A as a closed subalgebra of A_1. By denoting $a = (a, 0)$ if $a \in A$, we can write $(a, \lambda) = a + \lambda\delta$ and the projection $\chi_0(a + \lambda\delta) := \lambda$ is a character $\chi_0 \in \Delta(A_1)$.

If A is unitary, then the unit e cannot be the unit δ of A_1.

Exercise 8.2. Suppose that $A = \mathbf{C}_1$ is the commutative unitary Banach algebra obtained by adjoining a unit to $\mathbf{C}$ as in Exercise 8.1. Describe $\Delta(A)$ and the corresponding Gelfand transform.

Exercise 8.3. (a) Prove that the polynomials $P(z) = \sum_{n=0}^{N} c_n z^n$ are dense in the disc algebra $A(D)$ by showing that, if $f \in A(D)$ and

$$f_n(z) := f\Big(\frac{nz}{n+1}\Big),$$

then $f_n \to f$ uniformly on $\bar{D}$ and, if $\|f - f_n\| \leq \varepsilon/2$, there is a Taylor polynomial P of f_n such that $\|f_n - P\| \leq \varepsilon/2$.

Hence, polynomials P are not dense in $\mathcal{C}(\bar{D})$. Why is this not in contradiction to the Stone-Weierstrass theorem?

(b) Prove that the characters of $A(D)$ are the evaluations δ_z ($|z| \leq 1$) and that $z \in \bar{D} \mapsto \delta_z \in \Delta(A(D))$ is a homeomorphism.

(c) If $f_1, \dots, f_n \in A(D)$ have no common zeros, prove that there exist $g_1, \dots, g_n \in A(D)$ such that $f_1 g_1 + \cdots + f_n g_n = 1$.

Exercise 8.4. Show that, with the convolution product,

$$f * g(x) := \int_{\mathbf{R}} f(x-y) g(y)\, dy,$$

the Banach space $L^1(\mathbf{R})$ becomes a nonunitary Banach algebra.

Exercise 8.5. Show also that $L^1(\mathbf{T})$, the Banach space of all complex 1-periodic functions that are integrable on $(0, 1)$, with the convolution product

$$f * g(x) := \int_0^1 f(x-y) g(y)\, dy,$$

and the usual L^1 norm, is a nonunitary Banach algebra.

Exercise 8.6. Show that $\ell^1(\mathbf{Z})$, with the discrete convolution,

$$(u * v)[k] := \sum_{m=-\infty}^{+\infty} u[k-m] v[m],$$

is a unitary Banach algebra.

Exercise 8.7. Every unitary Banach algebra, A, can be considered a closed subalgebra of $\mathcal{L}(A)$ by means of the isometric homomorphism $a \mapsto L_a$, where $L_a(x) := ax$.

Exercise 8.8. In this exercise we want to present the Fourier transform on $L^1(\mathbf{R})$ as a special case of the Gelfand transform. To this end, consider the unitary commutative Banach algebra $L^1(\mathbf{R})_1$ obtained as in Exercise 8.1 by adjoining the unit to $L^1(\mathbf{R})$, which is a nonunitary convolution Banach algebra $L^1(\mathbf{R})$ (see Exercise 8.4).

(a) Prove that, if $\chi \in \Delta(L^1(\mathbf{R})_1) \setminus \{\chi_0\}$ and $\chi(u) = 1$ with $u \in L^1(\mathbf{R})$, then $\gamma_\chi(\alpha) := \chi(\tau_{-\alpha}u) = \chi([u(t+\alpha)])$ defines a function $\gamma_\chi : \mathbf{R} \to \mathbf{T} \subset \mathbf{C}$ which is continuous and such that $\gamma_\chi(\alpha+\beta) = \gamma_\chi(\alpha)\gamma_\chi(\beta)$.

(b) Prove that there exists a uniquely determined number $\xi_\chi \in \mathbf{R}$ such that $\gamma_\chi(\alpha) = e^{2\pi i \xi_\chi \alpha}$.

(c) Check that, if $\mathcal{G}f$ denotes the Gelfand transform of $f \in L^1(\mathbf{R})$, then $\mathcal{G}f(\chi) = \int_{\mathbf{R}} f(\alpha)e^{-2\pi i\alpha\xi_\chi}\, d\alpha = \mathcal{F}f(\xi_\chi)$.

Exercise 8.9. Let us consider the unitary Banach algebra $L^\infty(\Omega)$ of Example 8.2. The **essential range**, $f[\Omega]$, of $f \in L^\infty(\Omega)$ is the complement of the open set $\bigcup\{G;\, G \text{ open}, \mu(f^{-1}(G)) = 0\}$. Show that $f[\Omega]$ is the smallest closed subset F of $\mathbf{C}$ such that $\mu(f^{-1}(F^c)) = 0$, $\|f\|_\infty = \max\{|\lambda|;\, \lambda \in f[\Omega]\}$, and $f[\Omega] = \sigma(f)$.

Exercise 8.10. The algebra of **quaternions**, $\mathbf{H}$, is the real Banach space $\mathbf{R}^4$ endowed with the distributive product such that

$$1x = x,\ ij = -ji = k,\ jk = -kj = i,\ ki = -ik = j,\ i^2 = j^2 = k^2 = -1$$

if $x \in \mathbf{H}$, $1 = (1,0,0,0)$, $i = (0,1,0,0)$, $j = (0,0,1,0)$, and $k = (0,0,0,1)$, so that one can write $(a,b,c,d) = a + bi + cj + dk$.

Show that $\mathbf{H}$ is an algebra such that $\|xy\| = \|x\|\|y\|$ and that every nonzero element of A has an inverse.

Remark. It can be shown that every real Banach algebra which is a field is isomorphic to the reals, the complex numbers or the quaternions (cf. Rickart *General Theory of Banach Algebras*, [**?**, 1.7]). Hence, $\mathbf{C}$ is the only (complex) Banach algebra which is a field and $\mathbf{H}$ is the only real Banach algebra which is a noncommutative field.

Exercise 8.11. If $\chi : A \to \mathbf{C}$ is linear such that $\chi(ab) = \chi(a)\chi(b)$ and $\chi \neq 0$, then prove that $\chi(e) = 1$, so that χ is a character.

Exercise 8.12. Prove that, if $\mathcal{T}$ is a compact topology on $\Delta(A)$ and every function $\widehat{a}$ $(a \in A)$ is $\mathcal{T}$-continuous, then $\mathcal{T}$ is the Gelfand topology.

Exercise 8.13. Prove that the Gelfand transform is an isometric isomorphism from $\mathcal{C}(K)$ onto $\mathcal{C}(\Delta)$.

Exercise 8.14. Let U be the open unit disc of $\mathbf{C}$ and suppose Δ is the spectrum of $H^\infty(U)$. Prove that, through the embedding $U \hookrightarrow \Delta$, U is an open subset of Ω. Write

$$\Delta = D \cup \Big(\bigcup_{\xi \in \partial D} \Delta_\xi \Big),$$

as in (8.1), and prove that the fibers Δ_ξ ($|\xi| = 1$) are homeomorphic to one another.

Exercise 8.15 (Wiener algebra)**.** Show that the set of all 2π-periodic complex functions on $\mathbf{R}$

$$f(t) = \sum_{k=-\infty}^{+\infty} c_k e^{ikt} \qquad (\sum_{k=-\infty}^{+\infty} |c_k| < \infty),$$

with the usual operations and the norm $\|f\|_W := \sum_{k=-\infty}^{+\infty} |c_k|$, is a commutative unitary Banach algebra, W. Moreover prove that the characters of W are the evaluations δ_t on the different points $t \in \mathbf{R}$ and that, if $f \in W$ has no zeros, then $1/f \in W$.

Exercise 8.16. Every $f \in W$ is 2π-periodic and it can be identified as the function F on $\mathbf{T}$ such that $f(t) = F(e^{it})$. If $f(t) = \sum_k c_k e^{ikt}$, $F(z) = \sum_k c_k z^k$. In Exercice 8.15 we have seen that the δ_t ($t \in \mathbf{R}$) are the characters of W, but show that $\mathcal{G} : W \to \mathcal{C}(\mathbf{T})$, one-to-one and with $\|\widehat{f}\| \leq \|f\|_W$, is not an isometry and it is not exhaustive.

Exercise 8.17. Suppose A is a unitary Banach algebra and $a \in A$, and denote $M(U) = \sup_{\lambda \in U^c} \|R_a(\lambda)\|$. Prove that, if $U \subset \mathbf{C}$ is an open set and $\sigma_A(a) \subset U$, then $\sigma_A(b) \subset U$ whenever $\|b - a\| < \delta$ if $\delta \leq 1/M(U)$ (upper semi-continuity of σ_A).

Exercise 8.18. Let A be a commutative unitary Banach algebra. Prove that the Gelfand transform $\mathcal{G} : A \to \mathcal{C}(\Delta)$ is an isometry if and only if $\|a^2\| = \|a\|^2$ for every $a \in A$.

Show that in order for $\|a\|$ to coincide with the spectral radius $r(a)$, the condition $\|a^2\| = \|a\|^2$ is necessary and sufficient.

Remark. This condition characterizes when a Banach algebra A is a **uniform algebra**, meaning that A is a closed unitary subalgebra of $\mathcal{C}(K)$ for some compact topological space K.

Exercise 8.19. In the definition of an involution, show that property (e), $e^* = e$, is a consequence of (a)–(d). If $x \in A$ is invertible, prove that $(x^*)^{-1} = (x^{-1})^*$.

Exercise 8.20. With the involution $f \mapsto \bar{f}$, where $\bar{f}$ is the complex conjugate of f, show that $\mathcal{C}(K)$ is a commutative C^*-algebra. Similarly, show that $L^\infty(\Omega)$, with the involution $f \mapsto \bar{f}$, is also a commutative C^*-algebra.

Exercise 8.21. If $\{P_n\}_{n=1}^{\infty}$ is a sequence of orthogonal projections and their images are mutually orthogonal, then the series $\sum_{n=1}^{\infty} P_n$ is strongly convergent to the orthogonal projection on the closed linear hull $\bigoplus P_n(H)$ of the images of the projections P_n.

Exercise 8.22. Let $Ax = \sum_{k=1}^{\infty} \lambda_k (x, e_k)_H e_k$ be the spectral representation of a self-adjoint compact operator of H, and let $\{P_n\}_{n=1}^{\infty}$ be the sequence of the orthogonal projections on the different eigensubspaces

$$H_1 = [e_1, \ldots, e_{k(1)}],\ \ldots,\ H_n = [e_{k(n-1)+1}, \ldots, e_{k(n)}],\ \ldots$$

for the eigenvalues $\alpha_n = \lambda_{k(n-1)+1} = \ldots = \lambda_{k(n)}$ of A.

Show that we can write

$$A = \sum_{n=1}^{\infty} \alpha_n P_n$$

and prove that

$$E(B) = \sum_{\alpha_n \in B} P_n$$

is the resolution of the identity of the spectral resolution of A.

Exercise 8.23. Let μ be a Borel measure on a compact set $K \subset \mathbf{C}$ and let $H = L^2(\mu)$. Show that multiplication by characteristic functions of Borel sets in K, $E(B) := \chi_B \cdot$, is a spectral measure $E : \mathcal{B}_K \to L^2(\mu)$.

Exercise 8.24. If $E : \mathcal{B}_K \to \mathcal{L}(H)$ is a spectral measure, show that the null sets for the spectral measure have the following desirable properties:

(a) If $E(B_n) = 0$ $(n \in \mathbf{N})$, then $E(\bigcup_{n=1}^{\infty} B_n) = 0$.

(b) If $E(B_1) = 0$ and $B_2 \subset B_1$, then $E(B_2) = 0$.

Exercise 8.25. Show that the equivalences of Theorem 8.25 are untrue on the real Hilbert space $\mathbf{R}^2$.

Exercise 8.26. Show that every positive $T \in \mathcal{L}(H)$ in the sense of Theorem 8.25 has a unique positive square root.

Exercise 8.27. With the functional calculus, prove also that, if $T \in \mathcal{L}(H)$ is normal, then it can be written as

$$T = UP$$

with U unitary and P positive. This is the **polar decomposition** of a bounded normal operator in a complex Hilbert space.

References for further reading:

I. M. Gelfand, D. A. Raikov and G. E. Chilov, *Commutative Normed Rings.*

E. Hille and R. S. Phillips, *Functional Analysis and Semigroups.*

T. Kato, *Perturbation Theory for Linear Operators.*

P. D. Lax, *Functional Analysis.*

M. A. Naimark, *Normed Rings.*

M. Reed and B. Simon, *Methods of Modern Mathematical Physics.*

C. E. Rickart, *General Theory of Banach Algebras.*

F. Riesz and B. Sz. Nagy, *Leçons d'analyse fonctionelle.*

W. Rudin, *Functional Analysis.*

A. E. Taylor and D. C. Lay, *Introduction to Functional Analysis.*

K. Yosida, *Functional Analysis.*

Chapter 9

Unbounded operators in a Hilbert space

Up to this moment all of our linear operators have been bounded, but densely defined unbounded operators also occur naturally in connection with the foundations of quantum mechanics.

When in 1927 J. von Neumann[1] introduced axiomatically Hilbert spaces, he recognized the need to extend the spectral theory of self-adjoint operators from the bounded to the unbounded case and immediately started to obtain this extension, which was necessary for his presentation of the transformation theory of quantum mechanics created in 1925–1926 by Heisenberg and Schrödinger.[2]

The definition of unbounded self-adjoint operators on a Hilbert space requires a precise selection of the domain, the symmetry condition $(x, Ax)_H = (Ax, x)_H$ for a densely defined operator not being sufficient for A to be self-adjoint, since its spectrum has to be a subset of $\mathbf{R}$. The creators of quantum

[1] The Hungarian mathematician János (John) von Neumann is considered one of the foremost mathematicians of the 20th century: he was a pioneer of the application of operator theory to quantum mechanics, a member of the Manhattan Project, and a key figure in the development of game theory and of the concepts of cellular automata. Between 1926 and 1930 he taught in the University of Berlin. In 1930 he emigrated to the USA where he was invited to Princeton University and was one of the first four people selected for the faculty of the Institute for Advanced Study (1933–1957).

[2] The German physicist Werner Karl Heisenberg, in Göttingen, was one of the founders of quantum mechanics and the head of the German nuclear energy project; with Max Born and Pascual Jordan, Heisenberg formalized quantum mechanics in 1925 using matrix transformations. The Austrian physicist Erwin Rudolf Josef Alexander Schrödinger, while in Zurich, in 1926 derived what is now known as the Schrödinger wave equation, which is the basis of his development of quantum mechanics. Based on the Born statistical interpretation of quantum theory, P. Dirac and Jordan unified "matrix mechanics" and "wave mechanics" with their "transformation theory".

mechanics did not care about this and it was von Neumann himself who clarified the difference between a self-adjoint operator and a symmetric one.

In this chapter, with the Laplacian as a reference example, we include the Rellich theorem, showing that certain perturbations of self-adjoint operators are still self-adjoint, and the Friedrichs method of constructing a self-adjoint extension of many symmetric operators.

Then the spectral theory of bounded self-adjoint operators on a Hilbert space is extended to the unbounded case by means of the Cayley transform, which changes a self-adjoint operator T into a unitary operator U. The functional calculus of this operator allows us to define the spectral resolution of T.

We include a very short introduction on the principles of quantum mechanics, where an observable, such as position, momentum, and energy, is an unbounded self-adjoint operator, their eigenvalues are the observable values, and the spectral representing measure allows us to evaluate the observable in a given state in terms of the probability of belonging to a given set.[3]

Von Neumann's text "Mathematical Foundations of Quantum Mechanics" [**?**] is strongly recommended here for further reading: special attention is placed on motivation, detailed calculations and examples are given, and the thought processes of a great mathematician appear in a very transparent manner. More modern texts are available, but von Neumann's presentation contains in a lucid and very readable way the germ of his ideas on the subject.

In that book, for the first time most of the modern theory of Hilbert spaces is defined and elaborated, as well as "quantum mechanics in a unified representation which ... is mathematically correct". The author explains that, just as Newton mechanics was associated with infinitesimal calculus, quantum mechanics relies on the Hilbert theory of operators.

With von Neumann's work, quantum mechanics is Hilbert space analysis and, conversely, much of Hilbert space analysis is quantum mechanics.

9.1. Definitions and basic properties

Let H denote a complex linear space. We say that T is an operator on H if it is a linear mapping $T : \mathcal{D}(T) \to H$, defined on a linear subspace $\mathcal{D}(T)$ of H, which is called the **domain** of the operator.

[3]Surprisingly, in this way the atomic spectrum appears as Hilbert's spectrum of an operator. Hilbert himself was extremely surprised to learn that his spectrum could be interpreted as an atomic spectrum in quantum mechanics.

Example 9.1. The derivative operator $D : f \mapsto f'$ (distributional derivative) on $L^2(\mathbf{R})$ has

$$\mathcal{D}(D) = \Big\{ f \in L^2(\mathbf{R});\ f' \in L^2(\mathbf{R}) \Big\},$$

as its domain, which is the Sobolev space $H^1(\mathbf{R})$. This domain is dense in $L^2(\mathbf{R})$, since it contains $\mathcal{D}(\mathbf{R})$.

Example 9.2. As an operator on $L^2(\mathbf{R})$, the domain of **the position operator**, $Q : f(x) \mapsto xf(x)$, is

$$\mathcal{D}(Q) = \Big\{ f \in L^2(\mathbf{R});\ [xf(x)] \in L^2(\mathbf{R}) \Big\}.$$

It is unbounded, since $\|\chi_{(n,n+1)}\|_2 = 1$ and $\|Q\chi_{(n,n+1)}\|_2 \geq n$.

Recall that $f \in L^2(\mathbf{R})$ if and only if $\widehat{f} \in L^2(\mathbf{R})$ and both f and $x\widehat{f}(x)$ are in $L^2(\mathbf{R})$ if and only if $f, f' \in L^2(\mathbf{R})$. Thus, the Fourier transform is a unitary operator which maps $\mathcal{D}(Q)$ onto $H^1(\mathbf{R}) = \mathcal{D}(D)$ and changes $2\pi i Q$ into D. Conversely, $2\pi i Q = \mathcal{F}^{-1} D \mathcal{F}$ on $\mathcal{D}(Q)$.

Under these conditions it is said that $2\pi i D$ and Q are **unitarily equivalent**. Unitarily equivalent operators have the same spectral properties.

Of course, it follows that D is also unbounded (see Exercise 9.3).

We are interested in the **spectrum** of T. If for a complex number λ the operator $T - \lambda I : \mathcal{D}(T) \to H$ is bijective and $(T - \lambda I)^{-1} : H \to \mathcal{D}(T) \subset H$ is continuous, then we say that λ is a **regular point** for T.

The spectrum $\sigma(T)$ is the subset of $\mathbf{C}$ which consists of all nonregular points, that is, all complex numbers λ for which $T - \lambda I : \mathcal{D}(T) \to H$ does not have a continuous inverse. Thus $\lambda \in \sigma(T)$ when it is in one of the following disjoint sets:

(a) The **point spectrum** $\sigma_p(T)$, which is the set of the eigenvalues of T. That is, $\lambda \in \sigma_p(T)$ when $T - \lambda I : \mathcal{D}(T) \to H$ is not injective. In this case $(T - \lambda I)^{-1}$ does not exist.

(b) The **continuous spectrum** $\sigma_c(T)$, the set of all $\lambda \in \mathbf{C} \setminus \sigma_p(T)$ such that $T - \lambda I : \mathcal{D}(T) \to H$ is not exhaustive but $\overline{\operatorname{Im}(T - \lambda I)} = H$ and $(T - \lambda I)^{-1}$ is unbounded.

(c) The **residual spectrum** $\sigma_r(T)$, which consists of all $\lambda \in \mathbf{C} \setminus \sigma_p(T)$ such that $\overline{\operatorname{Im}(T - \lambda I)} \neq H$. Then $(T - \lambda I)^{-1}$ exists but is not densely defined.

The set $\sigma(T)^c$ of all regular points is called the **resolvent set**. Thus, $\lambda \in \sigma(T)^c$ when we have $(T - \lambda I)^{-1} \in \mathcal{L}(H)$.

The **resolvent** of T is again the function

$$R_T : \sigma(T)^c \to \mathcal{L}(H), \quad R_T(\lambda) := (T - \lambda I)^{-1}.$$

The spectrum of T is not necessarily a bounded subset of $\mathbf{C}$, but it is still closed and the resolvent function is analytic:

Theorem 9.3. *The set $\sigma(T)^c$ is an open subset of $\mathbf{C}$, and every point $\lambda_0 \in \sigma(T)^c$ has a neighborhood where*

$$R_T(\lambda) = -\sum_{k=0}^{\infty} (\lambda - \lambda_0)^k R_T(\lambda_0)^{k+1},$$

the sum of a convergent Neumann series.

Proof. Let us consider $\lambda = \lambda_0 + \mu$ such that $|\mu| < \|R_T(\lambda_0)\|$. The sum of the Neumann series

$$S(\mu) := \sum_{k=0}^{\infty} \mu^k R_T(\lambda_0)^{k+1} \qquad (|\mu| < 1/\|R_T(\lambda_0)\|)$$

will be the bounded inverse of $T - \lambda I$.

The condition $\|\mu R_T(\lambda_0)\| < 1$ ensures that the series is convergent, and it is easily checked that $(T - \lambda I)S(\mu) = I$:

$$(T - \lambda_0 I - \mu I) \sum_{k=0}^{N} \mu^k ((T - \lambda_0 I)^{-1})^{k+1} = I - (\mu R_T(\lambda_0))^{N+1} \to I,$$

and $S(\mu)$ commutes with T. $\square$

A **graph** is a linear subspace $F \subset H \times H$ such that, for every $x \in H$, the section $F_x := \{y;\ (x, y) \in F\}$ has at most one point, y, so that the first projection $\pi_1(x, y) = x$ is one-to-one on F. This means that $x \mapsto y$ ($y \in F_x$) is an operator T_F on H with $\mathcal{D}(T_F) = \{x \in H;\ F_x \neq \emptyset\}$ and $\mathcal{G}(T_F) = F$.

We write $S \subset T$ if the operator T is an extension of another operator S, that is, if $\mathcal{D}(S) \subset \mathcal{D}(T)$ and $T_{|\mathcal{D}(S)} = S$ or, equivalently, if $\mathcal{G}(S) \subset \mathcal{G}(T)$.

If $\mathcal{G}(T)$ is closed in $H \times H$, then we say that T is a **closed operator**. Also, T is said to be **closable** if it has a closed extension $\widetilde{T}$. This means that $\overline{\mathcal{G}(T)}$ is a graph, since, if $\bar{T}$ is a closed extension of T, $\overline{\mathcal{G}(T)} \subset \mathcal{G}(\widetilde{T})$ and ψ_1 is one-to-one on $\mathcal{G}(\bar{T})$, so that it is also one-to-one on $\overline{\mathcal{G}(T)}$. Conversely, if $\overline{\mathcal{G}(T)}$ is a graph, it is the graph of a closed extension of T, since $\mathcal{G}(T) \subset \overline{\mathcal{G}(T)}$.

If T is closable, then $\bar{T}$ will denote the closure of T; that is, $\bar{T} = T_{\overline{\mathcal{G}(T)}}$.

When defining operations with unbounded operators, the domains of the new operators are the intersections of the domains of the terms. Hence

$$\mathcal{D}(S \pm T) = \mathcal{D}(S) \cap \mathcal{D}(T) \ \text{ and } \ \mathcal{D}(ST) = \Big\{x \in \mathcal{D}(T);\ Tx \in \mathcal{D}(S)\Big\}.$$

Example 9.4. The domain of the commutator $[D,Q] = DQ - QD$ of the derivation operator with the position operator on $L^2(\mathbf{R})$ is $\mathcal{D}(DQ) \cap \mathcal{D}(QD)$, which contains $\mathcal{D}(\mathbf{R})$, a dense subspace of $L^2(\mathbf{R})$.

Since $D(xf(x)) - xDf(x) = f(x)$, the commutator $[D,Q]$ coincides with the identity operator on its domain, so that we simply write $[D,Q] = I$ and consider it as an operator on $L^2(\mathbf{R})$.

9.1.1. The adjoint. We will only be interested in **densely defined** operators, which are the operators T such that $\overline{\mathcal{D}(T)} = H$.

If T is densely defined, then every bounded linear form on $\mathcal{D}(T)$ has a unique extension to H, and from the Riesz representation Theorem **??** we know that it is of the type $(\cdot, z)_H$. This fact allows us to define the **adjoint** T^* of T. Its domain is defined as

$$\mathcal{D}(T^*) = \{y \in H;\, x \mapsto (Tx,y)_H \text{ is bounded on } \mathcal{D}(T)\}$$

and, if $y \in \mathcal{D}(T^*)$, $T^*y \in H$ is the unique element such that

$$(Tx,y)_H = (x, T^*y)_H \qquad (x \in \mathcal{D}(T)).$$

Hence, $y \in \mathcal{D}(T^*)$ if and only if $(Tx,y)_H = (x,y^*)_H$ for some $y^* \in H$, for all $x \in \mathcal{D}(T)$, and then $y^* = T^*y$.

Theorem 9.5. *Let T be densely defined. Then the following properties hold:*

(a) $(\lambda T)^* = \bar{\lambda} T^*$.

(b) $(I+T)^* = I + T^*$.

(c) *T^* is closed.*

(d) *If $T : \mathcal{D}(T) \to H$ is one-to-one with dense image, then T^* is also one-to-one and densely defined, and $(T^{-1})^* = (T^*)^{-1}$.*

Proof. Both (a) and (b) are easy exercises.

To show that the graph of T^* is closed, suppose that $(y_n, T^*y_n) \to (y,z)$ $(y_n \in \mathcal{D}(T^*))$. Then $(x, T^*y_n)_H \to (x,z)_H$ and $(Tx, y_n)_H \to (Tx,y)_H$ for every $x \in \mathcal{D}(T)$, with $(x, T^*y_n)_H = (Tx, y_n)_H$. Hence $(x,z)_H = (Tx,y)_H$ and $z = T^*y$, so that $(y,z) \in \mathcal{G}(T^*)$.

In (d) the inverse $T^{-1} : \operatorname{Im} T \to \mathcal{D}(T)$ is a well-defined operator with dense domain and image. We need to prove that $(T^*)^{-1}$ exists and coincides with $(T^{-1})^*$.

First note that $T^*y \in \mathcal{D}((T^{-1})^*)$ for every $y \in \mathcal{D}(T^*)$, since the linear form $x \mapsto (T^{-1}x, T^*y)_H = (x,y)_H$ on $\mathcal{D}(T^{-1})$ is bounded and T^*y is well-defined. Moreover $(T^{-1})^*T^*y = y$, so that $(T^{-1})^*T^* = I$ on $\mathcal{D}(T^*)$, $(T^*)^{-1} : \operatorname{Im} T^* \to \mathcal{D}(T^*)$, and

$$(T^*)^{-1} \subset (T^{-1})^*$$

since, for $y = (T^*)^{-1}z$ in $(T^{-1})^*T^*y = y$, we have $(T^{-1})^*z = (T^*)^{-1}z$.

To also prove that $(T^{-1})^* \subset (T^*)^{-1}$, let $x \in \mathcal{D}(T)$ and $y \in \mathcal{D}((T^*)^{-1})$. Then $Tx \in \operatorname{Im}(T) = \mathcal{D}(T^{-1})$ and

$$(Tx, (T^{-1})^* y)_H = (x, y)_H, \quad (Tx, (T^{-1})^* y)_H = (x, T^*(T^{-1})^* y)_H.$$

Thus, $(T^{-1})^* y \in \mathcal{D}(T^*)$ and $T^*(T^{-1})^* y = y$, so that $T^*(T^{-1})^* = I$ on $\mathcal{D}((T^*)^{-1}) = \operatorname{Im} T^*$, and $(T^*)^{-1} : \operatorname{Im}(T^*) \to \mathcal{D}(T^*)$ is bijective. □

It is useful to consider the "rotation operator" $G : H \times H \to H \times H$, such that $G(x, y) = (-y, x)$. It is an isometric isomorphism with respect to the norm $\|(x, y)\| := (\|x\|_H^2 + \|y\|_H^2)^{1/2}$ associated to the scalar product

$$((x, y), (x', y'))_{H\times H} := (x, x')_H + (y, y')_H,$$

which makes $H \times H$ a Hilbert space. Observe that $G^2 = -I$.

Theorem 9.6. *If T is closed and densely defined, then*

$$H \times H = G(\mathcal{G}(T)) \oplus \mathcal{G}(T^*) = \mathcal{G}(T) \oplus G(\mathcal{G}(T^*)),$$

orthogonal direct sums, T^* is also closed and densely defined, and $T^{**} = T$.

Proof. Let us first prove that $\mathcal{G}(T^*) = G(\mathcal{G}(T))^\perp$, showing the first equality, and that T^* is closed. Since $(y, z) \in \mathcal{G}(T^*)$ if and only if $(Tx, y)_H = (x, z)_H$ for every $x \in \mathcal{D}(T)$, we have

$$(G(x, Tx), (y, z))_{H\times H} = ((-Tx, x), (y, z))_{H\times H} = 0,$$

and this means that $(y, z) \in G(\mathcal{G}(T))^\perp$, so that $\mathcal{G}(T^*) = G(\mathcal{G}(T))^\perp$.

Also, since $G^2 = -I$,

$$H \times H = G(G(\mathcal{G}(T)) \oplus \mathcal{G}(T^*)) = \mathcal{G}(T) \oplus G(\mathcal{G}(T^*)).$$

If $(z, y)_H = 0$ for all $y \in \mathcal{D}(T^*)$, then $((0, z), (-T^* y, y))_{H\times H} = 0$. Hence, $(0, z) \in G(\mathcal{G}(T^*))^\perp = \mathcal{G}(T)$ and it follows that $z = T0 = 0$. Thus, $\mathcal{D}(T^*)$ is dense in H.

Finally, since also $H \times H = G(\mathcal{G}(T^*)) \oplus \mathcal{G}(T^{**})$ and $\mathcal{G}(T)$ is the orthogonal complement of $G(\mathcal{G}(T^*))$, we obtain the identity $T = T^{**}$. □

9.2. Unbounded self-adjoint operators

$T : \mathcal{D}(T) \subset H \to H$ is still a possibly unbounded linear operator on the complex Hilbert space H.

9.2.1. Self-adjoint operators. We say that the operator T is **symmetric** if it is densely defined and

$$(Tx, y)_H = (x, Ty)_H \qquad (x, y \in \mathcal{D}(T)).$$

Note that this condition means that $T \subset T^*$.

Theorem 9.7. *Every symmetric operator T is closable and its closure is T^{**}.*

Proof. Since T is symmetric, $T \subset T^*$ and, $\mathcal{G}(T^*)$ being closed,

$$\mathcal{G}(T) \subset \overline{\mathcal{G}(T)} \subset \mathcal{G}(T^*).$$

Hence $\overline{\mathcal{G}(T)}$ is a graph, T is closable, and $\overline{\mathcal{G}(T)}$ is the graph of $\bar{T}$. As a consequence, let us show that the domain of T^* is dense.

According to Theorem 9.6, $(x, y) \in \mathcal{G}(T^*)$ if and only if $(-y, x) \in \mathcal{G}(T)^\perp$, in $H \times H$. Hence,

$$\overline{\mathcal{G}(T)} = \mathcal{G}(T)^{\perp\perp} = \{(T^*x, -x);\, x \in \mathcal{D}(T^*)\}^\perp.$$

This subspace is not a graph if and only if $(y, z_1), (y, z_2) \in \overline{\mathcal{G}(T)}$ for two different points $z_1, z_2 \in H$; that is, $(0, z) \in \{(T^*x, -x);\, x \in \mathcal{D}(T^*)\}^\perp$ for some $z \neq 0$. Then $(z, x)_H = 0$ for all $x \in \mathcal{D}(T^*)$ which means that $0 \neq z \in \mathcal{D}(T^*)^\perp$, and it follows that $\overline{\mathcal{D}(T^*)} \neq H$.

Since $\overline{\mathcal{D}(T^*)} = H$, T^{**} is well-defined. We need to prove that

$$\overline{\mathcal{G}(T)} = \mathcal{G}(\bar{T}) = \{(T^*x, -x);\, x \in \mathcal{D}(T^*)\}^\perp$$

is $\mathcal{G}(T^{**})$. But $(v, u) \in \overline{\mathcal{G}(T)}$ if and only if $(T^*x, v)_H - (x, u)_H = 0$ for all $x \in \mathcal{D}(T^*)$; that is, $v \in \mathcal{D}(T^{**})$ and $u = T^{**}v$, which means that $(v, u) \in \mathcal{G}(T^{**})$. □

The operator T is called **self-adjoint** if it is densely defined and $T = T^*$, i.e., if it is symmetric and

$$\mathcal{D}(T^*) \subset \mathcal{D}(T),$$

this inclusion meaning that the existence of $y^* \in H$ such that $(Tx, y)_H = (x, y^*)_H$ for all $x \in \mathcal{D}(T)$ implies $y^* = Tx$.

Theorem 9.8. *If T is self-adjoint and S is a symmetric extension of T, then $S = T$. Hence T does not have any strict symmetric extension; it is "maximally symmetric".*

Proof. It is clear that $T = T^* \subset S$ and $S \subset S^*$, since S is symmetric. It follows from the definition of a self-adjoint operator that $T \subset S$ implies $S^* \subset T^*$. From $S \subset S^* \subset T \subset S$ we obtain the identity $S = T$. □

We are going to show that, in the unbounded case, the spectrum of a self-adjoint operator is also real. This property characterizes the closed symmetric operators that are self-adjoint.

First note that, if $T = T^*$, the point spectrum is real, since if $Tx = \lambda x$ and $0 \neq x \in \mathcal{D}(T)$, then

$$\bar{\lambda}(x,x)_H = (x,Tx)_H = (Tx,x)_H = \lambda(x,x)_H$$

and $\bar{\lambda} = \lambda$.

Theorem 9.9. *Suppose that T is self-adjoint. The following properties hold:*

(a) $\lambda \in \sigma(T)^c$ *if and only if* $\|Tx - \lambda x\|_H \geq c\|x\|_H$ *for all* $x \in \mathcal{D}(T)$, *for some constant* $c > 0$.

(b) *The spectrum* $\sigma(T)$ *is real and closed.*

(c) $\lambda \in \sigma(T)$ *if and only if* $Tx_n - \lambda x_n \to 0$ *for some sequence* $\{x_n\}$ *in* $\mathcal{D}(T)$ *such that* $\|x_n\|_H = 1$ (λ *is an* **approximate eigenvalue**).

(d) *The inequality* $\|R_T(\lambda)\| \leq 1/|\Im\lambda|$ *holds.*

Proof. (a) If $\lambda \in \sigma(T)^c$, then $R_T(\lambda) \in \mathcal{L}(H)$ and

$$\|x\|_H \leq \|R_T(\lambda)\|\|(T-\lambda I)x\|_H = c^{-1}\|(T-\lambda I)x\|_H.$$

Suppose now that $\|Tx - \lambda x\|_H \geq c\|x\|_H$ and let $M = \operatorname{Im}(T - \lambda I)$, so that we have $T - \lambda I : \mathcal{D}(T) \to M$ with continuous inverse. To prove that $M = H$, let us first show that M is dense in H.

If $z \in M^\perp$, then for every $Tx - \lambda x \in M$ we have

$$0 = (Tx - \lambda x, z)_H = (Tx,z)_H - \lambda(x,z)_H.$$

Hence $(Tx,z)_H = (x, \bar{\lambda}z)_H$ if $x \in \mathcal{D}(T)$, and then $z \in \mathcal{D}(T^*) = \mathcal{D}(T)$ and $Tz = \bar{\lambda}z$. Suppose $z \neq 0$, so that $\bar{\lambda} = \lambda$ and we arrive at $Tz - \lambda z = 0$ and $0 \neq z \in M$, a contradiction. Thus, $M^\perp = 0$ and M is dense.

To prove that M is closed in H, let $M \ni y_n = Tx_n - \lambda x_n \to y$. Then $\|x_p - x_q\| \leq c^{-1}\|y_p - y_q\|_H$, and there exist $x = \lim x_n \in H$ and $\lim_n Tx_n = y + \lambda x$. But T is closed, so that $Tx = y + \lambda x$ and $y \in M$.

(b) To show that every $\lambda = \alpha + i\beta \in \sigma(T)$ is real, observe that, if $x \in \mathcal{D}(T)$,

$$(Tx - \lambda x, x)_H = (Tx,x)_H - \lambda(x,x)_H, \quad \overline{(Tx - \lambda x, x)}_H = (Tx,x)_H - \bar{\lambda}(x,x)_H,$$

since $(Tx,x)_H \in \mathbf{R}$. Subtracting,

$$\overline{(Tx - \lambda x, x)}_H - (Tx - \lambda x, x)_H = 2i\beta\|x\|_H^2,$$

where $\overline{(Tx - \lambda x, x)}_H - (Tx - \lambda x, x)_H = -2i\operatorname{Im}(Tx - \lambda x, x)_H$. Hence,

$$|\beta|\|x\|_H^2 = |\operatorname{Im}(Tx - \lambda x)_H| \leq |(Tx - \lambda x, x)_H| \leq \|Tx - \lambda x\|_H\|x\|_H$$

and then $|\beta|\|x\|_H \le \|Tx - \lambda x\|_H$ if $x \in \mathcal{D}(T)$. As seen in the proof of (a), the assumption $\beta \neq 0$ would imply $\lambda \in \sigma(T)^c$.

(c) If $\lambda \in \sigma(T)$, the estimate in (a) does not hold and then, for every $c = 1/n$, we can choose $x_n \in \mathcal{D}(T)$ with norm one such that $\|Tx_n - \lambda x_n\|_H \le 1/n$ and λ is an approximate eigenvalue. Every approximate eigenvalue λ is in $\sigma(T)$, since, if $(T - \lambda I)^{-1}$ were bounded on H, then it would follow from $Tx_n - \lambda x_n \to 0$ that $x_n = (T - \lambda I)^{-1}(Tx_n - \lambda x_n) \to 0$, a contradiction to $\|x_n\|_H = 1$.

(d) If $y \in \mathcal{D}(T)$ and $\lambda = \Re\lambda + i\Im\lambda \notin \mathbf{R}$, then it follows that

$$\|(T - \lambda I)y\|_H^2 = (Ty - \lambda y, Ty - \lambda y)_H \ge ((\Im\lambda)y, (\Im\lambda)y)_H = |\Im\lambda|^2\|y\|_H^2.$$

If $x = (T - \lambda I)y \in H$, then $y = R_T(\lambda)x$ and $\|x\|_H^2 \ge |\Im\lambda|^2\|R_T(\lambda)x\|_H^2$; thus $|\Im\lambda|\|R_T(\lambda)\| \le 1$. □

The condition $\sigma(T) \subset \mathbf{R}$ is sufficient for a symmetric operator to be self-adjoint. In fact we have more:

Theorem 9.10. *Suppose that T is symmetric. If there exists $z \in \mathbf{C} \setminus \mathbf{R}$ such that $z, \bar{z} \in \sigma(T)^c$, then T is self-adjoint.*

Proof. Let us first show that $((T - zI)^{-1})^* = (T - \bar{z}I)^{-1}$, that is,

$$((T - zI)^{-1}x_1, x_2)_H = (x_1, (T - \bar{z}I)^{-1}x_2)_H.$$

We denote $(T - zI)^{-1}x_1 = y_1$ and $(T - \bar{z}I)^{-1}x_2 = y_2$. The desired identity means that $(y_1, (T - \bar{z}I)y_2)_H = ((T - zI)y_1, y_2)_H$ and it is true if $y_1, y_2 \in \mathcal{D}(T)$, since T is symmetric. But the images of $T - zI$ and $T - \bar{z}I$ are both the whole space H, so that $((T - zI)^{-1}x_1, x_2)_H = (x_1, (T - \bar{z}I)^{-1}x_2)_H$ holds for any $x_1, x_2 \in H$.

Now we can prove that $\mathcal{D}(T^*) \subset \mathcal{D}(T)$. Let $v \in \mathcal{D}(T^*)$ and $w = T^*v$, i.e.,

$$(Ty_1, v)_H = (y_1, w)_H \qquad (\forall y_1 \in \mathcal{D}(T)).$$

We subtract $z(y_1, v)_H$ to obtain

$$((T - zI)y_1, v)_H = (y_1, w - \bar{z}v)_H.$$

Still with the notation $(T - zI)^{-1}x_1 = y_1$ and $(T - \bar{z}I)^{-1}x_2 = y_2$, but now with $x_2 = w - \bar{z}v$, since $(x_1, v)_H = ((T - zI)y_1, v)_H = (y_1, w - \bar{z}v)_H$,

$$(x_1, v)_H = \big((T - zI)^{-1}x_1, w - \bar{z}v\big)_H = \big(x_1, (T - \bar{z}I)^{-1}(w - \bar{z}v)\big)_H \quad (\forall x_1 \in H).$$

Thus, $v = (T - \bar{z}I)^{-1}(w - \bar{z}v)$ and $v \in \operatorname{Im}(T - \bar{z})^{-1} = \mathcal{D}(A)$. □

In the preceding proof, we have only needed the existence of $z \notin \mathbf{R}$ such that $(T - zI)^{-1}$ and $(T - \bar{z}I)^{-1}$ are defined on H, but not their continuity.

Example 9.11. The position operator Q of Example 9.2 is self-adjoint and $\sigma(Q) = \mathbf{R}$, but it does not have an eigenvalue.

Obviously Q is symmetric, since, if $f, g, [xf(x)], [xg(x)] \in L^2(\mathbf{R})$, then

$$(Qf, g)_2 = \int_{\mathbf{R}} xf(x)\overline{g(x)}\, dx = \int_{\mathbf{R}} f(x)\overline{xg(x)}\, dx = (f, Qg)_2.$$

Let us show that $\mathcal{D}(Q^*) \subset \mathcal{D}(Q)$. If $g \in \mathcal{D}(Q^*)$, then there exists $g^* \in L^2(\mathbf{R})$ such that $(Qf, g)_2 = (f, g^*)_2$ for all $f \in \mathcal{D}(Q)$. So $\int_{\mathbf{R}} \varphi(x)\overline{xg(x)}\, dx = \int_{\mathbf{R}} \varphi(x)\overline{g^*(x)}\, dx$ if $\varphi \in \mathcal{D}(\mathbf{R})$, and then $g^*(x) = xg(x)$ a.e. on $\mathbf{R}$. I.e., $xg(x)$ is in $L^2(\mathbf{R})$ and $g \in \mathcal{D}(Q)$. Hence Q is self-adjoint, since it is symmetric and $\mathcal{D}(Q^*) \subset \mathcal{D}(Q)$.

If $\lambda \notin \sigma(Q)$, then $T = (Q - \lambda I)^{-1} \in \mathcal{L}(H)$ and, for every $g \in L^2(\mathbf{R})$, the equality $(Q - \lambda I)Tg = g$ implies that $(Tg)(x) = g(x)/(x - \lambda) \in L^2(\mathbf{R})$ and $\lambda \notin \mathbf{R}$, since, when $\lambda \in \mathbf{R}$ and $g := \chi_{(\lambda,b)}$, $g(x)/(x - \lambda) \notin L^2(\mathbf{R})$.

Example 9.12. The adjoint of the derivative operator D of Example 9.1 is $-D$, and iD is self-adjoint. The spectrum of iD is also $\mathbf{R}$ and it has no eigenvalues.

By means of the Fourier transform we can transfer the properties of Q. If $f, g \in H^1(\mathbf{R})$, then $(Df, g)_2 = (\widehat{Df}, \widehat{g})_2$. From $(Q\widehat{f}, \widehat{g})_2 = (\widehat{f}, Q\widehat{g})_2$ and $\widehat{Df}(x) = 2\pi i t\widehat{f}(x) = 2\pi i(Q\widehat{f})(x)$ we obtain

$$(Df, g)_2 = (2\pi i Q\widehat{f}, \widehat{g})_2 = (\widehat{f}, -2\pi i Q\widehat{g})_2 = (f, -Dg)_2$$

and $-D \subset D^*$. As in the case of Q, also $\mathcal{D}(D^*) \subset \mathcal{D}(D)$.

Furthermore, $(iD)^* = -iD^* = iD$ and $\sigma(iD) \subset \mathbf{R}$. If $T = (iD - \lambda I)^{-1}$ and $g \in L^2$, then an application of the Fourier transform to $(iD - \lambda I)Tg = g$ gives $-2\pi x\widehat{Tg}(x) - \lambda\widehat{Tg}(x) = \widehat{g}(x)$, and

$$\widehat{Tg}(x) = -\frac{\widehat{g}(x)}{2\pi x + \lambda}$$

has to lie in $L^2(\mathbf{R})$. So we arrive to $\lambda \notin \mathbf{R}$ by taking convenient functions $\widehat{g} = \chi_{(a,b)}$.

Example 9.13. The Laplace operator Δ of $L^2(\mathbf{R}^n)$ with domain $H^2(\mathbf{R}^n)$ is self-adjoint. Its spectrum is $\sigma(\Delta) = [0, \infty)$.

Recall that

$$\begin{aligned} H^2(\mathbf{R}^n) &= \Big\{u \in L^2(\mathbf{R}^n);\ D^\alpha u \in L^2(\mathbf{R}^n),\ |\alpha| \le 2\Big\} \\ &= \Big\{u \in L^2(\mathbf{R}^n);\ \int_{\mathbf{R}^n} |(1 + |\xi|^2)\widehat{u}(\xi)|^2\, d\xi < \infty\Big\} \end{aligned}$$

and Δ, with this domain, is symmetric: $(\Delta u, v)_2 = (u, \Delta v)_2$ follows from the Fourier transforms, since

$$\int_{\mathbf{R}^n} \widehat{u}(\xi)|\xi|^2\overline{\widehat{v}(\xi)}\, d\xi = \int_{\mathbf{R}^n} \widehat{u}(\xi)\overline{|\xi|^2\widehat{v}(\xi)}\, d\xi.$$

To prove that it is self-adjoint, let $u \in \mathcal{D}(\Delta^*) \subset L^2(\mathbf{R}^n)$. If $w \in L^2(\mathbf{R}^n)$ is such that

$$(\Delta v, u)_2 = (v, w)_2 \qquad (v \in H^2(\mathbf{R}^n)),$$

then, up to a nonzero multiplicative constant,

$$\int_{\mathbf{R}^n} \widehat{v}(\xi)|\xi|^2\overline{\widehat{u}(\xi)}\, d\xi = \int_{\mathbf{R}^n} \widehat{v}(\xi)\overline{\widehat{w}(\xi)}\, d\xi$$

for every $\widehat{v} \in H^2(\mathbf{R}^n)$, a dense subspace of $L^2(\mathbf{R}^n)$, and $|\xi|^2\widehat{u}(\xi) = c\widehat{w}(\xi)$, in $L^2(\mathbf{R}^n)$. Hence, $\int_{\mathbf{R}^n} |(1+|\xi|^2)\widehat{u}(\xi)|^2\, d\xi < \infty$ and $u \in H^1(\mathbf{R}^n) = \mathcal{D}(\Delta)$. Thus, $\mathcal{D}(\Delta^*) \subset \mathcal{D}(\Delta)$.

The Fourier transform, $\mathcal{F}$, is a unitary operator of $L^2(\mathbf{R}^n)$, so that the spectrum of Δ is the same as the spectrum of the multiplication operator $\mathcal{F}\Delta\mathcal{F}^{-1} = 4\pi^2|\xi|^2\cdot$, which is self-adjoint with domain

$$\Big\{f \in L^2(\mathbf{R}^n);\ \int_{\mathbf{R}^n} |(1+|\xi|^2)f(\xi)|^2\, d\xi < \infty\Big\},$$

and $\lambda \in \sigma(\mathcal{F}\Delta\mathcal{F}^{-1})^c$ if and only if the multiplication by $4\pi^2|\xi|^2 - \lambda$ has a continuous inverse on $L^2(\mathbf{R}^n)$, the multiplication by $1/(4\pi^2|\xi|^2 - \lambda)$. This means that $\lambda \neq 4\pi^2|\xi|^2$ for every $\xi \in \mathbf{R}^n$, i.e., $\lambda \notin [0, \infty)$.

An application of Theorem 9.10 shows that a perturbation of a self-adjoint operator with a "small" symmetric operator, is still self-adjoint. For a more precise statement of this fact, let us say that an operator S is **relatively bounded**, with constant α, with respect to another operator A if $\mathcal{D}(A) \subset \mathcal{D}(S)$ and there are two constants $\alpha, c \geq 0$ such that

$$\|Sx\|_H^2 \leq \alpha^2\|Ax\|_H^2 + c^2\|x\|_H^2 \qquad (x \in \mathcal{D}). \tag{9.1}$$

Let us check that this kind of estimate is equivalent to

$$\|Sx\|_H \leq \alpha'\|Ax\|_H + c'\|x\|_H \qquad (x \in \mathcal{D}) \tag{9.2}$$

and that we can take $\alpha' < 1$ if $\alpha < 1$ and $\alpha < 1$ if $\alpha' < 1$.

By completing the square, it is clear that (9.2) follows from (9.1) with $\alpha = \alpha'$ and $c = c'$. Also, from (9.2) we obtain (9.1) with $\alpha^2 = (1+\varepsilon^{-1})\alpha'^2$ and $c^2 = (1+\varepsilon)c'^2$, for any $\varepsilon > 0$, since $2\alpha'\|Ax\|_H c'\|x\|_H \leq \varepsilon^{-1}\alpha'^2\|Ax\|_H^2 + \varepsilon c'^2\|x\|_H^2$ and an easy substitution shows that

$$(\alpha'\|Ax\|_H + c'\|x\|_H)^2 \leq \alpha^2\|Ax\|_H^2 + c^2\|x\|_H^2.$$

Theorem 9.14 (Rellich[4]). *Let A be a self-adjoint operator and let S be symmetric, with the same domain $D \subset H$. If S is relatively bounded with constant α with respect to A, then $T = A+S$ is also self-adjoint with domain D.*

Proof. Let us first check that the symmetric operator T is closed. If $(x, y) \in \overline{\mathcal{G}(T)}$, then we choose $x_n \in D$ so that $x_n \to x$ and $Tx_n \to y$. From the hypothesis we have

$$\|Ax_n - Ax_m\|_H \leq \|Tx_n - Tx_m\|_H + \alpha\|Ax_n - Ax_m\|_H + c\|x_n - x_m\|_H,$$

which implies

$$\|Ax_n - Ax_m\|_H \leq \frac{1}{1-\alpha}\|Tx_n - Tx_m\|_H + \frac{c}{1-\alpha}\|x_n - x_m\|_H,$$

and there exists $z = \lim_n Ax_n$. But A is closed and $z = Ax$ with $x \in D$.

Moreover $\|Sx_n - Sx\|_H \leq \alpha\|A(x_n - x)\|_H + c\|x_n - x\|_H$ and $Sx_n \to Sx$. Hence $y = \lim_n Tx_n = Tx$ and $(x, y) \in \mathcal{G}(T)$.

The operators $T - zI$ ($z \in \mathbf{C}$), with domain D, are also closed. With Theorem 9.10 in hand, we only need to check that $\pm\lambda i \in \sigma(T)^c$ when $\lambda \in \mathbf{R}$ is large enough ($|\lambda| > c$).

To show that $T - \lambda i I$ is one-to-one if $\lambda \neq 0$, let

$$(T - \lambda i I)x = y \qquad (x \in D)$$

and note that the absolute values of the imaginary parts of both sides of

$$(Tx, x)_H - \lambda i (x, x)_H = (y, x)_H$$

are equal, so that $|\lambda|\|x\|_H^2 = |\Im(y, x)_H| \leq \|y\|_H\|x\|_H$ and

$$\|x\|_H \leq |\lambda|^{-1}\|y\|_H \qquad (x \in D).$$

Thus, $y = 0$ implies $x = 0$.

Let us prove now that $T - \lambda i I$ has a closed image. Let $y_n \to y$ with $y_n = (T - \lambda i I)x_n$. Then $\|x_n - x_m\|_H \leq |\lambda|^{-1}\|y_n - y_m\|_H$ and the limit $x = \lim x_n \in H$ exists. Since $(T - \lambda i I)x_n \to y$ and the graph of $T - \lambda i I$ is closed, $x \in D$ and $y = (T - \lambda i I)x \in \operatorname{Im}(T - \lambda i I)$.

Let us also show that $\operatorname{Im}(T - \lambda i I) = H$ by proving that the orthogonal is zero. Let $v \in H$ be such that

$$(Ax + Sx - \lambda i x, v)_H = 0 \qquad (x \in D).$$

[4]F. Rellich worked on the foundations of quantum mechanics and on partial differential equations, and his most important contributions, around 1940, refer to the perturbation of the spectrum of self-adjoint operators $A(\epsilon)$ which depend on a parameter ϵ. See also footnote **??** in Chapter 7.

Then $(A-\lambda i)(D)=H$, since $\lambda i \in \sigma(A)^c$. If $(A-\lambda iI)u=v$, let $x=u$, and then we obtain that

$$((A-\lambda i)u,(A-\lambda i)u)_H+(Su,(A-\lambda i)u)_H=0.$$

From the Cauchy-Schwarz inequality, $\|Au-\lambda iu\|_H^2 \le \|Su\|_H\|Au-\lambda iu\|_H$ and

$$\|Au-\lambda iu\|_H \le \|Su\|_H.$$

Since A is symmetric, $(Ay-\lambda iy, Ay-\lambda iy)_H=\|Ay\|_H^2+\lambda^2\|y\|_H^2$ and

$$\|Au\|^2+\lambda^2\|u\|_H^2=\|Au-\lambda iu\|_H^2 \le \|Su\|_H^2 \le \alpha^2\|Au\|_H^2+c^2\|u\|_H^2.$$

But, if $|\lambda|>c$, the condition $\alpha^2<1$ implies $u=0$, and then $v=0$.

We have proved that $(T-\lambda iI)^{-1}: H \to H$ is well-defined and closed, i.e., it is bounded. Hence, $\pm i\lambda \in \sigma(T)^c$ and the symmetric operator T is self-adjoint. □

Example 9.15. The operator $H=-\Delta-|x|^{-1}$ on $L^2(\mathbf{R}^3)$, with domain $H^2(\mathbf{R}^3)$, is self-adjoint.

Let $-|x|^{-1}=V_0(x)+V_1(x)$ with $V_0(x):=\chi_B(x)V(x)$ $(B=\{|x|\le 1\})$. Multiplication by the real function $|x|^{-1}$ is a symmetric operator whose domain contains $H^2(\mathbf{R}^3)$, the domain of $-\Delta$, since $V_0u \in L^2(\mathbf{R}^3)$ if $u \in H^2(\mathbf{R}^3)$, with

$$\|V_0u\|_2 \le \|V_0\|_\infty\|u\|_2=\|u\|_2$$

and $\|V_1u\|_2 \le \|V_1\|_2\|u\|_\infty$, where $\|u\|_\infty \le \|\widehat{u}\|_1$. To apply Theorem 9.14, we will show that multiplication by $-|x|^{-1}$ is relatively bounded with respect to $-\Delta$.

From the Cauchy-Schwarz inequality and from the relationship between the Fourier transform and the derivatives, we obtain

$$\Big(\int_{\mathbf{R}^3}|\widehat{u}(\xi)|\,d\xi\Big)^2 \le \int_{\mathbf{R}^3}\frac{d\xi}{(|\frac{2\pi}{\xi}|^2+\beta^2)^2}\|(-\Delta+\beta^2I)u\|_2^2=\frac{2\pi^3}{\xi}\|(-\Delta+\beta^2I)u\|_2^2.$$

From the inversion theorem we obtain that u is bounded and continuous, since it is the Fourier co-transform of the integrable function $\widehat{u}$. Then

$$\|V_1u\|_2 \le c(\beta^{-1/2}\|-\Delta u\|_2+\beta^{3/2}\|u\|_2) \qquad (u \in H^2(\mathbf{R}^3))$$

so that

$$\|Vu\|_2 \le c\beta^{-1/2}\|-\Delta u\|_2+(c\beta^{3/2}+1)\|u\|_2$$

and $c\beta^{-1/2}<1$ if β is large.

It follows from the Rellich theorem that H is self-adjoint with domain $H^2(\mathbf{R}^3)$.

9.2.2. Essentially self-adjoint operators. Very often, operators appear to be symmetric but they are not self-adjoint, and in order to apply the spectral theory, it will be useful to know whether they have a self-adjoint extension. Recall that a symmetric operator is closable and that a self-adjoint operator is always closed and maximally symmetic.

A symmetric operator is said to be **essentially self-adjoint** if its closure is self-adjoint. In this case, the closure is the unique self-adjoint extension of the operator.

Example 9.16. It follows from Example 9.13 that the Laplacian Δ, as an operator on $L^2(\mathbf{R}^n)$ with domain $\mathcal{S}(\mathbf{R}^n)$, is essentially self-adjoint. Its closure is again Δ, but with domain $H^2(\mathbf{R}^n)$.

Theorem 9.17. *If T is symmetric and a sequence $\{u_n\}_{n\in\mathbf{N}} \subset \mathcal{D}(T)$ is an orthonormal basis of H such that $Tu_n = \lambda_n u_n$ $(n \in \mathbf{N})$, then T is essentially self-adjoint and the spectrum of its self-adjoint extension $\bar{T}$ is $\sigma(\bar{T}) = \overline{\{\lambda_n;\, n \in \mathbf{N}\}}$.*

Proof. The eigenvalues λ_n are all real. Define

$$\mathcal{D}(\bar{T}) := \Big\{ \sum_{n=1}^{\infty} \alpha_n u_n;\ \sum_{n=1}^{\infty} |\alpha_n|^2 + \sum_{n=1}^{\infty} \lambda_n^2 |\alpha_n|^2 < \infty \Big\},$$

a linear subspace of H that contains $\mathcal{D}(T)$, since, if $x = \sum_{n=1}^{\infty} \alpha_n u_n \in \mathcal{D}(T)$ and $Tx = \sum_{n=1}^{\infty} \beta_n u_n \in H$, the Fourier coefficients α_n and β_n satisfy

$$\beta_n = (Tx, u_n)_H = (x, Tu_n)_H = \lambda_n (x, u_n)_H = \lambda_n \alpha_n$$

and $\{\alpha_n\}, \{\lambda_n \alpha_n\} \in \ell^2$.

We can define the operator $\bar{T}$ on $\mathcal{D}(\bar{T})$ by

$$\bar{T}\Big(\sum_{n=1}^{\infty} \alpha_n u_n \Big) := \sum_{n=1}^{\infty} \lambda_n \alpha_n u_n.$$

Let us show that $\bar{T}$ is a self-adjoint extension of T.

It is clear that $\bar{T}$ is symmetric, $T \subset \bar{T}$, and every λ_n is an eigenvalue of $\bar{T}$, so that $\overline{\{\lambda_n;\, n \in \mathbf{N}\}} \subset \sigma(\bar{T})$.

If $\lambda \notin \overline{\{\lambda_n;\, n \in \mathbf{N}\}}$, so that $|\lambda - \lambda_n| \geq \delta > 0$, then it follows that $\lambda \notin \sigma(\bar{T})$ since we can construct the inverse of

$$(T - \lambda I)\Big(\sum_{n=1}^{\infty} \alpha_n u_n \Big) = \sum_{n=1}^{\infty} \alpha_n (\lambda_n - \lambda) u_n$$

by defining

$$R\Big(\sum_{n=1}^{\infty} \alpha_n u_n \Big) := \sum_{n=1}^{\infty} \frac{\alpha_n}{\lambda_n - \lambda} u_n.$$

Indeed, we obtain an operator $R \in \mathcal{L}(H)$ ($\|R\| \leq 1/\delta$) which obviously is one-to-one and its image is $\mathcal{D}(\bar{T})$, since

$$\sum_{n=1}^{\infty} \Big| \frac{\alpha_n}{\lambda - \lambda_n} \Big|^2 \lambda_n^2 \leq |\alpha_n|^2 \Big(1 + \frac{\lambda}{\delta}\Big)^2 < \infty$$

and, moreover, we can associate to every $x = \sum_{n=1}^{\infty} \alpha_n u_n \in \mathcal{D}(\bar{T})$ the element

$$y = \sum_{n=1}^{\infty} \beta_n u_n = \sum_{n=1}^{\infty} \alpha_n (\lambda_n - \lambda) u_n \in H$$

such that $Ry = x$.

Also,

$$(\bar{T} - \lambda I) R \Big(\sum_{n=1}^{\infty} \alpha_n u_n \Big) = \sum_{n=1}^{\infty} - \frac{\alpha_n}{\lambda - \lambda_n} (\bar{T} - \lambda I) u_n = \sum_{n=1}^{\infty} \alpha_n u_n$$

and $R = (\bar{T} - \lambda I)^{-1}$.

To prove that $\bar{T}$ is self-adjoint, we need to see that $\mathcal{D}((\bar{T})^*) \subset \mathcal{D}(\bar{T})$.

If $x = \sum_{n=1}^{\infty} \alpha_n u_n \in \mathcal{D}((\bar{T})^*)$ and $y = (\bar{T})^* x$, then, for every $n \in \mathbf{N}$,

$$(y, u_n)_H = (x, \bar{T} u_n)_H = \lambda_n (x, u_n)_H = \lambda_n \alpha_n$$

and $\sum_{n=1}^{\infty} |\lambda_n \alpha_n|^2 < \infty$, i.e., $x \in \mathcal{D}(T)$.

Finally, to prove that $\bar{T}$ is the closure of T, consider

$$(x, \bar{T} y) = (\sum_{n=1}^{\infty} \alpha_n u_n, \sum_{n=1}^{\infty} \lambda_n \alpha_n u_n) \in \mathcal{G}(\bar{T}).$$

Then $x_N := \sum_{n=1}^{N} \alpha_n u_n \in \mathcal{D}(T)$ and

$$(x_N, T x_N) = (\sum_{n=1}^{N} \alpha_n u_n, \sum_{n=1}^{N} \lambda_n \alpha_n u_n) \to (x, \bar{T} x)$$

in $H \times H$, since $\{\alpha_n\}, \{\lambda_n \alpha_n\} \in \ell^2$. □

Remark 9.18. A symmetric operator T may have no self-adjoint extensions at all, or many self-adjoint extensions. According to Theorems 9.7 and 9.8, if T is essentially self-adjoint, T^{**} is the unique self-adjoint extension of T.

9.2.3. The Friedrichs extensions. A sufficient condition for a symmetric operator T to have self-adjoint extensions, known as the **Friedrichs extensions**, concerns the existence of a lower bound for the quadratic form $(Tx, x)_H$.

We say that T, symmetric, is **semi-bounded**[5] with constant c, if

$$c := \inf_{x\in\mathcal{D}(T),\, \|x\|_H=1} (Tx, x)_H > -\infty,$$

so that $(Tx, x)_H \geq c\|x\|_H^2$ for all $x \in \mathcal{D}(T)$.

In this case, for any $c' \in \mathbf{R}$, $T - c'I$ is also symmetric on the same domain and semi-bounded, with constant $c + c'$. If $\widetilde{T}$ is a self-adjoint extension of T, then $\widetilde{T} - c'I$ is a self-adjoint extension of $T - c'I$, and we will choose a convenient constant in our proofs. Let us denote

$$(x, y)_T := (Tx, y)_H,$$

a sesquilinear form on $\mathcal{D}(T)$ such that $(y, x)_T = \overline{(x, y)_T}$. If $c > 0$, then we have an inner product.

Theorem 9.19 (Friedrichs-Stone[6])**.** *If T is a semi-bounded symmetric operator, with constant c, then it has a self-adjoint extension $\widetilde{T}$ such that $(\widetilde{T}x, x)_H \geq c\|x\|_H^2$, if $x \in \mathcal{D}(\widetilde{T})$.*

Proof. We can suppose that $c = 1$, and then $(x, y)_T$ is a scalar product on $D = \mathcal{D}(T)$ which defines a norm $\|x\|_T = (x, x)_T^{1/2} \geq \|x\|_H$.

Let D_T be the $\|\cdot\|_T$-completion of D. Since $\|x\|_H \leq \|x\|_T$, every $\|\cdot\|_T$-Cauchy sequence $\{x_n\} \subset D$, which represents a point $\widetilde{x} \in D_T$, has a limit x in H, and we have a natural mapping $J : D_T \to H$, such that $J\widetilde{x} = x$.

This mapping J is one-to-one, since, if $Jy = 0$ and $x_n \to y$ in D_T, $\{x_n\} \subset D$ is also a Cauchy sequence in H and there exists $x = \lim x_n$ in H. Then $x = Jy = 0$ and, from the definition of $(y, x)_T$ and by the continuity of the scalar product, it follows that, for every $v \in D$,

$$(v, y)_T = \lim_n (v, x_n)_T = \lim_n (Tv, x_n)_H = (Tv, x)_H = 0.$$

But D is dense in D_T and $y = 0$.

We have $D = \mathcal{D}(T) \subset D_T \hookrightarrow H$ and, to define the Friedrichs extension $\widetilde{T}$ of T, we observe that, for every $u = (\cdot, y)_H \in H'$,

$$|u(x)| \leq \|x\|_H\|y\|_H \leq \|x\|_T\|y\|_H \qquad (x \in D_T)$$

and there exists a unique element $w \in D_T$ such that $u = (\cdot, w)_T$ on D_T. We define $\mathcal{D}(\widetilde{T})$ as the set of all these elements,

$$\mathcal{D}(\widetilde{T}) = \Big\{ w \in D_T;\ (\cdot, w)_T = (\cdot, y)_H \text{ on } D_T \text{ for some } y \in H \Big\},$$

[5] In 1929 J. von Neumann and also A. Wintner identified this class of operators that admit self-adjoint extensions.

[6] Kurt Otto Friedrichs (1901–1982) made contributions to the theory of partial differential equations, operators in Hilbert space, perturbation theory, and bifurcation theory. He published his extension theorem in Göttingen in 1934, and M. Stone did the same in New York in 1932.

i.e., $\mathcal{D}(\widetilde{T}) = \mathcal{D}(T^*) \cap D_T$ and $T^*w = y$ for a unique $w \in \mathcal{D}(\widetilde{T})$, for every $y \in H$. Next we define

$$\widetilde{T}w = y \text{ if } (\cdot, y)_H = (\cdot, w)_T \text{ over } D_T \qquad \big(w \in \mathcal{D}(\widetilde{T})\big),$$

so that $\widetilde{T}$ is the restriction of T^* to $\mathcal{D}(\widetilde{T}) = \mathcal{D}(T^*) \cap D_T$.

This new operator is a linear extension of T, since, for all $v \in D_T$,

$$(v, w)_T = (v, \widetilde{T}w)_H \qquad \big(w \in \mathcal{D}(\widetilde{T})\big) \tag{9.3}$$

and, if $y = Tx \in H$ with $x \in D$,

$$(v, y)_H = (v, Tx)_H = (Tv, x)_H = (v, x)_T \qquad (v \in D).$$

Thus, $x = w$ and $\widetilde{T}w = Tx$, i.e., $D \subset \mathcal{D}(\widetilde{T})$ and $T \subset \widetilde{T}$.

To show that $\widetilde{T}$ is symmetric, apply (9.3) to $w \in \mathcal{D}(\widetilde{T}) \subset D_T$. If $v, w \in \mathcal{D}(\widetilde{T})$, then $(w, v)_T = (w, \widetilde{T}v)_H$ and the scalar product is symmetric, so that $(\widetilde{T}w, v)_H = (w, \widetilde{T}v)_H$.

Observe that $\widetilde{T} : \mathcal{D}(\widetilde{T}) \to H$ is bijective, since, in our construction, since every $y \in H$, w was the unique solution of the equation $\widetilde{T}w = y$. Moreover, the closed graph theorem shows that $A := \widetilde{T}^{-1} : H \to \mathcal{D}(\widetilde{T}) \subset H$ is a bounded operator, since $y_n \to 0$ and $\widetilde{T}^{-1}y_n \to w$ imply

$$0 = \lim_n (\widetilde{T}^{-1}x, y_n)_H = (x, \widetilde{T}^{-1}y_n)_H = (y, w)_H$$

for every $x \in H$, and then $w = 0$. This bounded operator, being the inverse of a symmetric operator, is also symmetric, i.e., it is self-adjoint. But then, every $z \in \mathbf{C} \setminus \mathbf{R}$ is in $\sigma(A)^c$.

The identity $z^{-1}I - A^{-1} = A^{-1}(A - zI)z^{-1}$ shows that $z^{-1} \in \sigma(A^{-1})^c = \sigma(\widetilde{T})^c$, if $z \notin \mathbf{R}$. By Theorem 9.10, $\widetilde{T}$ is self-adjoint. □

9.3. Spectral representation of unbounded self-adjoint operators

$T : \mathcal{D}(T) \subset H \to H$ is still a possibly unbounded linear operator.

The functional calculus for a bounded normal operator T has been based on the spectral resolution

$$T = \int_{\sigma(T)} \lambda \, dE(\lambda),$$

where E represents a spectral measure on $\sigma(T)$. If f is bounded, then this representation allows us to define

$$f(T) = \int_{\sigma(T)} f(\lambda) \, dE(\lambda).$$

This functional calculus can be extended to unbounded functions, h, and then it can be used to set a spectral theory for unbounded self-adjoint operators. The last section of this chapter is devoted to the proof of the following result:

Theorem 9.20 (Spectral theorem). *For every self-adjoint operator T on H, there exists a unique spectral measure E on $\mathbf{R}$ which satisfies*

$$T = \int_{\mathbf{R}} t\, dE(t)$$

in the sense that

$$(Tx, y)_H = \int_{-\infty}^{+\infty} t\, dE_{x,y}(t) \qquad (x \in \mathcal{D}(T),\ y \in H).$$

If f is a Borel measurable function on $\mathbf{R}$, then a densely defined operator

$$f(T) = \int_{\mathbf{R}} f(t)\, dE(t)$$

is obtained such that

$$(f(T)x, y)_H = (\Phi_E(f)x, y)_H = \int_{-\infty}^{+\infty} f(t)\, dE_{x,y}(t) \qquad (x \in \mathcal{D}(f),\ y \in H),$$

where

$$\mathcal{D}(f) = \Big\{x \in H;\ \int_{-\infty}^{+\infty} |f(\lambda)|^2\, dE_{x,x} < \infty\Big\}.$$

For this functional calculus,

(a) $\|f(T)x\|_H^2 = \int_{\sigma(T)} |f|^2\, dE_{x,x}$ *if* $x \in \mathcal{D}(f(T))$,

(b) $f(T)h(T) \subset (fh)(T)$, $\mathcal{D}(f(T)h(T)) = \mathcal{D}(h(T)) \cap \mathcal{D}((fh)(T))$, *and*

(c) $f(T)^* = \bar{f}(T)$ *and* $f(T)^*f(T) = |f|^2(T) = f(T)f(T)^*$.

If f is bounded, then $\mathcal{D}(f) = H$ and $f(T)$ is a bounded normal operator. If f is real, then $f(T)$ is self-adjoint.

The following example will be useful in the next section.

Example 9.21. The spectral measure of the position operator of Examples 9.2 and 9.11,

$$Q = \int_{\mathbf{R}} t\, dE(t),$$

is $E(B) = \chi_B\cdot$ and $dE_{\varphi,\psi}(t) = \varphi(t)\overline{\psi(t)}\, dt$, i.e.,

$$\int_{\mathbf{R}} t\, dE_{\varphi,\psi}(t) = (Q\varphi, \psi)_2 = \int_{\mathbf{R}} t\, \varphi(t)\overline{\psi(t)}\, dt \qquad (\varphi \in \mathcal{D}(Q)).$$

This is proved by defining $F(B)\psi := \chi_B\psi$ for every Borel set $B \subset \mathbf{R}$; that is, $F(B) = \chi_B\cdot$, a multiplication operator. It is easy to check that $F : \mathcal{B}_{\mathbf{R}} \to \mathcal{L}(L^2(\mathbf{R}))$ is a spectral measure, and to show that $F = E$, we only need to see that

$$\int_{\mathbf{R}} t\varphi(t)\overline{\psi(t)}\,dt = \int_{\mathbf{R}} t\,dF_{\varphi,\psi}(t),$$

where $F_{\varphi,\psi}(B) = (F(B)\varphi,\psi)_2 = \int_{\mathbf{R}}(F(B)\varphi)(t)\overline{\psi(t)}\,dt$.

But the integral for the complex measure $F_{\varphi,\psi}$ is a Lebesgue-Stieltjes integral with the distribution function

$$F(t) = F_{\varphi,\psi}\big((-\infty,t]\big) = (\chi_{(-\infty,t]}\varphi,\psi)_2 = \int_{-\infty}^{t} \varphi(s)\overline{\psi(s)}\,ds,$$

and then $dF(t) = \varphi(t)\overline{\psi(t)}\,dt$.

The spectrum $\sigma(T)$ of a self-adjoint operator can be described in terms of its spectral measure E:

Theorem 9.22. *If $T = \int_{\mathbf{R}} t\,dE(t)$ is the spectral representation of a self-adjoint operator T, then*

(a) $\sigma(T) = \operatorname{supp} E$,

(b) $\sigma_p(T) = \{\lambda \in \mathbf{R};\ E\{\lambda\} \neq 0\}$, *and*

(c) $\operatorname{Im} E\{\lambda\}$ *is the eigenspace of every* $\lambda \in \sigma_p(T)$.

Proof. We will use the fact that

$$\|(T-\lambda I)x\|_H^2 = \int_{\mathbf{R}} (t-\lambda)^2\,dE_{x,x}(t) \qquad (x \in \mathcal{D}(T),\ \lambda \in \mathbf{R}),$$

which follows from Theorem 9.20(a).

(a) If $\lambda \notin \operatorname{supp} E$, then $E_{x,x}(\lambda-\varepsilon,\lambda+\varepsilon) = 0$ for some $\varepsilon > 0$, and

$$\|(T-\lambda I)x\|_H^2 = \int_{(\lambda-\varepsilon,\lambda+\varepsilon)} (t-\lambda)^2\,dE_{x,x}(t) \geq \varepsilon^2\|x\|_H^2,$$

which means that $\lambda \notin \sigma(T)$, by Theorem 9.9.

Conversely, if $\lambda \in \operatorname{supp} E$, then $E(\lambda-1/n,\lambda+1/n) \neq 0$ for every $n > 0$ and we can choose $0 \neq x_n \in \operatorname{Im} E(\lambda-1/n,\lambda+1/n)$. Then $\operatorname{supp} E_{x_n,x_n} \subset [\lambda-1/n,\lambda+1/n]$ since it follows from $V \cap (\lambda-1/n,\lambda+1/n) = \emptyset$ that $E(V)$ and $E(\lambda-1/n,\lambda+1/n)$ are orthogonal and $E_{x,x}(V) = (E(V)x,x)_H = 0$. Thus

$$\|(T-\lambda I)x_n\|_H^2 = \int_{\mathbf{R}} (t-\lambda)^2\,dE_{x_n,x_n}(t) \leq \frac{1}{n^2}\|x_n\|_H^2$$

and λ is an approximate eigenvalue.

(b) $Tx = \lambda x$ for $0 \neq x \in \mathcal{D}(T)$ if and only if $\int_{\mathbf{R}} (t-\lambda)^2 \, dE_{x,x}(t) = 0$, meaning that $E_{x,x}\{\lambda\} \neq 0$ and $E_{x,x}(\mathbf{R} \setminus \{\lambda\}) = 0$.

(c) The identity $E_{x,x}(\mathbf{R} \setminus \{\lambda\}) = 0$ means that $x = E\{\lambda\}(x)$ satisfies $Tx = \lambda x$. □

Since $E(B) = E(B \cap \operatorname{supp} E)$, in the spectral representation of the self-adjoint operator, T, $\mathbf{R}$ can be changed by $\operatorname{supp} E = \sigma(T)$; that is,

$$T = \int_{\mathbf{R}} t \, dE(t) = \int_{\sigma(T)} t \, dE(t),$$

and also

$$h(T) = \int_{\mathbf{R}} h \, dE = \int_{\sigma(T)} h \, dE(t).$$

As an application, we define the square root of a positive operator:

Theorem 9.23. *A self-adjoint operator T is positive ($(Tx, x) \geq 0$ for all $x \in \mathcal{D}(T)$) if and only if $\sigma(T) \subset [0, \infty)$. In this case there exists a unique self-adjoint operator R which is also positive and satisfies $R^2 = T$, so that $R = \sqrt{T}$, the square root of T.*

Proof. If $(Tx, x)_H \geq 0$ for every $x \in \mathcal{D}(T)$ and $\lambda > 0$, we have

$$\lambda \|x\|_H^2 \leq ((T + \lambda I)x, x)_H \leq \|(T + \lambda I)x\|_H \|x\|_H,$$

so that

$$\|(T + \lambda I)x\|_H \geq \lambda \|x\|_H \qquad (x \in \mathcal{D}(T)).$$

By Theorem 9.9 there exists $(T + \lambda I)^{-1} \in \mathcal{L}(H)$ and $-\lambda \notin \sigma(T)$.

Conversely, if $\sigma(T) \subset [0, \infty)$ and $x \in \mathcal{D}(T)$, then $\int_0^\infty t \, dE_{x,x}(t) \geq 0$. Moreover

$$(Tx, y)_H = \int_0^\infty t \, dE_{x,y}(t) \qquad (x \in \mathcal{D}(T),\ y \in H).$$

Define $R = f(T)$ with $f(t) = t^{1/2}$. Then $\mathcal{D}(R) = \{x;\ \int_0^\infty t \, dE_{x,x} < \infty\}$, which contains $\mathcal{D}(T) = \{x;\ \int_0^\infty t^2 \, dE_{x,x} < \infty\}$. Thus,

$$R = \sqrt{T} = \int_0^\infty t^{1/2} \, dE(t).$$

From Theorem 9.29(b), $R^2 = T$, since $\mathcal{D}(f^2) = \mathcal{D}(T) \subset \mathcal{D}(f)$.

To prove the uniqueness, suppose that we also have

$$S = \int_0^\infty t \, dF(t)$$

such that $S^2 = T$ and

$$T = \int_0^\infty t^2 \, dF(t).$$

With the substitution $\lambda = t^2$ we obtain a spectral measure $E'(\lambda) = F(\lambda^{1/2})$ such that $T = \int_0^\infty \lambda\, dE'(\lambda)$. From the uniqueness of the spectral measure, $E' = E$ and then $S = R$. $\square$

9.4. Unbounded operators in quantum mechanics

To show how unbounded self-adjoint operators are used in the fundamentals of quantum mechanics, we are going to start by studying the case of a single particle constrained to move along a line.

9.4.1. Position, momentum, and energy. In quantum mechanics, what matters about the **position** is the probability that the particle is in $[a,b] \subset \mathbf{R}$, and this probability is given by an integral

$$\int_a^b |\psi(x)|^2\, dx.$$

The density distribution $|\psi(x)|^2$ is defined by some $\psi \in L^2(\mathbf{R})$, which is called the state function, such that $\int_{\mathbf{R}} |\psi(x)|^2\, dx = \|\psi\|_2^2 = 1$ is the total probability. Here ψ is a complex-valued function and a complex factor α in ψ is meaningless ($|\alpha| = 1$ is needed to obtain $\|\psi\|_2 = 1$). There is a dependence on the time, t, which can be considered as a parameter.

The mean position of the particle will be

$$\mu_\psi = \int_{\mathbf{R}} x|\psi(x)|^2\, dx = \int_{\mathbf{R}} x\psi(x)\overline{\psi(x)}\, dx = \int_{\mathbf{R}} x\, dE_{\psi,\psi}(x)$$

with $dE_{\psi,\psi} = \psi(x)\overline{\psi(x)}\, dx$.

If Q denotes the position operator, $Q\varphi(x) = x\varphi(x)$, note that $\mu_\psi = (Q\psi,\psi)_2$.

The dispersion of the position with respect to its mean value is measured by the variance,

$$\mathrm{var}_\psi = \int_{\mathbf{R}} (x-\mu_\psi)^2|\psi(x)|^2\, dq = \int_{\mathbf{R}} x(x-\mu_\psi)^2\, dE_{\psi,\psi}(x) = ((Q-\mu_\psi I)^2\psi,\psi)_2.$$

Similarly, if $\int_{\mathbf{R}} |f(x)||\psi(x)|^2\, dx < \infty$, the mathematical expectation of f is

$$\int_{\mathbf{R}} f(x)|\psi(x)|^2\, dx = (f\psi,\psi)_2 = \int_{\mathbf{R}} f(x)\, dE_{\psi,\psi}(x). \tag{9.4}$$

The **momentum** of the particle is defined as mass $\times$ velocity:

$$p = m\dot{x}.$$

Note that, from the properties of the Fourier transform,
(9.5)
$$\int_{\mathbf{R}} \xi|\widehat{\psi}(\xi)|^2\,d\xi = \int_{\mathbf{R}} \xi\widehat{\psi}(\xi)\overline{\widehat{\psi}(\xi)}\,d\xi = \frac{1}{2\pi i}\int_{\mathbf{R}} \widehat{\psi}'(\xi)\overline{\widehat{\psi}(\xi)}\,d\xi = \frac{1}{2\pi i}(\psi',\psi)_2.$$
By assuming that the probability that $p \in [a,b]$ is given by
$$\frac{1}{h}\int_a^b \left|\widehat{\psi}\Big(\frac{p}{h}\Big)\right|^2 dp = \int_{a/h}^{b/h} |\widehat{\psi}(\xi)|^2\,d\xi,$$
where $h = 6.62607095(44)\cdot 10^{-34}\,J$·sec is the **Planck constant**,[7] the average value of p is
$$\frac{1}{h}\int_{\mathbf{R}} p\left|\widehat{\psi}\Big(\frac{p}{h}\Big)\right|^2 dp = h\int_{\mathbf{R}} \xi|\widehat{\psi}(\xi)|^2\,d\xi.$$
Here the Fourier transform can be avoided by considering the **momentum operator** P defined as
$$P = \frac{h}{2\pi i}D \qquad \Big(D = \frac{d}{dx}\Big),$$
since then, as noted in (9.5), this average is

$$h\int_{\mathbf{R}} \xi|\widehat{\psi}(\xi)|^2\,d\xi = (P\psi,\psi)_2. \tag{9.6}$$

If $\int_{\mathbf{R}} |f(h\xi)||\widehat{\psi}(\xi)|^2\,d\xi < \infty$, then we have the value
$$\mu_\psi(f) = \int_{\mathbf{R}} f(h\xi)|\widehat{\psi}(\xi)|^2\,d\xi$$
for the mathematical expectation of f, which in the case $f(p) = p^n$ is
$$\mu_\psi(p^n) = (P^n\psi,\psi)_2.$$

The **kinetic energy** is
$$T = \frac{p^2}{2m},$$
so that its mathematical expectation will be
$$\mu_\psi(T) = \frac{1}{2m}(P^2\psi,\psi)_2.$$
The potential energy is given by a real-valued function $V(x)$ and from (9.4) we obtain the value
$$\mu_\psi(V) = \int_{\mathbf{R}} V(x)|\psi(x)|^2\,dx = (V\psi,\psi)_2$$
for the mathematical expectation of V if $\int_{\mathbf{R}} |V(x)||\psi(x)|^2\,dx < \infty$.

[7] This is the value reported in October 2007 by the National Physical Laboratory for this constant, named in honor of Max Planck, considered to be the founder of quantum theory in 1901 when, in his description of the black-body radiation, he assumed that the electromagnetic energy could be emitted only in quantized form, $E = h\nu$, where ν is the frequency of the radiation.

The mathematical expectation is additive, so that the average of the total energy is

$$\mu_\psi(T+V) = \big(\frac{1}{2m}P^2\psi + V\psi, \psi\big)_2 = (H\psi, \psi)_2,$$

where

$$H = \frac{1}{2m}P^2 + V$$

is the **energy operator**, or **Hamiltonian**, of the particle.

9.4.2. States, observables, and Hamiltonian of a quantic system. As in the case of classical mechanics, the basic elements in the description of a general quantic system are those of **state** and **observable**.

Classical mechanics associates with a given system a phase space, so that for an N-particle system we have a $6N$-dimensional phase state.

Similarly, quantum mechanics associates with a given system a complex Hilbert space $\mathcal{H}$ as the **state space**, which is $L^2(\mathbf{R})$ in the case of a single particle on the line. In a quantum system the **observables** are self-adjoint operators, such as the position, momentum, and energy operators.

A quantum system, in the **Schrödinger picture**, is ruled by the following postulates:

Postulate 1: States and observables

A **state** of a physical system at time t is a line $[\psi] \subset \mathcal{H}$, which we represent by $\psi \in \mathcal{H}$ such that $\|\psi\|_{\mathcal{H}} = 1$.

A **wave function** is an $\mathcal{H}$-valued function of the time parameter $t \in \mathbf{R} \mapsto \psi(t) \in \mathcal{H}$. If $\psi(t)$ describes the state, then $c\psi(t)$, for any nonzero constant c, represents the same state.

The observable values of the system are magnitudes such as position, momentum, angular momentum, spin, charge, and energy that can be measured. They are associated to self-adjoint operators. In a quantic system, an **observable** is a time-independent[8] self-adjoint operator A on $\mathcal{H}$, which has a spectral representation

$$A = \int_{\mathbf{R}} \lambda \, dE(\lambda).$$

By the "superposition principle", all self-adjoint operators on $\mathcal{H}$ are assumed to be observable,[9] and all lines $[\psi] \subset \mathcal{H}$ are admissible states.

[8]In the Heisenberg picture of quantum mechanics, the observables are represented by time-dependent operator-valued functions $A(t)$ and the state ψ is time-independent.

[9]Here we are following the early assumptions of quantum mechanics, but the existence of "superselection rules" in quantum field theories indicated that this superposition principle lacks experimental support in relativistic quantum mechanics.

The elements of the spectrum, $\lambda \in \sigma(A)$, are the **observable values** of the observable A.

Postulate 2: Distribution of an observable in a given state

The values $\lambda \in \sigma(A)$ in a state ψ are observable in terms of a probability distribution P_ψ^A.

As in the case of the position operator Q for the single particle on $\mathbf{R}$, the observable $A = \int_{\mathbf{R}} \lambda \, dE(\lambda)$ on $\mathcal{H}$ is evaluated in a state ψ at a given time in terms of the probability $P_\psi^A(B)$ of belonging to a set $B \subset \mathbf{R}$ with respect to the distribution $dE_{\psi,\psi}(\lambda)$ (we are assuming that $\|\psi\|_{\mathcal{H}} = 1$), so that

$$P_\psi^A(B) = \int_B \lambda dE_{\psi,\psi}(\lambda) = (E(B)\psi, \psi)_{\mathcal{H}}$$

and the **mean value** is

$$\widehat{A}_\psi := \int_{\mathbf{R}} \lambda \, dE_{\psi,\psi}(\lambda) = (A\psi, \psi)_{\mathcal{H}}.$$

When $\psi \in \mathcal{D}(A)$, this mean value $\widehat{A}_\psi$ exists, since λ^2 is integrable with respect to the finite measure $E_{\psi,\psi}$, and also $\int_{\mathbf{R}} |\lambda| \, dE_{\psi,\psi}(\lambda) < \infty$.

In general, if $f \in L^2(E_{\psi,\psi})$,

$$\widehat{f(A)}_\psi = (f(A)\psi, \psi)_{\mathcal{H}},$$

is the **expected value** of f, the mean value with respect to $E_{\psi,\psi}$.

The **variance** of A in the state $\psi \in \mathcal{D}(A)$ is then

$$\mathrm{var}_\psi(A) = \int_{\mathbf{R}} (\lambda - \widehat{A}_\psi)^2 \, dE_{\psi,\psi}(\lambda) = ((A - \widehat{A}_\psi I)^2\psi, \psi)_{\mathcal{H}} = \|A\psi - \widehat{A}_\psi\psi\|_{\mathcal{H}}^2.$$

It is said that A **certainly takes** the value λ_0 in the state ψ if $\widehat{A}_\psi = \lambda_0$ and $\mathrm{var}_\psi(A) = 0$.

This means that ψ is an eigenvector of A with eigenvalue λ_0, since it follows from $A\psi = \lambda_0\psi$ that $\widehat{A}_\psi = (A\psi, \psi)_{\mathcal{H}} = \lambda_0$, and also

$$\mathrm{var}_\psi(A) = \|A\psi - \widehat{A}_\psi\psi\|_{\mathcal{H}}^2 = 0.$$

Conversely, $\mathrm{var}_\psi(A) = 0$ if and only if $A\psi - \widehat{A}_\psi\psi = 0$.

Postulate 3: Hamiltonians and the Schrödinger equation

There is an observable, H, the **Hamiltonian**, defining the evolution of the system

$$\psi(t) = U_t\psi_0,$$

where ψ_0 is the initial state and U_t is an operator defined as follows:

If h is the Planck constant and $g_t(\lambda) = e^{-\frac{it}{h}\lambda}$, a continuous function with its values in the unit circle, then using the functional calculus, we can define the unitary operators

$$U_t := g_t(H) \in \mathcal{L}(\mathcal{H}) \qquad (t \in \mathbf{R})$$

that satisfy the conditions

$$U_0 = I, \quad U_sU_t = U_{s+t}, \text{ and } \lim_{t\to s} \|U_tx - U_sx\|_{\mathcal{H}} = 0\,\forall x \in \mathcal{H},$$

since $g_t\bar{g}_t = 1$, $g_0 = 1$, $g_sg_t = g_{st}$, and, if $H = \int_{\sigma(H)} \lambda\, dE(\lambda)$ is the spectral representation of H, then the continuity property

$$\|U_t\psi - U_s\psi\|^2_{\mathcal{H}} = \int_{\sigma(H)} |e^{-\frac{it}{h}\lambda} - e^{-\frac{is}{h}\lambda}|^2\, dE_{\psi,\psi}(\lambda) \to 0 \quad \text{as} \quad t \to s$$

follows from the dominated convergence theorem.

Such a family of operators U_t is called a **strongly continuous one-parameter group of unitary operators**, and we say that $A = -(i/h)H$ is the **infinitesimal generator**.

It can be shown (**Stone's theorem**) that the converse is also true: every strongly continuous one-parameter group of unitary operators $\{U_t\}_{t\in\mathbf{R}}$ has a self-adjoint infinitesimal generator $A = -(i/h)H$; that is, $U_t = e^{-\frac{it}{h}H}$ for some self-adjoint operator H.

It is said that

$$U_t = e^{-\frac{it}{h}H}$$

is the **time-evolution operator** of the system.

It is worth noticing that, if $\psi \in \mathcal{D}(H)$, the function $t \mapsto U_t\psi$ is differentiable and

$$\frac{d}{dt}U_t\psi = U_tA\psi = AU_t\psi$$

at every point $t \in \mathbf{R}$. Indeed,

$$\frac{1}{s}(U_sU_t\psi - U_t\psi) = U_t\frac{1}{s}(U_s\psi - \psi)$$

and

$$\lim_{s\to 0} \frac{1}{s}(U_s\psi - \psi) = A\psi = -\frac{i}{h}H\psi, \tag{9.7}$$

since

$$\|\frac{1}{s}(U_s\psi - \psi) + \frac{i}{h}H\psi\|^2_{\mathcal{H}} = \int_{\sigma(A)} \left|\frac{e^{-is\lambda/h} - 1}{s} + \frac{i}{h}\lambda\right|^2 dE_{\psi,\psi}(\lambda) \to 0$$

as $s \to 0$, again by dominated convergence.

For a given initial state ψ_0, it is said that $\psi(t) = U_t\psi_0$ is the corresponding wave function.

If $\psi(t) \in \mathcal{D}(H)$ for every $t \in \mathbf{R}$, then the vector-valued function $t \mapsto \psi(t)$ is derivable and satisfies the **Schrödinger equation**[10]

$$ih\psi'(t) = H\psi(t),$$

since by (9.7)

$$\psi'(t) = \frac{d}{dt}(U_t\psi_0) = AU_t\psi_0 = -\frac{i}{h}H\psi(t).$$

In this way, from a given initial state, subsequent states can be calculated causally from the Schrödinger equation.[11]

9.4.3. The Heisenberg uncertainty principle and compatible observables. To illustrate the role of probabilities in the postulates, let us consider again the case of a single particle on $\mathbf{R}$. Recall that the momentum operator,

$$P\psi(q) = \frac{h}{2\pi i}\psi'(q),$$

is self-adjoint on $L^2(\mathbf{R})$ and with domain $H^1(\mathbf{R})$.

From Example 9.4 we know that the commutator of P and Q is bounded and

$$[P,Q] = PQ - QP = \frac{h}{2\pi i}I,$$

where $\mathcal{D}([P,Q]) = \mathcal{D}(PQ) \cap \mathcal{D}(QP)$, or extended to all $L^2(\mathbf{R})$ by continuity.

Lemma 9.24. *The commutator $C = [S,T] = ST - TS$ of two self-adjoint operators on $L^2(\mathbf{R})$ satisfies the estimate*

$$|\widehat{C}_\psi| \le 2\sqrt{\mathrm{var}_\psi(S)}\sqrt{\mathrm{var}_\psi(T)}$$

for every $\psi \in \mathcal{D}(C)$.

Proof. Obviously, $A = S - \widehat{S}_\psi I$ and $B = T - \widehat{T}_\psi I$ are self-adjoint (note that $\widehat{S}_\psi, \widehat{T}_\psi \in \mathbf{R}$) and $C = [A,B]$. From the definition of the expected value,

$$|\widehat{C}_\psi| \le |(B\psi, A\psi)_2| + |(A\psi, B\psi)_2| \le 2\|B\psi\|_2\|A\psi\|_2,$$

where, A being self-adjoint, $\|A\psi\|_2^2 = (A^2\psi,\psi)_2 = \mathrm{var}_\psi(S)$. Similarly, $\|B\psi\|_2^2 = \mathrm{var}_\psi(T)$. □

[10]E. Schrödinger published his equation and the spectral analysis of the hydrogen atom in a series of four papers in 1926, which where followed the same year by Max Born's interpretation of $\psi(t)$ as a probability density.

[11]We have assumed that the energy is constant and the Hamiltonian does not depend on t but, if the system interacts with another one, the Hamiltonian is an operator-valued function $H(t)$ of the time parameter. In the Schrödinger picture, all the observables except the Hamiltonian are time-invariant.

Theorem 9.25 (Uncertainty principle).

$$\sqrt{\mathrm{var}_\psi(Q)}\sqrt{\mathrm{var}_\psi(P)} \geq \frac{h}{4\pi}.$$

Proof. In the case $C = [P, Q]$, $|\widehat{C}_\psi| = |(h/2\pi i)\widehat{I}_\psi| = h/2\pi$ and we can apply Lemma 9.24. $\square$

The standard deviations $\sqrt{\mathrm{var}_\psi(Q)}$ and $\sqrt{\mathrm{var}_\psi(P)}$ measure the uncertainties of the position and momentum, and the uncertainty principle shows that both uncertainties cannot be arbitrarily small simultaneously. Position and moment are said to be **incompatible observables**.

It is a basic principle of all quantum theories that if n observables $A_1, \dots, A_n$ are compatible in the sense of admitting arbitrarily accurate simultaneous measurements, they must commute. However, since these operators are only densely defined, the commutators $[A_j, A_k]$ are not always densely defined. Moreover, the condition $AB = BA$ for two commuting operators is unsatisfactory; for example, taking it literally, $A0 \neq 0A$ if A is unbounded, but $A0 \subset 0A$ and $[A, 0] = 0$ on the dense domain of A.

This justifies saying that $A_j = \int_{\mathbf{R}} \lambda\, dE^j(\lambda)$ and $A_k = \int_{\mathbf{R}} \lambda\, dE^k(\lambda)$ **commute**, or that **their spectral measures commute**, if

$$[E^j(B_1), E^k(B_2)] = 0 \qquad (B_1, B_2 \in \mathcal{B}_{\mathbf{R}}). \tag{9.8}$$

If both A_j and A_k are bounded, then this requirement is equivalent to $[A_j, A_k] = 0$ (see Exercise 9.18).[12] For such commuting observables and a given (normalized) state ψ, there is a probability measure P_ψ on $\mathbf{R}^n$ so that

$$P_\psi(B_1 \times \cdots \times B_n) = (E^1(B_1) \cdots E^n(B_n)\psi, \psi)_{\mathcal{H}}$$

is the predicted probability that a measurement to determine the values $\lambda_1, \dots, \lambda_n$ of the observables $A_1, \dots, A_n$ will lie in $B = B_1 \times \cdots \times B_n$. See Exercise 9.19, where it is shown how a spectral measure E on $\mathbf{R}^n$ can be defined so that $dE_{\psi,\psi}$ is the distribution of this probability; with this spectral measure there is an associated functional calculus $f(A_1, \dots, A_n)$ of n commuting observables.

9.4.4. The harmonic oscillator. A heuristic recipe to determine a quantic system from a classical system of energy

$$T + V = \sum_{j=1}^{n} \frac{p_j^2}{2m_j} + V(q_1, \dots, q_1)$$

[12] But E. Nelson proved in 1959 that there exist essentially self-adjoint operators A_1 and A_2 with a common and invariant domain, so that $[A_1, A_2]$ is defined on this domain and $[A_1, A_2] = 0$ but with noncommuting spectral measures.

is to make a formal substitution of the generalized coordinates q_j by the position operators Q_j (multiplication by q_j) and every p_j by the corresponding momentum P_j. Then the Hamiltonian or energy operator should be a self-adjoint extension of

$$H = \sum_{j=1}^{n} \frac{P_j^2}{2m_j} + V(Q_1, \dots, Q_1).$$

For instance, in the case of the two-body problem under Coulomb force, which derives from the potential $-e/|x|$, $n = 3$ and the energy of the system is

$$E = T + V = \frac{1}{2m}|p|^2 - \frac{e^2}{|x|}.$$

Hence, in a convenient scale, $H = -\Delta - \frac{1}{|x|}$ is the possible candidate of the Hamiltonian of the hydrogen atom. In Example 9.15 we have seen that it is a self-adjoint operator with domain $H^2(\mathbf{R}^3)$.

With the help of his friend Hermann Weyl, Schrödinger calculated the eigenvalues of this operator. The coincidence of his results with the spectral lines of the hydrogen atom was considered important evidence for the validity of Schrödinger's model for quantum mechanics.

Several problems appear with this quantization process, such as finding the self-adjoint extension of H, determining the spectrum, and describing the evolution of the system for large values of t ("scattering").

Let us consider again the simple classical one-dimensional case of a single particle with mass m, now in a Newtonian field with potential V, so that

$$-\nabla V = F = \frac{d}{dt}(m\dot{q}),$$

q denoting the position. We have the linear momentum $p = m\dot{q}$, the kinetic energy $T = (1/2)m\dot{q}^2 = p^2/2m$, and the total energy $E = T + V$.

The **classical harmonic oscillator** corresponds to the special case of the field $F(q) = -m\omega^2 q$ on a particle bound to the origin by the potential

$$V(q) = m\frac{\omega^2}{2}q^2$$

if $q \in \mathbf{R}$ is the position variable. Hence, in this case,

$$E = T + V = \frac{1}{2}m\dot{q}^2 + m\frac{\omega^2}{2}q^2 = \frac{1}{2m}p^2 + m\frac{\omega^2}{2}q^2.$$

From Newton's second law, the initial state $\dot{q}(0) = 0$ and $q(0) = a > 0$ determines the state of the system at every time,

$$q = a\cos(\omega t).$$

The state space for the **quantic harmonic oscillator** is $L^2(\mathbf{R})$, and the position $Q = q\cdot$ and the momentum P are two observables. By making the announced substitutions, we obtain as a possible Hamiltonian the operator

$$H = \frac{1}{2m}P^2 + m\frac{\omega^2}{2}Q^2.$$

On the domain $\mathcal{S}(\mathbf{R})$, which is dense in $L^2(\mathbf{R})$, it is readily checked that

$$(H\varphi, \psi)_2 = (\varphi, H\psi)_2,$$

so that H is a symmetric operator. We will prove that it is essentially self-adjoint and the Hamiltonian will be its unique self-adjoint extension $\bar{H} = H^{**}$, which is also denoted H.

In coordinates,

$$H = -\frac{h^2}{2m \cdot 4\pi^2}\frac{d^2}{dq^2} + \frac{m\omega^2}{2}q^2$$

which after the substitution $x = aq$, with $a^2 = 2\pi m\omega/h$, can be written

$$H = \frac{h\omega}{2}\Big(x^2 - \frac{d^2}{dx^2}\Big).$$

Without loss of generality, we suppose $h\omega = 1$, and it will be useful to consider the action of

$$H = \frac{1}{2}\Big(x^2 - \frac{d^2}{dx^2}\Big)$$

on

$$\mathcal{F} := \{P(x)e^{-x^2/2};\ P \text{ polynomial}\},$$

the linear subspace of $\mathcal{S}(\mathbf{R})$ that has the functions $x^n e^{-x^2/2}$ as an algebraic basis.

Since $H(x^n e^{-x^2/2}) \in \mathcal{F}$, we have $H(\mathcal{F}) \subset \mathcal{F}$. Similarly, $A(\mathcal{F}) \subset \mathcal{F}$ and $B(\mathcal{F}) \subset \mathcal{F}$ if

$$A := \frac{1}{\sqrt{2}}\Big(x + \frac{d}{dx}\Big),$$

the **annihilation operator**, and

$$B := \frac{1}{\sqrt{2}}\Big(x - \frac{d}{dx}\Big),$$

the **creation operator**.

Theorem 9.26. *The subspace $\mathcal{F}$ of $\mathcal{S}(\mathbf{R})$ is dense in $L^2(\mathbf{R})$, and the Gram-Schmidt process applied to $\{x^n e^{-x^2/2}\}_{n=0}^{\infty}$ generates an orthonormal basis $\{\widetilde{\psi}_n\}_{n=0}^{\infty}$ of $L^2(\mathbf{R})$. The functions $\widetilde{\psi}_n$ are in the domain $\mathcal{S}(\mathbf{R})$ of H and they are eigenfunctions with eigenvalues $\lambda_n = n + 1/2$. According to Theorem 9.17, the operator H is essentially self-adjoint.*

Proof. On $\mathcal{F}$, a simple computation gives

$$H = BA + \frac{1}{2}I = AB - \frac{1}{2}I;$$

hence $HB = BAB + \frac{1}{2}B$ and $BH = BAB - \frac{1}{2}B$, so that

$$[H, B] = B.$$

Then, if $H\psi = \lambda\psi$ and $B\psi \neq 0$ with $\psi \in \mathcal{F}$, it follows that $\lambda + 1$ is also an eigenvalue of H, with the eigenfunction $B\psi$, since

$$H(B\psi) = B(H\psi) + B\psi = \lambda B\psi + B\psi = (\lambda + 1)B\psi.$$

For

$$\psi_0(x) := e^{-x^2/2},$$

we have $2H\psi_0(x) = x^2 e^{-x^2/2} - (e^{-x^2/2})'' = e^{-x^2/2}$, so that

$$H\psi_0 = \frac{1}{2}\psi_0$$

and ψ_0 is an eigenfunction with eigenvalue $1/2$.

We have $\sqrt{2}B\psi_0(x) = 2xe^{-x^2/2} \neq 0$ and, if we denote

$$\psi_n := (\sqrt{2}B)^n \psi_0 = \sqrt{2}B\psi_{n-1},$$

from the above remarks we obtain

$$H\psi_n = \Big(n + \frac{1}{2}\Big)\psi_n \qquad (n = 0, 1, 2, \dots)$$

and $\psi_n(x) = H_n(x)e^{-x^2/2}$. By induction over n, it follows that H_n is a polynomial with degree n. It is called a **Hermite polynomial**.

The functions ψ_n are mutually orthogonal, since they are eigenfunctions with different eigenvalues, and they generate $\mathcal{F}$.

To prove that $\mathcal{F}$ is a dense subspace of $L^2(\mathbf{R})$, let $f \in L^2(\mathbf{R})$ be such that $\int_{\mathbf{R}} f(x)x^n e^{-x^2/2}\,dx = (x^n e^{-x^2/2}, f(x))_2 = 0$ for all $n \in \mathbf{N}$. Then

$$F(z) := \int_{\mathbf{R}} f(x)e^{-x^2/2}e^{-2\pi i x z}\,dx$$

is defined and continuous on $\mathbf{C}$, and the Morera theorem shows that F is an entire function, with

$$F^{(n)}(z) = (-2\pi i)^n \int_{\mathbf{R}} x^n f(x)e^{-x^2/2}e^{-2\pi i x z}\,dx.$$

But $F^{(n)}(0) = (x^n e^{-x^2/2}, f(x))_2 = 0$ for all $n \in \mathbf{N}$, so that $F = 0$. From the Fourier inversion theorem we obtain $f(x)e^{-x^2/2} = 0$ and $f = 0$.

It follows that the eigenfunctions $\widetilde{\psi}_n := \|\psi_n\|_2^{-1}\psi_n$ of H are the elements of an orthonormal basis of $L^2(\mathbf{R})$, all of them contained in $\mathcal{S}(\mathbf{R})$, which is the domain of the essentially self-adjoint operator H. $\square$

Remark 9.27. In the general setting, for any mh,

$$H = \frac{h\omega}{2}\Big(x^2 - \frac{d^2}{dx^2}\Big),$$

and we have $H\widetilde{\psi}_n = h\omega(n + \frac{1}{2})\widetilde{\psi}_n$. Thus

$$\sigma(H) = \{h\omega/2, h\omega(1 + 1/2), h\omega(2 + 1/2), \ldots\}.$$

The wave functions $\widetilde{\psi}_n$ are known as the **bound states**, and the numbers are the **energy eigenvalues** of these bound states. The minimal energy is $h\omega/2$,[13] and $\widetilde{\psi}_0$ is the "**ground state**".

9.5. Appendix: Proof of the spectral theorem

The proof of Theorem 9.20 will be obtained in several steps. First, in Theorem 9.28, we define a functional calculus with bounded functions for spectral measures. Then this functional calculus will be extended to unbounded functions in Theorem 9.29. The final step will prove the spectral theorem for unbounded self-adjoint operators by the von Neuman method based on the use of the Cayley transform.

9.5.1. Functional calculus of a spectral measure. Our first step in the proof of the spectral theorem for unbounded self-adjoint operators will be to define a functional calculus associated to a general spectral measure

$$E : \mathcal{B}_K \to \mathcal{L}(H)$$

as the integral with respect to this operator-valued measure.

Denote by $L^\infty(E)$ the complex normed space of all E-essentially bounded complex functions (the functions coinciding E-a.e. being identified as usual) endowed with the natural operations and the norm

$$\|f\|_\infty = E\text{-}\sup|f|.$$

With the multiplication and complex conjugation, it becomes a commutative C^*-algebra, and the constant function 1 is the unit. Every $f \in L^\infty(E)$ has a bounded representative.

We always represent simple functions as

$$s = \sum_{n=1}^{N} \alpha_n \chi_{B_n} \in S(K),$$

where $\{B_1, \ldots, B_N\}$ is a partition of K. Since every bounded measurable function is the uniform limit of simple functions, $S(K)$ is dense in $L^\infty(E)$, and we will start by defining the integral of simple functions:

[13] Max Planck first applied his quantum postulate to the harmonic oscillator, but he assumed that the lowest level energy was 0 instead of $h\omega/2$. See footnote 7 in this chapter.

As in the scalar case,

$$\int s\,dE := \sum_{n=1}^{N} \alpha_n E(B_n) \in \mathcal{L}(H)$$

is well-defined and uniquely determined, independently of the representation of s, by the relation

$$(\Big(\int s\,dE\Big)x, y)_H = \int_K s\,dE_{x,y} \qquad (x, y \in H),$$

since $\int_K s\,dE_{x,y} = \sum_{n=1}^{N} \alpha_n (E(B_n)x, y)_H = (\sum_{n=1}^{N} \alpha_n E(B_n)x, y)_H$.

It is readily checked that this integral is clearly linear, $\int 1\,dE = I$, and $(\int s\,dE)^* = \int \bar{s}\,dE$.

It is also multiplicative,

$$\int st\,dE = \int s\,dE \int t\,dE = \int t\,dE \int s\,dE, \tag{9.9}$$

since for a second simple function t we can suppose that $t = \sum_{n=1}^{N} \beta_n \chi_{B_n}$, with the same sets B_n as in s, and then

$$\begin{aligned}
\int s\,dE \int t\,dE &= \sum_{n=1}^{N} \beta_n \Big(\int s\,dE\Big) E(B_n) \\
&= \sum_{n=1}^{N} \beta_n \alpha_n E(B_n) E(B_n) = \sum_{n=1}^{N} \alpha_n \beta_n E(B_n) \\
&= \int st\,dE.
\end{aligned}$$

Also

$$\Big\|\Big(\int s\,dE\Big)x\Big\|_H^2 = \int_K |s|^2\,dE_{x,x} \qquad (x \in H,\ s \in S(K))$$

since

$$\begin{aligned}
(\Big(\int s\,dE\Big)x, \Big(\int s\,dE\Big)x)_H &= (\Big(\int s\,dE\Big)^* \Big(\int s\,dE\Big)x, x)_H \\
&= (\Big(\int |s|^2\,dE\Big)x, x)_H.
\end{aligned}$$

This yields

$$\Big\|\int s\,dE\Big\| \le \|s\|_\infty$$

and, in fact, the integral is isometric. Indeed, if we choose n so that $\|s\|_\infty = |\alpha_n|$ with $E(B_n) \neq 0$ and $x \in \operatorname{Im} E(B_n)$, then

$$\Big(\int s\,dE\Big)x = \alpha_n E(B_n)x = \alpha_n x$$

and necessarily

$$\Big\|\Big(\int s\,dE\Big)x\Big\|_H = \|s\|_\infty.$$

Now the integral can be extended over $L^\infty(E)$ by continuity, since it is a bounded linear map from the dense vector subspace $S(K)$ of $L^\infty(E)$ to the Banach space $\mathcal{L}(H)$.

We will denote

$$\Phi_E(f) := \int f\,dE = \lim_n \int s_n\,dE$$

if $s_k \to f$ in $L^\infty(E)$ ($s_k \in S(K)$).

The identities $(\Phi_E(s_k)x, y)_H = \int_K s_k\,dE_{x,y}$ extend to

$$(\Phi_E(f)x, y)_H = \int_K f\,dE_{x,y}$$

by taking limits. All the properties of Φ_E contained in the following theorem are now obvious:

Theorem 9.28. *If $E : \mathcal{B}_K \to \mathcal{L}(H)$ is a spectral measure, then there is a unique homomorphism of C^*-algebras $\Phi_E : L^\infty(K) \to \mathcal{L}(H)$ such that*

$$(\Phi_E(f)x, y)_H = \int_K f\,dE_{x,y} \qquad (x, y \in H,\ f \in L^\infty(K)).$$

This homomorphism also satisfies

$$\|\Phi_E(f)x\|_H^2 = \int_K |f|^2\,dE_{x,x} \qquad (x \in H,\ f \in L^\infty(K)). \tag{9.10}$$

9.5.2. Unbounded functions of bounded normal operators. To extend the functional calculus $f(T) = \Phi_f(T)$ of a bounded normal operator with bounded functions to unbounded measurable functions h, we start by extending to unbounded functions the functional calculus of Theorem 9.28 for any spectral measure E:

Theorem 9.29. *Suppose K a locally compact subset of $\mathbf{C}$, $E : \mathcal{B}_K \to \mathcal{L}(H)$ a spectral measure, h a Borel measurable function on $K \subset \mathbf{C}$, and*

$$\mathcal{D}(h) := \Big\{x \in H;\ \int_K |h(\lambda)|^2\,dE_{x,x} < \infty\Big\}.$$

Then there is a unique linear operator $\Phi_E(h)$ on H, represented as

$$\Phi_E(h) = \int_K h\,dE,$$

with domain $\mathcal{D}(\Phi_E(h)) = \mathcal{D}(h)$ and such that

$$(\Phi_E(h)x, y)_H = \int_K h(\lambda)\,dE_{x,y}(\lambda) \qquad (x \in \mathcal{D}(h),\ y \in H).$$

This operator is densely defined and, if f and h are Borel mesurable functions on K, the following properties hold:

(a) $\|\Phi_E(h)x\|_H^2 = \int_K |h|^2\, dE_{x,x}$, *if* $x \in \mathcal{D}(h)$.

(b) $\Phi_E(f)\Phi_E(h) \subset \Phi_E(fh)$ *and* $\mathcal{D}(\Phi_E(f)\Phi_E(h)) = \mathcal{D}(h) \cap \mathcal{D}(fh)$.

(c) $\Phi_E(h)^* = \Phi_E(\bar{h})$ *and* $\Phi_E(h)^*\Phi_E(h) = \Phi_E(|h|^2) = \Phi_E(h)\Phi_E(h)^*$.

Proof. It is easy to check that $\mathcal{D}(h)$ is a linear subspace of H. For instance, $\|E(B)(x+y)\|_H^2 \le 2\|E(B)x\|_H^2 + 2\|E(B)y\|_H^2$ so that

$$E_{x+y,x+y}(B) \le 2E_{x,x}(B) + 2E_{y+y}(B)$$

and $\mathcal{D}(h) + \mathcal{D}(h) \subset \mathcal{D}(h)$.

This subspace is dense. Indeed, if $y \in H$, we consider

$$B_n := \big\{|h| \le n\big\} \uparrow K,$$

so that, from the strong σ-additivity of E,

$$y = E(K)y = \lim_n E(B_n)y,$$

where $x_n := E(B_n)y \in \mathcal{D}(h)$ since

$$E(B)x_n = E(B)E(B_n)x_n = E(B \cap B_n)x_n \qquad (B \subset K)$$

and $E_{x_n,x_n}(B) = E_{x_n,x_n}(B \cap B_n)$, the restriction of E_{x_n,x_n} to B_n, so that

$$\int_K |h|^2\, dE_{x_n,x_n} = \int_{B_n} |h|^2\, dE_{x_n,x_n} \le n^2\|x_n\|_H^2 < \infty.$$

If h is bounded, then let us also prove the estimate

$$\Big|\int_K h\, dE_{x,y}\Big| \le \int_K |h|\, d|E_{x,y}| \le \Big(\int_K |h|^2\, dE_{x,x}\Big)^{1/2}\|y\|_H < \infty, \tag{9.11}$$

where $|E_{x,y}|$ is the total variation of the Borel complex measure $E_{x,y}$.

From the polar representation of a complex measure (see Lemma **??**), we obtain a Borel measurable function ϱ such that $|\varrho| = 1$ and

$$\varrho h\, dE_{x,y} = |h|\, d|E_{x,y}|,$$

where $|E_{x,y}|$ denotes the total variation of $E_{x,y}$. Thus,

$$\begin{aligned}
\Big|\int_K h\, dE_{x,y}\Big| &\le \int_K |h|\, d|E_{x,y}| = \int_K \varrho h\, dE_{x,y} = (\Phi_E(\varrho h)x, y)_H \\
&\le \|\Phi_E(\varrho h)x\|_H\|y\|_H = \Big(\int_K |\varrho h|^2\, dE_{x,x}\Big)^{1/2}\|y\|_H \\
&= \Big(\int_K |h|^2\, dE_{x,x}\Big)^{1/2}\|y\|_H,
\end{aligned}$$

where in the second line we have used (9.10), and (9.11) holds.

When h is unbounded, to define $\Phi_E(h)x$ for every $x \in \mathcal{D}(h)$, we are going to show that $y \mapsto \int_K h\, dE_{x,y}$ is a bounded conjugate-linear form on H. Let us consider $h_n(z) = h(z)\chi_{B_n}(z) \to h(z)$ if $z \in K$, so that

$$\Big| \int_K h_n\, dE_{x,y} \Big| \le \Big(\int_K |h_n|^2\, dE_{x,x} \Big)^{1/2} \|y\|_H,$$

and by letting $n \to \infty$, we also obtain (9.11) for h in this unbounded case if $x \in \mathcal{D}(h)$.

Then the conjugate-linear functional $y \mapsto \int_K h\, dE_{x,y}$ is bounded with norm $\le (\int_K |h|^2\, dE_{x,x})^{1/2}$, and by the Riesz representation theorem there is a unique $\Phi_E(h)x \in H$ such that, for every $y \in H$,

$$(\Phi_E(h)x, y)_H = \int_K h(\lambda)\, dE_{x,y}(\lambda), \quad \|\Phi_E(h)x\|_H \le \Big(\int_K |h|^2\, dE_{x,x} \Big)^{1/2}.$$

The operator $\Phi_E(h)$ is linear, since $E_{x,y}$ is linear in x, and densely defined.

We know that (a) holds if h is bounded. If it is unbounded, then let $h_k = h\chi_{B_k}$ and observe that $\mathcal{D}(h - h_k) = \mathcal{D}(h)$. By dominated convergence,

$$\|\Phi_E(h)x - \Phi_E(h_k)x\|_H^2 = \|\Phi_E(h - h_k)x\|_H^2 \le \int_K |h - h_k|^2\, dE_{x,x} \to 0$$

as $k \to \infty$; according to Theorem 9.28, every h_k satisfies (a), which will follow for h by letting $k \to \infty$.

To prove (b) when f is bounded, we note that $\mathcal{D}(fh) \subset \mathcal{D}(h)$ and $dE_{x,\bar{\Phi}_E(f)z} = f\, dE_{x,z}$, since both complex measures coincide on every Borel set. It follows that, for every $z \in H$,

$$\begin{aligned} (\Phi_E(f)\Phi_E(h)x, z)_H &= (\Phi_E(h)x, \bar{\Phi}_E(f)z)_H = \int_K h\, dE_{x,\bar{\Phi}_E(f)z} \\ &= (\Phi_E(fh)x, z)_H \end{aligned}$$

and, if $x \in \mathcal{D}(h)$, we obtain from (a) that

$$\int_K |f|^2\, dE_{\Phi_E(h)x, \Phi_E(h)x} = \int_K |fh|^2\, dE_{x,x} \qquad (x \in \mathcal{D}(h)).$$

Hence, $\Phi_E(f)\Phi_E(h) \subset \Phi_E(fh)$.

If f is unbounded, then we take limits and

$$\int_K |f|^2\, dE_{\Phi_E(h)x, \Phi_E(h)x} = \int_K |fh|^2\, dE_{x,x} \qquad (x \in \mathcal{D}(h))$$

holds, so that $\Phi_E(h)x \in \mathcal{D}(f)$ if and only if $x \in \mathcal{D}(fh)$, and

$$\mathcal{D}(\Phi_E(f)\Phi_E(h)) = \{x \in \mathcal{D}(h);\, \Phi_E(h)x \in \mathcal{D}(f)\} = \mathcal{D}(h) \cap \mathcal{D}(fh),$$

as stated in (b).

Now let $x \in \mathcal{D}(h) \cap \mathcal{D}(fh)$ and consider the bounded functions $f_k = f\chi_{B_k}$, so that $f_k h \to fh$ in $L^2(E_{x,x})$. From (a) we know that $\Phi_E(f_k h)x \to \Phi_E(fh)x$,

$$\Phi_E(f)\Phi_E(h)x = \lim_k \Phi_E(f_k)\Phi_E(h)x = \lim_k \Phi_E(f_k h)x = \Phi_E(fh)x,$$

and (b) is true.

To prove (c), let $x, y \in \mathcal{D}(h) = \mathcal{D}(\bar{h})$. If $h_k = h\chi_{B_k}$, then

$$(\Phi_E(h)x, y)_H = \lim_k (\Phi_E(h_k)x, y)_H = \lim_k (x, \Phi_E(\bar{h}_k)y)_H = (x, \Phi_E(\bar{h})y)_H,$$

and it follows that $y \in \mathcal{D}(\Phi_E(h)^*)$ and $\bar{\Phi}_E(h) \subset \Phi_E(h)^*$. To finish the proof, let us show that $\mathcal{D}(\Phi_E(h)^*) \subset \mathcal{D}(h) = \mathcal{D}(\bar{h})$.

Let $z \in \mathcal{D}(\Phi_E(h)^*)$. We apply (b) to $h_k = h\chi_{B_k}$ and we have $\Phi_E(h_k) = \Phi_E(h)\Phi_E(\chi_{B_k})$ with $\Phi_E(\chi_{B_k})$ bounded and self-adjoint. Then

$$\begin{aligned} \Phi_E(\chi_{B_k})\Phi_E(h)^* &= \Phi_E(\chi_{B_k})^*\Phi_E(h)^* \subset (\Phi_E(h)\Phi_E(\chi_{B_k}))^* \\ &= \Phi_E(h_k)^* = \Phi_E(\bar{h}_k) \end{aligned}$$

and $\chi_{B_k}(\Phi_E(h)^*)z = \Phi_E(\bar{h}_k)z$. But $|\chi_k| \le 1$, so that

$$\int_K |h_k|^2 dE_{z,z} = \int_K |\chi_{B_k}|^2 \, dE_{\Phi_E(h)^*z, \Phi_E(h)^*z} \le E_{\Phi_E(h)^*z, \Phi_E(h)^*z}(K).$$

We obtain that $z \in \mathcal{D}(h)$ by letting $k \to \infty$.

The last part follows from (b), since $\mathcal{D}(\Phi_E(h\bar{h})) \subset \mathcal{D}(h)$. □

Remark 9.30. In Theorem 9.29, if $\Phi_E(B_0) = 0$, we can change K to $K \setminus B_0$:

$$(\Phi_E(h)x, y)_H := \int_{K \setminus B_0} h(\lambda) \, dE_{x,y}(\lambda) \qquad (x \in \mathcal{D}(\Phi_E(h)),\ y \in H)$$

if h is Borel measurable on $K \setminus B_0$.

If E is the spectral measure of a bounded normal operator T, then we write $h(T)$ for $\Phi_E(h)$, and then the results of Theorem 9.29 read

$$h(T) = \int_{\sigma(T)} h \, dE$$

on $\mathcal{D}(h) = \{x \in H;\ \|f\|^2_{E_{x,x}} < \infty\}$, in the sense that

$$(h(T)x, y)_H = \int_{\sigma(T)} h \, dE_{x,y} \qquad (x \in \mathcal{D}(h),\ y \in H).$$

Also

(a) $\|h(T)x\|^2_H = \int_{\sigma(T)} |h|^2 \, dE_{x,x}$ if $x \in \mathcal{D}(h(T))$,

(b) $f(T)h(T) \subset (fh)(T)$, $\mathcal{D}(f(T)h(T)) = \mathcal{D}(h(T)) \cap \mathcal{D}((fh)(T))$ with

$$f(T)h(T) = (fh)(T) \text{ if and only if } \mathcal{D}((fh)(T)) \subset \mathcal{D}(h(T)),$$

and

(c) $h(T)^* = \bar{h}(T)$ and $h(T)^*h(T) = |h|^2(T) = h(T)h(T)^*$.

9.5.3. The Cayley transform. We shall obtain a spectral representation theorem for self-adjoint operators using von Neumann's method of making a reduction to the case of unitary operators.

If T is a bounded self-adjoint operator on H, then the continuous functional calculus allows a direct definition of the Cayley transform of T as[14]

$$U = g(T) = (T - iI)(T + iI)^{-1},$$

where $g(t) = (t-i)/(t+i)$, a continuous bijection from $\mathbf{R}$ onto $\mathbf{S} \setminus \{1\}$, and it is a unitary operator (cf. Theorem 8.24).

Let us show that in fact this is also true for unbounded self-adjoint operators.

Let T be a self-adjoint operator on H. By the symmetry of T and from the identity $\|Ty \pm iy\|_H^2 = \|y\|_H^2 + \|Ty\|_H^2 \pm (iy, Ty)_H \pm (Ty, iy)_H$,

$$\|Ty \pm iy\|_H^2 = \|y\|_H^2 + \|Ty\|_H^2 \qquad \big(y \in \mathcal{D}(T)\big).$$

The operators $T \pm iI : \mathcal{D}(T) \to H$ are bijective and with continuous inverses, since $\pm i \in \sigma(T)^c$.

For every $x = Ty + iy \in \operatorname{Im}(T + iI) = H$ $(y \in \mathcal{D}(T))$, we define $Ux = U(Ty + iy) := Ty - iy$; that is,

$$Ux = (T - iI)(T + iI)^{-1}x \qquad (x \in H).$$

Then U is a bijective isometry of H, since $\|Ty + iy\|_H^2 = \|Ty - iy\|_H^2$ and $\operatorname{Im}(T \pm iI) = H$, and U is called the **Cayley transform** of T.

Lemma 9.31. *The Cayley transform*

$$U = (T - iI)(T + iI)^{-1}$$

of a self-adjoint operator T *is unitary,* $I - U$ *is one-to-one,* $\operatorname{Im}(I - U) = \mathcal{D}(T)$*, and*

$$T = i(I + U)(I - U)^{-1}$$

on $\mathcal{D}(T)$.

[14] Named after Arthur Cayley, this transform was originally described by Cayley (1846) as a mapping between skew-symmetric matrices and special orthogonal matrices. In complex analysis, the Cayley transform is the conformal mapping between the upper half-plane and the unit disc given by $g(z) = (z-i)/(z+i)$. It was J. von Neumann who, in 1929, first used it to map self-adjoint operators into unitary operators.

Proof. We have proved that U is unitary and, from the definition, $Ux = (T - iI)y$ if $x = (T + iI)y$ for every $y \in \mathcal{D}(T)$ and every $x \in H$. It follows that $(I + U)x = 2Ty$ and $(I - U)x = 2iy$, with $(I - U)(H) = \mathcal{D}(T)$. If $(I - U)x = 0$, then $y = 0$ and also $(I + U)x = 0$, so that a subtraction gives $2Ux = 0$, and $x = 0$. Finally, if $y \in \mathcal{D}(T)$, $2Ty = (I + U)(I - U)^{-1}(2iy)$. □

Remark 9.32. Since $I - U$ is one-to-one, 1 is not an eigenvalue of U.

9.5.4. Proof of Theorem 9.20: Let T be a self-adjoint operator on H. To construct the (unique) spectral measure E on $\sigma(T) \subset \mathbf{R}$ such that

$$T = \int_{\sigma(T)} t\, dE(t),$$

the Cayley transform U of T will help us to transfer the spectral representation of U to the spectral representation of T.

According to Theorem 8.24, the spectrum of U is a closed subset of the unit circle $\mathbf{S}$, and 1 is not an eigenvalue, so that the spectral measure E' of U satisfies $E'\{1\} = 0$, by Theorem 8.26. We can assume that it is defined on $\Omega = \mathbf{S} \setminus \{1\}$ and we have the functional calculus

$$f(U) = \int_{\sigma(U)} f(\lambda)\, dE'(\lambda) = \int_\Omega f(\lambda)\, dE'(\lambda) \qquad (f \in \mathbf{B}(\Omega)),$$

which was extended to unbounded functions in Subsection 9.5.2.

If $h(\lambda) := i(1 + \lambda)/(1 - \lambda)$ on Ω, then we also have

$$(h(U)x, y)_H = \int_\Omega h(\lambda)\, dE'_{x,y}(\lambda) \qquad (x \in \mathcal{D}(h(U)),\, y \in H),$$

with

$$\mathcal{D}(h(U)) = \{x \in H;\ \int_\Omega |h|^2\, dE'_{x,x} < \infty\}.$$

The operator $h(U)$ is self-adjoint, since h is real and $h(U)^* = \bar{h}(U) = h(U)$.

From the identity

$$h(\lambda)(1 - \lambda) = i(1 + \lambda),$$

an application of (b) in Theorem 9.29 gives

$$h(U)(I - U) = i(I + U),$$

since $\mathcal{D}(I - U) = H$. In particular, $\operatorname{Im}(I - U) \subset \mathcal{D}(h(U))$.

From the properties of the Cayley transform, $T = i(I + U)(I - U)^{-1}$, and then

$$T(I - U) = i(I + U), \quad \mathcal{D}(T) = \operatorname{Im}(I - U) \subset \mathcal{D}(h(U)),$$

so that $h(U)$ is a self-adjoint extension of the self-adjoint operator T. But, T being maximally symmetric, $T = h(U)$. This is,

$$(Tx, y)_H = \int_\Omega h(\lambda)\, dE'_{x,y}(\lambda) \qquad (x \in \mathcal{D}(T),\, y \in H).$$

The function $t = h(\lambda)$ is a homeomorphism between Ω and $\mathbf{R}$ that allows us to define $E(B) := E'(h^{-1}(B))$, and it is readily checked that E is a spectral measure on $\mathbf{R}$ such that

$$(Tx, y)_H = \int_{\mathbf{R}} t\, dE_{x,y}(t) \qquad (x \in \mathcal{D}(T),\, y \in H).$$

Conversely, if E is a spectral measure on $\mathbf{R}$ which satisfies

$$(Tx, y)_H = \int_{\mathbf{R}} t\, dE_{x,y}(t) \qquad (x \in \mathcal{D}(T),\, y \in H),$$

by defining $E'(B) := E(h(B))$, we obtain a spectral measure on Ω such that

$$(h(U)x, y)_H = \int_\Omega h(\lambda)\, dE'_{x,y}(\lambda) \qquad (x \in \mathcal{D}(h(U)),\, y \in H).$$

But $U = h^{-1}(h(U))$ and

$$(Ux, y)_H = \int_\Omega \lambda\, dE'_{x,y}(\lambda) \qquad (x, y \in H).$$

From the uniqueness of E' with this property, the uniqueness of E follows.

Of course, the functional calculus for the spectral measure E defines the functional calculus $f(T) = \int_{-\infty}^{+\infty} f\, dE$ for $T = \int_{-\infty}^{+\infty} \lambda\, dE(\lambda)$, and $f(T) = f(h(U))$.

9.6. Exercises

Exercise 9.1. Let $T : \mathcal{D}(T) \subset H \to H$ be a linear and bounded operator. Prove that T has a unique continuous extension on $\overline{\mathcal{D}(T)}$ and that it has a bounded linear extension to H. Show that this last extension is unique if and only if $\mathcal{D}(T)$ is dense in H.

Exercise 9.2. Prove that if T is a symmetric operator on a Hilbert space H and $\mathcal{D}(T) = H$, then T is bounded.

Exercise 9.3. Prove that the derivative operator D is unbounded on $L^2(\mathbf{R})$.

Exercise 9.4. If T is an unbounded densely defined linear operator on a Hilbert space, then prove that $(\operatorname{Im} T)^\perp = \operatorname{Ker} T^*$.

Exercise 9.5. If T is a linear operator on H and $\lambda \in \sigma(T)^c$, then prove that $\|R_T(\lambda)\| \geq 1/d(\lambda, \sigma(T))$.

Exercise 9.6. Show that, if T is a symmetric operator on H and $\operatorname{Im} T = H$, then T is self-adjoint.

Exercise 9.7. If T is an injective self-adjoint operator on $\mathcal{D}(T) \subset H$, then show that $\operatorname{Im} T = \mathcal{D}(T^{-1})$ is dense in H and that T^{-1} is also self-adjoint.

Exercise 9.8. Prove that the residual spectrum of a self-adjoint operator on a Hilbert space H is empty.

Exercise 9.9. Suppose A is a bounded self-adjoint operator on a Hilbert space H and let

$$A = \int_{\sigma(A)} \lambda \, dE(\lambda)$$

be the spectral representation of A. A vector $z \in H$ is said to be **cyclic** for A if the set $\{A^n z\}_{n=0}^{\infty}$ is total in H.

If A has a cyclic vector z and $\mu = E_{z,z}$, then prove that A is unitarily equivalent to the multiplication operator $M : f(t) \mapsto tf(t)$ of $L^2(\mu)$; that is, $M = U^{-1}AU$ where $U : L^2(\mu) \to H$ is unitary.

Exercise 9.10. Let

$$A = \int_{\sigma(A)} \lambda \, dE(\lambda)$$

be the spectral resolution of a bounded self-adjoint operator of H and denote

$$F(t) := E(-\infty, t] = E(\sigma(A) \cap (-\infty, t]).$$

Prove that the operator-valued function $F : \mathbf{R} \to \mathcal{L}(H)$ satisfies the following properties:

(a) If $s \leq t$, then $F(s) \leq F(t)$; that is, $(F(s)x, x)_H \leq (F(t)x, x)_H$ for every $x \in H$.

(b) $F(t) = 0$ if $t < m(A)$ and $F(t) = I$ if $t \geq M(A)$.

(c) $F(t+) = F(t)$; that is, $\lim_{s \downarrow t} F(s) = F(t)$ in $\mathcal{L}(H)$.

If $a < m(A)$ and $b > M(A)$, then show that with convergence in $\mathcal{L}(H)$

$$A = \int_a^b t \, dF(t) = \int_{m(A)+}^{M(A)} t \, dF(t) = \int_{\mathbf{R}} f \, dF(t)$$

as a Stieltjes integral.

Exercise 9.11. On $L^2(0,1)$, let $S = iD$ with domain $H^1(0,1)$. Prove the following facts:

(a) $\operatorname{Im} S = L^2(\mathbf{R})$.

(b) $S^* = iD$ with domain $H_0^1(0,1)$.

(c) S is a non-symmetric extension of iD with $\mathcal{D}(iD) = H^2(0,1)$.

Exercise 9.12. On $L^2(0,1)$, let $R = iD$ with domain $H_0^1(0,1)$ (i.e, S^* in Exercice 9.11). Prove the following facts:

(a) $\operatorname{Im} R = \{u \in L^2(\mathbf{R});\ \int_0^1 u(t)\,dt = 0\}$.

(b) $R^* = iD$ with domain $H^1(0,1)$ (i.e, $R^* = S$ of Exercice 9.11).

Exercise 9.13. As an application of Theorem 9.17, show that the operator $-D^2 = -d^2/dx^2$ in $L^2(0,1)$ with domain the C^∞ functions f on $[0,1]$ such that $f(0) = f(1) = 0$ is essentially self-adjoint.

Exercise 9.14. Show also that the operator $-D^2 = -d^2/dx^2$ in $L^2(0,1)$ with domain the C^∞ functions f on $[0,1]$ such that $f'(0) = f'(1) = 0$ is essentially self-adjoint.

Exercise 9.15. Prove that $-D^2 = -d^2/dx^2$ with domain $\mathcal{D}(0,1)$ is not an essentially self-adjoint operator in $L^2(0,1)$.

Exercise 9.16. Let V be a nonnegative continuous function on $[0,1]$. Then the differential operator $T = -d^2/dx^2 + V$ on $L^2(0,1)$ with domain $\mathcal{D}^2(0,1)$ has a self-adjoint Friedrichs extension.

Exercise 9.17. Let

$$Q_k\varphi(x) = x_k\varphi(x), \quad P_k\varphi = \frac{h}{2\pi i}\partial_k\varphi \qquad (1 \le k \le n)$$

represent the position and momentum operators on $L^2(\mathbf{R}^n)$.

Prove that they are unbounded self-adjoint operators whose commutators satisfy the relations

$$[Q_j, Q_k] = 0, \quad [P_j, P_k] = 0, \quad [P_j, Q_k] = \delta_{j,k}\frac{h}{2\pi i}I.$$

Note: These are called the **canonical commutation relations** satisfied by the system $\{Q_1, \dots, Q_n; P_1, \dots, P_n\}$ of $2n$ self-adjoint operators, and it is said that Q_k is **canonically conjugate** to P_k.

Exercise 9.18. Prove that, if A_1 and A_2 are two bounded self-adjoint operators in a Hilbert space, then $A_1A_2 = A_2A_1$ if and only if their spectral measures E^1 and E^2 commute as in (9.8): $E^1(B_1)E^2(B_2) = E^2(B_2)E^1(B_1)$ for all $B_1, B_2 \in \mathcal{B}_{\mathbf{R}}$.

Exercise 9.19. Let

$$A_1 = \int_{\mathbf{R}} \lambda\, dE^1(\lambda), \quad A_2 = \int_{\mathbf{R}} \lambda\, dE^1(\lambda)$$

be two self-adjoint operators in a Hilbert space H. If they commute (in the sense that their spectral measures commute), prove that there exists a unique spectral measure E on $\mathbf{R}^2$ such that

$$E(B_1 \times B_2) = E(B_1)E(B_2) \qquad (B_1, B_2 \in \mathcal{B}_{\mathbf{R}}).$$

In the case of the position operators $A_1 = Q_1$ and $A_2 = Q_2$ on $L^2(\mathbf{R}^2)$, show that $E(B) = \chi_B \cdot$ $(B \subset \mathcal{B}_{\mathbf{R}^2})$.

Exercise 9.20. Find the infinitesimal generator of the one-parameter group of unitary operators $U_t f(x) := f(x+t)$ on $L^2(\mathbf{R})$.

Exercise 9.21. Suppose that $g : \mathbf{R} \to \mathbf{R}$ is a continuous function. Describe the multiplication $g\cdot$ as a self-adjoint operator in $L^2(\mathbf{R})$ and $U_t f := e^{itgf}$ as a one-parameter group of unitary operators. Find the infinitesimal generator A of U_t $(t \in \mathbf{R})$.

References for further reading

E. Hille and R. S. Phillips, *Functional Analysis and Semigroups.*

T. Kato, *Perturbation Theory for Linear Operators.*

P. D. Lax, *Functional Analysis.*

E. H. Lieb and M. Loss, *Analysis.*

E. Prugovečki, *Quantum Mechanics in Hilbert Space.*

M. Reed and B. Simon, *Methods of Modern Mathematical Physics.*

F. Riesz and B. Sz. Nagy, *Leçons d'analyse fonctionelle.*

W. Rudin, *Functional Analysis.*

A. E. Taylor and D. C. Lay, *Introduction to Functional Analysis.*

J. von Neumann, *Mathematical Foundations of Quantum Mechanics.*

K. Yosida, *Functional Analysis.*